Lorenz J. Halbeisen

Combinatorial Set Theory

with a gentle introduction to forcing

November 17, 2011

Springer
Berlin Heidelberg NewYork
Hong Kong London
Milan Paris Tokyo
To Joringel,
Meredith, Andrin, and Salome
As the original design of casting Peals of Bells was in order to make pleasant Musick thereon; so the Notes in every Peal are formed apt for that end and purpose, every Peal of Bells being tun'd according to the principles of Musick.

Yet the Notes may be so placed in ringing that their Musick may be rendred much more pleasant: for in Musick there are Conconds, which indeed may be term'd the very life and soul of it.

For this Musical end were changes on Bells first practiceed, changes being nothing else but a moving and placing of the Notes in ringing; wherein 'tis to be observed as a general Rule, That every change must be made betwixt two notes that strike next to each other.

FABIAN STEDMAN
Campanologia, 1677

By the campanologist, the playing of tunes is considered to be a childish game; the proper use of bells is to work out mathematical permutations and combinations. His passion finds its satisfaction in mathematical completeness and mechanical perfection.

DOROTHY L. SAYERS
The Nine Tailors, 1934
Preface

This book provides a self-contained introduction to *Axiomatic Set Theory* with main focus on *Infinitary Combinatorics* and the *Forcing Technique*. The book is intended to be used as a textbook in undergraduate and graduate courses of various levels, as well as for self-study. To make the book valuable for experienced researchers also, some historical background and the sources of the main results have been provided in the *Notes*, and some topics for further studies are given in the section *Related Results*—where those containing open problems are marked with an asterisk.

The axioms of Set Theory *ZFC*, consisting of the axioms of *Zermelo-Fraenkel Set Theory* (denoted *ZF*) and the *Axiom of Choice*, are the foundation of Mathematics in the sense that essentially all Mathematics can be formalised within ZFC. On the other hand, Set Theory can also be considered as a mathematical theory, like Group Theory, rather than the basis for building general mathematical theories. This approach allows us to drop or modify axioms of ZFC in order to get, for example, a Set Theory without the *Axiom of Choice* (see Chapter 4) or in which just a weak form of the *Axiom of Choice* holds (see Chapter 7). In addition, we are also allowed to extend the axiomatic system ZFC in order to get, for example, a Set Theory in which, in addition to the ZFC axioms, we also have *Martin's Axiom* (see Chapter 13), which is a very powerful axiom with many applications for *Infinitary Combinatorics* as well as other fields of Mathematics. However, this approach prevents us from using any kind of Set Theory which goes beyond *ZFC*, which is used, for example, to prove the existence of a countable model of *ZFC* (see the *Łoś-Celome-Skolem Theorem* in Chapter 15).

Most of the results presented in this book are combinatorial results, in particular the results in *Ramsey Theory* (introduced in Chapter 2 and further developed in Chapter 11), or those results whose proofs have a combinatorial flavour. For example, we get results of the latter type if we work in Set Theory without the *Axiom of Choice*, since in the absence of the *Axiom of Choice*, the proofs must be constructive and therefore typically have a much more combinatorial flavour than proofs in *ZFC* (examples can be found in Chap-
ters 4 & 7). On the other hand, there are also elegant combinatorial proofs using the Axiom of Choice. An example is the proof in Chapter 6, where it is shown that one can divide the solid unit ball into five parts, such that one can build two solid unit balls out of these five parts — another such paradoxical result is given in Chapter 17, where it is shown that it might be possible in ZF to decompose a square into more parts than there are points on the square.

Even though the ZFC axiomatic system is the foundation of Mathematics, by Gödel’s Incompleteness Theorem — briefly discussed at the end of Chapter 3 — no axiomatic system of Mathematics is complete in the sense that every statement can either be proved or disproved; in other words, there are always statements which are independent of the axiomatic system. The main tool to show that a certain statement is independent of the axioms of Set Theory is Cohen’s Forcing Technique, which he originally developed in the early 1960s in order to show that there are models of ZF in which the Axiom of Choice fails (see Chapter 17) and that the Continuum Hypothesis is independent of ZFC (see Chapter 14). The Forcing Technique is introduced and discussed in great detail in Part II, and in Part III it is used to investigate combinatorial properties of the set of real numbers. This is done by comparing the Cardinal Characteristics of the Continuum introduced in Chapter 8.

The following table indicates which of the main topics appear in which chapter, where *** means that it is the main topic of that chapter, ** means that some new results in that topic are proved or at least that the topic is important for understanding certain proofs, and * means that the topic appears somewhere in that chapter, but not in an essential way:

<table>
<thead>
<tr>
<th>Chapter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forcing Technique</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axiom of Choice &amp; ZF</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramsey Theory</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardinal Characteristics</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For example, Ramsey’s Theorem, which is the nucleus of Ramsey Theory, is the main topic in Chapter 2, it is used in some proofs in Chapters 4 & 7, it is used as a choice principle in Chapter 5, it is related to two Cardinal Characteristics defined in Chapter 8, it is used to define what is called a Ramsey ultrafilter in Chapter 10, it is used in the proof of the Hales-Jewett Theorem in Chapter 11, and it is used to formulate a combinatorial feature of Mathias reals in
Preface

Chapter 24. Furthermore, one can see that Cardinal Characteristics are our main tool in Part III in the investigation of combinatorial properties of various forcing notions, even in the cases when — in Chapters 25 & 26 — the existence of Ramsey ultrafilters are investigated. Finally, in Chapter 27 we show how Cardinal Characteristics can be used to shed new light on a classical problem in Measure Theory. On the other hand, the Cardinal Characteristics are used to describe some combinatorial features of different forcing notions. In particular, it will be shown that the cardinal characteristic $\mathfrak{h}$ (introduced in Chapter 8 and investigated in Chapter 9) is closely related to Mathias forcing (introduced in Chapter 24), which is used in Chapter 25 to show that the existence of Ramsey ultrafilters is independent of ZFC.

I tried to write this book like a piece of music, not just writing note by note, but using various themes or voices — like Ramsey's Theorem and the cardinal characteristic $\mathfrak{h}$ — again and again in different combinations. In this undertaking, I was inspired by the English art of bell ringing and tried to base the order of the themes on Zarlino's introduction to the art of counterpoint.

Acknowledgement. First of all, I would like to thank Andreas Blass for his valuable remarks and comments, as well as for his numerous corrections, which improved the quality of the book substantially. Furthermore, I would like to thank my spouse Stephanie Halbeisen, not only for reading Chapters 1 & 12 and parts of Chapters 5 & 13, but also for her patience during the last seven years. I would also like to thank Dandolo Flumini for reading Chapters 2, 3, 13, 14, 15, Ioanna Dimitriou for reading Chapters 16 & 17, and Geróidín Diserens for reading Chapter 1 as well as the introductory comments of several chapters. Finally, I would like to thank Jörg Sixt, editor of Springer-Verlag, for making every effort to ensure that the book was published in the optimal style.

Winterthur, October 2011

Lorenz Halbeisen
# Contents

1 The Setting .................................................. 1

## Part I Topics in Combinatorial Set Theory

2 **Overture: Ramsey’s Theorem** ................................ 11
   The nucleus of Ramsey Theory .................................. 11
   Corollaries of Ramsey’s Theorem .............................. 14
   Generalisations of Ramsey’s Theorem ........................ 16
   Notes & Related Results ...................................... 21
   References ................................................... 25

3 **The Axioms of Zermelo-Fraenkel Set Theory** ........... 29
   Why axioms? .................................................. 29
   First-order logic in a nutshell ................................ 31
      Syntax: formulae, formal proofs, and consistency ....... 32
      Semantics: models, completeness, and independence .... 39
      Limits of first-order logic ................................ 42
   The axioms of Zermelo-Fraenkel set theory ................ 44
      Empty Set, Extensionality, Pairing, Union, Infinity ...... 45
      Separation, Power Set, Replacement, Foundation ......... 48
   Models of ZF ................................................ 58
   Cardinals in ZF .............................................. 60
   On the consistency of ZF ................................... 63
   Notes & Related Results ...................................... 64
   References ................................................... 71

4 **Cardinal Relations in ZF only** ................................ 79
   Basic cardinal relations ...................................... 80
   On the cardinals \(2^{\aleph_0}\) and \(\aleph_1\) .................. 85
   Ordinal numbers revisited .................................... 87
More cardinal relations ........................................... 92
fin(m) < 2^m ......................................................... 92
seq^{-1}(m) \neq 2^m \neq seq(m) .................................... 95
2^{2^m} + 2^{2^m} = 2^{2^m} ....................................... 100
Notes & Related Results ......................................... 104
References .......................................................... 108

5 The Axiom of Choice ............................................. 111
Zermelo's Axiom of Choice and its consistency with ZF .......... 111
Equivalent forms of the Axiom of Choice ......................... 113
Cardinal arithmetic in the presence of AC ....................... 122
Some weaker forms of the Axiom of Choice ..................... 127
The Prime Ideal Theorem and related statements ............... 127
König's Lemma and other choice principles .................... 134
Notes & Related Results ......................................... 137
References .......................................................... 147

6 How to Make Two Balls from One .......................... 153
Equidecomposability .............................................. 153
Hausdorff's Paradox .............................................. 154
Robinson's decomposition ....................................... 157
Notes & Related Results ......................................... 164
References .......................................................... 165

7 Models of Set Theory with Atoms ______________________ 167
Permutation models .............................................. 168
The basic Fraenkel model ....................................... 171
The second Fraenkel model .................................... 172
The ordered Mostowski model ................................. 174
The Prime Ideal Theorem revisited ............................ 177
Custom-built permutation models ............................... 180
Notes & Related Results ......................................... 184
References .......................................................... 186

8 Twelve Cardinals and their Relations ...................... 189
The cardinals \(\omega_1\) and \(\mathfrak{c}\) ............................. 190
The cardinal \(p\) .................................................. 190
The cardinals \(b\) and \(\mathfrak{d}\) ...................................... 191
The cardinals \(s\) and \(\mathfrak{r}\) ....................................... 192
The cardinals \(a\) and \(i\) .......................................... 194
The cardinals \(\mathfrak{p}_{\mathfrak{m}}\) and \(\mathfrak{h}_{\mathfrak{m}}\) ................. 198
The cardinal \(\mathfrak{b}\) ............................................... 200
Summary ............................................................ 203
Notes & Related Results ......................................... 204
References .......................................................... 207
9 The Shattering Number revisited ........................................ 211
The Ramsey property ....................................................... 211
The ideal of Ramsey-null sets .............................................. 213
The Ellentuck topology ................................................... 214
A generalised Suslin operation .......................................... 219
Notes & Related Results .................................................. 221
References ......................................................................... 222

10 Happy Families and their Relatives ................................. 225
Happy families ................................................................. 225
Ramsey ultrafilters ........................................................... 229
\(P\)-points and \(Q\)-points ............................................. 232
Ramsey families and \(P\)-families ...................................... 236
Notes & Related Results .................................................. 241
References ......................................................................... 243

11 Coda: A Dual Form of Ramsey’s Theorem ..................... 245
The Hales-Jewett Theorem ............................................... 245
Families of partitions ....................................................... 250
Carlson’s Lemma and the Partition Ramsey Theorem ........... 252
A weak form of the Halpern-Läuchli Theorem .................... 259
Notes & Related Results .................................................. 261
References ......................................................................... 264

Part II From Martin’s Axiom to Cohen’s Forcing

12 The Idea of Forcing ........................................................ 271

13 Martin’s Axiom ............................................................. 275
Filters on partially ordered sets ......................................... 275
Weaker forms of MA ......................................................... 279
Some consequences of MA(\(\beta\)-centred) .......................... 279
MA(\(\text{countable}\)) implies the existence of Ramsey ultrafilters . 281
Notes & Related Results .................................................. 283
References ......................................................................... 284

14 The Notion of Forcing ..................................................... 285
The language of forcing ..................................................... 285
Generic extensions ........................................................... 290
ZFC in generic models ...................................................... 292
Independence of \(\text{CH}:\) the gentle way ............................... 302
On the existence of generic filters .................................... 304
Notes ............................................................................... 305
References ......................................................................... 305
15 Models of finite fragments of Set Theory .......................... 307
   Basic model-theoretical facts .................................. 307
   The Reflection Principle ....................................... 308
   Countable transitive models of finite fragments of ZFC .... 312
   Notes & Related Results ....................................... 314
   References .................................................... 315

16 Proving Unprovability ............................................ 317
   Consistency and independence proofs: the proper way ....... 317
   The cardinality of the continuum ............................. 321
   Notes & Related Results ....................................... 322
   References .................................................... 322

17 Models in which AC fails ........................................ 325
   Symmetric submodels of generic extensions .................. 325
   Examples of symmetric models ................................ 328
      A model in which the reals cannot be well-ordered ....... 328
      A model in which every ultrafilter over \( \omega \) is principal . 330
      A model with a paradoxical decomposition of the real line .... 331
   Simulating permutation models by symmetric models ....... 334
   Notes & Related Results ....................................... 339
   References .................................................... 340

18 Combining Forcing Notions ...................................... 343
   Products ..................................................... 344
      General products of forcing notions ......................... 344
      Products of Cohen forcing ................................ 345
      A model in which \( a < \kappa \) ................................ 347
   Iterations ................................................... 349
      Two-step iterations ........................................ 349
      General iterations ....................................... 354
      A model in which \( i < \kappa \) ................................ 357
   Notes & Related Results ....................................... 360
   References .................................................... 361

19 Models in which \( p = c \) ..................................... 363
   A model in which \( p = c = \omega_2 \) .......................... 363
   On the consistency of MA + ¬CH .............................. 365
   \( p = c \) is preserved under adding a Cohen real .......... 366
   Notes & Related Results ....................................... 370
   References .................................................... 370
### Part III Combinatorics of Forcing Extensions

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>Properties of Forcing Extensions</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td>Dominating, splitting, bounded, and unbounded reals</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td>The Laver property and not adding Cohen reals</td>
<td>379</td>
</tr>
<tr>
<td></td>
<td>Proper forcing notions and preservation theorems</td>
<td>380</td>
</tr>
<tr>
<td></td>
<td>Notes &amp; Related Results</td>
<td>384</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>384</td>
</tr>
<tr>
<td>21</td>
<td>Cohen Forcing revisited</td>
<td>387</td>
</tr>
<tr>
<td></td>
<td>Properties of Cohen forcing:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cohen forcing adds unbounded but no dominating reals</td>
<td>387</td>
</tr>
<tr>
<td></td>
<td>Cohen forcing adds splitting reals</td>
<td>388</td>
</tr>
<tr>
<td></td>
<td>Cohen reals and the covering number of meagre sets</td>
<td>388</td>
</tr>
<tr>
<td></td>
<td>A model in which $a &lt; d = \text{cov}({\mathcal{M}})$</td>
<td>393</td>
</tr>
<tr>
<td></td>
<td>A model in which $s = b &lt; d$</td>
<td>394</td>
</tr>
<tr>
<td></td>
<td>Notes &amp; Related Results</td>
<td>395</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>397</td>
</tr>
<tr>
<td>22</td>
<td>Silver-Like Forcing Notions</td>
<td>399</td>
</tr>
<tr>
<td></td>
<td>Properties of Silver-like forcing:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silver-like forcing is proper and $\omega$-bounding</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>Silver-like forcing adds splitting reals</td>
<td>401</td>
</tr>
<tr>
<td></td>
<td>A model in which $\mathfrak{d} &lt; \mathfrak{r}$</td>
<td>401</td>
</tr>
<tr>
<td></td>
<td>Notes &amp; Related Results</td>
<td>402</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>403</td>
</tr>
<tr>
<td>23</td>
<td>Miller Forcing</td>
<td>405</td>
</tr>
<tr>
<td></td>
<td>Properties of Miller forcing:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Miller forcing is proper and adds unbounded reals</td>
<td>406</td>
</tr>
<tr>
<td></td>
<td>Miller forcing does not add splitting reals</td>
<td>407</td>
</tr>
<tr>
<td></td>
<td>Miller forcing preserves $P$-points</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td>A model in which $\mathfrak{r} &lt; \mathfrak{d}$</td>
<td>413</td>
</tr>
<tr>
<td></td>
<td>Notes &amp; Related Results</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>416</td>
</tr>
<tr>
<td>24</td>
<td>Mathias Forcing</td>
<td>417</td>
</tr>
<tr>
<td></td>
<td>Properties of Mathias forcing:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mathias forcing adds dominating reals</td>
<td>417</td>
</tr>
<tr>
<td></td>
<td>Mathias forcing is proper and has the Laver property</td>
<td>418</td>
</tr>
<tr>
<td></td>
<td>A model in which $p &lt; h$</td>
<td>422</td>
</tr>
<tr>
<td></td>
<td>Notes &amp; Related Results</td>
<td>424</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>425</td>
</tr>
</tbody>
</table>
Contents

25 On the Existence of Ramsey Ultrafilters .......................... 427
There may be a Ramsey ultrafilter and $\text{cov}(\mathcal{M}) < \varepsilon$ ....... 427
There may be no Ramsey ultrafilter and $\mathfrak{b} = \varepsilon$ ........... 429
Notes & Related Results .............................................. 438
References ............................................................. 439

26 Combinatorial Properties of Sets of Partitions ................. 441
A dual form of Mathias forcing ..................................... 441
A dual form of Ramsey ultrafilters ................................ 448
Notes & Related Results .............................................. 450
References ............................................................. 452

27 Suite ............................................................... 455
Prelude ............................................................... 455
Allemande ............................................................. 456
Courante ............................................................... 457
Sarabande ............................................................. 458
Gavotte I & II .......................................................... 459
Gigue ................................................................. 460
References ............................................................. 461

Index ................................................................. 463
Symbols ............................................................... 463
Names ................................................................. 466
Subjects ............................................................. 470
The Setting

For one cannot order or compose anything, or understand the nature of the composite, unless one knows first the things that must be ordered or combined, their nature, and their cause.

Gioseffo Zarlino
Le Institution Harmoniche, 1558

Combinatorics with all its various aspects is a broad field of Mathematics which has many applications in areas like Topology, Group Theory and even Analysis. A reason for its wide range of applications might be that Combinatorics is rather a way of thinking than a homogeneous theory, and consequently Combinatorics is quite difficult to define. Nevertheless, let us start with a definition of Combinatorics which will be suitable for our purpose:

Combinatorics is the branch of Mathematics which studies collections of objects that satisfy certain criteria, and is in particular concerned with deciding how large or how small such collections might be.

Below we give a few examples which should illustrate some aspects of infinitary Combinatorics. At the same time, we present the main topics of this book, which are the Axiom of Choice, Ramsey Theory, cardinal characteristics of the continuum, and forcing.

Let us start with an example from Graph Theory: A graph is a set of vertices, where some pairs of vertices are connected by an edge. Connected pairs of vertices are called neighbours. A graph is infinite if it has an infinite number of vertices. A tree is a cycle-free (i.e., one cannot walk in proper cycles along edges), connected (i.e., any two vertices are connected by a path of edges) graph, where one of its vertices is designated as the root. A tree
is finitely branching if every vertex has only a finite number of neighbours. Furthermore, a branch through a tree is a maximal edge-path beginning at the root, in which no edge appears twice.

Now we are ready to state König's Lemma, which is often used implicitly in fields like Combinatorics, Topology, and many other branches of Mathematics.

**König's Lemma:** Every infinite, finitely branching tree contains an infinite branch.

At first glance, this result looks straightforward and one would construct an infinite branch as follows: Let $v_0$ be the root. Since the tree is infinite but finitely branching, there must be a neighbour of $v_0$ from which we reach infinitely many vertices without going back to $v_0$. Let $v_1$ be such a neighbour of $v_0$. Again, since we reach infinitely many vertices from $v_1$ (without going back to $v_1$) and the tree is finitely branching, there must be a neighbour of $v_1$, say $v_2$, from which we reach infinitely many vertices without going back to $v_2$. Proceeding in this way, we finally get the infinite branch $(v_0, v_1, v_2, \ldots)$.

Let us now have a closer look at this proof. Firstly, in order to prove that the set of neighbours of $v_0$ from which we reach infinitely many vertices without going back to $v_0$ is not empty, we need an infinite version of the so-called Pigeon-Hole Principle. The Pigeon-Hole Principle can be seen as the fundamental principle of Combinatorics.

**Pigeon-Hole Principle:** If $n + 1$ pigeons roost in $n$ holes, then at least two pigeons must share a hole. More prosaically: If $m$ objects are coloured with $n$ colours and $m > n$, then at least two objects have the same colour.

An infinite version of the Pigeon-Hole Principle reads as follows:

**Infinite Pigeon-Hole Principle:** If infinitely many objects are coloured with finitely many colours, then infinitely many objects have the same colour.

Using the Infinite Pigeon-Hole Principle we are now sure that the set of neighbours of $v_0$ from which we reach infinitely many vertices without going back to $v_0$ is not empty. However, the next problem we face is which element we should choose from that non-empty set. If the vertices are ordered in some way, then we can choose the first element with respect to that order, but otherwise, we would need some kind of choice function which selects infinitely often (and this is the crucial point!) one vertex from a given non-empty set of vertices. Such a choice function is guaranteed by the Axiom of Choice, denoted AC, which is discussed in Chapter 5.
Axiom of Choice: For every family $\mathcal{F}$ of non-empty sets, there is a function $f$ — called a choice function — which selects one element from each member of $\mathcal{F}$ (i.e., for each $x \in \mathcal{F}$, $f(x) \in x$); or equivalently, every Cartesian product of non-empty sets is non-empty.

The Axiom of Choice is one of the main topics of this book: In Chapter 3, the axioms of Zermelo-Fraenkel Set Theory (i.e., the usual axioms of Set Theory except AC) are introduced. In Chapter 4 we shall introduce the reader to Zermelo-Fraenkel Set Theory and show how combinatorics can, to some extent, replace the Axiom of Choice. Subsequently, the Axiom of Choice (and some of its weaker forms) is introduced in Chapter 5. From then on, we always work in Zermelo-Fraenkel Set Theory with the Axiom of Choice — even in the case as in Chapters 7 & 17 when we construct models of Set Theory in which AC fails.

Now, let us turn back to König’s Lemma. In order to prove König’s Lemma we do not need full AC, since it would be enough if every family of non-empty finite sets had a choice function — the family would consist of all subsets of neighbours of vertices. However, as we will see later, even this weaker form of AC is a proper axiom and is independent of the other axioms of Set Theory (cf. Proposition 7.7). Thus, depending on the axioms of Set Theory we start with, AC — as well as some weakened forms of it — may fail, and consequently, König’s Lemma may become unprovable. On the other hand, as we will see in Chapter 5, König’s Lemma may be used as a non-trivial choice principle.

Thus, this first example shows that — with respect to our definition of Combinatorics given above — some “objects satisfying certain criteria,” may, but need not, exist.

The next example can be seen as a problem in infinitary Extremal Combinatorics. The word “extremal” describes the nature of problems dealt with in this field and refers to the second part of our definition of Combinatorics, namely “how large or how small collections satisfying certain criteria might be.”

If the objects considered are infinite, then the answer, how large or how small certain sets are, depends again on the underlying axioms of Set Theory, as the next example shows.

Reaping Families: A family $\mathcal{R}$ of infinite subsets of the natural numbers $\mathbb{N}$ is said to be reaping if for every colouring of $\mathbb{N}$ with two colours there exists a monochromatic set in the family $\mathcal{R}$.

For example, the set of all infinite subsets of $\mathbb{N}$ is such a family. The reaping number $\tau$ — a so-called cardinal characteristic of the continuum — is the smallest cardinality (i.e., size) of a reaping family. In general, a cardinal characteristic of the continuum is typically defined as the smallest cardinality of a subset of a given set $S$ which has certain combinatorial properties, where $S$ is of the same cardinality as the continuum $\mathbb{R}$. 
Consider the cardinal characteristic \( \tau \) (i.e., the size of the smallest reaping family). Since \( \tau \) is a well-defined cardinality we can ask: How large is \( \tau \)? Can it be countable? Is it always equal to the cardinality of the continuum?

Let us just show that a reaping family can never be countable: Let \( \mathcal{A} = \{ A_i : i \in \mathbb{N} \} \) be any countable family of infinite subsets of \( \mathbb{N} \). For each \( i \in \mathbb{N} \), pick \( n_i \) and \( m_i \) from the set \( A_i \) in such a way that, at the end, for all \( i \) we have \( n_i < m_i < n_{i+1} \). Now we colour all \( n_i \)'s blue and all the other numbers red. For this colouring, there is no monochromatic set in \( \mathcal{A} \), and hence, \( \mathcal{A} \) cannot be a reaping family. The Continuum Hypothesis, denoted \( \text{CH} \), states that every subset of the continuum \( \mathbb{R} \) is either countable or of cardinality \( \mathfrak{c} \), where \( \mathfrak{c} \) denotes the cardinality of \( \mathbb{R} \). Thus, if we assume \( \text{CH} \), then any reaping family is of cardinality \( \mathfrak{c} \). The same holds if we assume Martin’s Axiom which will be introduced in Chapter 13.

On the other hand, with the forcing technique— invented by Paul Cohen in the early 1960s — one can show that the axioms of Set Theory do not decide whether or not the cardinals \( \tau \) and \( \mathfrak{c} \) are equal. The forcing technique is introduced in Part II and a model in which \( \tau < \mathfrak{c} \) is given in Chapter 18.

Thus, the second example shows that — depending on the additional axioms of Set Theory we start with — we can get different answers when we try to "decide how large or how small certain collections might be."

Many more cardinal characteristics like \( \text{hom} \) and \( \text{par} \) (see below) are introduced in Chapter 8. Possible (i.e., consistent) relations between these cardinals are investigated in Part II and more systematically in Part III where the cardinal characteristics are also used to distinguish the combinatorial features of certain forcing notions.

Another field of Combinatorics is the so-called Ramsey Theory, and since many results in this work rely on Ramsey-type theorems, let us give a brief description of Ramsey Theory.

Loosely speaking, Ramsey Theory (which can be seen as a part of extremal Combinatorics) is the branch of Combinatorics which deals with structures preserved under partitions, or colourings. Typically, one looks at the following kind of question: If a particular object (e.g., algebraic, geometric or combinatorial) is arbitrarily coloured with finitely many colours, what kinds of monochromatic structures can we find?

For example, van der Waerden’s Theorem, which will be proved in Chapter 11, tells us that for any positive integers \( r \) and \( n \), there is a positive integer \( N \) such that for every \( r \)-colouring of the set \( \{0,1,\ldots,N\} \) we find always a monochromatic (non-constant) arithmetic progression of length \( n \).

Even though van der Waerden’s Theorem is one of the earliest results in Ramsey Theory, the most famous result in Ramsey Theory is surely Ramsey’s Theorem (which will be discussed in detail in the next chapter):
**Ramsey's Theorem:** Let $n$ be any positive integer. If we colour all $n$-element subsets of $\mathbb{N}$ with finitely many colours, then there exists an infinite subset of $\mathbb{N}$ all of whose $n$-element subsets have the same colour.

There is also a finite version of Ramsey's Theorem which gives an answer to problems like the following:

*How many people must be invited to a party in order to make sure that three of them mutually shook hands on a previous occasion or three of them mutually did not shake hands on a previous occasion?*

It is quite easy to show that at least six people must be invited. On the other hand, if we ask how many people must get invited such that there are five people who all mutually shook hands or did not shake hands on a previous occasion, then the precise number is not known — but it is conjectured that it is sufficient to invite 43 people.

As we shall see later, Ramsey's Theorem has many — sometimes unexpected — applications. For example, if we work in Set Theory without $\mathrm{AC}$, then Ramsey's Theorem can help to construct a choice function, as we will see in Chapter 4. Sometimes we get Ramsey-type (or anti-Ramsey-type) results even for partitions into infinitely many classes (i.e., using infinitely many colours). For example, one can show that there is a colouring of the points in the Euclidean plane with countably many colours, such that no two points of any "copy of the rationals" have the same colour. This result can be seen as an anti-Ramsey-type theorem (since we are far away from "monochromatic structures"), and it shows that Ramsey-type theorems cannot be generalised arbitrarily. However, concerning Ramsey's Theorem, we can ask for a "nice" family $\mathcal{F}$ of infinite subsets of $\mathbb{N}$, such that for every colouring of the $n$-element subsets of $\mathbb{N}$ with finitely many colours, there exists a homogeneous set in the family $\mathcal{F}$, where an infinite set $x \subseteq \mathbb{N}$ is called homogeneous if all $n$-element subsets of $x$ have the same colour. Now, "nice" could mean "as small as possible" but also "being an ultrafilter." In the former case, this leads to the *homogeneous number* $\dot{\text{hom}}$, which is the smallest cardinality of a family $\mathcal{F}$ which contains a homogeneous set for every 2-colouring of the 2-element subsets of $\mathbb{N}$. One can show that $\dot{\text{hom}}$ is uncountable and — like for the reaping number — that the axioms of Set Theory do not decide whether or not $\dot{\text{hom}}$ is equal to $\aleph$. The latter case, where "nice" means "being an ultrafilter," leads to so-called Ramsey *ultrafilters*. It is not difficult to show that Ramsey ultrafilters exist if one assumes $\mathrm{CH}$ or Martin's Axiom (see Chapter 10), but on the other hand, the axioms of Set Theory alone do not imply the existence of Ramsey ultrafilters (see Proposition 25.11). A somewhat anti-Ramsey-type question would be to ask how many 2-colourings of the 2-element subsets of $\mathbb{N}$ we need to make sure that no single infinite subset of $\mathbb{N}$ is almost homogeneous for all these colourings, where a set $H$ is called almost homogeneous if there is a finite set $K$ such that $H \setminus K$ is homogeneous. This question leads to the *partition number* $\text{par}$. Again, $\text{par}$ is uncountable and the
axioms of Set Theory do not decide whether or not $\pi$ is equal to $\frac{1}{\sqrt{6}}$ (see for example Chapter 18).

**Ramsey's Theorem**, as well as Ramsey Theory in general, play an important role throughout this book. Especially in all chapters of Part I, except for Chapter 3, we shall meet — sometimes unexpectedly — **Ramsey’s Theorem** in one form or other.

**Notes**

**Gioseffo Zarlino.** All citations of Zarlino (1517–1590) are taken from Part III of his book entitled *Le Institutions Harmoniche* (cf. [1]). This section of Zarlino's *Institutions* is concerned primarily with the art of counterpoint, which is, according to Zarlino, the concordance or agreement born of a body with diverse parts, its various melodic lines accommodated to the total composition, arranged so that voices are separated by commensurable, harmonious intervals. The word "counterpoint" presumably originated at the beginning of the 14th century and was derived from "punctus contra punctum," i.e., point against point or note against note. Zarlino himself was an Italian music theorist and composer. While he composed a number of masses, motets and madrigals, his principal claim to fame is as a music theorist. For example, Zarlino was ahead of his time in proposing that the octave should be divided into twelve equal semitones — for the lute, that is to say; he advocated a practice in the 16th century which was universally adopted three centuries later. He also advocated equal temperament for keyboard instruments and just intonation for unaccompanied vocal music and strings — a system which has been successfully practiced up to the present day. Furthermore, Zarlino arranged the modes in a different order of succession, beginning with the Ionian mode instead of the Dorian mode. This arrangement seems almost to have been dictated by a prophetic anticipation of the change which was to lead to the abandonment of the modes in favor of a newer tonality, for his series begins with a form which corresponds exactly with our modern major mode and ends with the prototype of the descending minor scale of modern music. (For the terminology of music theory we refer the interested reader to Benson [2].)

Zarlino's most notable student was the music theorist and composer Vincenzo Galilei, the father of Galileo Galilei.

**König’s Lemma and Ramsey’s Theorem.** A proof of König’s Lemma can be found in König’s book on Graph Theory [3, VI, §2, Satz 6], where he called the result *Unendlichkeitsschema*. As a first application of the *Unendlichkeitsschema* he proved the following theorem of de la Vallée Poussin: If $E$ is a subset of the open unit interval $(0,1)$ which is closed in $\mathbb{R}$ and $I$ is a set of open intervals covering $E$, then there is a natural number $n$, such that if one partitions $(0,1)$ into $2^n$ intervals of length $2^{-n}$, each of these intervals containing a point of $E$ is contained in an interval of $I$. Using the *Unendlichkeitsschema*, König also showed that van der Waerden's Theorem is equivalent to the following statement: If the positive integers are finitely coloured, then there are arbitrarily long monochromatic arithmetic progressions. In a similar way we will use König’s Lemma to derive the Finite Ramsey Theorem from Ramsey’s Theorem (cf. Corollary 23).
References

At first glance, König’s Lemma and Ramsey’s Theorem seem to be quite unrelated statements. In fact, König’s Lemma is a proper (but rather weak) choice principle, whereas Ramsey’s Theorem is a very powerful combinatorial tool. However, as we shall see in Chapter 5, Ramsey’s Theorem can also be considered as a proper choice principle which turns out to be even stronger than König’s Lemma (see Theorem 5.17).

References

Part I

Topics in Combinatorial Set Theory
Overture: Ramsey’s Theorem

*Musicians in the past, as well as the best of the moderns, believed that a counterpoint or other musical composition should begin on a perfect consonance, that is, a unison, fifth, octave, or compound of one of these.*

Giuseppe Zarlino

*Le Istitutioni Harmoniche, 1558*

The Nucleus of Ramsey Theory

Most of this text is concerned with sets of subsets of the natural numbers, so, let us start there. The set \{0, 1, 2, \ldots\} of natural numbers (or of non-negative integers) is denoted by \( \omega \). It is convenient to consider a natural number \( n \) as an \( n \)-element subset of \( \omega \), namely as the set of all numbers smaller than \( n \), so, \( n = \{ k \in \omega : k < n \} \). In particular, \( 0 = \emptyset \), where \( \emptyset \) is the empty set. For any \( n \in \omega \) and any set \( S \), let \([S]^n\) denote the set of all \( n \)-element subsets of \( S \) (e.g., \([S]^0 = \{ \emptyset \})\). Further, the set of all finite subsets of a set \( S \) is denoted by \([S]^{< \omega}\).

For a finite set \( S \) let \(|S|\) denote the number of elements in \( S \), also called the cardinality of \( S \).

A set \( S \) is called countable if there is an enumeration of \( S \), i.e., if \( S = \emptyset \) or \( S = \{ x_i : i \in \omega \} \). In particular, every finite set is countable. However, when we say that a set is countable we usually mean that it is a countably infinite set. For any set \( S \), \([S]^{\omega}\) denotes the set of all countably infinite subsets of \( S \), in particular, since every infinite subset of \( \omega \) is countable, \([\omega]^{\omega}\) is the set of all infinite subsets of \( \omega \).

Let \( S \) be an arbitrary non-empty set. A binary relation “\( \sim \)” on \( S \) is an equivalence relation if it is

- reflexive (i.e., for all \( x \in S : x \sim x \)),

\( S \in \{\emptyset\} \subset \omega \).
• symmetric (i.e., for all $x, y \in S$: $x \sim y \leftrightarrow y \sim x$), and
• transitive (i.e., for all $x, y, z \in S$: $x \sim y \land y \sim z \rightarrow x \sim z$).

The equivalence class of an element $x \in S$, denoted $[x]$, is the set $\{ y \in S : x \sim y \}$. We would like to recall the fact that, since $\sim$ is an equivalence relation, for any $x, y \in S$ we have either $[x] = [y]$ or $[x] \cap [y] = \emptyset$. A set $A \subseteq S$ is a set of representatives if for each equivalence class $[x]$ we have $|A \cap [x]| = 1$; in other words, $A$ has exactly one element in common with each equivalence class. It is worth mentioning that in general, the existence of a set of representatives relies on the Axiom of Choice (see Chapter 5).

For sets $A$ and $B$, let $\mathcal{A}B$ denote the set of all functions $f : A \to B$. For $f \in \mathcal{A}B$ and $S \subseteq A$ let $f[S] := \{ f(x) : x \in S \}$ and let $f|_S \in \mathcal{S}B$ (the restriction of $f$ to $S$) be such that for all $x \in S$, $f(x) = f|_S(x)$.

Further, for sets $A$ and $B$, let the set-theoretic difference of $A$ and $B$ be the set $A \setminus B := \{ a \in A : a \notin B \}$.

For some positive $n \in \omega$, let us colour all $n$-element subsets of $\omega$ with three colours, say red, blue, and yellow. In other words, each $n$-element set of natural numbers $\{ k_1, \ldots, k_n \}$ is coloured either red, or blue, or yellow. Now one can ask whether there is an infinite subset $H$ of $\omega$ such that all its $n$-element subsets have the same colour (i.e., $[H]^n$ is monochromatic). Such a set we would call homogeneous (for the given colouring). In the terminology above, this question reads as follows: Given any colouring (i.e., function) $\pi : [\omega]^n \to 3$, where $3 = \{ 0, 1, 2 \}$, does there exist a set $H \in [\omega]^{\omega}$ such that $\pi|[\mu]^n$ is constant? Alternatively, one can define an equivalence relation $\sim$ on $[\omega]^n$ by stipulating $x \sim y$ iff $\pi(x) = \pi(y)$ and ask whether there exists a set $H \in [\omega]^{\omega}$ such that $[H]^n$ is included in one equivalence class. The answer to this question is given by Ramsey’s Theorem 2.1 below, but before we state and prove this theorem, let us say a few words about its background.

Ramsey proved his theorem in order to investigate a problem in formal logic, namely the problem of finding a regular procedure to determine the truth or falsity of a given logical formula in the language of First-Order Logic, which is also the language of Set Theory (cf. Chapter 3). However, Ramsey’s Theorem is a purely combinatorial statement and was the nucleus — but not the earliest result — of a whole combinatorial theory, the so-called Ramsey Theory. We would also like to mention that Ramsey’s original theorem, which will be discussed later, is somewhat stronger than the theorem stated below but is, like König’s Lemma, not provable without assuming some form of the Axiom of Choice (see Proposition 7.8).

**Theorem 2.1 (Ramsey’s Theorem).** For any number $n \in \omega$, for any positive number $r \in \omega$, for any $S \in [\omega]^\omega$, and for any colouring $\pi : [S]^n \to r$, there is always an $H \in [S]^\omega$ such that $H$ is homogeneous for $\pi$, i.e., the set $[H]^n$ is monochromatic.

Before we prove Ramsey’s Theorem, let us consider a few examples: In the first example we colour the set of prime numbers $\mathbb{P}$ with two colours.
A **Wieferich prime** is a prime number \( p \) such that \( p^2 \mid 2^{p-1} - 1 \), denoted \( p^2 \mid 2^{p-1} - 1 \) — recall that by **Fermat’s Little Theorem** we have \( p \mid 2^{p-1} - 1 \) for any prime \( p \). Now, define the 2-colouring \( \pi_1 \) of \( \mathbb{P} \) by stipulating

\[
\pi_1(p) = \begin{cases} 
0 & \text{if } p \text{ is a Wieferich prime,} \\
1 & \text{otherwise.} 
\end{cases}
\]

Let \( H_0 = \{ p \in \mathbb{P} : p^2 \mid 2^{p-1} - 1 \} \) and \( H_1 = \mathbb{P} \setminus H_0 \). The only numbers which are known to belong to \( H_0 \) are 1093 and 3511. On the other hand, it is not known whether \( H_1 \) is infinite. However, by the **Infinite Pigeon-Hole Principle** we know that at least one of the two sets \( H_0 \) and \( H_1 \) is infinite, which gives us a homogeneous set for \( \pi_1 \).

As a second example, define the 2-colouring \( \pi_2 \) of the set of 2-element subsets of \( \{ 7l : l \in \omega \} \) by stipulating

\[
\pi_2(\{ n, m \}) = \begin{cases} 
0 & \text{if } n^m + m^n + 1 \text{ is prime,} \\
1 & \text{otherwise.} 
\end{cases}
\]

An easy calculation modulo 3 shows that the set \( H = \{ 42k + 14 : k \in \omega \} \subseteq \{ 7l : l \in \omega \} \) is homogeneous for \( \pi_2 \); in fact, for all \( \{ n, m \} \in [H]^2 \) we have \( 3 \mid (n^m + m^n + 1) \).

Before we give a third example, we prove the following special case of **Ramsey’s Theorem**.

**Proposition 2.2.** For any positive number \( r \in \omega \), for any \( S \subseteq [\omega]^\omega \), and for any colouring \( \pi : [S]^2 \to r \), there is always an \( H \in [S]^\omega \) such that \( |H|^2 \) is monochromatic.

**Proof.** The proof is in fact just a consequence of the **Infinite Pigeon-Hole Principle**; firstly, the **Infinite Pigeon-Hole Principle** is used to construct homogeneous sets for certain 2-colourings \( \tau \) and then it is used to show the existence of a homogeneous set for \( \pi \).

Let \( S_0 = S \) and let \( a_0 = \min(S_0) \). Define the \( r \)-colouring \( \tau_0 : S_0 \setminus \{ a_0 \} \to r \) by stipulating \( \tau_0(b) := \pi(\{ a_0, b \}) \). By the **Infinite Pigeon-Hole Principle** there is an infinite set \( S_1 \subseteq S_0 \setminus \{ a_0 \} \) such that \( \tau_0|_{S_1} \) is constant (i.e., \( \tau_0|_{S_1} \) is a constant function) and let \( \rho_0 := \tau_0(b) \), where \( b \) is any member of \( S_1 \). Now, let \( a_1 = \min(S_1) \) and define the \( r \)-colouring \( \tau_1 : S_1 \setminus \{ a_1 \} \to r \) by stipulating \( \tau_1(b) := \pi(\{ a_1, b \}) \). Again we find an infinite set \( S_2 \subseteq S_1 \setminus \{ a_1 \} \) such that \( \tau_1|_{S_2} \) is constant and let \( \rho_1 := \tau_1(b) \), where \( b \) is any member of \( S_2 \). Proceeding this way we finally get infinite sequences \( a_0 < a_1 < \ldots < a_n < \ldots \) and \( \rho_0, \rho_1, \ldots \). Notice that by construction, for all \( n \in \omega \) and all \( k > n \) we have \( \pi(\{ a_n, a_k \}) = \tau_n(a_k) = \rho_n \). Define the \( r \)-colouring \( \tau : \{ a_n : n \in \omega \} \to r \) by stipulating \( \tau(a_n) := \rho_n \). Again by the **Infinite Pigeon-Hole Principle** there is an infinite set \( H \subseteq \{ a_n : n \in \omega \} \) such that \( \tau|_H \) is constant, which implies that \( H \) is homogeneous for \( \pi \), i.e., \([H]^2\) is monochromatic.

\(-1\)
As a third example, consider the 17-colouring \( \pi_3 \) of the set of 9-element subsets of \( P \) defined by stipulating
\[
\pi_3(\{p_1, \ldots, p_9\}) = c \iff p_1 \cdot p_2 \cdot \ldots \cdot p_9 \equiv c \mod 17.
\]
For \( 0 \leq k \leq 16 \) let \( P_k = \{p \in P : p \equiv k \mod 17\} \). Then, by Dirichlet’s theorem on primes in arithmetic progression, \( P_k \) is infinite whenever \( \gcd(k, 17) = 1 \), i.e., for all positive numbers \( k \leq 16 \). Thus, by an easy calculation modulo 17 we get that for \( 1 \leq k \leq 16 \), \( P_k \) is homogeneous for \( \pi_3 \).

Now we give a complete proof of \textsc{Ramsey’s Theorem 2.1}:

\textbf{Proof of \textsc{Ramsey’s Theorem}.} The proof is by induction on \( n \). For \( n = 2 \) we get \textsc{Proposition 2.2}. So, we assume that the statement is true for \( n \geq 2 \) and prove it for \( n + 1 \). Let \( \pi : [\omega]^{n+1} \to r \) be any \( r \)-colouring of \([\omega]^{n+1}\). For each integer \( a \in \omega \) let \( \pi_a \) be the \( r \)-colouring of \([\omega \setminus \{a\}]^n\) defined as follows:
\[
\pi_a(x) = \pi(x \cup \{a\})
\]
By induction hypothesis, for each \( S' \in [\omega]^n \) and for each \( a \in S' \) there is an \( H_a^{S'} \in [S' \setminus \{a\}]^n \) such that \( H_a^{S'} \) is homogeneous for \( \pi_a \). Construct now an infinite sequence \( a_0 < a_1 < \ldots < a_i < \ldots \) of natural numbers and an infinite sequence \( S_0 \supseteq S_1 \supseteq \ldots \supseteq S_i \supseteq \ldots \) of infinite subsets of \( \omega \) as follows: Let \( S_0 = S \) and \( a_0 = \min(S) \), and in general let
\[
S_{i+1} = H_{a_i}^{S_i}, \text{ and } a_{i+1} = \min\{a \in S_{i+1} : a > a_i\}.
\]
It is clear that for each \( i \in \omega \), the set \([\{a_m : m > i\}]^n\) is monochromatic for \( \pi_{a_i} \); let \( \tau(a_i) \) be its colour (i.e., \( \tau \) is a colouring of \( \{a_i : i \in \omega\} \) with at most \( r \) colours). By the \textsc{Infinite Pigeon-Hole Principle} there is an \( H \subseteq \{a_i : i \in \omega\} \) such that \( \tau \) is constant on \( H \), which implies that \( \pi|_{[H]^{n+1}} \) is constant, too. Indeed, for any \( x_0 < \ldots < x_n \) in \( H \) we have \( \pi(\{x_0, \ldots, x_n\}) = \pi_{x_0}(\{x_1, \ldots, x_n\}) = \tau(x_0) \), which completes the proof.

\textbf{Corollaries of \textsc{Ramsey’s Theorem}}

In finite Combinatorics, the most important consequence of \textsc{Ramsey’s Theorem 2.1} is its finite version:

\textbf{Corollary 2.3 (Finite Ramsey Theorem).} \textit{For all} \( m, n, r \in \omega \), \textit{where} \( r \geq 1 \) \textit{and} \( n \leq m \), \textit{there exists an} \( N \in \omega \), \textit{where} \( N \geq m \), \textit{such that for every colouring of} \([N]^{m}\) \textit{with} \( r \) \textit{colours, there exists a set} \( H \in [N]^m \), \textit{all of whose} \( n \)-\textit{element subsets have the same colour}.

\textbf{Proof.} Assume towards a contradiction that the \textsc{Finite Ramsey Theorem} fails. So, there are \( m, n, r \in \omega \), \textit{where} \( r \geq 1 \) \textit{and} \( n \leq m \), \textit{such that for all} \( N \in \omega \)
Corollaries of Ramsey’s Theorem

with \( N \geq m \) there is a colouring \( \pi_N : [N]^n \to r \) such that no \( H \in [N]^m \) is homogeneous, i.e., \([H]^n \) is not monochromatic. We shall construct an \( r \)-colouring \( \pi \) of \([\omega]^n \) such that no infinite subset of \( \omega \) is homogeneous for \( \pi \), contradicting Ramsey’s Theorem. The \( r \)-colouring \( \pi \) will be induced by an infinite branch through a finitely branching tree, where the infinite branch is obtained by König’s Lemma. Thus, we first need an infinite, finitely branching tree. For this, consider the following graph \( G \): The vertex set of \( G \) consists of \( \emptyset \) and all colourings \( \pi_N : [N]^n \to r \), where \( N \geq m \), such that no \( H \in [N]^m \) is homogeneous for \( \pi_N \). There is an edge between \( \emptyset \) and each \( r \)-colouring \( \pi_m \) of \([m]^n \), and there is an edge between the colourings \( \pi_N \) and \( \pi_{N+1} \) if \( \pi_N \equiv \pi_{N+1} \mid_N \) (i.e., for all \( x \in [N]^n \), \( \pi_{N+1}(x) = \pi_N(x) \)). In particular, there is no edge between two different \( r \)-colouring of \([N]^n \). By our assumption, the graph \( G \) is infinite. Further, by construction, it is cycle-free, connected, finitely branching, and has a root, namely \( \emptyset \). In other words, \( G \) is an infinite, finitely branching tree and therefore, by König’s Lemma, contains an infinite branch of \( r \)-colourings, say \((\emptyset, \pi_m, \pi_{m+1}, \ldots, \pi_{m+i}, \ldots)\), where for all \( i, j \in \omega \), the colouring \( \pi_{m+i+j} \) is an extension of the colouring \( \pi_{m+i} \).

At this point we would like to mention that since for any \( N \in \omega \) the set of all \( r \)-colourings of \([N]^n \) can be ordered, for example lexicographically, we do not need any non-trivial form of the Axiom of Choice to construct an infinite branch.

Now, the infinite branch \((\emptyset, \pi_m, \pi_{m+1}, \ldots)\) induces an \( r \)-colouring \( \pi \) of \([\omega]^n \) such that no \( m \)-element subset of \( \omega \) is homogeneous. In particular, there is no infinite set \( H \in [\omega]^m \) such that \( \pi \mid_H \) is constant, which is a contradiction to Ramsey’s Theorem 2.1 and completes the proof.

The following corollary is a geometrical consequence of the Finite Ramsey Theorem 2.3:

**Corollary 2.4.** For every positive integer \( n \) there exists an \( N \in \omega \) with the following property: If \( P \) is a set of \( N \) points in the Euclidean plane without three collinear points, then \( P \) contains \( n \) points which form the vertices of a convex \( n \)-gon.

**Proof.** By the Finite Ramsey Theorem 2.3, let \( N \) be such that for every \( 2 \)-colouring of \([N]^3 \) there is a set \( H \in [N]^n \) such that \([H]^3 \) is monochromatic. Now let \( N \) points in the plane be given, and number them from 1 to \( N \) in an arbitrary but fixed way. Colour a triple \((i, j, k)\), where \( i < j < k \), red, if travelling from \( i \) to \( j \) to \( k \) is in clockwise direction; otherwise, colour it blue. By the choice of \( N \), there are \( n \) ordered points so that every triple has the same colour (i.e., orientation) from which one verifies easily (e.g., by considering the convex hull of the \( n \) points) that these points form the vertices of a convex \( n \)-gon.

The following theorem — discovered more than a decade before Ramsey’s Theorem — is perhaps the earliest result in Ramsey Theory:
Corollary 2.5 (Schur’s Theorem). If the positive integers are finitely coloured (i.e., coloured with finitely many colours), then there are three distinct positive integers \( x, y, z \) of the same colour, with \( x + y = z \).

Proof. Let \( r \) be a positive integer and let \( \pi \) be any \( r \)-colouring of \( \omega \setminus \{0\} \). Let \( N \in \omega \) be such that for every \( r \)-colouring of \( [N]^2 \) there is a homogeneous 3-element subset of \( N \). Define the colouring \( \pi^* : [N]^2 \to r \) by stipulating \( \pi^*(i, j) = \pi((i - j)) \), where \( |i - j| \) is the modulus or absolute value of the difference \( i - j \). Since \( N \) contains a homogeneous 3-element subset (for \( \pi^* \)), there is a triple \( 0 \leq i < j < k < N \) such that \( \pi^*(i, j) = \pi^*(j, k) = \pi^*(i, k) \), which implies that the numbers \( x = j - i, y = k - j, \) and \( z = k - i \), have the same colour, and in addition we have \( x + y = z \).

The next result is a purely number-theoretical result and follows quite easily from Ramsey’s Theorem. However, somewhat surprisingly, it is unprovable in Number Theory, or more precisely, in Peano Arithmetic (which will be discussed in Chapter 3). Before we can state the corollary, we have to introduce the following notion: A non-empty set \( S \subseteq \omega \) is called large if \( S \) has more than \( \min(S) \) elements. Further, for \( n, m \in \omega \) let \([n, m] := \{i \in \omega : n \leq i \leq m\}\).

Corollary 2.6. For all \( n, k, r \in \omega \) with \( r \geq 1 \), there is an \( m \in \omega \) such that for any \( r \)-colouring of \( [n, m]^k \), there exists a large homogeneous set.

Proof. Let \( n, k, r \in \omega \), where \( r \geq 1 \), be some arbitrary but fixed numbers. Let \( \pi : [\omega \setminus n]^k \to r \) be any \( r \)-colouring of the \( k \)-element subsets of \( \{i \in \omega : i \geq n\} \). By Ramsey’s Theorem 2.1 there exists an infinite homogeneous set \( H \subseteq [\omega \setminus n]^\omega \). Let \( a = \min(H) \) and let \( S \) denote the least \( a + 1 \) elements of \( H \). Then \( S \) is large and \( [S]^k \) is monochromatic.

The existence of a finite number \( m \) with the required properties now follows — using König’s Lemma — in the very same way as the Finite Ramsey Theorem followed from Ramsey’s Theorem (see the proof of the Finite Ramsey Theorem 23).

Generalisations of Ramsey’s Theorem

Even though Ramsey’s theorems are very powerful combinatorial results, they can still be generalised. The following result will be used later in Chapter 7 in order to prove that the Prime Ideal Theorem — introduced in Chapter 5 — holds in the ordered Mostowski permutation model (but it will not be used anywhere else in this book).

In order to illustrate the next theorem, as well as to show that it is optimal to some extent, we consider the following two examples: Firstly, define the 2-colouring \( \pi_1 \) of \( [\omega]^2 \times [\omega]^3 \times [\omega]^3 \) by stipulating

\[
\pi_1(\{x_1, x_2\}, \{y_1, y_2, y_3\}, \{z_1\}) = \begin{cases} 
1 \text{ if } 2^{x_1-x_2} + 13^{y_1-y_2}y_3 + 17z_1 - 3 \text{ is prime,} \\
0 \text{ otherwise.}
\end{cases}
\]
Let \( H_1 = \{3 \cdot k : k \in \omega\} \), \( H_2 = \{2 \cdot k : k \in \omega\} \), and \( H_3 = \{6 \cdot k : k \in \omega\} \). Then an easy calculation modulo 7 shows that \([H_1]^2 \times [H_2]^3 \times [H_3]^1\) is an infinite monochromatic set.

Secondly, define the 2-colouring \( \pi_2 \) of \([\omega]^1 \times [\omega]^1\) by stipulating

\[
\pi_2(\{x\}, \{y\}) = \begin{cases} 
1 & \text{if } x < y, \\
0 & \text{otherwise.}
\end{cases}
\]

It is easy to see that whenever \( H_1 \) and \( H_2 \) are infinite subsets of \( \omega \), then \([H_1]^1 \times [H_2]^1\) is not monochromatic; on the other hand, we easily find arbitrarily large finite sets \( M_1, M_2 \subseteq \omega \) such that \([M_1]^1 \times [M_2]^1\) is monochromatic.

Thus, if \([\omega]^{n_1} \times \ldots \times [\omega]^{n_l}\) is coloured with \( r \) colours, then, in general, we cannot expect to find infinite subsets of \( \omega \), say \( H_1, \ldots, H_l \), such that \([H_1]^{n_1} \times \ldots \times [H_l]^{n_l}\) is monochromatic; but we always find arbitrarily large finite subsets of \( \omega \).

**Theorem 2.7.** Let \( r, l, n_1, \ldots, n_l \in \omega \) with \( r \geq 1 \) be given. For every \( m \in \omega \) with \( m \geq \max\{n_1, \ldots, n_l\} \) there is some \( N \in \omega \) such that whenever \([N]^{n_1} \times \ldots \times [N]^{n_l}\) is coloured with \( r \) colours, then there are \( M_1, \ldots, M_l \in [N]^m\) such that \([M_1]^{n_1} \times \ldots \times [M_l]^{n_l}\) is monochromatic.

**Proof.** The proof is by induction on \( l \) and the induction step uses a so-called product-argument. For \( l = 1 \) the statement is equivalent to the Finite Ramsey Theorem 2.3. So, assume that the statement is true for \( l \geq 1 \) and let us prove it for \( l + 1 \). By induction hypothesis, for every \( r \geq 1 \) there is an \( N_l \) (depending on \( r \)) such that for every \( r \)-colouring of \([N_l]^{n_1} \times \ldots \times [N_l]^{n_l}\) there are \( M_1, \ldots, M_l \in [N_l]^m\) such that \([M_1]^{n_1} \times \ldots \times [M_l]^{n_l}\) is monochromatic. Now, the crucial idea in order to apply the Finite Ramsey Theorem is to consider the coloured \( l \)-tuples in \( ([N_l]^m)^l \) as new colours. More precisely, let \( u_l \) be the number of different \( l \)-tuples in \( ([N_l]^m)^l \) and let \( r_l := u_l \cdot r \). Notice that each colour in \( r_l \) corresponds to a pair \((t, c)\), where \( t \) is an \( l \)-tuple in \( ([N_l]^m)^l \) and \( c \) is one of \( r \) colours. Notice also that \( r_l \) is very large compared to \( r \). Now, by the Finite Ramsey Theorem 2.3, there is a number \( N_{l+1} \in \omega \) such that whenever \([N_{l+1}]^{n_{l+1}}\) is coloured with \( r_l \) colours, then there exists an \( M_{l+1} \in [N_{l+1}]^m\) such that \([M_{l+1}]^{n_{l+1}}\) is monochromatic. Let \( N = \max\{N_l, N_{l+1}\} \) and let \( \pi \) be any \( r \)-colouring of \([N]^{n_1} \times \ldots \times [N]^{n_l} \times [N]^{n_{l+1}}\). For every \( F \in [N]^{n_{l+1}} \) let \( \pi^F \) be the \( r \)-colouring of \([N]^{n_1} \times \ldots \times [N]^{n_l} \times [N]^{n_{l+1}}\) defined by stipulating

\[
\pi^F(X) = \pi((X, F)).
\]

By the definition of \( N \), for every \( F \in [N]^{n_{l+1}} \) there is a lexicographically first \( l \)-tuple \((M_1^F, \ldots, M_l^F) \in ([N]^m)^l \) such that \([M_1^F]^{n_1} \times \ldots \times [M_l^F]^{n_l}\) is monochromatic for \( \pi^F \). By definition of \( r_l \) we can define an \( r_l \)-colouring \( \pi_{l+1} \) on \([N]^{n_{l+1}}\) as follows: Every set \( F \in [N]^{n_{l+1}} \) is coloured according to the \( l \)-tuple \( t = (M_1^F, \ldots, M_l^F) \) (which can be encoded as one of \( u_l \) numbers) and the
colour \( c = \pi^F(X) \), where \( X \) is any element of the set \( [M^F]^n \times \ldots \times [M^F]^n \); because \( [M^F]^n \times \ldots \times [M^F]^n \) is monochromatic for \( \pi^F \), \( c \) is well-defined and one of \( r \) colours. In other words, for every \( F \in \mathbb{N}^{m+1} \), \( \pi_{i+1}(F) \) correspond to a pair \((t, c)\), where \( t \in (\mathbb{N}^m)^l \) and \( c \) is one of \( r \) colours. Finally, by definition of \( N \), there is a set \( M_{i+1} \in [N]^m \) such that \( [M^F]^n \times \ldots \times [M^F]^n \) is monochromatic for \( \pi_{i+1} \), which implies that for all \( F, F_1, F_2 \in [M^F]^n \) we have:

- \( [M^F]^n \times \ldots \times [M^F]^n \) is monochromatic for \( \pi^F \),
- \( (M_1^{F_1}, \ldots, M_i^{F_i}) = (M_i^{F_1}, \ldots, M_i^{F_2}) \),
- and restricted to the set \( [M^F]^n \times \ldots \times [M^F]^n \), the colourings \( \pi_i^{F_1} \) and \( \pi_i^{F_2} \) are identical.

Hence, there are \( M_1, \ldots, M_{i+1} \in [N]^m \) such that \( \pi_i|_{[M^F]^n \times \ldots \times [M^F]^n} \) is constant, which completes the proof.

A very strong generalisation of Ramsey’s Theorem in terms of partitions is the Partition Ramsey Theorem 11.4. However, since the proof of this generalisation is quite involved, we postpone the discussion of that result until Chapter 11 and consider now some other possible generalisations of Ramsey’s Theorem: Firstly one could finitely colour all finite subsets of \( \omega \), secondly one could colour \( [\omega]^n \) with infinitely many colours, and finally, one could finitely colour all the infinite subsets of \( \omega \). However, below we shall see that none of these generalisations works, but first, let us consider Ramsey’s original theorem, which is — at least in the absence of the Axiom of Choice — also a generalisation of Ramsey’s Theorem.

### Ramsey’s Original Theorem

The theorem which Ramsey proved originally is somewhat stronger than what we proved above. In our terminology, it states as follows:

**Ramsey’s Original Theorem.** For any infinite set \( A \), for any number \( n \in \omega \), for any positive number \( r \in \omega \), and for any colouring \( \pi : [A]^n \to r \), there is an infinite set \( H \subseteq A \) such that \( [H]^n \) is monochromatic.

Notice that the difference is just that the infinite set \( A \) is not necessarily a subset of \( \omega \), and therefore, it does not necessarily contain a countable infinite subset. However, this difference is crucial, since one can show that, like König’s lemma, this statement is not provable without assuming some form of the Axiom of Choice (AC). On the other hand, if one has AC, then every infinite set has a countably infinite subset, and so Ramsey’s Theorem implies the original version. Ramsey was aware of this fact and stated explicitly that he is assuming the axiom of selection (i.e., AC). Even though we do not need full AC in order to prove Ramsey’s Original Theorem, there is no way to avoid some non-trivial kind of choice, since there are models of Set Theory in which Ramsey’s Original Theorem fails (cf. Proposition 7.8). Consequently,
Ramsey’s Original Theorem can be used as a choice principle, which will be discussed in Chapter 5.

Finite Colourings of $[\omega]^{<\omega}$

Assume we have coloured all the finite subsets of $\omega$ with two colours, say red and blue. Can we be sure that there is an infinite subset of $\omega$ such that all its finite subsets have the same colour? The answer to this question is negative and it is not hard to find a counterexample (e.g., colour a set $x \in [\omega]^{<\omega}$ blue, if $|x|$ is even; otherwise, colour it red).

Thus, let us ask for slightly less. Is there at least an infinite subset of $\omega$ such that for each $n \in \omega$, all its $n$-element subsets have the same colour? The answer to this question is also negative: Colour a non-empty set $x \in [\omega]^{<\omega}$ red, if $x$ has more than $\min(x)$ elements (i.e., $x$ is large); otherwise, colour it blue. Now, let $I$ be an infinite subset of $\omega$ and let $n = \min(I)$. We leave it as an exercise to the reader to verify that $[I]^{n+1}$ is dichromatic.

The picture changes if we are asking just for an almost homogeneous sets: An infinite set $H \subseteq \omega$ is called almost homogeneous for a colouring $\pi : [\omega]^n \to r$ (where $n \in \omega$ and $r$ is a positive integer), if there is a finite set $K \subseteq \omega$ such that $H \setminus K$ is homogeneous for $\pi$. Now, for a positive integer $r$ consider any colouring $\pi : [\omega]^{<\omega} \to r$. Then, for each $n \in \omega$, $\pi|_{[\omega]^n}$ is a colouring $\pi_n : [\omega]^n \to r$. Is there an infinite set $H \subseteq \omega$ which is almost homogeneous for all $\pi_n$’s simultaneously? The answer to this question is affirmative and is given by the following result.

Proposition 2.8. Let $\{r_k : k \in \omega\}$ and $\{n_k : k \in \omega\}$ be two (possibly finite) sets of positive integers, and for each $k \in \omega$ let $\pi_k : [\omega]^{n_k} \to r_k$ be a colouring. Then there exists an infinite set $H \subseteq \omega$ which is almost homogeneous for each $\pi_k$ ($k \in \omega$).

Proof. A first attempt to construct the required almost homogeneous set would be to start with an $I_0 \in [\omega]^{<\omega}$ which is homogeneous for $\pi_0$, then take an $I_1 \in [I_0]^{<\omega}$ which is homogeneous for $\pi_1$, et cetera, and finally take the intersection of all the $I_k$’s. Even though this attempt fails — since it is very likely that we end up with the empty set — it is the right direction. In fact, if the intersection of the $I_k$’s would be non-empty, it would be homogeneous for all $\pi_k$’s, which is more than what is required. In order to end up with an infinite set we just have to modify the above approach — the trick, which is used almost always when the word “almost” is involved, is called diagonalisation.

The proof is by induction on $k$: By Ramsey’s Theorem 2.1 there exists an $H_0 \in [\omega]^{<\omega}$ which is homogeneous for $\pi_0$. Assume we have already constructed $H_k \in [\omega]^{<\omega}$ (for some $k \ge 0$) such that $H_k$ is homogeneous for $\pi_k$. Let $a_k = \min(H_k)$ and let $S_k = H_k \setminus \{a_k\}$. Then, again by Ramsey’s Theorem 2.1, there exists an $H_{k+1} \in [S_k]^{<\omega}$ such that $H_{k+1}$ is homogeneous for $\pi_{k+1}$. Let $H = \{a_k : k \in \omega\}$. Then, by construction, for every $k \in \omega$ we have that
$H \setminus \{a_0, \ldots, a_{k-1}\}$ is homogeneous for $\pi_k$, which implies that $H$ is almost homogeneous for all $\pi_k$'s simultaneously.

Now we could ask what is the least number of 2-colourings of 2-element subsets of $\omega$ we need in order to make sure that no single infinite subset of $\omega$ is almost homogeneous for all colourings simultaneously? By Proposition 2.8 we know that countably many colourings are not sufficient, but as we will see later, the axioms of Set Theory do not decide how large this number is (cf. Chapter 18).

The dual question would be as follows: How large must a family of infinite subsets of $\omega$ be, in order to make sure that for each 2-colouring of the 2-element subsets of $\omega$ we find a set in the family which is homogeneous for this colouring? Again, the axioms of Set Theory do not decide how large this number is (cf. Chapter 18).

**Going to the Infinite**

There are two parameters involved in a colouring $\pi : [\omega]^n \to r$, namely $n$ and $r$. Let first consider the case when $n = 2$ and $r = \omega$. In this case, we obviously cannot hope for any infinite homogeneous or almost homogeneous set. However, there are still infinite subsets of $\omega$ which are homogeneous in a broader sense which leads to the Canonical Ramsey Theorem. Even though the Canonical Ramsey Theorem is a proper generalisation of Ramsey’s Theorem, we will not discuss it here (but see Related Result 0).

In the case when $n = \omega$ and $r = 2$ we cannot hope for an infinite homogeneous set, as the following example illustrates (compare this result with Chapter 5 | Related Result 38):

*In the presence of the Axiom of Choice there is a 2-colouring of $[\omega]^\omega$ such that there is no infinite set, all whose infinite subsets have the same colour.*

The idea is to construct (or more precisely, to prove the existence of) a colouring of $[\omega]^\omega$ with say red and blue in such a way that whenever an infinite set $x \subseteq [\omega]^\omega$ is coloured blue, then for each $a \in x$, $x \setminus \{a\}$ is coloured red, and vice versa.

For this, define an equivalence relation on $[\omega]^\omega$ as follows: for $x, y \in [\omega]^\omega$ let

$$x \sim y \iff x \triangle y \text{ is finite}$$

where $x \triangle y = (x \setminus y) \cup (y \setminus x)$ is the symmetric difference of $x$ and $y$. It is easily checked that the relation “$\sim$” is indeed an equivalence relation on $[\omega]^\omega$.

Further, let $\mathcal{A} \subseteq [\omega]^\omega$ be any set of representatives, i.e., $\mathcal{A}$ has exactly one element in common with each equivalence class. Since the existence of the set $\mathcal{A}$ relies on the Axiom of Choice, the given proof is not entirely constructive.

Colour now an infinite set $x \in [\omega]^\omega$ blue, if $|x \triangle r_x|$ is even, where $r_x \in (\mathcal{A} \cap [x]^{\omega})$; otherwise, colour it red. Since two sets $x, y \in [\omega]^\omega$ with finite symmetric difference are always equivalent, every infinite subset of $\omega$ must
contain blue as well as red coloured infinite subsets.

So, there is a colouring \( \pi : [\omega]^{\omega} \to \{0, 1\} \) such that for no \( x \in [\omega]^{\omega} \), \( \pi|_x^{\omega} \) is constant. On the other hand, if the colouring is not too sophisticated we may find a homogeneous set: For \( \mathcal{A} \subseteq [\omega]^{\omega} \) define \( \pi_{\mathcal{A}} : [\omega]^{\omega} \to \{0, 1\} \) by stipulating \( \pi_{\mathcal{A}}(x) = 1 \) iff \( x \in \mathcal{A} \). Now we say that the set \( \mathcal{A} \subseteq [\omega]^{\omega} \) has the **Ramsey property** if there exists an \( x_0 \in [\omega]^{\omega} \) such that \( \pi_{\mathcal{A}}|_x^{\omega} \) is constant.

In other words, \( \mathcal{A} \subseteq [\omega]^{\omega} \) has the Ramsey property if and only if there exists an \( x_0 \in [\omega]^{\omega} \) such that either \( [x_0]^{\omega} \subseteq \mathcal{A} \) or \( [x_0]^{\omega} \cap \mathcal{A} = \emptyset \). The Ramsey property is related to the cardinal \( \mathfrak{b} \) (cf. Chapter 8) and will be discussed in Chapter 9.

A slightly weaker property than the Ramsey property is the so-called **doughnut property**: If \( a \) and \( b \) are subsets of \( \omega \) such that \( b \setminus a \) is infinite, then we call the set \( [a, b]^{\omega} := \{ x \in [\omega]^{\omega} : a \subseteq x \subseteq b \} \) a **doughnut**. (Why such sets are called “doughnuts” is left to the reader’s imagination.) Now, a set \( \mathcal{A} \subseteq [\omega]^{\omega} \) is said to have the doughnut property if there exists a doughnut \( [a, b]^{\omega} \) (for some \( a \) and \( b \)) such that either \( [a, b]^{\omega} \subseteq \mathcal{A} \) or \( [a, b]^{\omega} \cap \mathcal{A} = \emptyset \). Obviously, every set with the Ramsey property has also the doughnut property (consider doughnuts of the form \( \{\emptyset, b\}^{\omega} \)). On the other hand, it is not difficult to show that, in the presence of the Axiom of Choice, there are sets with the doughnut property which fail to have the Ramsey property (just modify the example given above).

**Notes**

**Ramsey’s Theorem.** Frank Plumpton Ramsey (1903-1930), the elder brother of Arthur Michael Ramsey (who was Archbishop of Canterbury from 1961 to 1974), proved his famous theorem in [34] and the part of the volume in which his article appeared was issued on the 16th of December in 1929, but the volume itself belongs to the years 1929 and 1930 (which caused some confusion about the year Ramsey’s article was actually published). However, Ramsey submitted his paper already in November 1928. For Ramsey’s paper and its relation to First-Order Logic, as well as for an introduction to Ramsey Theory in general, we refer the reader to the classical textbook by Graham, Rothschild, and Spencer [16] (for Ramsey’s other papers on Logic see [35]). In [34], **Ramsey’s Theorem 2.1** appears as **Theorem A** and the **Finite Ramsey Theorem 2.3** is proved as a corollary and appears as **Theorem B**. Although **Ramsey’s Theorem** is accurately attributed to Ramsey, its popularisation stems from the classical paper of Erdős and Szekeres [9], where they proved (independently of Ramsey) **Corollary 2.4** — which can be seen as a variant of the **Finite Ramsey Theorem 2.3** in a geometrical context (see also Morris and Soifer [27]). The elegant proof we gave for **Corollary 2.4** is due to Tarsi (cf. Lewin [25] or Graham, Rothschild, and Spencer [16, p. 28]).

**Schröder’s Theorem.** Schur’s original paper [36] was motivated by **Fermat’s Last Theorem**, and he actually proved the following result: For all natural numbers \( m \), if \( p \) is prime and sufficiently large, then the equation \( x^m + y^m = z^m \) has a non-zero solution in the integers modulo \( p \). A proof of this theorem can also be found in
Graham, Rothschild, and Spencer [16, Section 3.1]. For some historical background
and for the early development of Ramsey Theory (before Ramsey) see Feferman [38].

The Paris-Harrington Result. As mentioned above, Corollary 2.6 is true but
unprovable in Peano Arithmetic (also called First-Order Arithmetic). This result was
the first natural example of such a statement and is due to Paris and Harrington [31]
(see also Graham, Rothschild, and Spencer [16, Section 6.3]). For other statements
of that type see Paris [30].

It is worth mentioning that Peano Arithmetic is, in a suitable sense, equivalent
to Zermelo-Fraenkel Set Theory with the Axiom of Infinity replaced by its negation,
which is a reasonable formalisation of standard combinatorial reasoning about finite
sets.

Rado’s generalisation of the Finite Ramsey Theorem. Theorem 2.7, which
is the only proper generalisation of the Finite Ramsey Theorem shown in this
book so far, is due to Rado [32] (see also page 113, Problems 4 & 5 of Jech [23]).

Ramsey sets and doughnuts. Even though the Ramsey property and the doughnut
property look very similar, there are sets which have the Ramsey property, but
which fail to have the doughnut property. For the relation between the doughnut
property and other regularity properties see for example Halbeisen [18] or Brendle,
Halbeisen, and Löwe [4] (see also Chapter 9 Related Result 60).

Related Results

0. Canonical Ramsey Theorem. The following result, known as the Canonical
Ramsey Theorem, is due to Erdős and Rado (cf. [8, Theorem I]): Whenever
we have a colouring \( \pi \) of \( \omega^n \), for some \( n \in \omega \), with an arbitrary (e.g., infinite)
set of colours, there exist an infinite set \( H \subseteq \omega \) and a set \( I \subseteq \{1, 2, \ldots, n\} \) such
that for any ordered \( n \)-element subsets \( \{k_1 < \ldots < k_n\} \), \( \{l_1 < \ldots < l_n\} \in [H]^n \)
we have \( \pi(\{k_1, \ldots, k_n\}) = \pi(\{l_1, \ldots, l_n\}) \iff k_i = l_i \), for all \( i \in I \). The \( 2^n \)
possible choices for \( I \) correspond to the so-called canonical colourings of \( \omega^n \).
As an example let us consider the case when \( n = 2 \): Let \( \pi \) be an arbitrary
colouring of \( \omega^2 \) and let \( H \in [\omega]^n \) and \( I \subseteq \{1, 2\} \) be as above. Then we are in
exactly one of the following four cases for all \( \{k_1 < k_2\}, \{l_1 < l_2\} \in [H]^2 \) (cf. [8,
Theorem II]):

1. If \( I = \emptyset \), then \( \pi(\{k_1, k_2\}) = \pi(\{l_1, l_2\}) \).
2. If \( I = \{1, 2\} \), then \( \pi(\{k_1, k_2\}) = \pi(\{l_1, l_2\}) \iff \{k_1, k_2\} = \{l_1, l_2\} \).
3. If \( I = \{1\} \), then \( \pi(\{k_1, k_2\}) = \pi(\{l_1, l_2\}) \iff k_1 = l_1 \).
4. If \( I = \{2\} \), then \( \pi(\{k_1, k_2\}) = \pi(\{l_1, l_2\}) \iff k_2 = l_2 \).

Obviously, if \( \pi \) is a finite colouring of \( \omega^n \), then we are always in case (1),
which gives us just Ramsey’s Theorem 2.1.

1. Ramsey numbers. The least number of people that must be invited to a party,
in order to make sure that \( n \) of them mutually shook hands before or \( m \) of them
mutually did not shake hands before, is denoted by \( R(n,m) \), and the numbers
\( R(n,m) \) are called Ramsey numbers. Notice that by the Finite Ramsey
Theorem, Ramsey numbers \( R(n,m) \) exist for all integers \( n,m \in \omega \). Very few
Ramsey numbers are actually known. It is easy to show that \( R(2,3) = 3 \) (in
2. Monochromatic triangles in $K_6$-free graphs. Erdős and Hajnal [10] asked for a graph which contains no $K_6$ (i.e., no complete graph on 6 vertices) but has the property that whenever its edges are 2-coloured there must be a monochromatic triangle. A minimal example for such a graph was provided by Graham [14]: On the one hand he showed that if a 5-cycle is deleted from a $K_6$, then the resulting graph contains no $K_6$ and has the property that whenever its edges are 2-coloured there is a monochromatic triangle. On the other hand, if a graph on 7 vertices contains no $K_6$, then there is a 2-colouring of the edges with no monochromatic triangle.

3. Hindman's Theorem. If $F \subseteq [\omega]^<\omega$, then we write $\Sigma F$ for $\Sigma_{n\in F} a$, where as usual we define $\Sigma \emptyset := 0$. Hindman's Theorem states that if $\omega$ is finitely coloured, then there is an $x \in [\omega]^\omega$ such that $\{ \Sigma F : F \subseteq [x]^<\omega \land F \neq \emptyset \}$ is monochromatic (cf. Hindman [21, Theorem 3.1] or Hindman and Strauss [22, Corollary 5.10] where references to alternative proofs are given on page 102). Using Hindman's Theorem as a strong Pigeon-Hole Principle, Milliken proved in [25] a strengthened version of Ramsey's Theorem 2.1 which includes Hindman's Theorem as well as Ramsey's Theorem 2.1. Since Milliken's result was proved independently by Taylor (cf. [39]), it is usually called Milliken-Taylor Theorem. In order to state this result we have to introduce some notation. Two finite sets $K_1, K_2 \subseteq \omega$ are said to be unmeshed if $\max(K_1) < \min(K_2)$ or $\max(K_2) < \min(K_1)$. If $I$ and $H$ are two sets of pairwise unmeshed finite subsets of $\omega$ and every member of $I$ is the union of (finitely many) members of $H$, then we write $I \subseteq H$. Further, let $(\omega)^n$ denote the set of all infinite sets of pairwise unmeshed finite subsets of $\omega$, and for $H \in (\omega)^n$ let $(H)^n := \{ I : |I| = n \land I \subseteq H \}$. Now, the Milliken-Taylor Theorem states as follows: If all the $n$-element sets of pairwise unmeshed finite subsets of $\omega$ are finitely coloured, then there exists an $H \in (\omega)^\omega$ such that $(H)^n$ is monochromatic.

4. Colourings of the plane. Erdős [7] proved that there is a colouring of the Euclidean plane with countably many colours, such that any two points at a rational distance have different colours. This result was strengthened by Komjáth [24] in the following way: Let $Q$ be the set of rational numbers and let $Q := \{ (q, 0) : q \in Q \}$ be a copy of the rationals in the Euclidean plane. Then there exists a colouring of the Euclidean plane with countably many colours, such that for any rigid motion $\sigma$ of the plane, every colour occurs in $\sigma(Q) = \{ \sigma(p) : p \in Q \}$ exactly once.

5. Finite colourings of $Q$. If we colour the rational numbers $Q$ with finitely many colours, is there always an infinite homogeneous set which is order-isomorphic to $Q$? In general, this is not the case: Let $\{ q_n : n \in \omega \}$ be an enumeration of $Q$ (see Chapter 4, in particular Related Result 14) and colour a pair $\{ q_i, q_j \}$ blue if $q_i < q_j \iff i < j$, otherwise, colour it red. Then it is easy to see that an infinite homogeneous set which is order-isomorphic to $Q$ would yield an infinite decreasing sequence of natural numbers, which is obviously not possible. On the other hand, for every positive integer $n \in \omega$ there is a smallest number
\( t_n \in \omega \) such that if \( [Q]^n \) is finitely coloured then there is an infinite set \( X \subseteq Q \) which is order-isomorphic to \( Q \) such that \([X]^n\) is coloured with at most \( t_n \) colours. For this see Devlin [6] or Vulisanovic [41], where it is shown that such numbers exist and that the sequence of numbers \( t_n \) coincides with the so-called tangent numbers (cf. Sloane [37, A000122]). In particular, \( t_1 = 1 \) and for \( n \geq 2 \),
\[
t_n = \sum_{i=1}^{n-1} \binom{n-1}{i} t_{n-i}.
\]

6. Symmetry and colourings. Banakh and Protasov investigated in [2] the following problem: Is it true that for every \( n \)-colouring of the group \( Z^n \) there exists an infinite monochromatic subset of \( Z^n \) which is symmetric with respect to a central reflection. It turns out that the answer is always positive (for all \( n \)). However, there exists a 4-colouring of \( Z^n \) without infinite, symmetric, monochromatic set. For more general results we refer the reader to Banakh, Verbitski, and Vorobets [3].

7. Wieferich primes. The so-called Wieferich primes were first introduced by Wieferich in [42] in relation to Fermat’s Last Theorem. As mentioned above, the only known Wieferich primes (less than \( 1.25 \times 10^{15} \)) are 1093 and 3511 (found in 1913 and 1922 respectively). It is not known if there are infinitely many primes of this type, even though it is conjectured that this is the case (see for example Halbeisen and Hungerbühler [19]). Moreover, it is not even known whether there are infinitely many non-Wieferich primes — although it is very likely to be the case.

8. Sums and products. As a consequence of Ramsey’s Theorem we get that if \( \omega \) is finitely coloured, then there are infinite sequences of positive integers \((x_0, x_1, \ldots, x_k, \ldots)\) and \((y_0, y_1, \ldots, y_k, \ldots)\) such that \( \{x_i + x_j : i, j \in \omega \wedge i < j\} \) as well as \( \{y_i \cdot y_j : i, j \in \omega \wedge i < j\} \) is monochromatic (but not necessarily of the same colour). On the other hand, it is known (cf. Hindman and Strauss [22, Chapter 17.2]) that one can colour the positive integers with finitely many colours in such a way that there is no infinite sequence \((x_0, x_1, \ldots, x_k, \ldots)\) such that \( \{x_i + x_j : i, j \in \omega \wedge i < j\} \cup \{x_i \cdot x_j : i, j \in \omega \wedge i < j\} \) is monochromatic.

9. The graph of pairwise sums and products. One can show that if \( \omega \) is 2-coloured, then there are infinitely many pairs of distinct positive integers \( x \) and \( y \) such that \( x + y \) has the same colour as \( x \cdot y \). For this consider the graph on \( \omega \) with \( n \) joined to \( m \) if for some distinct \( x, y \in \omega \) we have \( x + y = n \) and \( x \cdot y = m \). Now, notice that it is enough to show that this so-called graph of pairwise sums and products contains infinitely many triangles (cf. Halbeisen [17]).

Suppose now that \( \omega \) is finitely coloured. Are there two distinct positive integers \( x \) and \( y \) such that \( x + y \) has the same colour as \( x \cdot y \)? This problem — which is equivalent to asking whether the chromatic number of the graph of pairwise sums and products is finite or infinite — is still open (cf. Hindman and Strauss [22, Question 17.18]). A partial result is given in Halbeisen [17], where it is shown that such numbers \( x \) and \( y \) exist if \( \omega \) is 3-coloured.

10. Problems in Ramsey Theory. For a variety of open problems from Ramsey Theory we refer the reader to Graham [15] (it might be worth mentioning that Graham is offering modest rewards for most of the presented problems).

11. Applications of Ramsey Theory to Banach Space Theory. There are many — and sometimes quite unexpected — applications of Ramsey Theory to Banach Space
Theory (see for example Odell [28], Gowers [13], or Argyros and Todorčević [1]). Let us mention just the following two applications:

An unexpected application of Ramsey Theory to Banach Space Theory is due to Brunel and Sucheston [3]: If $x_1, x_2, \ldots$ is an infinite normalised basic sequence in a Banach space $X$ and $\varepsilon_n \searrow 0$ (a sequence of positive real numbers which tends to 0), then one can find an infinite subsequence $y_1, y_2, \ldots$ of $x_1, x_2, \ldots$ which has the following property: For any positive $n \in \omega$, any sequence of scalars $(a_1, \ldots, a_n) \in [-1,1]^n$ and any natural numbers $n \leq i_0 < \ldots < i_{n-1}$ and $n \leq j_0 < \ldots < j_{n-1}$ we have

$$\left\| \sum_{k=1}^{n} a_k y_{i_k} \right\| - \left\| \sum_{k=1}^{n} a_k y_{j_k} \right\| < \varepsilon_n .$$

The limit $\left\| \sum_{k=1}^{n} a_k \tilde{e}_k \right\|$ we obtain for each finite sequence $(a_1, \ldots, a_n) \in [-1,1]^n$ leads to the sequence $\tilde{e}_1, \tilde{e}_2, \ldots$, and the Banach space generated by $\tilde{e}_1, \tilde{e}_2, \ldots$ is called a spreading model of $X$. The notion of spreading models was generalised (e.g., using the Miliken-Taylor Theorem) and investigated by Halbeisen and Odell in [30].

Another example is due to Gowers [11, 12] (see also Todorčević [40, Section 2.3]), who discovered the long sought Block Ramsey Theorem—a genuinely new Ramsey-type result—for Banach spaces, which he used to prove his famous Dichotomy Theorem (see also Gowers [33, Section 5] or Odell [29]): Every Banach space $X$ contains a subspace $Y$ which either has an unconditional basis or is hereditarily indecomposable (i.e., $Y$ contains no subspaces having a non-trivial complemented subspace).

References

6. Dennis Devlin, Some partition theorems and ultrafilters on $\omega$, Ph.D. thesis (1979), Dartmouth College, Hanover (USA).
References


36. Issai Schur, Über die Kongruenz $x^m + y^m = z^m (mod p)$, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 25 (1916), 114–117.


The Axioms of Zermelo-Fraenkel Set Theory

Every mathematical science relies upon demonstration rather than argument and opinion. Certain principles, called premises, are granted, and a demonstration is made which resolves everything easily and clearly. To arrive at such a demonstration the means must be found for making it accessible to our judgment. Mathematicians, understanding this, devised signs, not separate from matter except in essence, yet distant from it. These were points, lines, planes, solids, numbers, and countless other characters, which are depicted on paper with certain colours, and they used these in place of the things symbolised.

GIOSEFFO ZARLINO
Le Istituzioni Harmoniche, 1558

Why Axioms?

In the middle and late 19th century, members of the then small mathematical community began to look for a rigorous foundation of Mathematics. In accordance with the Euclidean model for reason, the ideal foundation consists of a few simple, clear principles, so-called axioms, on which the rest of knowledge can be built via firm and reliable thoughts free of contradictions. However, at the time it was not clear what assumptions should be made and what operations should be allowed in mathematical reasoning.

At the beginning of the last book of Politeia, Plato develops his theory of ideas. Translated into the mathematical setting, Plato's theory of ideas reads as follows: Even though there may be more than one human approach to Mathematics, there is only one idea of Mathematics (i.e., a unique mathematical world), and from this idea alone we can attain real knowledge — all
human approaches are just opinions. In particular, the mathematical world already exists and is just waiting to be discovered. So, from a Platonic point of view it would make sense to search for the unique set of true axioms for Set Theory — also because the axioms of Set Theory are supposed to describe the world of “real” Mathematics.

However, if we consider Set Theory as a mathematical discipline, then, like in any other field of Mathematics, there is no true axiom system, and moreover, we are even allowed to weaken the axioms or to extend them by additional assumptions in order to get weaker or stronger theories. This is done for example in Group Theory in order to study semigroups or monoids, or to focus on abelian groups.

It is often the case that a mathematical theory is developed long before its formal axiomatisation, and in rare instances, mathematical theories were already partially developed before mathematicians were aware of them, which happened with Group Theory: Around the year 1600 in England it was discovered that by altering the fittings around each bell in a bell tower, it was possible for each ringer to maintain precise control of when his (there were no female ringers then) bell sounded. This enabled the ringers to ring the bells in any particular order, and either maintain that order or permute the order in a precise way. (For technical reasons, not every permutation is allowed. In fact, just products of mutually disjoint elementary transpositions may be used, that means that two bells can exchange their places only if they are adjacent.) So, in the first half of the 17th century the ringers tried to continuously change the order of the bells for as long as possible, while not repeating any particular order, and return to rounds at the end. This game evolved into a challenge to ring the bells in every possible order, without any repeats, and return to rounds at the end. Thus, bell-ringers began to investigate permutations and Stedman’s work *Campanologia* (Cambridge, 1677) can fairly be said to be the first work in which Group Theory was successfully applied to a “musical” situation and consequently, Stedman can be regarded as the first group theorist. This also shows that permutations — the prototype of finite groups — were first studied in the 17th century in the context of the change-ringing, and therefore had a practical application long before they were used in Lagrange’s work of 1770–1771 on the theory of algebraic equations.

Let us now turn back to Set Theory. The history of Set Theory is rather different from the history of most other areas of Mathematics. Usually a long process can be traced in which ideas evolve until an ultimate flash of inspiration, often by a number of mathematicians almost simultaneously, produces a discovery of major importance. Set Theory however is the creation of only one person, namely of Georg Cantor (1845–1918), who first discovered that infinite sets may have different sizes, i.e. cardinalities. In fact, the birth of Set Theory dates to 1873 when Cantor proved that the set of real numbers is uncountable. Until then, no one envisioned the possibility that infinities come in different sizes, and moreover, mathematicians had no use for the actual in-
finite" — in contrast to the potential infinite, as it is introduced by Aristotle in *Physics* Book III. The difference between actual and potential infinite is that the latter just means "unlimited" or "arbitrarily large" (e.g., there are arbitrarily large — and therefore arbitrarily many — prime numbers), whereas the former means that there are infinite objects which actually exist (e.g., there exists a set containing all, i.e., infinitely many, prime numbers). Moreover, Cantor also showed that for every infinite set, there is a set of larger cardinality, which implies that there is no largest set. Cantor never introduced formal axioms for Set Theory, even though he was tacitly using most of the axioms introduced later by Zermelo and Fraenkel. However, Cantor considered a set as any collection of well-distinguished objects of our mind, which leads directly to Russell’s Paradox: Firstly, the collection of all sets is a set which is a member of itself. Secondly, the set of negative natural numbers is empty, and hence cannot be a member of itself (otherwise, it would not be empty). Now, call a set *x* good if *x* is not a member of itself and let *C* be the collection of all sets which are good. Is *C*, as a set, good or not? If *C* is good, then *C* is not a member of itself, but since *C* contains all sets which are good, *C* is a member of *C*, a contradiction. Otherwise, if *C* is a member of itself, then *C* must be good, again a contradiction. In order to avoid this paradox we have to exclude the collection *C* from being a set, but then, we have to give reasons why certain collections are sets and others are not. The axiomatic way to do this is described by Zermelo as follows: *Starting with the historically grown Set Theory, one has to search for the principles required for the foundations of this mathematical discipline. In solving the problem we must, on the one hand, restrict these principles sufficiently to exclude all contradictions and, on the other hand, take them sufficiently wide to retain all the features of this theory.*

The principles, which are called axioms, will tell us how to get new sets from already existing ones. In fact, most of the axioms of Set Theory are constructive to some extent, i.e., they tell us how new sets are constructed from already existing ones and what elements they contain.

However, before we state the axioms of Set Theory we would like to introduce informally the formal language in which these axioms will be formulated.

**First-Order Logic in a Nutshell**

First-Order Logic is the system of Symbolic Logic concerned not only to represent the logical relations between sentences or propositions as wholes (like *Propositional Logic*), but also to consider their internal structure in terms of subject and predicate. First-Order Logic can be consider as a kind of language which is distinguished from higher-order languages in that it does not allow quantification over subsets of the domain of discourse or other objects of higher type. Nevertheless, First-Order Logic is strong enough to formalise all of Set Theory and thereby virtually all of Mathematics. In other words,
First-Order Logic is an abstract language that in one particular case is the
language of Group Theory, and in another case is the language of Set Theory.
The goal of this brief introduction to First-Order Logic is to illustrate and
summarise some of the basic concepts of this language and to show how it is
applied to fields like Group Theory and Peano Arithmetic (two theories which
will accompany us for a while).

Syntax: Formulae, Formal Proofs, and Consistency

Like any other written language, First-Order Logic is based on an alphabet,
which consists of the following symbols:

(a) **Variables** such as \(v_0, v_1, x, y, \ldots\) which are place holders for objects of the
domain under consideration (which can for example be the elements of a group,
natural numbers, or sets).

(b) **Logical operators** which are “¬” (not), “∧” (and), “∨” (or), “→” (implies),
and “↔” (if and only if, abbreviated iff).

(c) **Logical quantifiers** which are the existential quantifier “∃” (there is or
there exists) and the universal quantifier “∀” (for all or for each), where
quantification is restricted to objects only and not to formulae or sets of
objects (but the objects themselves may be sets).

(d) **Equality symbol** “=” , which stands for the particular binary equality
relation.

(e) **Constant symbols** like the number 0 in Peano Arithmetic, or the neutral
element \(e\) in Group Theory. Constant symbols stand for fixed individual
objects in the domain.

(f) **Function symbols** such as \(f\) (the operation in Group Theory), or \(+, \ldots, \ast\)
(the operations in Peano Arithmetic). Function symbols stand for fixed
functions taking objects as arguments and returning objects as values.
With each function symbol we associate a positive natural number, its
called “arity” (e.g., “\(\ast\)” is a 2-ary or binary function, and the successor
operation “\(+\)” is a 1-ary or unary function).

(g) **Relation symbols** or **predicate constants** (such as \(\in\) in Set Theory)
stand for fixed relations between (or properties of) objects in the domain.
Again we associate an “arity” with each relation symbol (e.g., “\(\in\)” is a
binary relation).

The symbols in (a)–(d) form the core of the alphabet and are called **logical
symbols**. The symbols in (e)–(g) depend on the specific topic we are investigat-ing and are called **non-logical symbols**. The set of non-logical symbols
which are used in order to formalise a certain mathematical theory is called the
**language** of this theory, denoted by \(\mathcal{L}\), and **formulae** which are for-
mulated in a language \(\mathcal{L}\) are usually called \(\mathcal{L}\)-formulae. For example if we
investigate groups, then the only non-logical symbols we use are “\(e\)” and “\(\ast\),”
thus, \(\mathcal{L} = \{e, \ast\}\) is the language of Group Theory.
A first step towards a proper language is to build words (i.e., terms) with these symbols.

**Terms:**

(T1) Each variable is a term.

(T2) Each constant symbol is a term.

(T3) If $t_1, \ldots, t_n$ are terms and $F$ is an $n$-ary function symbol, then $F(t_1 \cdots t_n)$ is a term.

It is convenient to use auxiliary symbols like brackets in order to make terms, relations, and other expressions easier to read. For example we usually write $F(t_1, \ldots, t_n)$ rather than $F(t_1 \cdots t_n)$.

To some extent, terms correspond to words, since they denote objects of the domain under consideration. Like real words, they are not statements and cannot express or describe possible relations between objects. So, the next step is to build sentences (i.e., formulae) with these terms.

**Formulae:**

(F1) If $t_1$ and $t_2$ are terms, then $t_1 = t_2$ is a formula.

(F2) If $t_1, \ldots, t_n$ are terms and $R$ is an $n$-ary relation symbol, then $R(t_1 \cdots t_n)$ is a formula.

(F3) If $\varphi$ is a formula, then $\neg \varphi$ is a formula.

(F4) If $\varphi$ and $\psi$ are formulae, then $(\varphi \land \psi)$, $(\varphi \lor \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are formulae. (To avoid the use of brackets one could write these formulae for example in Polish notation, i.e., $\land \varphi \psi$, $\lor \varphi \psi$, et cetera.)

(F5) If $\varphi$ is a formula and $x$ a variable, then $\exists x \varphi$ and $\forall x \varphi$ are formulae.

Formule of the form (F1) or (F2) are the most basic expressions we have, and since every formula is a logical connection or a quantification of these formulae, they are called **atomic formulae**.

For binary relations $R$ it is convenient to write $xRy$ instead of $R(x, y)$. For example we write $x \in y$ instead of $\in(x, y)$, and we write $x \notin y$ rather than $\neg(x \in y)$.

If a formula $\varphi$ is of the form $\exists x \psi$ or of the form $\forall x \psi$ (for some formula $\psi$) and $x$ occurs in $\psi$, then we say that $x$ is in the range of a logical quantifier. A variable $x$ occurring at a particular place in a formula $\varphi$ is either in the range of a logical quantifier or it is not in the range of any logical quantifier. In the former case this particular instance of the variable $x$ is **bound** in $\varphi$, and in the latter case it is **free** in $\varphi$. Notice that it is possible that a certain variable occurs in a given formula bound as well as free (e.g., in $\exists z(x = z) \land \forall x(x = y)$, the variable $x$ is both bound and free, whereas $z$ is just bound and $y$ is just free). However, one can always rename the bound variables occurring in a given formula $\varphi$ such that each variable in $\varphi$ is either bound or free. For formulae $\varphi$, the set of variables occurring free in $\varphi$ is denoted by $\text{free}(\varphi)$. A
formula \( \varphi \) is a **sentence** if it contains no free variables (i.e., \( \text{free}(\varphi) = \emptyset \)). For example \( \forall x (x = x) \) is a sentence but \( (x = x) \) is not.

Sometimes it is useful to indicate explicitly which variables occur free in a given formula \( \varphi \), and for this we usually write \( \varphi(x_1, \ldots, x_n) \) to indicate that \( \{x_1, \ldots, x_n\} \subseteq \text{free}(\varphi) \).

If \( \varphi(x) \) is a formula (i.e., \( x \in \text{free}(\varphi) \)), and \( t \) a term, then \( \varphi(x/t) \) is the formula we get after replacing all free instances of \( x \) by \( t \). A so-called **substitution** \( \varphi(x/t) \) is **admissible** iff no free occurrence of \( x \) in \( \varphi \) is in the range of a quantifier that binds any variable contained in \( t \) (i.e., for each variable \( v \) appearing in \( t \), no place where \( x \) occurs free in \( \varphi \) is in the range of “\( \exists v \)” or “\( \forall v \)”).

So far we have letters, and we can build words and sentences. However, these sentences are just strings of symbols without any inherent meaning. Later we shall interpret formulae in the intuitively natural way by giving the symbols the intended meaning (e.g., “\( \land \)” meaning “and”, “\( \forall x \)” meaning “for all \( x \)”, *et cetera*). But before we shall do so, let us stay a little bit longer on the syntactical side — nevertheless, one should consider the formulae also from a semantical point of view.

Below we shall label certain formulae or types of formulae as **axioms**, which are used in connection with **inference rules** in order to derive further formulae. From a semantical point of view we can think of axioms as “true” statements from which we deduce or prove further results. We distinguish two types of axioms, namely **logical axioms** and **non-logical axioms** (which will be discussed later). A **logical axiom** is a sentence or formula \( \varphi \) which is universally valid (i.e., \( \varphi \) is true in any possible universe, no matter how the variables, constants, *et cetera*, occurring in \( \varphi \) are interpreted). Usually one takes as logical axioms some minimal set of formulae that is sufficient for deriving all universally valid formulae (such a set is given below).

If a symbol is involved in an axiom which stands for an arbitrary relation, function, or even for a first-order formula, then we usually consider the statement as an **axiom schema** rather than a single axiom, since each instance of the symbol represents a single axiom. The following list of axiom schemata is a system of logical axioms.

Let \( \varphi, \varphi_1, \varphi_2, \) and \( \psi \), be arbitrary first-order formulae:

\[
\begin{align*}
L_1 & : \varphi \rightarrow (\psi \rightarrow \varphi) \\
L_2 & : (\psi \rightarrow (\varphi_1 \rightarrow \varphi_2)) \rightarrow ((\psi \rightarrow \varphi_1) \rightarrow (\psi \rightarrow \varphi_2)) \\
L_3 & : (\varphi \land \psi) \rightarrow \varphi \\
L_4 & : (\varphi \land \psi) \rightarrow \psi \\
L_5 & : \varphi \rightarrow (\psi \rightarrow (\psi \land \varphi)) \\
L_6 & : \varphi \rightarrow (\varphi \lor \psi) \\
L_7 & : \psi \rightarrow (\varphi \lor \psi) \\
L_8 & : (\varphi_1 \rightarrow \varphi_3) \rightarrow ((\varphi_2 \rightarrow \varphi_3) \rightarrow ((\varphi_1 \lor \varphi_2) \rightarrow \varphi_3)) \\
L_9 & : (\varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow \neg \psi) \rightarrow \neg \varphi)
\end{align*}
\]
Syntax: formulae, formal proofs, and consistency

L_{10}: \neg \varphi \rightarrow (\varphi \rightarrow \psi)
L_{11}: \varphi \lor \neg \varphi

If \( t \) is a term and the substitution \( \varphi(x/t) \) is admissible, then:

L_{12}: \forall x \varphi(x) \rightarrow \varphi(t)
L_{13}: \varphi(t) \rightarrow \exists x \varphi(x)

If \( \psi \) is a formula such that \( x \notin \text{free}(\psi) \), then:

L_{14}: \forall x (\psi \rightarrow \varphi(x)) \rightarrow (\psi \rightarrow \forall x \varphi(x))
L_{15}: \forall x (\varphi(x) \rightarrow \psi) \rightarrow (\exists x \varphi(x) \rightarrow \psi)

What is not covered yet is the symbol "=" so, let us have a closer look at
the binary equality relation. The defining properties of equality can already
be found in Book VII, Chapter 1 of Aristotle’s *Topics*, where one of the rules
to decide whether two things are the same is as follows: ... you should look
at every possible predicate of each of the two terms and at the things of which
they are predicated and see whether there is any discrepancy anywhere. For
anything which is predicated of the one ought also to be predicated of the other,
and of anything of which the one is a predicate the other also ought to be a
predicate.

In our formal system, the binary equality relation is defined by the following
three axioms.

If \( t, t_1, \ldots, t_n, t'_1, \ldots, t'_n \) are any terms, \( R \) an \( n \)-ary relation symbol \( (e.g., \) the
binary relation symbol “=”), and \( F \) an \( n \)-ary function symbol, then:

L_{16}: t = t
L_{17}: (t_1 = t'_1 \land \cdots \land t_n = t'_n) \rightarrow (R(t_1, \ldots, t_n) \rightarrow R(t'_1, \ldots, t'_n))
L_{18}: (t_1 = t'_1 \land \cdots \land t_n = t'_n) \rightarrow (F(t_1, \ldots, t_n) = F(t'_1, \ldots, t'_n))

Finally, we define the logical operator “\( \leftrightarrow \)” by stipulating

\[ \varphi \leftrightarrow \psi \iff (\varphi \rightarrow \psi) \land (\psi \rightarrow \varphi), \]

i.e., \( \varphi \leftrightarrow \psi \) is just an abbreviation for \( (\varphi \rightarrow \psi) \land (\psi \rightarrow \varphi) \).

This completes the list of our logical axioms. In addition to these axioms, we
are allowed to state arbitrarily many theory-specific assumptions, so-called
non-logical axioms. Such axioms are for example the three axioms of *Group
Theory*, denoted \( \text{GT} \), or the axioms of *Peano Arithmetic*, denoted \( \text{PA} \).

\textbf{GT}: The language of Group Theory is \( \mathcal{L}_{\text{GT}} = \{ e, \cdot, - \} \), where “\( e \)” is a constant
symbol and “\( - \)” is a binary function symbol.

\textbf{GT}_0: \forall x \forall y \forall z \in (x \cdot (y \cdot z)) = (x \cdot y) \cdot z \quad (i.e., “\cdot” is associative)

\textbf{GT}_1: \forall x \in (e \cdot x = x) \quad (i.e., “\( e \)” is a left-neutral element)

\textbf{GT}_2: \forall x \exists y \in (y \cdot x = e) \quad (i.e., every element has a left-inverse)
PA: The language of Peano Arithmetic is $\mathcal{L}_\text{PA} = \{0, s, +, \cdot\}$, where “0” is a constant symbol, “s” is a unary function symbol, and “+” and “.” are binary function symbols.

PA$_1$: $\forall x (s(x) \neq 0)$
PA$_2$: $\forall x \forall y (s(x) = s(y) \rightarrow x = y)$
PA$_3$: $\forall x (x + 0 = x)$
PA$_4$: $\forall x \forall y (x + s(y) = s(x + y))$
PA$_5$: $\forall x (x \cdot 0 = 0)$
PA$_6$: $\forall x \forall y (x \cdot s(y) = (x \cdot y) + x)$

If $\varphi$ is any $\mathcal{L}_\text{PA}$-formula with $x \in \text{free}(\varphi)$, then:
PA$_7$: $(\varphi(0) \land \forall x (\varphi(x) \rightarrow \varphi(s(x)))) \rightarrow \forall x \varphi(x)$.

It is often convenient to add certain defined symbols to a given language so that the expressions get shorter or at least are easier to read. For example in Peano Arithmetic — which is an axiomatic system for the natural numbers — we usually replace the expression $s(0)$ with 1 and consequently $s(x)$ by $x + 1$. Probably, we would like to introduce an ordering “$<$” on the natural numbers. We can do this by stipulating

$$1 := s(0), \quad x < y \iff \exists z ((x + z) + 1 = y).$$

We usually use “:=” to define constants or functions, and “$\iff$” to define relations. Obviously, all that can be expressed in the language $\mathcal{L}_\text{PA} \cup \{1, <\}$ can also be expressed in $\mathcal{L}_\text{PA}$.

So far we have a set of logical and non-logical axioms in a certain language and can define, if we wish, as many new constants, functions, and relations as we like. However, we are still not able to deduce anything from the given axioms, since we have neither inference rules nor the notion of formal proof.

Surprisingly, just two inference rules are sufficient, namely:

**Modus Ponens:** $\quad \frac{\varphi \rightarrow \psi, \varphi}{\psi}$

and **Generalisation:** $\quad \frac{\varphi}{\forall x \varphi}$

In the former case we say that $\psi$ is obtained from $\varphi \rightarrow \psi$ and $\varphi$ by Modus Ponens, and in the latter case we say that $\forall x \varphi$ (where $x$ can be any variable) is obtained from $\varphi$ by Generalisation.

Using these two inference rules, we are able to define the notion of formal proof: Let $T$ be a possibly empty set of non-logical axioms (usually sentences), formulated in a certain language $\mathcal{L}$. An $\mathcal{L}$-formula $\psi$ is provable from $T$ (or provable in $T$), denoted $T \vdash \psi$, if there is a finite sequence $\varphi_1, \ldots, \varphi_n$ of $\mathcal{L}$-formulae such that $\varphi_n$ is equal to $\psi$ (i.e., the formulae $\varphi_n$ and $\psi$ are identical), and for all $i$ with $1 \leq i \leq n$ we have:
Syntax: formulae, formal proofs, and consistency

- \( \varphi_i \) is a logical axiom, or
- \( \varphi_i \in T \), or
- there are \( j, k < i \) such that \( \varphi_j \) is equal to the formula \( \varphi_k \rightarrow \varphi_i \), or
- there is a \( j < i \) such that \( \varphi_i \) is equal to the formula \( \forall x \varphi_j \).

If a formula \( \psi \) is not provable in \( T \), i.e., if there is no formal proof for \( \psi \), then we write \( T \not\vdash \psi \).

Formal proofs, even of very simple statements, can get quite long and tricky. So, before we give an example of a formal proof, let us state a theorem which allows us to simplify formal proofs:

**Theorem 3.1 (Deduction Theorem).** If \( \{ \psi_1, \ldots, \psi_n \} \cup \{ \varphi_1, \ldots, \varphi_k \} \vdash \varphi \), where **Generalisation** is not applied to the free variables of the formulae \( \varphi_1, \ldots, \varphi_k \) (e.g., if these formulae are sentences), then

\[ \{ \psi_1, \ldots, \psi_n \} \vdash (\varphi_1 \land \ldots \land \varphi_k) \rightarrow \varphi . \]

Now, as an example of a formal proof let us show the equality relation is symmetric. We first work with \( T_{x=y} \) consisting only of the formula \( x = y \), and show that \( T_{x=y} \vdash y = x \), in other words we show that \( \{ x = y \} \vdash y = x \):

\[
\begin{align*}
\varphi_1: & \quad (x = y \land x = x) \rightarrow (x = x \rightarrow y = x) & \text{instance of } L_{17} \\
\varphi_2: & \quad (x = y \land x = x) \rightarrow x = x & \text{instance of } L_4 \\
\varphi_3: & \quad \varphi_1 \rightarrow (\varphi_2 \rightarrow ((x = y \land x = x) \rightarrow y = x)) & \text{instance of } L_2 \\
\varphi_4: & \quad \varphi_2 \rightarrow ((x = y \land x = x) \rightarrow y = x) & \text{from } \varphi_3 \text{ and } \varphi_1 \\
\varphi_5: & \quad (x = y \land x = x) \rightarrow y = x & \text{by Modus Ponens} \\
\varphi_6: & \quad x = x & \text{instance of } L_{16} \\
\varphi_7: & \quad x = y & (x = y) \in T_{x=y} \\
\varphi_8: & \quad x = x \rightarrow (x = y \rightarrow (x = y \land x = x)) & \text{instance of } L_5 \\
\varphi_9: & \quad x = y \rightarrow (x = y \land x = x) & \text{from } \varphi_8 \text{ and } \varphi_6 \\
\varphi_{10}: & \quad x = y \land x = x & \text{by Modus Ponens} \\
\varphi_{11}: & \quad y = x & \text{from } \varphi_5 \text{ and } \varphi_7 \\
\end{align*}
\]

Thus, we have \( \{ x = y \} \vdash y = x \), and by the **Deduction Theorem 3.1** we get that \( \vdash x = y \rightarrow y = x \), and finally, by **Generalisation** we get

\[ \vdash \forall x \forall y (x = y \rightarrow y = x) . \]
We leave it as an exercise to the reader to show that the equality relation is also transitive, and since the equality relation is also reflexive (by L₁₀), it is an equivalence relation.

Furthermore, we say that two formulae \( \varphi \) and \( \psi \) are **equivalent**, denoted \( \varphi \equiv \psi \), if \( \vdash \varphi \iff \psi \). In other words, if \( \varphi \equiv \psi \), then —from a logical point of view— \( \varphi \) and \( \psi \) state exactly the same, and therefore we could call \( \varphi \iff \psi \) a tautology, which means *saying the same thing twice*. However, in Logic, a formula \( \varphi \) is a **tautology** if \( \vdash \varphi \). Thus, the formulae \( \varphi \) and \( \psi \) are equivalent if and only if \( \varphi \iff \psi \) is a tautology.

A few examples:

- \( \varphi \lor \psi \equiv \psi \lor \varphi \), \( \varphi \land \psi \equiv \psi \land \varphi \) This shows that “\( \lor \)” and “\( \land \)” are commutative (up to equivalence). Moreover, “\( \lor \)” and “\( \land \)” are (up to equivalence) also associative—a fact which we tacitly used already.
- \( \neg \varphi \equiv \varphi \), \( (\varphi \lor \psi) \equiv \neg (\neg \varphi \land \neg \psi) \) This shows for example how “\( \lor \)” can be replaced with “\( \neg \)” and “\( \land \)”.
- \( (\varphi \rightarrow \psi) \equiv (\neg \varphi \lor \psi) \) This shows how the logical operator “\( \rightarrow \)” can be replaced with “\( \neg \)” and “\( \lor \)”.
- \( \forall x \varphi \equiv \neg \exists x \neg \varphi \) This shows how “\( \forall \)” can be replaced with “\( \neg \)” and “\( \exists \)”.

Thus, some of the logical operators are redundant and we could work for example with just “\( \neg \)” and “\( \land \)” and “\( \lor \)”. However, it is more convenient to use all of them.

Let \( T \) be a set of \( \mathcal{L} \)-formulae. We say that \( T \) is **consistent**, denoted \( \text{Con}(T) \), if there is no \( \mathcal{L} \)-formula \( \varphi \) such that \( T \vdash (\varphi \land \neg \varphi) \), otherwise \( T \) is called **inconsistent**, denoted \( \neg \text{Con}(T) \).

**Proposition 3.2.** Let \( T \) be a set of \( \mathcal{L} \)-formulae.

(a) If \( \neg \text{Con}(T) \), then for every \( \mathcal{L} \)-formula \( \psi \) we have \( T \vdash \psi \).

(b) If \( \text{Con}(T) \) and \( T \vdash \varphi \) for some \( \mathcal{L} \)-formula \( \varphi \), then \( T \vdash \neg \varphi \).

**Proof.** (a) Let \( \psi \) be any \( \mathcal{L} \)-formula and assume that \( T \vdash (\varphi \land \neg \varphi) \) for some \( \mathcal{L} \)-formula \( \varphi \). Then \( T \vdash \psi \):

\[
\begin{align*}
\varphi_1 & : \varphi \land \neg \varphi & \text{provable from } T \text{ by assumption} \\
\varphi_2 & : (\varphi \land \neg \varphi) \rightarrow \varphi & \text{instance of } \text{L}_3 \\
\varphi_3 & : \varphi & \text{from } \varphi_2 \text{ and } \varphi_1 \text{ by Modus Ponens} \\
\varphi_4 & : (\varphi \land \neg \varphi) \rightarrow \neg \varphi & \text{instance of } \text{L}_4 \\
\varphi_5 & : \neg \varphi & \text{from } \varphi_4 \text{ and } \varphi_1 \text{ by Modus Ponens} \\
\varphi_6 & : \neg \varphi \rightarrow (\varphi \rightarrow \psi) & \text{instance of } \text{L}_{10} \\
\varphi_7 & : \varphi \rightarrow \psi & \text{from } \varphi_6 \text{ and } \varphi_5 \text{ by Modus Ponens} \\
\varphi_8 & : \psi & \text{from } \varphi_7 \text{ and } \varphi_5 \text{ by Modus Ponens}
\end{align*}
\]
(b) Assume that $T \vdash \varphi$ and $T \vdash \neg \varphi$. Then $T \vdash (\varphi \land \neg \varphi)$, i.e., $\neg \text{Con}(T)$:

| $\varphi_1$: $\varphi$ | provable from $T$ by assumption |
| $\varphi_2$: $\neg \varphi$ | provable from $T$ by assumption |
| $\varphi_3$: $\varphi \rightarrow (\neg \varphi \rightarrow (\varphi \land \neg \varphi))$ | instance of $L_5$ |
| $\varphi_4$: $\neg \varphi \rightarrow (\varphi \land \neg \varphi)$ | from $\varphi_3$ and $\varphi_1$ by Modus Ponens |
| $\varphi_5$: $\varphi \land \neg \varphi$ | from $\varphi_4$ and $\varphi_2$ by Modus Ponens |

Notice that Proposition 3.2(a) implies that from an inconsistent set of axioms $T$ one can prove everything and $T$ would be completely useless. So, if we design a set of axioms $T$, we have to make sure that $T$ is consistent. However, as we shall see later, in many cases this task is impossible.

**Semantics: Models, Completeness, and Independence**

Let $T$ be any set of $\mathcal{L}$-formulae (for some language $\mathcal{L}$). There are two different ways to approach $T$, namely the syntactical and the semantical way. The above presented syntactical approach considers the set $T$ just as a set of well-formed formulae — regardless of their intended sense or meaning — from which we can prove some other formulae.

On the other hand, we can consider $T$ also from a semantical point of view by interpreting the symbols of the language $\mathcal{L}$ in a reasonable way, and then seeking for a model in which all formulae of $T$ are true. To be more precise, we first have to define how models are built and what “true” means:

Let $\mathcal{L}$ be an arbitrary but fixed language. An $\mathcal{L}$-structure $\mathfrak{A}$ consists of a (non-empty) set or collection $A$, called the domain of $\mathfrak{A}$, together with a mapping which assigns to each constant symbol $c \in \mathcal{L}$ an element $c^A$ of $A$, to each $n$-ary relation symbol $R \in \mathcal{L}$ a set of $n$-tuples $R^A$ of elements of $A$, and to each $n$-ary function symbol $F \in \mathcal{L}$ a function $F^A$ from $n$-tuples of $A$ to $A$. Further, the interpretation of variables is given by a so-called assignment: An assignment in an $\mathcal{L}$-structure $\mathfrak{A}$ is a mapping $j$ which assigns to each variable an element of the domain $A$. Finally, an $\mathcal{L}$-interpretation $I$ is a pair $(\mathfrak{A}, j)$ consisting of an $\mathcal{L}$-structure $\mathfrak{A}$ and an assignment $j$ in $\mathfrak{A}$. For a variable $x$, an element $a \in A$, and an assignment $j$ in $\mathfrak{A}$ we define the assignment $j^A_x$ by stipulating

$$j^A_x(y) = \begin{cases} a & \text{if } y = x, \\ j(y) & \text{otherwise}. \end{cases}$$

Further, for an interpretation $I = (\mathfrak{A}, j)$ let $I^A_x := (\mathfrak{A}, j^A_x)$. 

We associate with every interpretation $I = (\mathcal{A}, j)$ and every term $t$ an element $I(t)$ from the domain $A$ as follows:

- For a variable $x$ let $I(x) := j(x)$.
- For a constant symbol $c \in \mathcal{L}$ let $I(c) := c^A$.
- For an $n$-ary function symbol $F \in \mathcal{L}$ and terms $t_1, \ldots, t_n$ let
  \[
  I(F(t_1, \ldots, t_n)) := F^A(I(t_1), \ldots, I(t_n)) .
  \]

Now, we are able to define precisely the notion of a formula $\varphi$ being true under an interpretation $I = (\mathcal{A}, j)$, in which case we write $I \models \varphi$ and say that $\varphi$ holds in $I$. The definition is by induction on the complexity of the formula $\varphi$ (where it is enough to consider formulae containing—besides terms and relations—just the logical operators “¬” and “∧”, and the logical quantifier “∃”):

- If $\varphi$ is of the form $t_1 = t_2$, then
  \[I \models t_1 = t_2 \iff I(t_1) \text{ is the same element as } I(t_2) .\]
- If $\varphi$ is of the form $R(t_1, \ldots, t_n)$, then
  \[I \models R(t_1, \ldots, t_n) \iff (I(t_1), \ldots, I(t_n)) \text{ belongs to } R^A .\]
- If $\varphi$ is of the form $\neg \psi$, then
  \[I \models \neg \psi \iff \text{it is not the case that } I \models \psi .\]
- If $\varphi$ is of the form $\exists x \psi$, then
  \[I \models \exists x \psi \iff \text{there is an element } a \in A \text{ such that } I^a_2 \models \psi .\]
- If $\varphi$ is of the form $\psi_1 \land \psi_2$, then
  \[I \models \psi_1 \land \psi_2 \iff I \models \psi_1 \text{ and } I \models \psi_2 .\]

Notice that since the domain of $I$ is non-empty we always have $I \models \exists x(x = x)$.

Now, let $T$ be an arbitrary set of $\mathcal{L}$-formulae. Then an $\mathcal{L}$-structure $\mathcal{A}$ is a model of $T$ if for every assignment $j$ in $\mathcal{A}$ and for each formula $\varphi \in T$ we have $(\mathcal{A}, j) \models \varphi$, i.e., $\varphi$ holds in the $\mathcal{L}$-interpretation $I = (\mathcal{A}, j)$. We usually denote models by bold letters like $M, N, V$, et cetera. Instead of saying “$M$ is a model of $T$” we just write $M \models T$. If $\varphi$ fails in $M$, then we write $M \not\models \varphi$, which is equivalent to $M \models \neg \varphi$ (this is because for any $\mathcal{L}$-formula $\varphi$ we have either $M \models \varphi$ or $M \models \neg \varphi$).

For example $S_7$ (i.e., the set of all permutations of seven different items) is a model of $GT$, where the interpretation of the binary operation is composition and the neutral element is interpreted as the identity permutation. In this case, the elements of the domain of $S_7$ can be real and can even be heard, namely
when the seven items are seven bells and a peal of for example Stedman Triples consisting of all 5040 permutations of the seven bells is rung— which happens quite often, since Stedman Triples are very popular with change-ringers. However, the objects of models of mathematical theories usually do not belong to our physical world and are not more real than for example the number zero or the empty set.

The following two theorems, which we state without proofs, are the main connections between the syntactical and the semantical approach to first-order theories. On the one hand, the Soundness Theorem 3.3 just tells us that our deduction system is sound, i.e., if a sentence \( \varphi \) is provable from \( T \) then \( \varphi \) is true in each model of \( T \). On the other hand, Gödel’s Completeness Theorem 3.4 tells us that our deduction system is even complete, i.e., every sentence which is true in all models of \( T \) is provable from \( T \). As a consequence we get that \( T \vdash \varphi \) if and only if \( \varphi \) is true in each model of \( T \). In particular, if \( T \) is empty, this implies that every tautology (i.e., universally valid formula) is provable.

**Theorem 3.3 (Soundness Theorem).** Let \( T \) be a set of \( \mathcal{L} \)-sentences and let \( \varphi \) be any \( \mathcal{L} \)-sentence. If \( T \vdash \varphi \), then in any model \( M \) such that \( M \models T \) we have \( M \models \varphi \).

**Theorem 3.4 (Gödel’s Completeness Theorem).** Let \( T \) be a set of \( \mathcal{L} \)-sentences and let \( \varphi \) be any \( \mathcal{L} \)-sentence. Then \( T \vdash \varphi \) or there is a model \( M \) such that \( M \models T \cup \{ \lnot \varphi \} \). In other words, if for every model \( M \models T \) we have \( M \models \varphi \), then \( T \vdash \varphi \). (Notice that this does not imply the existence of a model of \( T \).)

One of the main consequences of Gödel’s Completeness Theorem 3.4 is that formal proofs—which are usually quite long and involved—can be replaced with informal ones: Let \( T \) be a consistent set of \( \mathcal{L} \)-formulae and let \( \varphi \) be any \( \mathcal{L} \)-sentence. Then, by Gödel’s Completeness Theorem 3.4, in order to show that \( T \vdash \varphi \) it is enough to show that \( M \models \varphi \) whenever \( M \models \Gamma \). In fact, we would take an arbitrary model \( M \) of \( T \) and show that \( M \models \varphi \).

As an example let us show that \( GT \vdash (y \cdot x = a) \rightarrow (x \cdot y = a) \): Firstly, let \( G \) be a model of \( GT \), with domain \( G \), and let \( x \) and \( y \) be any elements of \( G \). By \( GT \), we know that every element of \( G \) has a left-inverse. In particular, \( y \) has a left-inverse, say \( \hat{y} \), and we have \( \hat{y} \cdot y = a \). By \( GT \), we have \( x \cdot y = (\hat{y} \cdot y) \cdot (x \cdot y) \), and by \( GT \) we get \( (\hat{y} \cdot y) \cdot (x \cdot y) = \hat{y} \cdot ((\hat{y} \cdot y) \cdot y) \). Now, if \( y \cdot x = a \), then we have \( x \cdot y = \hat{y} \cdot y \) and consequently we get \( x \cdot y = a \). Notice that we tacitly used that the equality relation is symmetric and transitive.

We leave it as an exercise to the reader to find the corresponding formal proof of this basic result in Group Theory. In a similar way one can show that every left-neutral element is also a right-neutral element (called neutral element) and that there is just one neutral element in a group.

The following result, which is a consequence of Gödel’s Completeness Theorem 3.4, shows that every consistent set of formulae has a model.
Proposition 3.5. Let $T$ be any set of $L$-formulae. Then $\text{Con}(T)$ if and only if $T$ has a model.

Proof. ($\Rightarrow$) If $T$ has no model, then, by Gödel’s Completeness Theorem 3.4, for every $L$-formula $\varphi$ we have $T \vdash \varphi$ (otherwise, there would be a model of $T \cup \{\neg \varphi\}$, and in particular for $T$). So, for $\varphi$ being $\varphi \land \neg \varphi$ we get $T \vdash (\varphi \land \neg \varphi)$, hence $T$ is inconsistent.

($\Leftarrow$) If $T$ is inconsistent, then, by Proposition 3.2.(a), for every $L$-formula $\varphi$ we have $T \vdash \varphi$, in particular, $T \vdash \varphi \land \neg \varphi$. Now, the Soundness Theorem 3.3 implies that in all models $M \models T$ we have $M \models \varphi \land \neg \varphi$; thus, there are no models of $T$.

A set of sentences $T$ is usually called a theory. A consistent theory $T$ (in a certain language $L$) is said to be complete if for every $L$-sentence $\varphi$, either $T \vdash \varphi$ or $T \vdash \neg \varphi$. If $T$ is not complete, we say that $T$ is incomplete.

The following result is an immediate consequence of Proposition 3.5.

Corollary 3.6. Every consistent theory is contained in a complete theory.

Proof. Let $T$ be a theory in the language $L$. If $T$ is consistent, then it has a model, say $M$. Now let $\overline{T}$ be the set of all $L$-sentences $\varphi$ such that $M \models \varphi$. Obviously, $\overline{T}$ is a complete theory which contains $T$.

Let $T$ be a set of $L$-formulae and let $\varphi$ be any $L$-formula not contained in $T$. $\varphi$ is said to be consistent relative to $T$ (or that $\varphi$ is consistent with $T$) if $\text{Con}(T)$ implies $\text{Con}(T \cup \{\varphi\})$ (later we usually write $T + \varphi$ instead of $T \cup \{\varphi\}$). If both $\varphi$ and $\neg \varphi$ are consistent with $T$, then $\varphi$ is said to be independent of $T$. In other words, if $\text{Con}(T)$, then $\varphi$ is independent of $T$ if neither $T \vdash \varphi$ nor $T \vdash \neg \varphi$. By Gödel’s Completeness Theorem 3.4 we get that if $\text{Con}(T)$ and $\varphi$ is independent of $T$, then there are models $M_1$ and $M_2$ of $T$ such that $M_1 \models \varphi$ and $M_2 \models \neg \varphi$. A typical example of a statement which is independent of $GT$ is $\forall x \forall y (x \neq y \lor y \neq y)$ (i.e., the binary operation is commutative), and indeed, there are abelian as well as non-abelian groups.

In order to prove that a certain statement $\varphi$ is independent of a given (consistent) theory $T$, one could try to find two different models of $T$ such that $\varphi$ holds in one model and fails in the other. However, this task is quite difficult, in particular if one cannot prove that $T$ has a model at all (as it happens for Set Theory).

Limits of First-Order Logic

We begin this section with a useful result, called Compactness Theorem. On the one hand, it is just a consequence of the fact that formal proofs are finite (i.e., finite sequences of formulae). On the other hand, the Compactness Theorem is the main tool to prove that a certain sentence (or a set of sentences) is consistent with a given theory. In particular, the Compactness
Theorem 3.7 (Compactness Theorem). Let T be an arbitrary set of L-formulae. Then T is consistent if and only if every finite subset Φ of T is consistent.

Proof. Obviously, if T is consistent, then every finite subset Φ of T must be consistent. On the other hand, if T is inconsistent, then there is a formula φ such that T ⊨ φ ∧ ¬φ. In other words, there is a proof of φ ∧ ¬φ from T. Now, since every proof is finite, there are only finitely many formulae of T involved in this proof, and if Φ is this finite set of formulae, then Φ ⊨ φ ∧ ¬φ, which shows that Φ, a finite subset of T, is inconsistent.

A simple application of the Compactness Theorem 3.7 shows that if PA is consistent, then there is more than one model of PA (i.e., beside the intended model of natural numbers with domain ℕ, there are also so-called non-standard models of PA with larger domains):

Firstly we extend the language LPA = {0, s, +, ·} by adding a new constant symbol n. Secondly we extend PA by adding the formulae

\[ n \neq 0, \quad n \neq s(0), \quad n \neq s(s(0)), \ldots \]

and let Ψ be the set of these formulae. Now, if PA has a model N with domain say ℕ, and Φ is any finite subset of Ψ, then, by interpreting n in a suitable way, N is also a model of PA∪Ψ, which implies that PA∪Ψ is consistent. Thus, by the Compactness Theorem 3.7, PA∪Ψ is also consistent and therefore has a model, say N. Now, N \models PA∪Ψ, but since n is different from every standard natural number of the form s(s(...s(0)...)), the domain of N must be essentially different from N (since it contains a kind of infinite number, whereas all standard natural numbers are finite).

This example shows that we cannot axiomatise Peano Arithmetic in First-Order Logic in such a way that all the models we get have essentially the same domain ℕ.

By Proposition 3.5 we know that a set of first-order formulae T is consistent if and only if it has a model, i.e., there is a model M such that M \models T. So, in order to prove for example that the axioms of Set Theory are consistent we only have to find a single model in which all these axioms hold. However, as a consequence of the following theorems—which we state again without proof—this turns out to be impossible (at least if one restricts oneself to methods formalisable in Set Theory).

Theorem 3.8 (Gödel's Incompleteness Theorem). Let T be a consistent set of first-order L-formulae which is sufficiently strong to define the concept of natural numbers and to prove certain basic arithmetical facts (e.g.,
PA is such a theory, but also slightly weaker theories would suffice). Then there is always an $\mathcal{L}$-sentence $\varphi$ which is independent of $T$, i.e., neither $T \vdash \varphi$ nor $T \vdash \neg \varphi$ (or in other words, there are models $M_1$ and $M_2$ of $T$ such that $M_1 \models \varphi$ and $M_2 \models \neg \varphi$).

In particular we get that there are number-theoretic statements which can neither be proved nor disproved in PA (i.e., the theory PA is incomplete). Moreover, the following consequence of Gödel’s Incompleteness Theorem 3.4 shows that even the consistency of PA can be proved with number-theoretical methods.

**Theorem 3.9 (Gödel’s Second Incompleteness Theorem).** Let $T$ be a set of first-order $\mathcal{L}$-formulae. Then the statement $\text{Con}(T)$, which says that $T \not\vdash \varphi \land \neg \varphi$ for some $\mathcal{L}$-formula $\varphi$, can be formulated as a number-theoretic sentence $\text{Con}_T$. Now, if $T$ is consistent and is sufficiently strong to define the concept of natural numbers and to prove certain basic arithmetical facts, then $T \not\vdash \text{Con}_T$, i.e., $T$ cannot prove its own consistency. In particular, $PA \not\vdash \text{Con}_PA$.

On the one hand, Gödel’s Incompleteness Theorem tells us that in any theory $T$ which is sufficiently strong, there are always statements which are independent of $T$ (i.e., which can neither be proved nor disproved in $T$). On the other hand, statements which are independent of a given theory (e.g., of Set Theory or of Peano Arithmetic) are often very interesting, since they say something unexpected, but in a language we can understand. From this point of view it is good to have Gödel’s Incompleteness Theorem which guarantees the existence of such statements in theories like Set Theory or Peano Arithmetic.

In Part II we shall present a technique with which we can prove the independence of certain set-theoretical statements from the axioms of Set Theory, which are introduced and discussed below.

**The Axioms of Zermelo-Fraenkel Set Theory**

In 1905, Zermelo began to axiomatise Set Theory and in 1908 he published his first axiomatic system consisting of seven axioms. In 1922, Fraenkel and Skolem independently improved and extended Zermelo’s original axiomatic system, and the final version was presented again by Zermelo in 1930. In this chapter we give the resulting axiomatic system called Zermelo-Fraenkel Set Theory, denoted ZF, which contains all axioms of Set Theory except the Axiom of Choice, which will be introduced and discussed in Chapter 5. Alongside the axioms of Set Theory we develop the theory of ordinals and give various notations which will be used throughout this book.

The language of Set Theory contains only one non-logical symbol, namely the binary **membership relation**, denoted by $\in$, and there exists just one
type of objects, namely sets. In other words, every object in the domain is a set and there are no other objects than sets. However, to make life easier, instead of \( \in (a, b) \) we write \( a \in b \) (or on rare occasions also \( b \ni a \)) and say that “\( a \) is an element of \( b \)”, or that “\( a \) belongs to \( b \)”. Later we will extend the language of Set Theory by defining some constants (like “\( \emptyset \)” and “\( \omega \)”), relations (like “\( \subseteq \)”), and operations (like the power set operation “\( \mathcal{P} \)”), but in fact, all that can be formulated in Set Theory, can be written as a formula containing only the non-logical relation “\( \in \)” (but for obvious reasons, we will usually not do so).

0. The Axiom of Empty Set

\[ \exists x \forall z (z \notin x) \]

This axiom not only postulates the existence of a set without any elements, i.e., an empty set, it also shows that the set-theoretic universe is non-empty, because it contains at least an empty set (of course, the logical axioms L_{16} and L_{13} already incorporate this fact).

1. The Axiom of Extensionality

\[ \forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y) \]

This axiom says that any sets \( x \) and \( y \) having the same elements are equal. Notice that the converse — which is \( x = y \) implies that \( x \) and \( y \) have the same elements — is just a consequence of the logical axiom L_{17}.

The Axiom of Extensionality also shows that the empty set, postulated by the Axiom of Empty Set, is unique. For assume that there are two empty sets \( x_0 \) and \( x_1 \), then we have \( \forall z (z \notin x_0 \land z \notin x_1) \), which implies that \( \forall z (z \in x_0 \leftrightarrow z \in x_1) \), and therefore, \( x_0 = x_1 \).

Let us introduce the following notation: If \( \varphi(x) \) is any first-order formula with free variable \( x \) (i.e., \( x \) occurs at a particular place in the formula \( \varphi \) where it is not in the range of any logical quantifier), then

\[ \exists x \varphi(x) \iff \exists x (\varphi(x) \land \forall z (\varphi(z) \rightarrow z = x)) \]

With this definition we can reformulate the Axiom of Empty Set as follows:

\[ \exists x \forall z (z \notin x) \]

and this unique empty set is denoted by \( \emptyset \).

We say that \( y \) is a subset of \( x \), denoted \( y \subseteq x \), if \( \forall z (z \in y \rightarrow z \in x) \). Notice that the empty set is a subset of every set. If \( y \) is a proper subset of \( x \), i.e., \( y \subseteq x \) and \( y \neq x \), then this is sometimes denoted by \( y \subset x \).

One of the most important concepts in Set Theory is the notion of ordinal number, which can be seen as a transfinite extension of the natural numbers.
In order to define the concept of ordinal numbers, we have to give first some definitions: Let \( z \in x \). Then \( z \) is called an \( \in \)-minimal element of \( x \), if 
\[
\forall y (y \notin z \lor y \notin x),
\]
or equivalently, \( \forall y (y \in z \rightarrow y \notin x) \). A set \( x \) is ordered 
by \( \in \) if for any sets \( y_1, y_2 \in x \) we have \( y_1 \in y_2 \) or \( y_1 = y_2 \) or \( y_1 \not\in y_2 \), 
but we do not require the three cases to be mutually exclusive. Now, a set 
\( x \) is called well-ordered by \( \in \) if it is ordered by \( \in \) and every non-empty 
subset of \( x \) has an \( \in \)-minimal element. Further, a set \( x \) is called transitive 
if \( \forall y (y \in x \rightarrow y \subseteq x) \). Notice that if \( x \) is transitive and \( z \in y \in x \), then this 
implies \( z \in x \). A set is called an ordinal number, or just an ordinal, if 
it is transitive and well-ordered by \( \in \). Ordinal numbers are usually denoted by 
Greek letters like \( \alpha, \beta, \gamma, \lambda \), \textit{et cetera}, and the collection of all ordinal numbers 
is denoted by \( \Omega \). We will see later, when we know more properties of ordinals, 
that \( \Omega \) is not a set. However, we can consider “\( \alpha \in \Omega \)” just as an abbreviation 
for “\( \alpha \) is an ordinal”, and thus, there is no harm in using the symbol \( \Omega \) in this 
way, even though \( \Omega \) is not an object of the set-theoretic universe.

**Fact 3.10.** If \( \alpha \in \Omega \), then either \( \alpha = \emptyset \) or \( \emptyset \in \alpha \).

**Proof.** Since \( \alpha \in \Omega \), \( \alpha \) is well-ordered by \( \in \). Thus, either \( \alpha = \emptyset \), or, since 
\( \alpha \subseteq \alpha \), \( \alpha \) contains an \( \in \)-minimal element, say \( x_0 \). Now, by transitivity of \( \alpha \), 
for all \( z \in x_0 \) we have \( z \in \alpha \), and since \( x_0 \in \alpha \)-minimal we get \( x_0 = \emptyset \). \( \Box \)

Notice that until now, we cannot prove the existence of any ordinal — or even 
of any set — beside the empty set, postulated by the Axiom of Empty Set. This 
will change with the following axiom.

2. **The Axiom of Pairing**

\[
\forall x \forall y \exists u (u = \{x, y\})
\]

where \( \{x, y\} \) denotes the set which contains just the elements \( x \) and \( y \). In 
order to write this axiom in a more formal way, let us introduce the following 
notation: If \( \varphi(z) \) is any first-order formula with free variable \( z \), and \( x \) is any 
set, then

\[
\forall z \in x (\varphi(z)) \iff \forall z ((z \in x) \rightarrow \varphi(z)),
\]

and similarly

\[
\exists z \in x (\varphi(z)) \iff \exists z ((z \in x) \land \varphi(z)).
\]

More formally the Axiom of Pairing reads as follows:

\[
\forall x \forall y \exists u (x \in u \land y \in u \land \forall z \in u (z = x \lor z = y))
\]

If in the above formula we set \( x = y \), then \( u = \{x, x\} \), which is, by the 
Axiom of Extensionality, the same as \( \{x\} \). Thus, by the Axiom of Pairing, if 
\( x \) is a set, then also \( \{x\} \) is a set. Starting with \( \emptyset \), an iterated application of 
the Axiom of Pairing yields for example the sets \( \emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots \).
Union

and \( \{\emptyset, \emptyset\}, \{\emptyset, \emptyset, \emptyset\}, \ldots \). Among these sets, \( \emptyset, \{\emptyset\}, \) and \( \{\emptyset, \emptyset\} \) are ordinals, but for example \( \{\emptyset\} \) is not an ordinal.

So far, we did not exclude the possibility that a set may be an element of itself, and in fact, we need the Axiom of Foundation in order to do so. However, we can already show that no ordinal is an element of itself:

**Fact 3.11.** If \( \alpha \in \Omega \), then \( \alpha \not\in \alpha \).

*Proof.* Assume towards a contradiction that \( \alpha \in \alpha \). Then \( \{\alpha\} \) is a non-empty subset of \( \alpha \) and therefore contains an \( \epsilon \)-minimal element. Now, since \( \{\alpha\} \) just contains the element \( \alpha \), the \( \epsilon \)-minimal element of \( \{\alpha\} \) must be \( \alpha \), but on the other hand, \( \alpha \in \alpha \) implies that \( \alpha \) is not \( \epsilon \)-minimal, a contradiction. \( \square \)

For any sets \( x \) and \( y \), the Axiom of Extensionality implies that \( \{x, y\} = \{y, x\} \). So, it does not matter in which order the elements of a 2-element set are written down. However, with the Axiom of Pairing we can easily define ordered pairs, denoted \( \langle x, y \rangle \), as follows:

\[
\langle x, y \rangle = \{\{x\}, \{x, y\}\}
\]

Notice that \( \langle x, y \rangle = \langle x', y' \rangle \) iff \( x = x' \) and \( y = y' \), and further notice that this definition also makes sense in the case when \( x = y \) — at least as long as we know that \( \{\{x\}\} \) is supposed to denote an ordered pair. By a similar trick, one can also define ordered triples by stipulating for example \( \langle x, y, z \rangle := \langle x, \langle y, z \rangle \rangle \), ordered quadruples, et cetera, but the notation becomes hard to read and it requires additional methods to distinguish for example between ordered pairs and ordered triples. However, when we have more axioms at hand we can define arbitrary tuples more elegantly.

### 3. The Axiom of Union

\[
\forall x \exists u \forall z (z \in u \iff \exists w \in x (z \in w))
\]

More informally, for all sets \( x \) there exists the union of \( x \), denoted \( \bigcup x \), consisting of all sets which belong to a member of \( x \).

For sets \( x \) and \( y \), let \( x \cup y := \bigcup \{x, y\} \) denote the **union** of \( x \) and \( y \). Notice that \( x = \bigcup \{x\} \). For \( x \cup y \), where \( x \) and \( y \) are disjoint (i.e., do not have any common elements) we sometimes write \( x \cup y \), and for \( x = \{y_i : i \in I\} \) we sometimes write \( \bigcup_{i \in I} y_i \) instead of \( \bigcup x \).

Now, with the Axiom of Union and the Axiom of Pairing, and by stipulating \( x+1 := x \cup \{x\} \), we can for example build the following sets (which are in fact ordinals): \( 0 := \emptyset \), \( 1 := 0+1 = 0 \cup \{0\} = \{0\} \), \( 2 := 1+1 = 1 \cup \{1\} = \{0, 1\} \), \( 3 := 2+1 = 2 \cup \{2\} = \{0, 1, 2\} \), and so on. In particular, if a set \( x \) of this type is already defined, we get that \( x+1 = \{0, 1, 2, \ldots, x\} \). This construction leads to the following definition:

A set \( x \) such that \( \forall y (y \in x \rightarrow (y \cup \{y\}) \in x) \) is called **inductive**. On the one
hand, $\emptyset$ is inductive. On the other hand, we cannot prove the existence of a non-empty inductive set without the aid of the following axiom.

4. **The Axiom of Infinity**

$$\exists I (\emptyset \in I \land \forall y \in I ((y \cup \{y\}) \in I))$$

More informally, the Axiom of Infinity postulates the existence of a non-empty inductive set containing $\emptyset$. All the sets $0, 1, 2, \ldots$ constructed above — which we recognise as natural numbers — must belong to every inductive set and in fact, the “smallest” inductive set contains just these sets.

5. **The Axiom Schema of Separation**

For each first-order formula $\varphi(z, p_1, \ldots, p_n)$ with $\text{free}(\varphi) \subseteq \{z, p_1, \ldots, p_n\}$, the following formula is an axiom:

$$\forall x \forall p_1 \ldots \forall p_n \exists y \forall z (z \in y \leftrightarrow (z \in x \land \varphi(z, p_1, \ldots, p_n)))$$

Informally, for each set $x$ and every first-order formula $\varphi(z)$, $\{z \in x : \varphi(z)\}$ is a set.

One can think of the sets $p_1, \ldots, p_n$ as parameters of $\varphi$, which are usually some fixed sets. For example for $\varphi(z, p) \equiv z \in p$ we get that for any sets $x$ and $p$ there exists a set $y$ such that $z \in y \leftrightarrow (z \in x \land z \in p)$. In other words, for any sets $x_0$ and $x_1$, the collection of all sets which belong to both, $x_0$ and $x_1$, is a set. This set is called the intersection of $x_0$ and $x_1$ and is denoted by $x_0 \cap x_1$. In general, for non-empty sets $x$ we define

$$\bigcap x = \{z \in \bigcup x : \forall y \in x (z \in y)\}$$

which is the intersection of all sets which belong to $x$. (In order to see that $\bigcap x$ is a set, let $\varphi(z, x) \equiv \forall y \in x (z \in y)$ and apply the Axiom Schema of Separation to $\bigcup x$.) Notice also that $x \cap y = \bigcap \{x, y\}$. Furthermore, for $x = \{y_i : i \in I\}$ we sometimes write $\bigcap_{i \in I} y_i$ instead of $\bigcap x$. Another example is when $\varphi(z, p) \equiv z \not\in p$. In this case, for $p = y$, we get that $\{z \in x : z \not\in y\}$ is a set, denoted $x \setminus y$, which is called the set-theoretic difference of $x$ and $y$.

Let us now turn back to ordinal numbers:

**Theorem 3.12.** (a) If $\alpha, \beta \in \Omega$, then $\alpha \in \beta$ or $\alpha = \beta$ or $\alpha \supset \beta$, where these three cases are mutually exclusive.

(b) If $\alpha \in \beta \in \Omega$, then $\alpha \in \Omega$.

(c) If $\alpha \in \Omega$, then also $(\alpha \cup \{\alpha\}) \in \Omega$.

(d) $\Omega$ is transitive and is well-ordered by $\in$, or more precisely, $\Omega$ is transitive, is ordered by $\in$, and every non-empty class $C \subseteq \Omega$ has an $\in$-minimal element.
Proof. (a) Firstly, notice that by Fact 3.11 the three cases \( \alpha \in \beta, \alpha = \beta, \alpha \ni \beta \) are mutually exclusive.

Let \( \alpha, \beta \in \Omega \) be given. If \( \alpha = \beta \), then we are done. So, let us assume that \( \alpha \neq \beta \). Without loss of generality we may assume that \( \alpha \setminus \beta \neq \emptyset \).

We first show that \( \alpha \cap \beta \) is the \( \varepsilon \)-minimal element of \( \alpha \setminus \beta \). Let \( \gamma \) be an \( \varepsilon \)-minimal element of \( \alpha \setminus \beta \). Since \( \alpha \) is transitive and \( \gamma \in \alpha \), \( \forall u \in \gamma \to u \in \alpha \), and since \( \gamma \) is an \( \varepsilon \)-minimal element of \( \alpha \setminus \beta \), \( \forall u \in \gamma \to u \in \beta \), which implies \( \gamma \subseteq \alpha \cap \beta \). On the other hand, if there is a \( w \in (\alpha \cap \beta) \setminus \gamma \), then, since \( \alpha \) is ordered by \( \in \) and \( \gamma \neq w \) (\( \gamma \notin \beta \ni w \)), we must have \( \gamma \in w \), and since \( \beta \) is transitive and \( w \in \beta \), this implies that \( \gamma \in \beta \), which contradicts the fact that \( \gamma \in (\alpha \setminus \beta) \). Hence, \( \gamma = \alpha \cap \beta \) is the \( \varepsilon \)-minimal element of \( \alpha \setminus \beta \). Now, if also \( \beta \setminus \alpha \neq \emptyset \), then we would get that \( \alpha \cap \beta \) is also the \( \varepsilon \)-minimal element of \( \beta \setminus \alpha \), which is obviously a contradiction.

Thus, \( \alpha \setminus \beta \neq \emptyset \) implies that \( \beta \setminus \alpha = \emptyset \), or in other words, \( \beta \subseteq \alpha \), which is the same as saying \( \beta = \alpha \cap \beta \). Consequently we get that \( \beta \) is the \( \varepsilon \)-minimal element of \( \alpha \setminus \beta \), in particular, \( \beta \in \alpha \).

(b) Let \( \alpha \in \beta \in \Omega \). Since \( \beta \) is transitive, \( \alpha \) is ordered by \( \in \). So, it remains to show that \( \alpha \) is transitive and well-ordered by \( \in \).

well-ordered by \( \in \): Because \( \beta \) is transitive, every subset of \( \alpha \) is also a subset of \( \beta \) and consequently contains an \( \varepsilon \)-minimal element.

transitive: Let \( \delta \in \gamma \in \alpha \). We have to show that \( \delta \in \alpha \). Since \( \beta \) is transitive, \( \delta \in \beta \), and since \( \beta \) is ordered by \( \in \), we have either \( \delta \in \alpha \) or \( \delta = \alpha \) or \( \alpha \in \delta \). If \( \delta \in \alpha \), we are done, and if \( \delta = \alpha \) or \( \alpha \in \delta \), then the set \( \{\alpha, \gamma, \delta\} \subseteq \beta \) does not have an \( \varepsilon \)-minimal element, which contradicts the fact that \( \beta \) is well-ordered by \( \in \).

(c) We have to show that \( \alpha \cup \{\alpha\} \) is transitive and well-ordered by \( \in \).

transitive: If \( \beta \in (\alpha \cup \{\alpha\}) \), then either \( \beta \in \alpha \) or \( \beta = \alpha \), and in both cases we have \( \beta \subseteq (\alpha \cup \{\alpha\}) \).

well-ordered by \( \in \): Since \( \alpha \) is an ordinal, \( \alpha \cup \{\alpha\} \) is ordered by \( \in \). Let now \( x \subseteq (\alpha \cup \{\alpha\}) \) be a non-empty set. If \( x = \{\alpha\} \), then \( \alpha \) is obviously an \( \varepsilon \)-minimal element of \( x \). Otherwise, \( x \cap \alpha \neq \emptyset \), and since \( \alpha \in \Omega \), \( x \cap \alpha \) has an \( \varepsilon \)-minimal element, say \( \gamma \). Since \( \alpha \) is transitive we have \( x \cap \gamma = \emptyset \) (otherwise, \( \gamma \) would not be \( \varepsilon \)-minimal in \( x \cap \alpha \)), which implies that \( \gamma \) is \( \varepsilon \)-minimal in \( \gamma \).

(d) \( \Omega \) is transitive and ordered by \( \in \): This is part (b) and part (a) respectively. \( \Omega \) is well-ordered by \( \in \): Let \( C \subseteq \Omega \) be a non-empty class of ordinals. If \( C = \{\alpha\} \) for some \( \alpha \in \Omega \), then \( \alpha \) is the \( \varepsilon \)-minimal element of \( C \). Otherwise, \( C \) contains an ordinal \( \delta_0 \) such that \( \delta_0 \cap C \neq \emptyset \) and let \( x := \delta_0 \cap C \). Then \( x \) is a non-empty set of ordinals. Now, let \( \alpha \in x \) and let \( \gamma \) be an \( \varepsilon \)-minimal element of \( x \cap (\alpha \cup \{\alpha\}) \). By definition, \( \gamma \in (\alpha \cup \{\alpha\}) \), and since \( (\alpha \cup \{\alpha\}) \in \Omega \), \( \gamma \subseteq (\alpha \cup \{\alpha\}) \). Thus, every ordinal \( \gamma' \in \gamma \) belongs to \( \alpha \cup \{\alpha\} \), but by the definition of \( \gamma \), \( \gamma' \) cannot belong to \( x \cap (\alpha \cup \{\alpha\}) \), which implies that \( \gamma \) is also \( \varepsilon \)-minimal in \( x \), and consequently in \( C \).

\( \Box \)
By **Theorem 3.12.(d)** we get that $\Omega$ is transitive and well-ordered by $\epsilon$. Thus, if $\Omega$ would be a set, $\Omega$ would be an ordinal number and therefore would belong to itself, but this is a contradiction to **Fact 3.11**.

In general, a collection of sets, satisfying for example a certain formula, which is not necessarily a set is called a **class**. For example $\Omega$ is a class which is not a set (it consists of all transitive sets which are well-ordered by $\epsilon$). Even though proper classes (i.e., classes which are not sets) do not belong to the set-theoretic universe, it is sometimes convenient to handle them like sets, e.g., taking intersections or extracting certain subsets or subclasses from them.

By **Theorem 3.12.(c)** we know that if $\alpha \in \Omega$, then also $(\alpha \cup \{\alpha\}) \in \Omega$. Now, for ordinals $\alpha \in \Omega$ let $\alpha + 1 := \alpha \cup \{\alpha\}$. Part (a) of the following result — which is just a consequence of **Theorem 3.12** — motivates this notation.

**Corollary 3.13.** (a) If $\alpha, \beta \in \Omega$ and $\alpha \in \beta$, then $\alpha + 1 \subseteq \beta$. In other words, $\alpha + 1$ is the least ordinal which contains $\alpha$.

(b) For every ordinal $\alpha \in \Omega$ we have either $\alpha = \bigcup \alpha$ or there exists $\beta \in \Omega$ such that $\alpha = \beta + 1$.

**Proof.** (a) Assume $\alpha \in \beta$, then $\{\alpha\} \subseteq \beta$, and since $\beta$ is transitive, we also have $\alpha \subseteq \beta$; thus, $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq \beta$.

(b) Since $\alpha$ is transitive, $\bigcup \alpha \subseteq \alpha$. Thus, if $\alpha \neq \bigcup \alpha$, then $\alpha \setminus \bigcup \alpha \neq \emptyset$. Let $\beta$ be $\epsilon$-minimal in $\alpha \setminus \bigcup \alpha$. Then $\beta \in \alpha$ and $\beta + 1 \in \Omega$, and by part (a) we have $\beta + 1 \subseteq \alpha$. On the one hand, $\alpha \in \beta + 1$ would imply that $\alpha \in \alpha$, a contradiction to **Fact 3.11**. On the other hand, $\beta + 1 \in \alpha$ would imply that $\beta \in \bigcup \alpha$, which contradicts the choice of $\beta$. Thus, we must have $\beta + 1 = \alpha$. \hfill \Box

This leads to the following definitions: An ordinal $\alpha$ is called a **successor ordinal** if there exists an ordinal $\beta$ such that $\alpha = \beta + 1$; otherwise, it is called a **limit ordinal**. In particular, $\emptyset$ (or equivalently $\emptyset$) is a limit ordinal.

We are now ready to define the set of **natural numbers** $\omega$, which will turn out to be the least non-empty limit ordinal. By the Axiom of Infinity we know that there exists an inductive set $I$. Below we show that there exists also a smallest inductive set. For this, let $I_\Omega = I \cap \Omega$; more precisely,

$$I_\Omega = \{\alpha \in I : \alpha \text{ is an ordinal}\}.$$ 

Then $I_\Omega$ is a set of ordinals and by **Theorem 3.12.(c)**, $I_\Omega$ is even an inductive set. Now, if there exists no $\alpha \in I_\Omega$ such that $\alpha$ is non-empty and inductive, let $\omega := I_\Omega$, otherwise, define

$$\omega = \bigcap \{\alpha \in I_\Omega : \emptyset \in \alpha \text{ and } \alpha \text{ is inductive}\}.$$ 

By definition, $\emptyset \in \omega$ and for all $\beta \in \omega$ we have $\beta + 1 \in \omega$, i.e., $\omega$ is inductive and contains $\emptyset$. In particular, $\bigcup \omega = \omega$, which shows that $\omega$ is a limit ordinal. Again by definition, $\omega$ does not properly contain any inductive set which contains $\emptyset$. 
In particular, \( \omega \) does not contain any limit ordinal other than \( \emptyset \) (since such an ordinal would be an inductive set containing \( \emptyset \)), and therefore, \( \omega \) is the smallest non-empty limit ordinal.

The ordinals belonging to \( \omega \) are called **natural numbers**. One can also define natural numbers inductively as we have done above: \( 0 := \emptyset \), and for any natural number \( n \), \( n + 1 := n \cup \{ n \} = \{ 0, 1, 2, \ldots, n \} \). Notice that each natural number \( n \) is the set \( \{ k \in \omega : k < n \} \), where \( k < n \iff k \in n \). Further notice that since \( \omega \) is the smallest non-empty limit ordinal, all natural numbers except \( 0 \) are successor ordinals. Now, a set \( A \) is called **finite** if there exists a bijection between \( A \) and a natural number \( n \in \omega \); otherwise, \( A \) is called **infinite**. Thus, all natural numbers are finite and \( \omega \) is the smallest infinite (i.e., not finite) ordinal number.

The following theorem is a consequence of the fact that \( \Omega \) is transitive and well-ordered by \( \in \) (which is just Theorem 3.12(d)).

**Theorem 3.14 (Transfinite Induction Theorem).** Let \( C \subseteq \Omega \) be a class of ordinals and assume that:

(a) if \( \alpha \in C \), then \( \alpha + 1 \in C \),
(b) if \( \alpha \) is a limit ordinal and \( \forall \beta \in \alpha (\beta \in C) \), then \( \alpha \in C \).

Then \( C \) is the class of all ordinals. (Notice that by (b) we have \( 0 \in C \), in particular, \( C \neq \emptyset \).)

**Proof.** Assume towards a contradiction that \( C \neq \Omega \) and let \( \alpha_0 \) be the \( \in \)-minimal ordinal which does not belong to \( C \) (such an ordinal exists by Theorem 3.12(d)). Now, \( \alpha_0 \) can be *neither* a successor ordinal, since this would contradict (a), *nor* a limit ordinal, since this would contradict (b). Thus, \( \alpha_0 \) does not exist which implies that \( \Omega \setminus C = \emptyset \), i.e., \( C = \Omega \). \( \dashv \)

The following result is just a reformulation of the Transfinite Induction Theorem.

**Corollary 3.15.** For any first-order formula \( \varphi(x) \) with free variable \( x \) we have

\[
\forall \alpha \in \Omega \left( \forall \beta \in \alpha (\varphi(\beta)) \rightarrow \varphi(\alpha) \right) \rightarrow \forall \alpha \in \Omega (\varphi(\alpha)).
\]

**Proof.** Let \( C \subseteq \Omega \) be the class of all ordinals \( \alpha \in \Omega \) such that \( \varphi(\alpha) \) holds and apply the Transfinite Induction Theorem 3.14. \( \dashv \)

When some form of Corollary 3.15 is involved we usually do not mention the corresponding formula \( \varphi \) and just say “by induction on . . .” or “by transfinite induction”.

**6. The Axiom of Power Set**

\[
\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)
\]
Informally, the **Axiom of Power Set** states that for each set $x$ there is a set $\mathcal{P}(x)$, called the **power set** of $x$, which consists of all subsets of $x$.

With the Axiom of Power Set (and other axioms like the Axiom of Union or the Axiom Schema of Separation) we can now define notions like functions, relations, and sequences: Let $A$ and $B$ be arbitrary sets. Then

$$A \times B = \{ (x, y) : x \in A \land y \in B \}$$

where $\langle x, y \rangle = \{ \{ x \}, \{ x, y \} \}$; thus, $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$. Further, let

$$A^B = \{ f \subseteq A \times B : \forall x \in A \exists y \in B (\langle x, y \rangle \in f) \}.$$

An element $f \in A^B$, usually denoted by $f : A \to B$, is called a **function** or **mapping** from $A$ to $B$, where $A$ is called the **domain** of $f$, denoted dom$(f)$.

For $f : A \to B$ we usually write $f(x) = y$ instead of $\langle x, y \rangle \in f$. If $S$ is a set, then the **image** of $S$ under $f$ is denoted by $f[S] = \{ f(x) : x \in S \}$ and $f|_S = \{ (x, y) \in f : x \in S \}$ is the restriction of $f$ to $S$. Furthermore, for a function $f : A \to B$, $f[A]$ is called the **range** of $f$, denoted ran$(f)$.

A function $f : A \to B$ is **surjective**, or onto, if $\forall y \in B \exists x \in A (f(x) = y)$. We sometimes emphasise the fact that $f$ is surjective by writing $f : A \twoheadrightarrow B$.

A function $f : A \to B$ is **injective**, also called one-to-one, if we have $\forall x_1 \in A \forall x_2 \in A (f(x_1) = f(x_2) \rightarrow x_1 = x_2)$. To emphasise the fact that $f$ is injective we sometimes write $f : A \hookrightarrow B$.

A function $f : A \to B$ is **bijective** if it is injective and surjective. If $f : A \to B$ is bijective, then

$$\forall y \in B \exists! x \in A ((x, y) \in f)$$

and therefore,

$$f^{-1} := \{ (y, x) : (x, y) \in f \} \in B^A$$

is a function which is even bijective. So, if there is a bijective function from $A$ to $B$, then there is also one from $B$ to $A$ and we sometimes just say that there is a **bijection** between $A$ and $B$. Notice that if $f : A \to B$ is injective, then $f$ is a bijection between $A$ and $f[A]$.

Let $x$ be any non-empty set and assume that for each $i \in x$ we have assigned a set $A_i$. For $A = \bigcup_{i \in x} A_i$, where $\bigcup_{i \in x} A_i := \bigcup\{ A_i : i \in x \}$, the set

$$\prod_{i \in x} A_i = \{ f \in {}^xA : \forall i \in x (f(i) \in A_i) \}$$

is called the **Cartesian product** of the sets $A_i (i \in x)$. Notice that if all sets $A_i$ are equal to a given set $A$, then $\prod_{i \in x} A_i = {}^xA$. If $x = n$ for some $n \in \omega$, in abuse of notation we also write $A^n$ instead of $^nA$ by identifying $^nA$ with the set

$$A^n = A \times \ldots \times A \quad {n \text{-times}}.$$
Similarly, for \( \alpha \in \Omega \) we sometimes identify a function \( f \in \mathcal{P} \) with the sequence \((f(0), f(1), \ldots, f(\beta), \ldots)_\alpha\) of length \( \alpha \), and vice versa. Sequences (of length \( \alpha \)) are usually denoted by using angled brackets (and by using \( \alpha \) as a subscript), e.g., \((s_0, \ldots, s_\beta, \ldots)_\alpha\) or \((s_\beta : \beta < \alpha)\).

For any set \( A \) and any \( n \in \omega \), a set \( R \subseteq A^n \) is called an \textit{n-ary relation} on \( A \). If \( n = 2 \), then \( R \subseteq A \times A \) is also called a \textit{binary relation}. A binary relation \( R \) on \( A \) is a \textbf{well-ordering} of \( A \), if there is an ordinal \( \alpha \in \Omega \) and a bijection \( f : A \rightarrow \alpha \) such that

\[
R(x, y) \iff f(x) \in f(y).
\]

For any set \( A \), let \( \text{seq}(A) \) be the set of all finite sequences which can be formed with elements of \( A \), or more formally:

\[
\text{seq}(A) = \bigcup_{n \in \omega} A^n.
\]

Furthermore, let \( \text{seq}^{-1}(A) \) be those sequences of \( \text{seq}(A) \) in which no element appears twice. Again more formally, this reads as follows:

\[
\text{seq}^{-1}(A) = \{ \sigma \in \text{seq}(A) : \sigma \text{ is injective} \}
\]

The last notion we introduce in this section is the notion of cardinality: Two sets \( A \) and \( B \) are said to have the same \textbf{cardinality}, denoted \( |A| = |B| \), if there is a bijection between \( A \) and \( B \). Notice that cardinality equality is an equivalence relation. For example \( |\omega \times \omega| = |\omega| \), e.g., define the bijection \( f : \omega \times \omega \rightarrow \omega \) by stipulating \( f((n, m)) = m + \frac{1}{2} \cdot (n + m)(n + m + 1) \).

If \( |A| = |B'| \) for some \( B' \subseteq B \), then the cardinality of \( A \) is less than or equal to the cardinality of \( B \), denoted \( |A| \leq |B| \). Notice that \( |A| \leq |B| \) if there is an injection from \( A \) into \( B \). Finally, if \( |A| \neq |B| \) but \( |A| \leq |B| \), then cardinality of \( A \) is said to be strictly less than the cardinality of \( B \), denoted \( |A| < |B| \). Notice that the relation “\( \leq \)” is reflexive and transitive. The notation suggests that \( |A| \leq |B| \) and \( |B| \leq |A| \) implies \( |A| = |B| \). This is indeed the case and a consequence of the following result.

**Lemma 3.16.** Let \( A_0, A_1, A \) be sets such that \( A_0 \subseteq A_1 \subseteq A \). If \( |A| = |A_0| \), then \( |A| = |A_1| \).

**Proof.** If \( A_1 = A \) or \( A_1 = A_0 \), then the statement is trivial. So, let us assume that \( A_0 \subseteq A_1 \subseteq A \) and let \( C = A \setminus A_1 \), i.e., \( A \setminus C = A_1 \). Further, let \( f : A \rightarrow A_0 \) be a bijection and define \( g : \mathcal{P}(A) \rightarrow \mathcal{P}(A_0) \) by stipulating \( g(D) := f[D] \). Let \( \varphi(z, p_1, p_2, p_3) \) be the following formula:

\[
z \in p_1 \land (0, p_2) \in z \land \forall n \in \omega \exists u \exists v (\langle n, u \rangle \in z \land \langle u, v \rangle \in p_3 \land \langle n + 1, v \rangle \in z)
\]

By the Axiom Schema of Separation, for \( x = p_1 = \omega \mathcal{P}(A) \), \( p_2 = C \), and \( p_3 = g \), there exists a set \( y \) such that \( z \in y \iff (z \in \omega \mathcal{P}(A) \land \varphi(z, \omega \mathcal{P}(A), C, g)) \). By induction on \( n \) and by assembling the various partial functions produced
by the induction into a single function, one gets that \( y \) contains just a single function, say \( z_0 : \omega \rightarrow \mathcal{P}(A) \). In fact, \( z_0(0) = C \) and for all \( n \in \omega \) we have \( z_0(n+1) = f[z_0(n)] \). Now, let
\[
\bar{C} = \bigcup \{ z_0(n) : n \in \omega \}
\]
and define the function \( \bar{f} : A \rightarrow A \) by stipulating
\[
\bar{f}(x) = \begin{cases} 
  f(x) & x \in \bar{C}, \\
  x & \text{otherwise.}
\end{cases}
\]
By definition of \( \bar{f} \) and since \( f \) is a bijection which maps \( C \) into \( A_0 \), \( \bar{f}[C] = C \setminus C \). Moreover, the function \( \bar{f} \) is injective. To see this, let \( x, y \in A \) be distinct and consider the following three cases:

1. If \( x, y \in \bar{C} \), then \( \bar{f}(x) = f(x) \) and \( \bar{f}(y) = f(y) \), and since \( f \) is injective we get \( \bar{f}(x) \neq \bar{f}(y) \).
2. If \( x, y \in A \setminus \bar{C} \), then \( \bar{f}(x) = x \) and \( \bar{f}(y) = y \), and hence, \( \bar{f}(x) \neq \bar{f}(y) \).
3. If \( x \in \bar{C} \) and \( y \in A \setminus \bar{C} \), then \( \bar{f}(x) = f(x) \in \bar{C} \) and \( \bar{f}(y) = y \notin \bar{C} \), and therefore, \( \bar{f}(x) \neq \bar{f}(y) \).

We already know that \( \bar{f}[C] = C \setminus C \) and by definition we have \( \bar{f}[A \setminus \bar{C}] = A \setminus \bar{C} \).

Hence,
\[
\bar{f}[A] = (A \setminus C) \cup (C \setminus C) = A \setminus C = A_1
\]
which shows that \(|A| = |A_1|\). \( \square \)

**Theorem 3.17 (Cantor-Bernstein Theorem).** Let \( A \) and \( B \) be any sets. If \(|A| \leq |B|\) and \(|B| \leq |A|\), then \(|A| = |B|\).

**Proof.** Let \( f : A \leftrightarrow B \) be a one-to-one mapping from \( A \) into \( B \), and \( g : B \leftrightarrow A \) be a one-to-one mapping from \( B \) into \( A \). Further, let \( A_0 := (g \cdot f)[A] \) and \( A_1 := g[B] \). Then \(|A_0| = |A|\) and \( A_0 \subseteq A_1 \subseteq A \), hence, by Lemma 3.16, \(|A| = |A_1|\), and since \(|A_1| = |B|\) we have \(|A| = |B|\). \( \square \)

As an application of the Cantor-Bernstein Theorem 3.17 let us show that the set of real numbers, denoted by \( \mathbb{R} \), has the same cardinality as \( \mathcal{P}(\omega) \).

**Proposition 3.18.** \(|\mathbb{R}| = |\mathcal{P}(\omega)|\).

**Proof.** Cantor showed that every real number \( r > 1 \) can be written in a unique way as a product of the form
\[
r = \prod_{n \in \omega} \left( 1 + \frac{1}{q_n} \right)
\]
where all \( q_n \)'s are positive integers and for all \( n \in \omega \) we have \( q_{n+1} \geq q_n^2 \). Such products are called Cantor products. So, for every real number \( r > 1 \) there
exists a unique infinite sequence \( q_0(r), q_1(r), \ldots, q_n(r), \ldots \) of positive integers with \( q_{n+1} \geq q_n^2 \) (for all \( n \in \omega \)) such that \( r = \prod_{n \in \omega} (1 + \frac{1}{q_n}) \).

Let us first show that \( |R| \leq |\mathcal{P}(\omega)| \): For \( r \in R \) let

\[
f(r) = \left\{ \sum_{j \leq n} q_j(r)(2^j + 1) : n \in \omega \right\}.
\]

Define the function \( h : R \to \mathbb{R} \) by stipulating \( h(x) := 1 + e^x \), where \( e \) is the Euler number and \( e^x = \sum_{n \in \omega} (x^n/n!) \). Then \( h \) is a bijection between \( \mathbb{R} \) and the set of real numbers \( r > 1 \). We leave it as an exercise to the reader to verify that the composition \( f \circ h \) is an injective mapping from \( \mathbb{R} \) into \( \mathcal{P}(\omega) \).

To see that \( |\mathcal{P}(\omega)| \leq |R| \), consider for example the function \( g(x) = \sum_{n \in \omega} 3^{-n} \),

where \( g(0) := 0 \), which is obviously a injective mapping from \( \mathcal{P}(\omega) \) into \( R \) (or more precisely, into the interval \( [0, \frac{1}{2}] \)).

So, by the Cantor-Bernstein Theorem 3.17, \( |R| = |\mathcal{P}(\omega)| \). \( \Box \)

7. The Axiom Schema of Replacement

For every first-order formula \( \varphi(x, y, p) \) with \( \text{free}(\varphi) \subseteq \{x, y, p\} \), where \( p \) can be an ordered \( n \)-tuple of parameters, the following formula is an axiom:

\[
\forall A \forall p (\forall x \in A \exists y \varphi(x, y, p) \to \exists B \forall x \in A \exists y \in B \varphi(x, y, p))
\]

In other words, for every set \( A \) and for each class function \( F \) (i.e., a certain class of ordered pairs of sets) defined on \( A \), \( F[A] = \{F(x) : x \in A\} \) is a set.

Or even more informally, images of sets under functions are sets.

The Axiom Schema of Replacement is needed to build sets like \( \{\mathcal{P}^n(\omega) : n \in \omega\} \), where \( \mathcal{P}^0(\omega) := \omega \) and \( \mathcal{P}^{n+1}(\omega) := \mathcal{P}(\mathcal{P}^n(\omega)) \).

Another application of the Axiom Schema of Replacement is the following result, which will be used for example to define ordinal addition (see Theorem 3.20) or to build the cumulative hierarchy of sets (see Theorem 3.22).

Theorem 3.19 (Transfinite Recursion Theorem). Let \( F \) be a class function which is defined for all sets. Then there is a unique class function \( G \) defined on \( \Omega \) such that for each \( \alpha \in \Omega \) we have

\[
G(\alpha) = F(G|_{\alpha}), \quad \text{where} \quad G|_{\alpha} = \{\langle \beta, G(\beta) \rangle : \beta \in \alpha\}.
\]

Proof. If such a class function \( G \) exists, then, by the Axiom Schema of Replacement, for every ordinal \( \alpha \), \( \text{ran}(G|_{\alpha}) \) is a set, and consequently, \( G|_{\alpha} \) is a function with \( \text{dom}(G|_{\alpha}) = \alpha \). This leads to the following definition: For \( \delta \in \Omega \), a function \( g \) with \( \text{dom}(g) = \delta \) is called a \( \delta \)-approximation if

\[
\forall \beta \in \delta (g(\beta) = F(g|_{\beta})).
\]

In other words, \( g \) is an \( \delta \)-approximation if and only if \( g \) has the following properties:
(a) If $\beta + 1 \in \delta$, then $g(\beta + 1) = F(g|_\beta \cup \{\langle \beta, g(\beta) \rangle \})$.
(b) If $\beta \in \delta$ is a limit ordinal, then $g(\beta) = F(g|_\beta)$.

In particular, by (b) we get $g(0) = F(\emptyset)$. For example $g_1 = \{\langle 0, F(\emptyset) \rangle \}$ is a $1$-approximation; in fact, $g_1$ is the unique $1$-approximation. Similarly, $g_2 = \{\langle 0, F(\emptyset) \rangle, \langle 1, F(\{0, F(\emptyset)\}) \rangle \}$ is the unique $2$-approximation.

Firstly, notice that for all ordinals $\delta$ and $\delta'$, if $g$ is an $\delta$-approximation and $g'$ is an $\delta'$-approximation, then $g|_{\delta \cap \delta'} = g'|_{\delta \cap \delta'}$. Otherwise, there would be a smallest ordinal $\delta_0$ such that $g(\delta_0) \neq g'(\delta_0)$, but by (a) and (b), $\delta_0$ would be neither a successor ordinal nor a limit ordinal.

Secondly, notice that for each ordinal $\delta$ there exists a $\delta$-approximation. Otherwise, by Theorem 3.12.(d), there would be a smallest ordinal $\delta_0$ such that there is no $\delta_0$-approximation. In particular, for each $\delta \in \delta_0$ there would be a $\delta$-approximation, and by the Axiom Schema of Replacement, the collection of all $\delta$-approximations (for $\delta \in \delta_0$) is a set, where the union of this set is a $\delta'$-approximation for some $\delta' \in \Omega$. Now, if $\delta_0$ is a limit ordinal, then $\delta' = \delta_0$ and we get a $\delta_0$-approximation, and if $\delta_0$ is a successor ordinal, then $\delta_0 = \delta' + 1$ and we get a $\delta_0$-approximation by (a). So, in both cases we get a contradiction to our assumption that there is no $\delta_0$-approximation.

Now, for each $\alpha \in \Omega$ define $G(\alpha) := g(\alpha)$, where $g$ is the $\delta$-approximation for any $\delta$ such that $\alpha \in \delta$.

By transfinite recursion we are able to define addition, multiplication, and exponentiation of arbitrary ordinal numbers:

**Ordinal Addition:** For arbitrary ordinals $\alpha \in \Omega$ we define:
(a) $\alpha + 0 := \alpha$.
(b) $\alpha + (\beta + 1) := (\alpha + \beta) + 1$, for all $\beta \in \Omega$.
(c) If $\beta \in \Omega$ is non-empty and a limit ordinal, then $\alpha + \beta := \bigcup_{\delta \in \beta} (\alpha + \delta)$.

Notice that addition of ordinals is in general not commutative (e.g., $1 + \omega = \omega \neq \omega + 1$).

**Ordinal Multiplication:** For arbitrary ordinals $\alpha \in \Omega$ we define:
(a) $\alpha \cdot 0 := 0$.
(b) $\alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha$, for all $\beta \in \Omega$.
(c) If $\beta \in \Omega$ is a limit ordinal, then $\alpha \cdot \beta := \bigcup_{\delta \in \beta} (\alpha \cdot \delta)$.

Notice that multiplication of ordinals is in general not commutative (e.g., $2 \cdot \omega = \omega \neq \omega + \omega = \omega \cdot 2$).
**Ordinal Exponentiation:** For arbitrary ordinals \( \alpha \in \Omega \) we define:

(a) \( \alpha^0 := 1. \)
(b) \( \alpha^{\beta+1} := \alpha^\beta \cdot \alpha \), for all \( \beta \in \Omega. \)
(c) If \( \beta \in \Omega \) is non-empty and a limit ordinal, then \( \alpha^\beta := \bigcup_{\delta \in \beta} (\alpha^\delta). \)

Notice that for example \( 2^\omega = \omega \), which should not be confused with cardinal exponentiation defined in Chapter 5.

**Theorem 3.20. Addition, multiplication, and exponentiation of ordinals are proper binary operations on \( \Omega. \)**

**Proof.** We just prove it for addition (the proof for the other operations is similar): For each \( \alpha \in \Omega \) define a class function \( F_\alpha \) by stipulating \( F_\alpha(x) := \emptyset \) if \( x \) is not a function; and if \( x \) is a function, then let

\[
F_\alpha(x) = \begin{cases} 
\alpha & \text{if } x = \emptyset, \\
x(\beta) \cup \{x(\beta)\} & \text{if } \text{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega, \\
\bigcup_{\delta \in \beta} x(\delta) & \text{if } \text{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal}, \\
\emptyset & \text{otherwise.}
\end{cases}
\]

By the Transfinite Recursion Theorem 3.19, for each \( \alpha \in \Omega \) there is a unique class function \( G_\alpha \) defined on \( \Omega \) such that for each \( \beta \in \Omega \) we have \( G_\alpha(\beta) = F_\alpha(G_\alpha|_\beta) \), and in particular we get \( G_\alpha(\beta) = \alpha + \beta. \) -\|

Even though addition and multiplication of ordinals are not commutative, they are still associative.

**Proposition 3.21. Addition and multiplication of ordinals defined as above are associative operations.**

**Proof.** We have to show that for all \( \alpha, \beta, \gamma \in \Omega \), \( (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \) and \( (\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma) \). We give the proof just for addition and leave the proof for multiplication as an exercise to the reader.

Let \( \alpha \) and \( \beta \) be arbitrary ordinals. The proof is by induction on \( \gamma \in \Omega. \)

For \( \gamma = 0 \) we obviously have \( \alpha + (\beta + 0) = \alpha + \beta = \alpha + (\beta + 0). \) Now, let us assume that \((\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)\) for some \( \gamma \). Then:

\[
\begin{align*}
(\alpha + \beta) + (\gamma + 1) &= ((\alpha + \beta) + \gamma) + 1 & \text{(by definition of "+")} \\
&= (\alpha + (\beta + \gamma)) + 1 & \text{(by our assumption)} \\
&= \alpha + ((\beta + \gamma) + 1) & \text{(by definition of "+")} \\
&= \alpha + (\beta + (\gamma + 1)) & \text{(by definition of "+")}
\end{align*}
\]

Finally, let \( \gamma \) be a limit ordinal. Notice first that \( \alpha + (\beta + \gamma) = \bigcup_{\delta \in (\beta + \gamma)} \alpha + \delta = \bigcup_{\delta \in (\beta + \gamma') \in (\beta + \gamma)} \alpha + (\beta + \gamma'). \) Thus, if \( (\alpha + \beta) + \gamma' = \alpha + (\beta + \gamma') \) for all \( \gamma' \in \gamma \), then

\[
(\alpha + \beta) + \gamma = \bigcup_{\gamma' \in \gamma} (\alpha + \beta) + \gamma' = \bigcup_{\gamma' \in \gamma} \alpha + (\beta + \gamma') = \alpha + (\beta + \gamma).
\]
8. The Axiom of Foundation

\[ \forall x (\exists z (z \in x) \rightarrow \exists y \in x (y \cap x = \emptyset)) \]

As a consequence of the Axiom of Foundation we get that there is no infinite descending sequence \( x_0 \ni x_1 \ni x_2 \ni \cdots \) since otherwise, the set \( \{x_0, x_1, x_2, \ldots\} \) would contradict the Axiom of Foundation. In particular, there is no set \( x \) such that \( x \in x \) and there are also no cycles like \( x_0 \ni x_1 \ni \cdots \ni x_n \ni x_0 \). As a matter of fact we would like to mention that if one assumes the Axiom of Choice, then the non-existence of such infinite descending sequences can be proved to be equivalent to the Axiom of Foundation.

The axiom system containing the axioms 0–8 is called Zermelo-Fraenkel Set Theory and is denoted by ZF. In fact, ZF contains all axioms of Set Theory except the Axiom of Choice.

Even though the Axiom of Foundation is irrelevant outside Set Theory, it is extremely useful in the metamatematics of Set Theory, since it allows us to arrange all sets in a cumulative hierarchy and let us define cardinalities as sets.

Models of ZF

By induction on \( \alpha \in \Omega \), define the following sets:

\[
V_0 = \emptyset \\
V_\alpha = \bigcup_{\beta < \alpha} V_\beta \quad \text{if } \alpha \text{ is a limit ordinal} \\
V_{\alpha+1} = \mathcal{P}(V_\alpha)
\]

and let

\[
V = \bigcup_{\alpha \in \Omega} V_\alpha .
\]

Notice that by construction, for each \( \alpha \in \Omega \), \( V_\alpha \) is a set. Again by induction on \( \alpha \in \Omega \) one can easily show that the sets \( V_\alpha \) have the following properties:

- Each \( V_\alpha \) is transitive.
- If \( \alpha \in \beta \), then \( V_\alpha \subsetneq V_\beta \).
- \( \alpha \subseteq V_\alpha \) and \( \alpha \in V_{\alpha+1} \).

These facts are visualised by the following figure:
Before we can prove that the class $V$, called the **cumulative hierarchy**, contains all set, we have to introduce the notion of transitive closure: Let $S$ be an arbitrary set. By induction on $n \in \omega$ define

$$S_0 = S, \quad S_{n+1} = \bigcup S_n,$$

and finally

$$TC(S) = \bigcup_{n \in \omega} S_n$$

where $\bigcup_{n \in \omega} S_n := \bigcup \{S_n : n \in \omega\}$. For example $x_1 \in S_1$ iff $\exists x_0 \in S_0 (x_0 \ni x_1)$, and in general, $x_{n+1} \in S_{n+1}$ iff $\exists x_0 \in S_0 \ldots \exists x_n \in S_n (x_0 \ni x_1 \ni \ldots \ni x_{n+1})$. Notice that by the Axiom of Foundation, every descending sequence of the form $x_0 \ni x_1 \ni \ldots$ must be finite.

By construction, $TC(S)$ is transitive, i.e., $x \in TC(S)$ implies $x \subseteq TC(S)$, and we further have $S \subseteq TC(S)$. Moreover, since every transitive set $T$ must satisfy $\bigcup T \subseteq T$, it follows that the set $TC(S)$ is the smallest transitive set which contains $S$. Thus,

$$TC(S) = \bigcap \{T : T \supseteq S \text{ and } T \text{ is transitive}\}$$

and consequently the set $TC(S)$ is called the **transitive closure** of $S$.

**Theorem 3.22.** For every set $x$ there is an ordinal $\alpha$ such that $x \in V_\alpha$. In particular, the class $V$ is equal to the set-theoretic universe.

**Proof.** Assume towards a contradiction that there exists a set $x$ which does not belong to $V$. Let $\bar{x} := TC(\{x\})$ and let $w := \{z \in \bar{x} : z \notin V\}$, i.e., $w = \bar{x} \setminus \{z' \in \bar{x} : \exists \alpha \in \Omega (z' \in V_\alpha)\}$. Since $x \in w$ we have $w \neq \emptyset$, and by the Axiom of Foundation there is a $z_0 \in w$ such that $(z_0 \cap w) = \emptyset$. Since $z_0 \in w$ we have $z_0 \notin V$, which implies that $z_0 \neq \emptyset$, but for all $u \in z_0$ there is a least ordinal $\alpha_u$ such that $u \in V_{\alpha_u}$. By the Axiom Schema of Replacement, \(\{\alpha_u : u \in z_0\}\) is a set, and moreover, $\alpha = \bigcup \{\alpha_u : u \in z_0\} \in \Omega$. This implies that $z_0 \subseteq V_\alpha$ and consequently we get $z_0 \in V_{\alpha+1}$, which contradicts the fact that $z_0 \notin V$ and completes the proof. \(\blacksquare\)
It is natural to ask whether there exists some kind of upper bound or ceiling for the set-theoretic universe $V$ or if there exists arbitrarily large sets. In order to address this question, we have to introduce the notion of cardinal numbers.

Cardinals in ZF

Let $A$ be an arbitrary set. The cardinality of $A$, denoted $|A|$, could be defined as the class of all sets $B$ which have the same cardinality as $A$ (i.e., for which there exists a bijection between $A$ and $B$), but this would have the disadvantage that except for $A = \emptyset$, $|A|$ would not belong to the set-theoretic universe. However, with the Axiom of Foundation the cardinality of a set $A$ can be defined as a proper set:

$$|A| = \{B \in V_{\beta_0} : \text{there exists a bijection between } B \text{ and } A\}$$

where $\beta_0$ is the least ordinal number for which there is a $B \in V_{\beta_0}$ such that $B$ has the same cardinality as $A$. Notice that for example $|\emptyset| = \{\emptyset\}$, where $\{\emptyset\} \subseteq V_1$ (in this case, $\beta_0 = 1$). The set $|A|$ is called a cardinal number, or just a cardinal. Notice that $A$ is not necessarily a member of $|A|$. Further notice that $|A| = |B|$ iff there is a bijection between $A$ and $B$, and as above we write $|A| \leq |B|$ if $|A| = |B'|$ for some $B' \subseteq B$. Cardinal numbers are usually denoted by Fraktur letters like $\aleph$ and $\beth$. A cardinal number is finite if it is the cardinality of a natural number $n \in \omega$, otherwise it is infinite. Finite cardinals are usually denoted by letters like $n, m, \ldots$. An infinite cardinal which contains a well-orderable set is traditionally called an aleph and consequently denoted by an “$\aleph$”, e.g., $\aleph_0 := |\omega|$. The following fact summarises some simple properties of alephs.

**FACT 3.23.** All sets which belong to an aleph can be well-ordered and the cardinality of any ordinal is an aleph. Further, for any ordinals $\alpha, \beta \in \Omega$ we have $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$ or $|\alpha| > |\beta|$, and these three cases are mutually exclusive.

A non-empty set $A$ is called uncountable if there is no enumeration of the elements of $A$, or equivalently, no mapping from $\omega$ to $A$ is surjective.

By the Axiom of Infinity we know that there is an infinite set and we have seen that there is even a smallest infinite ordinal, namely $\omega$, which is of course a countable set. Now, the question arises whether every infinite set is countable. We answer this question in two steps: First we show that the set of real numbers is uncountable, and then we show that in general, for every set $A$ there exists a set of strictly greater cardinality than $A$ — which implies that there is no largest cardinal.

**PROPOSITION 3.24.** The set of real numbers is uncountable.
Cardinals in ZF

Proof. By Proposition 3.18 we already know that there is a bijection between $R$ and $\mathcal{P}(\omega)$. Further we have $|\mathcal{P}(\omega)| = |\omega|^2$. Indeed, for every $x \in \mathcal{P}(\omega)$ let $\chi_x \in \omega^2$ be such that

$$\chi_x(n) = \begin{cases} 1 & \text{if } n \in x, \\ 0 & \text{otherwise.} \end{cases}$$

So, it is enough to show that no mapping from $\omega$ to $\omega^2$ is surjective. Let

$$g : \omega \longrightarrow \omega^2$$

$$n \longmapsto f_n$$

be any mapping from $\omega$ to $\omega^2$. Define the function $f \in \omega^2$ by stipulating

$$f(n) = 1 - f_n(n).$$

Then for each $n \in \omega$ we have $f(n) \neq f_n(n)$, so $f$ is distinct from every function $f_n$ ($n \in \omega$), which shows that $g$ is not surjective. \hfill \Box

For cardinals $m = |A|$ let $2^m := |\mathcal{P}(A)|$. By modifying the proof above we can show the following result:

**Theorem 3.25 (Cantor’s Theorem).** For every cardinal $m$, $2^m > m$.

Proof. Let $A \in m$ be arbitrary. It is enough to show that there is an injection from $A$ into $\mathcal{P}(A)$, but there is no surjection from $A$ onto $\mathcal{P}(A)$.

Firstly, the function

$$f : A \longrightarrow \mathcal{P}(A)$$

$$x \longmapsto \{x\}$$

is obviously injective, and therefore we get $m \leq 2^m$.

Secondly, let $g : A \rightarrow \mathcal{P}(A)$ be an arbitrary function. Consider the set

$$A' = \{x \in A : x \notin g(x)\}.$$

As a subset of $A$, the set $A'$ is an element of $\mathcal{P}(A)$. If there would be an $x_0 \in A$ such that $g(x_0) = A'$, then $x_0 \in A' \iff x_0 \notin g(x_0)$, but since $g(x_0) = A'$, $x_0 \notin g(x_0) \iff x_0 \notin A'$. Thus, $x_0 \in A' \iff x_0 \notin A'$, which is obviously a contradiction and shows that $g$ is not surjective. \hfill \Box

As an immediate consequence of Cantor’s Theorem 3.25 we get that there are arbitrarily large cardinal numbers. Before we show that there are also arbitrarily large ordinal numbers, let us summarise some basic facts about well-orderings: Recall that a binary relation $R \subseteq A \times A$ is a well-ordering of $A$, if there is an $\alpha \in \Omega$ and a bijection $f : A \rightarrow \alpha$ such that $R(x, y) \iff f(x) \in f(y)$.

The following proposition is crucial in order to define the order type of a well-ordering.
Proposition 3.26. If \( \alpha, \beta \in \Omega \) and \( f : \alpha \to \beta \) is a bijection such that for all \( \gamma_1 \in \gamma_2 \in \alpha \) we have \( f(\gamma_1) \in f(\gamma_2) \), then \( \alpha = \beta \).

Proof. If \( \alpha \neq \beta \), then, by Theorem 3.12(a), we have either \( \alpha \in \beta \) or \( \beta \in \alpha \). Without loss of generality we assume that \( \alpha \in \beta \). Thus, there is a \( \eta \in \beta \setminus \alpha \). Since \( f \) is a bijection, there is a \( \gamma \in \alpha \) such that \( f(\gamma) = \eta \), and since \( \eta \notin \alpha \), \( f(\gamma) \neq \eta \) — in fact, \( f(\gamma) \in \eta \). Let \( \gamma_0 \) be the \( \epsilon \)-minimal ordinal in \( \alpha \) such that \( f(\gamma_0) \neq \gamma_0 \), in particular, \( f|_{\gamma_0} \) is the identity. The situation we have is illustrated by the following figure:

\[
\begin{array}{c}
\alpha \\
\gamma_0 \\
\delta \\
\emptyset
\end{array}
\]

Since \( f(\delta) = \delta \) for all \( \delta \in \gamma_0 \). \( \gamma_0 \in f(\gamma_0) \). Let \( \delta_0 = f^{-1}(\gamma_0) \). By the definition of \( \gamma_0 \) we have \( \gamma_0 \in \delta_0 \), which implies \( f(\gamma_0) \in f(\delta_0) \), or equivalently, \( f(\gamma_0) \in \gamma_0 \), a contradiction.

As an immediate consequence we get that each well-ordering \( R \) of \( A \) corresponds to exactly one ordinal, called the order type of \( R \), denoted o.t.\((R)\), such that there exists a bijection \( f : A \to o.t.(R) \) with the property that for all \( a_1, a_2 \in A \) we have \( a_1 R a_2 \iff f(a_1) \in f(a_2) \). Indeed, for every \( b \in A \) define \( A_b = \{ a \in A : a R b \} \) and let \( f : A \to \Omega \) such that for each \( b \in A \) there exists a unique ordinal \( \beta \) such that \( f[A_b] = \beta \); then \( o.t.(R) = f[A] \). Moreover, by Theorem 3.12(a), if \( R_1 \) and \( R_2 \) are well-orderings of any two subsets of \( A \), then we have \( o.t.(R_1) \in o.t.(R_2) \) or \( o.t.(R_1) = o.t.(R_2) \) or \( o.t.(R_1) \supset o.t.(R_2) \), where the three cases are mutually exclusive.

Theorem 3.27 (Hartogs’ Theorem). For every cardinal \( m \) there is a smallest aleph, denoted \( \aleph(m) \), such that \( \aleph(m) \notin m \).

Proof. Let \( A \in m \) be arbitrary and let \( B \subseteq \mathcal{P}(A \times A) \) be the set of all well-orderings of subsets of \( A \). For every \( R \in B \), o.t.\((R)\) is an ordinal, and for every
On the consistency of ZF

$R \in \mathcal{R}$ and any $\beta \in \text{o.t.}(R)$ there is an $R' \in \mathcal{R}$ such that $\text{o.t.}(R') = \beta$, which shows that

$$\alpha = \{ \text{o.t.}(R) : R \in \mathcal{R} \}$$

is an ordinal. By definition, for every $\beta \in \alpha$ there is a well-ordering $R_S$ of some $S \subseteq A$ such that $\text{o.t.}(R_S) = \beta$, which implies that $|\beta| \leq |A|$. On the other hand, $|\alpha| \leq |A|$ would imply that $\alpha \subseteq A$, which is obviously a contradiction. Let $\aleph(\alpha) := |\alpha|$, then $\aleph(\alpha) \not\subseteq \alpha$ and for each $\aleph \prec \aleph(\alpha)$ we have $\aleph \leq \alpha$. —

**Corollary 3.28.** For every ordinal number $\alpha$ and for every cardinal number $\aleph$, there exists an ordinal number $\beta$ such that $|\beta| > |\alpha|$ and $|\beta| \not\subseteq \aleph$.

**Proof.** For the first inequality let $\alpha \in \Omega$ and let $\aleph = |\alpha|$. By Hartogs’ Theorem 3.27 there is an aleph, namely $\aleph(n)$, such that $\aleph(n) \not\subseteq \aleph$. Now, since $\aleph$ and $\aleph(n)$ both contain well-ordered sets we have $\aleph < \aleph(n)$. Let $\aleph \in \aleph(n)$ be a well-ordered set and let $\beta$ be the order type of $\aleph$. Then $\aleph(\aleph(n)) = |\beta| > |\alpha| = \aleph$.

For the second inequality let $\beta$ be the order type of a well-ordered set which belongs to $\aleph(\alpha)$; then $|\beta| \not\subseteq \aleph$. —

**On the Consistency of ZF**

Zermelo writes in [118, p. 262] that he was not able to show that the seven axioms for Set Theory given in that article are consistent. Even though it is essential to know whether a theory is consistent or not, by Gödel’s Second Incompleteness Theorem 3.9 we know that for a sufficiently strong consistent theory, there is no way to prove its consistency within this theory.

To apply this result for Set Theory, we first have to show that ZF is “sufficiently strong”. In other words, we have to show that ZF is strong enough to define the concept of natural numbers and to prove certain basic arithmetical facts. We do this by showing that $\omega \models \text{PA}$. Firstly, Proposition 3.21 shows that addition and multiplication is associative. Secondly, by replacing $\Omega$ with $\omega$ in Corollary 3.15 we get the Induction Schema for natural numbers:

**Proposition 3.29 (Induction Schema).** If $\varphi(0)$ and $\varphi(n) \rightarrow \varphi(n+1)$ for all $n \in \omega$, then we have $\varphi(n)$ for all $n \in \omega$.

Hence, every model of ZF contains a model of PA (i.e., if ZF is consistent, then so is PA). However, by Gödel’s Second Incompleteness Theorem 3.9, if ZF is consistent (what we believe or at least assume), then ZF cannot prove its own consistency (i.e., cannot provide a model for itself). In other words, there is no mathematical proof for the consistency of ZF within ZF, which means that there is no way to construct or to define a model of ZF without the aid of some concepts that go beyond what is provided in ordinary Mathematics. More formally, any proof for $\text{Con(ZF)}$ has to be carried out in some theory $T$ which contains some information that is not in ZF, and whose consistency cannot be proved within $T$. 
To sum up, either ZF is inconsistent — which is hopefully not the case — or any proof of the consistency of ZF has to be carried out in a theory whose consistency is not provable within that theory.

Notes

Some of the papers mentioned below, or at least their translation into English, can be found in the collection [109] edited by van Heijenoort (whose biography is written by Feferman [39]).

Milestones in Logic. Before we discuss the development of Set Theory, let us give a brief overview of the history of Logic (see Bochenski [11] for a comprehensive problem history of formal logic, providing also large quotes from historical texts).

Organon. Aristotle’s logical treatises contain the earliest formal study of Logic (i.e., of Propositional Logic, which is concerned about logical relations between propositions as wholes) and consequently he is commonly considered the first logician. Aristotle’s logical works were grouped together by the ancient commentators under the title Organon, consisting of Categories, On Interpretation, Prior Analytics, Posterior Analytics, Topics, and On Sophistical Refutations. Aristotle’s work was so outstanding and ahead of his time that nothing significant had been added to his views during the following two millennia.

The Laws of Thought. In 1854, Boole published in An Investigation of the Laws of Thought [15] (see also [14]) a new approach to Logic by reducing it to a kind of algebra and thereby incorporated Logics into Mathematics. Boole noticed that Aristotle’s Logic was essentially dealing with classes of objects and he further observed that these classes can be denoted by symbols like $x, y, z$, subject to the ordinary rules of algebra, with the following interpretations.

(a) $xy$ denotes the class of members of $x$ which are also members of $y$.
(b) If $x$ and $y$ have no members in common, then $x + y$ denotes the class of objects which belong either to $x$ or to $y$.
(c) $1 - x$ denotes all the objects not belonging to the class $x$.
(d) $x = 0$ means that the class $x$ has no members.

However, Boole’s Logic was still Propositional Logic, but just 25 years later this weakness was eliminated.

Begriffsschrift. In 1879, Frege published in his Begriffsschrift [42] the most important advance in Logic since Aristotle. In this work, Frege presented for the first time what we would recognise today as a logical system with negation, implication, universal quantification, logical axioms, et cetera. Even though Frege’s achievement in Logic was a major step towards First-Order Logic, his work had led to some contradictions — discovered by Russell — and further steps had to be taken.

Peano Arithmetic. Written in Latin, [89] was Peano’s first attempt at an axiomatisation of Mathematics — and in particular of Arithmetic — in a symbolic language. The initial arithmetic notions are number, one, successor, is equal to, and nine axioms are stated concerning these notions. (Today, “=” belongs to the underlying language of Logic, and so, Peano’s axioms dealing with equality become redundant; further, we start the natural numbers with zero, rather than one.) Concerning the problem whether the natural numbers can be considered as symbols without inherent meaning, we refer the reader to the discussion between Muller [83].
and Bernays [6]. For Peano’s work in Logic, and in particular for the development of the axioms for natural numbers, we refer the reader to Jourdain [67, pp. 270–314] (where one can also find some comments by Peano) and to Wang [111]. According to Jourdain (cf. [67, p. 273]), Peano [89] succeeded in writing out wholly in symbols the propositions and proofs of a complete treatise on the arithmetic of positive numbers. However, in the arithmetical demonstrations, Peano made extensive use of Grassmann’s work [54], and in fundamental questions of arithmetic as well as in the theory of logical functions, he used Dedekind’s work [21]. The main feature of Wang’s paper [111] is the printing of a letter (mentioned by Noether on page 490 of [25]) from Dedekind to a headmaster in Hamburg, dated 27 February, 1890. In that letter, Dedekind points out the appearance of non-standard models of axioms for natural numbers (see Kaye [71]) and explains how one could avoid such unintended models by using his Kettentheorie (i.e., concept of chains) which he developed in [24]. He also refers to Frege’s works [42, 43] and notes that Frege’s method of defining a kind of “successor relation” agrees in essence with his concept of chains.

Principia Mathematica. One of these steps was taken by Russell and Whitehead in their Principia Mathematica [113], which is a three-volume work on the foundations of Mathematics, published between 1910 and 1913. It is an attempt to derive all mathematical truths from a well-defined set of axioms and inference rules in symbolic logic. The main inspiration and motivation for the Principia Mathematica was Frege’s earlier work on Logic, especially the contradictions discovered by Russell (as mentioned above). The questions remained whether a contradiction could also be derived from the axioms given in the Principia Mathematica, and whether there exists a mathematical statement which could neither be proven nor disproven in the system (for Russell’s search for truth we refer the reader to Dostadis and Papadimitriou [27]). It took another twenty odd years until these questions were answered by Gödel’s Incompleteness Theorem, but before, the logical axioms had to be settled.

Grundzüge der theoretischen Logik. In 1938, Ackermann and Hilbert published in their Grundzüge der theoretischen Logik [66] to some extent the final version of logical axioms (for the development of these axioms see Hilbert [61, 62, 64]).

Our approach to First-Order Logic is partially taken from the first few sections of the hyper-textbook for students by Detlovs and Podnieks (these sections are an extended translation of the corresponding chapters of Detlovs [26]). For other rules of inference see for example Hermes [60] or Ehlinghaus, Flum, and Thomas [28, 29].

Über die Vollständigkeit des Logikkalküls. Gödel proved the Completeness Theorem in his doctoral dissertation Über die Vollständigkeit des Logikkalküls [16] which was completed in 1929. In 1930, he published the same material as in the doctoral dissertation in a rewritten and shortened form in [47]. The standard proof for Gödel’s Completeness Theorem is Henkin’s proof, which can be found in [58] (see also [59]) as well as in most other textbooks on Logic. A slightly different approach can be found for example in Kleene [72, §72].

Über formal unentscheidbare Sätze der Principia Mathematica. In 1930, Gödel announced in [18] his Incompleteness Theorem (published later in [49]), which is probably the most famous theorem in Logic. The theorem as it is stated above is Satz VI of [49], and Gödel’s Second Incompleteness Theorem 3.9, which is in fact a consequence of the proof of that theorem, is Satz XI of [49]. Gödel’s Incompleteness Theorem 3.4 is discussed in great detail in Mostowski [82] (see also Goldstern and Judja [53, Chapter 4]); and for a different proof of Gödel’s Incompleteness Theorem, not just a different version of Gödel’s proof, see Put-
nam [95]. For more historical background—as well as for Gödel’s platonism—we refer the reader to Goldstein [51].

Now, let us discuss the development of Set Theory: To some extent, Set Theory is the theory of infinite sets; but, what is the infinite and does it exist?

The infinite. As mentioned before, there are two different kinds of infinite, namely the actual infinite and the potential infinite. To illustrate the difference, let us consider the collection of prime numbers. Euclid proved that for any prime number \( p \) there is a prime number \( p' \) which is larger than \( p \) (see [31, Book IX]). This shows that there are arbitrarily many prime numbers, and therefore, the collection of primes is “potentially” infinite. However, he did not claim that the collection of all prime numbers as a whole “actually” exists. (The difference between actual and potential infinity is discussed in greater detail for example in Bernays [7, Teil II].)

Two quite similar attempts to prove the objective existence of the (actual) infinite are due to Bolzano [12, 13, §13] and Dedekind [24, §§5, No. 66], and both are similar to the approach suggested in Plato’s Parmenides [94, 132a-b] (for a philosophical view to the notion of infinity we refer the reader to Mancosu [78]). However, Russell [99, Chapter XIII, p. 139 ff.] (see also [101, Chapter XII]) shows that these attempts must fail. Moreover, he demonstrates that the infinite is neither self-contradictory nor demonstrable logically and writes that we must conclude that nothing can be known a priori as to whether the number of things in the world is finite or infinite. The conclusion is, therefore, to adopt a Leibnizian phraseology, that some of the possible worlds are finite, some infinite, and we have no means of knowing to which of these two kinds our actual world belongs. The axiom of infinity will be true in some possible worlds and false in others; whether it is true or false in this world, we cannot tell (cf. [99, p. 14]).

If the infinite exists, the problem still remains how one would recognise infinite sets, or in other words, how one would define the predicate “infinite”. Dedekind provided a definition in [24, §§5, No. 64], which is—as Schröder [103, p. 303 ff.] pointed out—equivalent to the definition given three years earlier by Peirce (cf. [91, p. 202] or [5, p. 51]). However, the fact that an infinite set can be mapped injectively into a proper subset of itself—which is the key idea of Dedekind’s definition of infinite sets—was already discovered and clearly explained about 250 years earlier by Galilei (see [45, First Day]). Another definition of the infinite—which will be compared with Dedekind’s definition in Chapter 7—can be found in von Neumann [86, p. 736]. More definitions of finiteness, as well as their dependencies, can be found for example in Lévy [75] and in Spišák and Vojtáš [106].

Birth of Set Theory. As mentioned above, the birth of Set Theory dates to 1873 when Cantor proved that the set of real numbers is uncountable. One could even argue that the exact birth date is 7 December 1873, the date of Cantor’s letter to Dedekind informing him of his discovery.

Cantor’s first proof that there is no bijection between the set of real numbers and the set of natural numbers used an argument with nested intervals (cf. [18, §2] or [23, p. 117]). Later, he improved the result by showing that \( 2^n > n \) for every cardinal \( n \) (cf. [20] or [23, III, §8]), which is nowadays called \textit{Cantor’s Theorem}. The argument used in the proof of \textbf{Proposition 3.24}—which is in fact just a special case of \textit{Cantor’s Theorem}—is sometimes called \textit{Cantor’s diagonal argument}. The word “diagonal” comes from the diagonal process used in the proofs of \textbf{Proposition 3.24} and \textit{Cantor’s Theorem}. The diagonal process is a technique of constructing a
new member of a set of lists which is distinct from all members of a given list. This is done by arranging first the list as a matrix, whose diagonal gives information about the $x^\text{th}$ term of the $x^\text{th}$ row of the matrix. Then, by changing each term of the diagonal, we get a new list which is distinct from every row of the matrix (see also Kleene [72, §2]).

For a brief biography of Cantor and for the development of Set Theory see for example Fraenkel [41], Schoenflies [102], and Kanamori [68].

**Russell’s Paradox.** The fact that a naïve approach to the notion of “set” leads to contradictions was discovered by Russell in June 1901 while he was working on his Principles of Mathematics [101] (see also Grattan-Guinness [53]). When Russell published his discovery, other mathematicians and set-theorists like Zermelo (see [115, footnote p. 118f.] or Rang and Thomas [96]) had already been aware of this antinomy, which — according to Hilbert — had a downright catastrophic effect when it became known throughout the world of Mathematics (cf. [63, p. 169] or [65, p. 190]). However, Russell was the first to discuss the contradiction at length in his published works, the first to attempt to formulate solutions and the first to appreciate fully its importance. For example the entire Chapter X of [101] was dedicated to discussing this paradox (in particular see [101, Chapter X, §102]). In order to prevent the emergence of antinomies and paradoxes in Set Theory and in Logic in general, Russell developed in [101, Appendix B] (see also [98]) his theory of logical types which rules out self-reference. According to this theory, self-referential statements are neither true nor false, but meaningless.

Russell’s Paradox as well as some other antinomies can also be found in Fraenkel, Bar-Hillel, and Lévy [36, Chapter I].

**Axiomatisation of Set Theory.** In 1908, Zermelo published in [118] his first axiomatic system consisting of seven axioms, which he called:

1. **Axiom der Bestimmtheit**
   which corresponds to the Axiom of Extensionality

2. **Axiom der Elementarmengen**
   which includes the Axiom of Empty Set as well as the Axiom of Pairing

3. **Axiom der Aussonderung**
   which corresponds to the Axiom Schema of Separation

4. **Axiom der Potenzmenge**
   which corresponds to the Axiom of Power Set

5. **Axiom der Vereinigung**
   which corresponds to the Axiom of Union

6. **Axiom der Auswahl**
   which corresponds to the Axiom of Choice

7. **Axiom des Unendlichen**
   which corresponds to the Axiom of Infinity

In 1930, Zermelo presented in [116] his second axiomatic system, which he called ZF-system, in which he incorporated ideas of Fraenkel [38], Skolem [104], and von Neumann [85, 86, 88]. (see also Zermelo [114]). In fact, he added the Axiom Schema of
Replacement and the Axiom of Foundation to his former system, cancelled the Axiom of Infinity (since he thought that it does not belong to the general theory of sets), and did not mention explicitly the Axiom of Choice (because of its different character and since he considered it as a general logical principle). For Zermelo’s published work in Set Theory, described and analysed in its historical context, see Zermelo [117], Kanamori [20], and Ebbinghaus [39].

The need for the Axiom Schema of Replacement was first noticed by Fraenkel (see [117, p. 23]) who introduced a certain form of it in [38] (another form of it he gave in [37, Definition 2, p. 158]). However, the present form was introduced by von Neumann [87] (see the note below on the Transfinite Recursion Theorem). As a matter of fact we would like to mention that the Axiom Schema of Replacement was already used implicitly by Cantor in 1899 (cf. [23, p. 444, line 3]). Beside Fraenkel, also Skolem realised that Zermelo’s first axiomatic system was not sufficient to provide a complete foundation for the usual theory of sets and introduced — independently of Fraenkel — in 1922 the Axiom Schema of Replacement (see [104] or [105, p. 145f.]). In [104], he also gave a proper definition of the notion “definite proposition” and, based on a theorem of Löwenheim [77], he discovered the following fact [105, p. 139] (stated in Chapter 15 as LOEVENHEIM-SKOLEM THEOREM 15.1): If the axioms are consistent, there exists a domain in which the axioms hold and whose elements can all be enumerated by means of the positive finite integers. At a first glance this looks strange, since we know for example that the set of real numbers is uncountable. However, this so-called SKOLEM PARADOX — which we will meet in a slightly different form in Chapter 15 — is not a paradox in the sense of an antinomy, it is just a somewhat unexpected feature of formal systems (see also Kleene [72, p. 426f.] and von Plato [110]).

Concerning the terminology we would like to mention that the definition of ordered pairs given above was introduced by Kuratowski [74, Définition V, p. 174] (compare with Hausdorff [57, p. 32] and see also Kanamori [69, §3]), and that the infinite set which corresponds to \( \omega = \{ \emptyset, \{ \emptyset \}, \{ \emptyset, \{ \emptyset \} \}, \{ \emptyset, \{ \emptyset, \{ \emptyset \} \} \} \ldots \) was introduced by von Neumann [84]. For more historical background see Bachmann [4] or Fraenkel [8, Part I], and for a brief discussion of the axiom systems of von Neumann, Bernays, and Gödel see Fraenkel [8, Part I, Section 7].

The Axiom of Foundation. As mentioned above, Zermelo introduced this axiom in his second axiomatisation of Set Theory in 1930, but it goes back to von Neumann (cf. [85, p. 230] and [88, p. 233]), and in fact, the idea can already be found in Mirimanoff [80, 81]: For example in [80, p. 21] he calls a set \( x \) regular (French “ordinaire”) if every descending sequence \( x \ni x_1 \ni x_2 \ni \ldots \) is finite. However, he did not postulate the regularity of sets as an axiom, but if one would do so, one would get the Axiom of Regularity saying that every set is regular. Now, as a consequence of the Axiom of Foundation we get the fact that there are no infinite descending sequences of the form \( x_1 \ni x_2 \ni \ldots \ni x_i \ldots \), which just tells us that every set is regular. Thus, the Axiom of Foundation implies the Axiom of Regularity. The converse is not true, unless we assume some non-trivial form of the Axiom of Choice (see Mendelson [79]). As a matter of fact we would like to mention that Zermelo, when he formulated the Axiom of Foundation in [116], gave both definitions and just mentioned (without proof) that they are equivalent.

Ordinal numbers. The theory of ordinals was first developed in an axiomatic way by von Neumann in [84] (see also [85, 86, 87]). For an alternative axiomatic approach
to ordinals, independently of ordered sets and types, see Tarski [108] or Lindenbaum and Tarski [76]. For some more definitions of ordinals see Bachmann [4, p. 24].

The Transfinite Recursion Theorem. The Transfinite Recursion Theorem was first formulated and proved by von Neumann [87], who also pointed out that, beside the axioms of Zermelo, also the Axiom Schema of Replacement has to be used. Even though a certain form of the Axiom Schema of Replacement was already given by Fraenkel (see above), von Neumann showed that Fraenkel’s notion of function is not sufficient to prove the Transfinite Recursion Theorem. Moreover, he showed (cf. [87, 1.3]) that Fraenkel’s version of the Axiom Schema of Replacement given in [99, 1.1] follows from the other axioms given there (see also Fraenkel’s note [40]).

The Cantor-Bernstein Theorem. This theorem, unfortunately also known as Schröder-Bernstein Theorem, was first stated and proved by Cantor (cf. [19, VIII.4] or [23, p. 413], and [21, §2, Satz B] or [23, p. 285]). In order to prove this theorem, Cantor used the Trichotomy of Cardinals, which is—as we will see in Chapter 5—equivalent to the Axiom of Choice (see also [23, p. 351, Ann. 2]). An alternative proof, avoiding any form of the Axiom of Choice, was found by Bernstein, who was initially a student of Cantor’s. Bernstein presented his proof around Easter 1897 in one of Cantor’s seminars in Halle, and the result was published in 1898 in Borel [16, p. 103–106] (see Related Result 12). About the same time, Schröder gave a similar proof in [103] (submitted May 1896), but unfortunately, Schröder’s proof was flawed by an irreparable error. While other mathematicians regarded his proof as correct, Korselt wrote to Schröder about the error in 1902. In his reply, Schröder admitted his mistake which he had already found some time ago but did not have the opportunity to make public. A few weeks later, Korselt submitted the paper [73]—which appeared almost a decade later—with a proof of the Cantor-Bernstein Theorem which is quite different from the one given by Bernstein. A proof of the Cantor-Bernstein Theorem, similar to Korselt’s proof, was found in 1906 independently by Peano [90] and Zermelo (see [118, footnote p. 272 f.]). However, they could not know that they had just rediscovered the proof that had already been obtained twice by Dedekind in 1887 and 1897, since Dedekind’s proof—in our terminology given above—was not published until 1932 (see [23, LXII&Er. p. 448] and [23, p. 449]).

Cantor products. Motivated by a result due to Euler on partition numbers (cf. [32, Caput XVI]), Cantor showed in [17] (see also [23, pp. 43–50]) that every real number \( r > 1 \) can be written in a unique way as a product of the form \( \prod_{n \in \omega} (1 + \frac{1}{q_n}) \), where all \( q_n \)'s are positive integers and \( q_{n+1} \geq q_n^2 \). He also showed that \( r = \prod_{n \in \omega} (1 + \frac{1}{q_n}) \) is rational if and only if there is an \( m \in \omega \) such that for all \( n \geq m \) we have \( q_{n+1} = q_n^2 \), and further he gave the representation of the square roots of some small natural numbers. For example, the \( q_n \)'s in the representation of \( \sqrt{2} \) are \( q_0 = 3 \) and \( q_{n+1} = 2q_n^2 - 1 \). More about Cantor products can be found for example in Perron [92, §§35].

Cardinal numbers. The concept of cardinal number is one of the most fundamental concepts in Set Theory. Cantor describes cardinal numbers as follows (cf. [21, §1] or [23, p. 282f.]): The general concept which with the aid of our active intelligence results from a set \( M \), when we abstract from the nature of its various elements and from the order of their being given, we call the "power" or "cardinal number" of \( M \).
This double abstraction suggests his notation "\(\mathbb{M}\)" for the cardinality of \(M\). As mentioned above, one can define the cardinal number of a set \(M\) as an object \(\mathbb{M}\) which consists of all those sets (including \(M\) itself) which have the same cardinality as \(M\). This approach, which was for example taken by Frege (cf. [43, 44]), and Russell (cf. [97, p. 378] or [98, Section IX, p. 256]), has the advantage that it can be carried out in naïve Set Theory (see also Kleene [72, p. 9]). However, it has the disadvantage that for every non-empty set \(M\), the object \(\mathbb{M}\) is a proper class and therefore does not belong to the set-theoretic universe.

**Hartog's Theorem.** The proof of Hartog's Theorem is taken from Hartogs [56]. In that paper, Hartog's main motivation was to find a proof for Zermelo's Well-Ordering Principle which does not make use of the Axiom of Choice. However, since the Well-Ordering Principle and the Axiom of Choice are equivalent, he had to assume something similar, which he had done assuming explicitly Trichotomy of Cardinals. These principles will be discussed in greater detail in Chapter 3.

In 1935, Hartogs was forced to retire from his position in Munich, where he committed suicide in August 1943 because he could not bear any longer the continuous humiliations by the Nazis.

**Related Results**

12. **Bernstein's proof of the Cantor-Bernstein Theorem.** Below we sketch out Bernstein's proof of the Cantor-Bernstein Theorem as it was published by Borel in [16, p. 104 ff.]. Let \(A\) and \(B\) be two arbitrary sets and let \(f : A \hookrightarrow B\) and \(g : B \hookrightarrow A\) two injections. Further, let \(A_0 := A\), \(B_0 := g[B]\), and for \(n \in \omega\) let \(A_{n+1} := (g \circ f)[A_n]\) and \(B_{n+1} := (g \circ f)[B_n]\); finally let \(D := \bigcap_{n \in \omega} A_n\).

We get the following picture:
It is not hard to verify that the sets $A_n$ and $B_n$ have the following properties:

(a) $A_0 = D \cup (A_0 \setminus B_0) \cup (B_0 \setminus A_1) \cup (A_1 \setminus B_1) \cup (B_1 \setminus A_2) \cup \ldots$

(b) $B_0 = D \cup (B_0 \setminus A_1) \cup (A_1 \setminus B_1) \cup (B_1 \setminus A_2) \cup (A_2 \setminus B_2) \cup \ldots$

(c) For all $n \in \omega$, $|A_n \setminus B_n| = |A_{n+1} \setminus B_{n+1}|$.

Since the sets $(A_n \setminus B_n)$, $(B_n \setminus A_{n+1})$, and $D$, are pairwise disjoint, by (c) and by regrouping the representation of $B_0$ in (b), we get $|A_0| = |B_0|$.

References

1. Aristotle, Origanon, Athens, published by Andronikos of Rhodos around 40 B.C.
2. ______, Physics, Athens, published by Andronikos of Rhodos around 40 B.C.
6. ______, Die Philosophie der Mathematik und die Hilbertsche Beweistheorie, Blätter für deutsche Philosophie, vol. 4 (1930), 326-367 (also published in [9]).
18. , Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen, Journal für die reine und angewandte Mathematik (Crelle), vol. 77 (1874), 258–262.


24. Richard Dedekind, Was sind und was sollen die Zahlen, Friedrich Vieweg & Sohn, Braunschweig, 1888 (see also [25, pp. 335–390])


31. Euclid, “Elements”, Alexandria, written around 300 B.C.

32. Leonhard Euler, Introduction in analysin infinitorum (Tomus Primus), Marcum-Michaeli Bouquet & Socios, Lausanne, 1748 (see [34], [33] for a translation into English/German).


References

45. Kurt Gödel, *Über die Vollständigkeit des Logikkalküls*, Dissertation (1929), University of Vienna (Austria), (reprinted and translated into English in [30]).
52. ______, Kurt Gödel, Jahrhundertmathematiker und großer Entdecker, [translated into German by Thorsten Schmidt], Piper, München and Zürich, 2006.


References


78. Paolo Mancosu, *Measuring the size of infinite collections of natural numbers: was Cantor's theory of infinite number inevitable?*, *The Review of Symbolic Logic*, vol. 4 (2009), 612–646.


86. __________, *Die Axiomatisierung der Mengenlehre*, *Mathematische Zeitschrift*, vol. 27 (1928), 609–752.


93. Plato, Politeia, Athens, Book I written around 390 B.C., Books II–X written around 375 B.C.
94. _____, Parmenides, Athens, written around 370 B.C.
98. _____, Mathematical logic as based on the theory of types, American Journal of Mathematics, vol. 30 (1908), 222–262 (also published in [109]).
100. _____, Einführung in die mathematische Philosophie, [translated into German by E. J. Gunzel and W. Gordon], Drei Masken Verlag, München, 1923.
107. Fabian Stedman, Campanologia; or the Art of Ringing Improved, W. Godbid, London, 1677 [reprint: Christopher Groome 1990].
114. Ernst Zermelo, Bericht an die Notgemeinschaft der Deutschen Wissenschaft über meine Forschungen betreffend die Grundlagen der Mathematik, typescript, 5 pp., with appendices, 2 pp., dated 3 December 1930, Universitätsarchiv Freiburg, Zermelo Nachlass, part of C 129/140 (see [117] for a translation into English).
Cardinal Relations in ZF only

To some it may appear novel that I include the fourth among the consonances, because practicing musicians have until now relegated it to the dissonances. Hence I must emphasise that the fourth is actually not a dissonance but a consonance.

Gioseffo Zarlino
Le istitutioni harmoniche, 1558

In the previous chapter we introduced cardinal numbers as certain sets, which contain only sets of the same cardinality. Cardinal numbers in Zermelo-Fraenkel Set Theory are traditionally denoted by Fraktur letters like $m$ and $n$. However, the cardinality of a given set $A$ is denoted by $|A|$. If $|A| = m$, then we say that $A$ is of cardinality $m$. Recall that for cardinals $m = |A|$, $2^m := |\mathcal{P}(A)|$, in particular $2^\omega = |\mathcal{P}(\omega)|$.

Recall that a set $A$ is finite if there exists a bijection between $A$ and a natural number $n \in \omega$. Now, a cardinal number $m$ is finite if $m$ contains a finite set — recall that $|\emptyset| = \{\emptyset\}$. Finite cardinal numbers are usually denoted like elements of $\omega$, i.e., by letters like $n, m, k$ et cetera. In other words, for $n \in \omega$ we usually do not distinguish between the ordinal number $n$ and the cardinal number $n$. Finally, a cardinal number is infinite if it is not finite. Recall that an infinite cardinal which contains a well-orderable set is called an aleph and that alephs are denoted by $\aleph$'s, e.g., $\aleph_0 := |\omega|$. A cardinal $m$ is called transfinite or Dedekind-infinite if $\aleph_0 \leq m$. Notice that transfinite cardinals are always infinite. If the cardinality of a set $A$ is transfinite, then $A$ is called transfinite. Notice that for each transfinite set $A$ there is an injection from $\omega$ into $A$. Sets or cardinals which are not transfinite are called D-finite or Dedekind-finite. Notice that every finite set is D-finite, but as we will see later, D-finite sets are not necessarily finite. For other notions of finiteness see Related Result 13.
Basic Cardinal Relations

Below we show some relations between cardinals which can be proved in ZF.
We start with some simple facts.

Fact 4.1. \( \aleph_0 = |\mathcal{P}| = |\mathbb{Z}| = |\mathbb{Z}^2| = |\mathbb{Q}| \), where \( \mathcal{P} \) denotes the set of prime numbers, \( \mathbb{Z} \) denotes the set of integers, and \( \mathbb{Q} \) denotes the set of rational numbers.

Proof. By definition we have \( \aleph_0 = |\omega| \). Further, \( |\mathcal{P}| \leq |\omega| \leq |\mathbb{Z}| \leq |\mathbb{Q}| \), and since every reduced rational number \( \frac{p}{q} \) corresponds to an ordered pair \( \langle p, q \rangle \) of integers we also have \( |\mathbb{Q}| \leq |\mathbb{Z}^2| \). Thus, by the Cantor-Bernstein Theorem 3.17 it is enough to show that the set \( \mathcal{P} \) is transfinite and to find an injection from \( \mathbb{Z}^2 \) into \( \omega \). That \( \mathcal{P} \) is transfinite follows from the fact that \( \mathcal{P} \) is an infinite, well-orderable set; and to construct an injection \( f : \mathbb{Z}^2 \rightarrow \omega \) we define for example first \( g : \mathbb{P} \times \mathbb{Z} \rightarrow \omega \) by stipulating \( g(p, z) := \max\{1, p^2\} \) and then let \( f(\langle x, y \rangle) := g(2, x) \cdot g(3, -x) \cdot g(5, y) \cdot g(7, -y) \).

For an arbitrary set \( A \) let \( \text{fin}(A) \) denote the set of all finite subsets of \( A \). Notice that \( \text{fin}(A) = \mathcal{P}(A) \) if and only if \( A \) is finite. Further, recall that \( \text{seq}(A) \) denotes the set of all finite sequences which can be formed with elements of \( A \) and that \( \text{seq}^{-1}(A) \) be those sequences of \( \text{seq}(A) \) in which no element appears twice. Further, recall that \( |A|^2 \) is the set of all 2-element subsets of \( A \).

Fact 4.2. \( \aleph_0 = |\omega|^2 = |\text{fin}(\omega)| = |\text{seq}^{-1}(\omega)| = |\text{seq}(\omega)| = |A| \), where \( A \) denotes the set of algebraic numbers, which is the set of all real numbers which are roots of polynomials with integer coefficients.

Proof. Since every finite subset of \( \omega \) corresponds to a strictly increasing finite sequence of elements of \( \omega \) we obviously have \( \aleph_0 \leq |\omega|^2 \leq |\text{fin}(\omega)| \leq |\text{seq}^{-1}(\omega)| \leq |\text{seq}(\omega)| \). By the Cantor-Bernstein Theorem 3.17, in order to prove that \( |\text{seq}(\omega)| = \aleph_0 \) it is enough to find an injection from \( \text{seq}(\omega) \) into \( \omega \). Let \( \mathcal{P} = \{p_i : i \in \omega\} \) be such that for all \( i, j \in \omega \), \( i < j \rightarrow p_i < p_j \), and define \( f : \text{seq}(\omega) \rightarrow \omega \) by stipulating

\[
  f(\langle a_0, a_1, \ldots, a_n \rangle) := p_0^{a_0+1} \cdot p_1^{a_1+1} \cdots p_n^{a_n+1}.
\]

Then, by unique factorisation of integers, \( f \) is injective. Now, let us consider the set \( A \): A polynomial \( p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \) with integer coefficients has at most \( n \) different real roots say \( r_0 < r_1 < \ldots r_k \) where \( k < n \), and since there exists a bijection \( g \) between \( \mathbb{Z} \) and \( \omega \) (by Fact 4.1), we can define a mapping \( h_{p(x)} \) which assigns to each root \( r_i \) of \( p(x) \) an element of \( \text{seq}(\omega) \) by stipulating

\[
  h_{p(x)}(r_i) = \langle g(a_0), \ldots, g(a_n), i \rangle,
\]

and define \( H : A \rightarrow \omega \) by stipulating

\[
  H(r) = \min \{ h_{p(x)}(r_i) : p(r) = 0 \land r = r_i \}.
\]

This shows that \( |A| \leq \aleph_0 \) and completes the proof. \( \square \)
By Proposition 3.18 we know that $|\mathbb{R}| = 2^{\aleph_0}$ and by Cantor's Theorem 3.25 we get that $\mathbb{N}_0 < 2^{\aleph_0}$, hence, the set of reals is uncountable (cf. Proposition 3.24). The following result gives a few examples of sets of the same cardinality as $\mathbb{R}$.

**Fact 4.3.** $|\mathbb{N}_0, \varepsilon)| = |\mathbb{R}| = |\omega^2| = |\mathbb{R} \times \mathbb{R} = |\omega^\omega| = |\mathbb{R} \times \mathbb{R}| = |C[0, 1]| = |\mathbb{R} \setminus \mathbb{A}|$, where for $\varepsilon > 0$, $[0, \varepsilon) = \{r \in \mathbb{R} : 0 \leq r \leq \varepsilon\}$, and $C[0, 1]$ denotes the set of continuous functions from $[0, 1]$ to $\mathbb{R}$.

**Proof.** The function $\varepsilon \cdot (\arctan(x) + \frac{\pi}{2})/\pi$ is a bijection between $\mathbb{R}$ and the open interval $(0, \varepsilon)$, thus, by the Cantor-Bernstein Theorem 3.17 we get $|\mathbb{N}_0, \varepsilon)| = |\mathbb{R}|$.

Since the function $h: \omega^\omega \to \mathcal{P}(\omega)$ defined by stipulating $h(f) := \{n \in \omega : f(n) = 1\}$ is bijective, and since $|\mathbb{R}| = |\mathcal{P}(\omega)|$, we get $|\mathbb{R}| = |\omega^\omega|$. Recall that there is a bijection $g: \omega \times \omega \to \omega$, e.g., let $g((n, m)) := m + \frac{1}{2}(n + m)(n + m + 1)$. In order to show that $|\omega^\omega| = |\mathbb{R}|$ it is enough to show that there is a bijection between $\omega^\omega$ and $\mathcal{P}(\omega)$. Now, there is a one-to-one correspondence between functions $h \in \omega^\omega$ and sets $X \in \mathcal{P}(\omega \times \omega)$ by $(a, b) \in X \iff b \in h(a)$. Thus, the function

$$
\mathcal{P}(\omega \times \omega) \longrightarrow \mathcal{P}(\omega)
$$

induces a bijection between $\omega^\omega$ and $\mathcal{P}(\omega)$, hence, $|\omega^\omega|$ and $|\mathbb{R}|$, and since $|\mathbb{R}| \leq |\mathbb{R} \times \mathbb{R}| \leq |\omega^\omega|$ and $|\mathbb{R}| = |\omega^2| \leq |\omega^\omega| \leq |\omega^\omega|$, we finally get $|\mathbb{R}| = |\omega^2| = |\omega^\omega| = |\mathbb{R} \times \mathbb{R}| = |\omega^\omega|$.

To see that $|\mathbb{R}| = |C[0, 1]|$, notice first that a continuous function from $[0, 1]$ to $\mathbb{R}$ is defined by its values on $\mathbb{Q} \cap [0, 1]$. By Fact 4.1 there is a bijection between $\mathbb{Q} \cap [0, 1]$ and $\omega$, and consequently there is a one-to-one correspondence between functions in $C[0, 1]$ and some functions in $\omega^\omega$ which shows that $|C[0, 1]| \leq |\omega^\omega|$. Since $|\omega^\omega| = |\mathbb{R}|$ and since we obviously have $|\mathbb{R}| \leq |C[0, 1]|$, by the Cantor-Bernstein Theorem 3.17 we finally get $|C[0, 1]| = |\mathbb{R}|$.

By Fact 4.2, $|\mathbb{A}| = \mathbb{N}_0$ and we leave it as an exercise to the reader to show that $|\mathbb{R} \setminus \mathbb{A}| = |\mathbb{R}|$ for all countable sets $A \subseteq \mathbb{R}$. At this point we would like to mention that the reals $\mathbb{R} \setminus \mathbb{A}$ are called transcendental numbers; thus, all but countably many reals are transcendental.

Let us now turn our attention to arbitrary cardinalities and let us prove that whenever we can embed $\omega$ into $\mathcal{P}(A)$. Then we can also embed $\mathbb{R}$ into $\mathcal{P}(A)$.

**Proposition 4.4.** If $\mathbb{N}_0 \leq 2^m$, then $2^{\aleph_0} \leq 2^m$.

**Proof.** Let $A$ be an arbitrary set of cardinality $m$. Because $\mathbb{N}_0 \leq 2^m$ there is an injection $f_0 : \omega \hookrightarrow \mathcal{P}(A)$. Define an equivalence relation on $A$ by stipulating $x \sim y \iff \forall n \in \omega \left(x \in f_0(n) \iff y \in f_0(n)\right)$. 

and let \([x]^- := \{y \in A : y \sim x\}\). For \(x \in A\) let \(g_x := \{n \in \omega : x \in f_0(n)\}\). Then for every \(x \in A\) we have \(g_x \subseteq \omega\) and \(g_x = g_y\) if \([x]^- = [y]^-\). We can consider the set \(g_x\) as a function from \(\omega\) to \(\{0, 1\}\) by \(g_x(n) = 0\) if \(x \in f_0(n)\) and \(g_x(n) = 1\) if \(x \notin f_0(n)\). Now we define an ordering \(\prec\) on the set \(\{g_x : x \in A\}\) by stipulating

\[
g_x \prec g_y \iff \exists n \in \omega \left(g_x(n) < g_y(n) \land \forall k \in n \left(g_x(k) = g_y(k)\right)\right).
\]

Notice that for all \(x, y \in A\) such that \(g_x \neq g_y\) we have either \(g_x < g_y\) or \(g_y < g_x\). Let \(P^0_n := \{g_x : g_x(n) = 0\}\). Then for each \(n \in \omega\), \(P^0_n \subseteq \omega^2\). Obviously, the relation \(\prec\) defines an ordering on each \(P^0_n\). We consider the following two cases:

If for each \(n \in \omega\), \(P^0_n\) is well-ordered by \(\prec\), then we can easily well-order the infinite set \(\bigcup_{n \in \omega} P^0_n\) and construct a countably infinite set \(\{g_{x_i} : i \in \omega\}\) such that for all distinct \(i, j \in \omega\), \(g_{x_i} \neq g_{x_j}\). If we define \(q_i := \{x \in A : g_x = g_{x_i}\}\), then the set \(Q := \{q_i : i \in \omega\}\) is a countable infinite set of pairwise disjoint subsets of \(A\).

If not every \(P^0_n\) is well-ordered by \(\prec\), there exists a least \(m \in \omega\) such that \(P^0_m\) is not well-ordered by \(\prec\) and we can define

\[
S_0 = \bigcup \{S \subseteq P^0_m : S \text{ has no } \prec\text{-minimal element}\}.
\]

By definition of \(S_0 \subseteq P^0_m\), \(S_0\) has no \(\prec\)-minimal element, too. For \(k \in \omega\) we define \(S_{k+1}\) as follows: If \(S_k \cap P^0_{m+k+1} = \emptyset\), then \(S_{k+1} := S_k\); otherwise, \(S_{k+1} := S_k \cap P^0_{m+k+1}\). By construction, for every \(k \in \omega\), \(S_k \neq \emptyset\) and \(S_k\) is not well-ordered by \(\prec\). This implies that for every \(k \in \omega\) there exists an \(l > k\) such that \(S_l\) is a proper subset of \(S_k\). Now let \(S_{k_1}, S_{k_2}, \ldots\) be such that for all \(i < j\) we have \(S_{k_i} \setminus S_{k_j} \neq \emptyset\) and let \(q_i := \{x \in A : g_x \in (S_{k_i} \setminus S_{k_{i+1}})\}\). Then the set \(Q := \{q_i : i \in \omega\}\) is again a countable infinite set of pairwise disjoint subsets of \(A\).

Thus, in both cases the cardinality of \(\mathcal{P}(Q)\) is \(2^{\aleph_0}\), and since the function

\[
\mathcal{P}(Q) \rightarrow \mathcal{P}(A)
\]

\[
X \mapsto \bigcup X
\]

is injective we finally have \(2^{\aleph_0} \leq 2^m\).

It is now time to define addition and multiplication of cardinals. Let \(m\) and \(n\) be cardinals and let \(A\) and \(B\) be disjoint sets of cardinality \(m\) and \(n\) respectively. Then we define the sum and product of \(m\) and \(n\) as follows:

\[
m + n = |A \cup B|
\]

\[
m \cdot n = |A \times B|
\]
Furthermore, let $2m := m + m$ and $m^2 := m \cdot m$. We leave it as an exercise to the reader to show that for any cardinals $m$, $n$ and $p$ we have for example:

$$m + n = n + m, \quad m \cdot n = n \cdot m$$

$$m \leq n \rightarrow p + m \leq p + n, \quad m \leq n \rightarrow p \cdot m \leq p \cdot n$$

$$2^{m+n} = 2^m \cdot 2^n, \quad 2^{m \cdot n} = \left(2^m\right)^n$$

For example to show that $2^{m+n} = 2^m \cdot 2^n$, define $f : \mathcal{P}(A \cup B) \to \mathcal{P}(A) \times \mathcal{P}(B)$ by stipulating $f(S) := (S \cap A, S \cap B)$.

The following fact is just an easy consequence of the definition of ordered pairs.

**Fact 4.5.** For any cardinal $m$, $m^2 \leq 2^m$.

**Proof.** Let $A$ be a set of cardinality $m$. Any $(a, b) \in A \times A$ can be written in the form $\{a, \{a, b\}\}$, which is obviously an element of $\mathcal{P}(\mathcal{P}(A))$.

Let $m$ be a cardinal and let $A$ be a set of cardinality $m$. Then we define $\text{fin}(m) := |\text{fin}(A)|$ and $|m|^2 := |\mathcal{P}(A)|$. Notice that for all cardinals $m > 2$ we have $m \leq |m|^2 \leq \text{fin}(m)$. We leave it as an exercise to the reader to show that $\aleph_0 \leq m^2 \rightarrow \aleph_0 \leq m$; however, $\aleph_0 \leq |m|^2 \rightarrow \aleph_0 \leq m$ is not provable in ZF (see Theorem 7.6 (b)).

As mentioned above, an infinite set can be D-finite and moreover, even the power set of an infinite set can be D-finite. However, for every infinite cardinal $m$, $2^{\text{fin}(m)}$ is transfinite (notice that $2^{\text{fin}(m)} \leq 2^{2m}$).

**Fact 4.6.** If $m$ is an infinite cardinal, then $2^{\aleph_0} \leq 2^{\text{fin}(m)}$, in particular $2^{\text{fin}(m)}$ is transfinite.

**Proof.** Let $A$ be an arbitrary infinite set of cardinality $m$. For every $n \in \omega$ let $X_n := \{x \subseteq A : |x| = n\}$. Then for any $n \in \omega$, $X_n \in \mathcal{P}(\text{fin}(A))$. For any two distinct integers $n, m \in \omega$ we get $X_n \neq X_m$. This shows that $\aleph_0 \leq 2^{\text{fin}(m)}$, and hence, by Proposition 4.4, $2^{\aleph_0} \leq 2^{\text{fin}(m)}$.

The following result is an immediate consequence of Fact 4.6 (see Theorem 4.28 for a stronger result).

**Fact 4.7.** If $m$ is an infinite cardinal, then $2^{2^m} + 2^{2^m} = 2^{2^m}$.

**Proof.** Notice that

$$2^{2^m} + 2^{2^m} = 2 \cdot 2^{2^m} = 2^{2^m + 1},$$

and since $2^{2^m}$ is transfinite, $2^{2^m} + 1 = 2^{2^m}$. 
For arbitrary sets $A$ and $B$ we write $|A| \leq^* |B|$ if either $A = \emptyset$ or there is a surjection from $B$ onto $A$. Similarly we write $m \leq^* n$ if there are sets $A \in m$ and $B \in n$ such that $|A| \leq^* |B|$. Notice that cardinal relation “$\leq^*$” is reflexive and transitive, and that $m \leq n \rightarrow m \leq^* n$. We leave it as an exercise to the reader to show that for all cardinals $m$, $[m]^2 \leq^* m^2$ (compare this result with Proposition 7.18). However, in ZF, $|A| \leq^* |B|$ and $|B| \leq^* |A|$ does not imply $|A| = |B|$ (see Chapter 7 for counterexamples). On the other hand, we have the following

**Fact 4.8.** If $m \leq^* n$, then $2^m \leq 2^n$. Moreover, if $m \leq^* n$, then $m \leq n$.

**Proof.** Let the sets $A$ and $B$ be of cardinality $m$ and $n$ respectively. Since $m \leq^* n$ there is a surjection $g : B \rightarrow A$. Let $f : \mathcal{P}(A) \rightarrow \mathcal{P}(B)$ by stipulating $f(X) := \{y \in B : g(y) \in X\}$. Then $f$ is injective which shows that $|\mathcal{P}(A)| \leq |\mathcal{P}(B)|$.

Now, let $S$ be a set of cardinality $\aleph$ and let $R_S \subseteq S \times S$ be a well-ordering of $S$. Further, let $g : S \rightarrow A$ (where $|A| = m$) be a surjection. Then $f : A \rightarrow S$, where $f(a)$ is the $R_S$-minimal element of $\{s \in S : g(s) = a\}$ is obviously an injection.

Recall that by Hartogs’ Theorem 3.27, for any cardinal $m$ there is a smallest $\aleph$, denoted $\aleph(m)$, such that $\aleph(m) \not= m$.

**Fact 4.9.** If $m$ is an infinite cardinal, then $\aleph(m) \leq^* 2^{m^2}$.

**Proof.** Let $A$ be a set of cardinality $m$. Any binary relation $R$ on $A$ corresponds to a subset $X_R$ of $A \times A$ by stipulating $(a_0, a_1) \in X_R \iff R(a_0, a_1)$. Thus, we get that the cardinality of the set of binary relations on $A$ is less than or equal to $2^{m^2}$. Further, let $S$ be a well-orderable set of cardinality $\aleph(m)$, let $R$ be a well-ordering of $S$, and let $\alpha = \text{o.t.}(R)$ be the order type of $R$. Then $|\alpha| = |S| = \alpha(m)$. Define $f : \mathcal{P}(A \times A) \rightarrow \alpha$ by stipulating

$$f(X) = \begin{cases} \emptyset & \text{if } X \text{ is not a well-ordering of a subset of } A, \\ \text{o.t.}(X) & \text{otherwise.} \end{cases}$$

By the proof of Hartogs’ Theorem 3.27, for every $\beta \in \alpha$ there is a well-ordering $R$ of a subset of $A$ such that $\text{o.t.}(R) = \beta$, hence, $f$ is surjective.

In the proof of Cantor’s Theorem 3.25 it is in fact shown that for all cardinals $m$, $2^m \not< m$. On the other hand, we obviously have $2^m \leq^* m^2$ in the case when $m \leq 4$; however, it is not known whether $2^m \leq^* m^2 \rightarrow m \leq 4$ is provable in ZF (see Related Result 21).

The situation is different when we replace “$\leq^*$” by “$\leq$”. By Cantor’s Theorem 3.25 we know that $m < 2^m$, thus, $2^m \not< m$. Moreover, $2^m \leq m^2 \rightarrow m \leq 4$ (see Theorem 4.20), but we have to postpone the proof until we can compute the cardinality of products of infinite ordinal numbers. However, let us first investigate the cardinality of the continuum $\mathbb{R}$. 
On the Cardinals $2^\aleph_0$ and $\aleph_1$

By Hartog's Theorem 3.27 we know that for any cardinal $m$ (e.g., $m = \aleph_0$) there is a smallest $\aleph$, denoted $\aleph(m)$, such that $\aleph(m) \not\subseteq m$. Now let $\aleph_1 := \aleph(\aleph_0)$. Then $\aleph_1$ contains an uncountable well-orderable set, say $A$, such that every subset of $A$ of cardinality strictly less than $A$ is countable. Let $\alpha$ be the order type of a well-ordering of $A$. Then, since $|\alpha| = \aleph_1$, $\alpha$ is an uncountable ordinal. Now, if $\alpha \setminus \{ \beta \in \alpha : |\beta| = \aleph_0 \} = \emptyset$, then $\alpha$ is the least uncountable ordinal which is usually denoted $\omega_1$. Otherwise, the non-empty set $\alpha \setminus \{ \beta \in \alpha : |\beta| = \aleph_0 \}$, as a set of ordinals, has an $\omega$-minimal element, say $\gamma$. Then $\gamma$ is the least uncountable ordinal, i.e., $\gamma = \omega_1$. In particular, we find $|\omega_1| = \aleph_1$, and for all $\beta \in \omega_1$ we have $|\beta| = \aleph_0$.

If $2^{\aleph_0}$ would be an aleph, then we would have $\aleph_1 \leq 2^{\aleph_0}$, (notice that $\aleph_0 < 2^{\aleph_0}$ and that $\aleph_0 < \aleph_1$). Now, the Continuum Hypothesis, denoted CH, states that $2^{\aleph_0} = \aleph_1$. In particular, if $2^{\aleph_0}$ is an aleph then CH is equivalent to saying that every subset of $\aleph$ is either countable or of cardinality $2^{\aleph_0}$.

In Chapter 16 we shall see that CH is independent of ZF, thus, neither ZF $\vdash$ CH nor ZF $\vdash \neg$CH. Below we investigate the relationship between the cardinals $2^{\aleph_0}$ and $\aleph_1$. In order to construct a surjection from $\aleph$ onto $\omega_1$—even though there might be no injection from $\omega_1$ into $\aleph$—we prove first the following result:

**Lemma 4.10.** For every ordinal $\alpha \in \omega_1$ there is a set of rationals $Q_\alpha \subseteq \mathbb{Q} \cap (0, 1)$ and a bijection $h_\alpha : \alpha \to Q_\alpha$ such that for all $\beta, \beta' \in \alpha$, $\beta \in \beta' \iff h_\alpha(\beta) < h_\alpha(\beta')$.

**Proof.** Let $\alpha$ be an arbitrary but fixed ordinal in $\omega_1$. For $\alpha = 0$ let $Q_0 := \emptyset$ and we are done; and if $0 \neq \alpha \in \omega$ (i.e., if $\alpha$ is finite), then for $n \in \omega$ we define $h_\alpha(n) := 1 - 1/(n + 2)$. If $\alpha$ is infinite we proceed as follows. Firstly let

$$
\omega \rightarrow \alpha
$$

$$
n \rightarrow \beta_n
$$

and

$$
\omega \rightarrow \mathbb{Q} \cap (0, 1)
$$

$$
n \rightarrow q_n
$$

be two bijections (notice that the sets $\alpha$ and $\mathbb{Q} \cap (0, 1)$ are both countably infinite). Since $\{ \beta_n : n \in \omega \} = \alpha$, it is enough to define $h_\alpha(\beta_n)$ for all $n \in \omega$ which is done by induction: $h_\alpha(\beta_0) := q_0$ and if $h_\alpha(\beta_k)$ is defined for all $k \in n$, then $h_\alpha(\beta_n) = q_\mu(n)$

where

$$
\mu(n) = \min \{ m \in \omega : \forall k \in n (q_m \leq h_\alpha(\beta_k) \leftrightarrow \beta_n \in \beta_k) \}.
$$

Further, let $Q_\alpha := h_\alpha[\alpha]$. Then by induction one can show that $h_\alpha$ and $Q_\alpha$ have the required properties (the details are left to the reader). $\Box$
Theorem 4.11. $\aleph_1 \leq^* 2^{\aleph_0}$.

Proof. It is enough to construct a surjection from the open interval $(0,1)$ onto $\omega_1$. Firstly notice that every real $r \in (0,1)$ can be written uniquely as

$$r = \sum_{n \in \omega} r_n \cdot 2^{-(n+1)}$$

where for all $n \in \omega$, $r_n \in \{0,1\}$, and infinitely many $r_n$’s are equal to 0. On the other hand, for every function $f \in {}^{\omega}2$ such that $\{n \in \omega : f(n) = 0\}$ is infinite there exists a unique real $r = \sum_{n \in \omega} f(n) \cdot 2^{-(n+1)}$ in $(0,1)$. Secondly, for $r \in (0,1)$ let $Q_r = \{q_n : r_{2n} = 1\}$ where the function which maps $n$ to $q_n$ is a bijection between $\omega$ and $Q \cap (0,1)$. If $Q_r$ is well-ordered by “$<$”, then let $\eta(r)$ be the order type of $(Q_r, <)$; otherwise, let $\eta(r) = 0$. Since the set of rational numbers is countable, $\eta$ is a function from $(0,1)$ to $\omega_1$. Moreover, the function $\eta$ is even surjective. Indeed, by Lemma 4.10 we know that for any $\alpha \in \omega_1$ there is a set of rational numbers $Q_\alpha \subseteq Q \cap (0,1)$ such that the order type of $(Q_\alpha, <)$ is equal to $\alpha$. Thus, for

$$r = \sum_{n \in N(Q_\alpha)} 2^{-(2n+1)}$$

we have $r \in (0,1)$ and $\eta(r) = \alpha$, and since $\alpha \in \omega_1$ was arbitrary this shows that $\eta$ is surjective.

In contrast to Theorem 4.11 the existence of an injection from $\omega_1$ into $\mathbb{R}$ is not provable in ZF, i.e., $\aleph_1 \not\leq 2^{\aleph_0}$ is consistent with ZF. For example there is no such injection in the case when the reals can be written as a countable union of countable sets (for the consistency of this statement with ZF see Chapter 17).

Proposition 4.12. If the set of real numbers is a countable union of countable sets, then $\aleph_1 \not\leq 2^{\aleph_0}$.

Proof. By Fact 4.3, $|\mathbb{R}| = |\omega^\omega|$. Thus, if $\mathbb{R}$ is a countable union of countable sets, then we also have $\omega \mathbb{R} = \bigcup_{n \in \omega} F_n$ where each $F_n$ is countable. The proof is by contraposition: Under the assumption that there is an injection $j : \omega_1 \hookrightarrow \mathbb{R}$ we show that $\omega \mathbb{R} \neq \bigcup_{n \in \omega} F_n$. Consider the function

$$G : \omega \rightarrow \mathcal{P}(\mathbb{R})$$

$$n \mapsto \{r \in \mathbb{R} : \exists f \in F_n \exists k \in \omega (f(k) = r)\}.$$ 

For each $n \in \omega$ we have $|G(n)| \leq \aleph_0$ and we can define $h : \omega \rightarrow \mathbb{R}$ by stipulating

$$h(n) := j(\alpha_n)$$

where $\alpha_n = \min \{\beta \in \omega_1 : j(\beta) \notin G(n)\}$. 


By definition $h \in \omega \mathbb{R}$, but on the other hand, $h$ does not belong to any set $F_n$ (for $n \in \omega$); since otherwise we would have $h(n) \in G(n)$ which contradicts the definition of $h(n)$. Thus, $h \notin \bigcup_{n \in \omega} F_n$ which shows that $\omega \mathbb{R}$ — and consequently $\mathbb{R}$ — cannot be covered by countably many countable sets.

As a consequence of Proposition 4.12 one can show that if $\mathbb{R}$ is a countable union of countable sets, then $\mathbb{R}$ can be partitioned into strictly more parts than real numbers exist, where a partition of $\mathbb{R}$ is a set $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R})$ such that $\bigcup \mathcal{R} = \mathbb{R}$ and for any distinct $x, y \in \mathcal{R}$, $x \cap y = \emptyset$.

Corollary 4.13. If the set of real numbers is a countable union of countable sets, then there exists a partition $\mathcal{R}$ of $\mathbb{R}$ such that $|\mathcal{R}| > |\mathbb{R}|$.

Proof. By Fact 4.3 and the Cantor-Bernstein Theorem 3.17 there exists a bijection between $\mathbb{R} \setminus (0, 1)$ and $\mathbb{R}$, and by Theorem 4.11 there exists a surjection from $(0, 1)$ onto $\omega_1$. Thus, there is a surjection $f : \mathbb{R} \to \mathbb{R} \setminus \omega_1$, and with $f$ we can define an equivalence relation "~" on $\mathbb{R}$ by stipulating $x \sim y \iff f(x) = f(y)$. Let $\mathcal{R} = \{[x] : x \in \mathbb{R}\}$. Then $\mathcal{R}$ is a partition of $\mathbb{R}$ and we have $|\mathcal{R}| = \aleph_1 + 2^{\aleph_0}$. By Proposition 4.12, $\aleph_1 \notin 2^{\aleph_0}$ and consequently $\aleph_1 + 2^{\aleph_0} \notin 2^{\aleph_0}$, and since $2^{\aleph_0} \leq \aleph_1 + 2^{\aleph_0}$ we have $2^{\aleph_0} < \aleph_1 + 2^{\aleph_0}$, in particular, $|\mathbb{R}| < |\mathcal{R}|$.

Ordinal Numbers revisited

In the previous chapter we have defined addition, multiplication, and exponentiation of ordinal numbers. Using these arithmetical operations we can show that every ordinal number can be uniquely represented in a standardised form, but first let us introduce some terminology: For ordinals $\alpha, \beta \in \Omega$ we will write $\beta < \alpha$ instead of $\beta \in \alpha$ and consequently we define $\beta \leq \alpha$ $\iff$ $\beta \in \alpha \lor \beta = \alpha$. Further notice that if $\beta \leq \alpha$, then there is a unique ordinal, denoted $\alpha - \beta$, such that $\beta + (\alpha - \beta) = \alpha$.

Lemma 4.14. For every ordinal $\alpha > 0$ there exists a unique ordinal $\alpha_0$ such that $\omega^{\alpha_0} \leq \alpha$ and $\omega^{\alpha_0 + 1} > \alpha$.

Proof. Firstly notice that by the rules of ordinal exponentiation, for $\gamma < \gamma'$ we have $\omega^\gamma < \omega^{\gamma'} \cdot \omega \leq \omega^{\gamma'} \cdot \omega^{\gamma'-\gamma} = \omega^{\gamma'}$. In particular, for any ordinal $\alpha_0$ we have $\omega^{\alpha_0} < \omega^{\alpha_0 + 1}$. Secondly notice that for all ordinals $\alpha$ we have $\omega^\alpha \geq \alpha$, hence, $\omega^{\alpha+1} > \omega^\alpha \geq \alpha$. Now, since $\alpha + 1$ is well-ordered by "$<" and $\omega^{\alpha+1} > \alpha \geq \omega^\alpha$, there is a unique least ordinal $\beta \leq \alpha + 1$ such that $\omega^\beta > \alpha$. It remains to show that $\beta$ is a successor ordinal, i.e., $\beta = \alpha_0 + 1$ for some $\alpha_0$. Indeed, if $\beta$ would be a limit ordinal, then $\omega^\beta = \bigcup_{\gamma \in \beta} \omega^\gamma$, and by definition of $\beta$ we would have $\omega^{\gamma} \leq \alpha$ (for all $\gamma \in \beta$). Since $\omega^{\gamma+1} > \omega^\gamma$ and since $\beta$ is a limit ordinal, this would imply that $\omega^\gamma \in \alpha$ whenever $\gamma \in \beta$ and consequently $\omega^\beta \leq \alpha$, whereas $\omega^\beta > \alpha$, a contradiction.
Lemma 4.15. Let \( \alpha \geq \omega \) be an infinite ordinal. Then there exist a positive integer \( k_0 \) and ordinals \( \alpha' \) and \( \alpha_0 \) where \( \alpha' < \omega^{\alpha_0} \) such that \( \alpha = \omega^{\alpha_0} \cdot k_0 + \alpha' \).
Moreover, the ordinals \( k_0, \alpha_0, \) and \( \alpha' \) are uniquely determined by \( \alpha \).

Proof. Let \( \alpha_0 \) be as in Lemma 4.14. Then \( \omega^{\alpha_0} \leq \alpha \) and \( \omega^{\alpha_0+1} > \alpha \). By a similar argument as in the proof of Lemma 4.14, this implies that there are positive integers \( k \) such that \( \omega^{\alpha_0} \cdot k > \alpha \). Let \( k_0 \) be the least integer such that \( \omega^{\alpha_0} \cdot (k_0 + 1) > \alpha \); then \( 1 \leq k_0 < \omega \) (notice that \( \omega^{\alpha_0} = \omega^{\alpha_0} \cdot 1 \leq \alpha \)). Finally, let \( \alpha' = (\alpha - \omega^{\alpha_0} \cdot k_0) \). Then \( \omega^{\alpha_0} \cdot k_0 + \alpha' = \alpha \) and since \( \omega^{\alpha_0} \cdot (k_0 + 1) = \omega^{\alpha_0} \cdot k_0 + \omega^{\alpha_0} > \alpha \), \( \alpha' < \omega^{\alpha_0} \). We leave it as an exercise to the reader to show that \( k_0, \alpha_0, \) and \( \alpha' \) are uniquely determined by \( \alpha \).

Now we are ready to prove the following result:

Theorem 4.16 (Cantor’s Normal Form Theorem). Every ordinal number \( \alpha > 0 \) can be uniquely represented in the form
\[
\alpha = \omega^{\alpha_0} \cdot k_0 + \omega^{\alpha_1} \cdot k_1 + \ldots + \omega^{\alpha_n} \cdot k_n
\]
where \( n + 1 \) and \( k_0, k_1, \ldots, k_n \) are positive integers and the ordinal exponents satisfy \( \alpha \geq \alpha_0 > \alpha_1 > \alpha_2 > \ldots > \alpha_n \geq 0 \).

Proof. By an iterative application of Lemma 4.15 we get
\[
\alpha = \omega^{\alpha_0} \cdot k_0 + \alpha'
\]
\[
\alpha' = \omega^{\alpha_1} \cdot k_1 + \alpha''
\]
\[
\alpha'' = \omega^{\alpha_2} \cdot k_2 + \alpha'''
\]
and so on.

The form \( \alpha = \omega^{\alpha_0} \cdot k_0 + \ldots + \omega^{\alpha_n} \cdot k_n \) is called the Cantor normal form of \( \alpha \), denoted \( \text{cnf}(\alpha) \). Notice that by Cantor’s Normal Form Theorem 4.16, every ordinal number can be written in a unique way in Cantor normal form.

For \( \alpha = \omega^{\alpha_0} \cdot k_0 + \ldots + \omega^{\alpha_n} \cdot k_n \) let \( \text{cnf}_0(\alpha) := \omega^{\alpha_0} \cdot k_0 \). The next lemma will be used to show that for every infinite ordinal \( \alpha \), there is a bijection between \( \alpha \) and \( \text{cnf}_0(\alpha) \).

Lemma 4.17. If \( \alpha_0, \alpha_1, k_0, k_1 \) are ordinals, where \( \alpha_0 > \alpha_1 \) and \( 0 < k_0, k_1 < \omega \), then
\[
\omega^{\alpha_1} \cdot k_1 + \omega^{\alpha_0} \cdot k_0 = \omega^{\alpha_0} \cdot k_0.
\]
Proof. By distributivity we get \( \omega^{\alpha_1} \cdot k_1 + \omega^{\alpha_0} \cdot k_0 = \omega^{\alpha_1} \cdot (k_1 + \omega^{\alpha_0 - \alpha_1} \cdot k_0) \), and since \( k_1 + \omega = \omega \) we get \( k_1 + \omega^{\alpha_0 - \alpha_1} \cdot k_0 = \omega^{\alpha_0 - \alpha_1} \cdot k_0 \). Thus, \( \omega^{\alpha_1} \cdot (k_1 + \omega^{\alpha_0 - \alpha_1} \cdot k_0) = \omega^{\alpha_1} \cdot k_1 + \omega^{\alpha_0} \cdot k_0 \).

Lemma 4.18. For each ordinal \( \alpha > 0 \) there exists a bijection between \( \alpha \) and \( \text{cnf}(\alpha) \).

Proof. Let \( \text{cnf}(\alpha) = \omega^{\alpha_n} \cdot k_n + \omega^{\alpha_{n-1}} \cdot k_{n-1} + \ldots + \omega^{\alpha_0} \cdot k_0 \) and define the “reverse Cantor normal form” of \( \alpha \), denoted \( \overline{\text{cnf}}(\alpha) \), by

\[
\overline{\text{cnf}}(\alpha) = \omega^{\alpha_n} \cdot k_n + \omega^{\alpha_{n-1}} \cdot k_{n-1} + \ldots + \omega^{\alpha_0} \cdot k_0.
\]

If \( \alpha < \omega \), then \( \alpha_0 = 0 \), hence, \( \alpha = \omega^{\alpha_0} \cdot k_0 = k_0 \) and therefore \( \alpha = \text{cnf}_0(\alpha) \).

If \( \alpha \geq \omega \), then by an iterative application of Lemma 4.17 we get \( \overline{\text{cnf}}(\alpha) = \omega^{\alpha_n} \cdot k_0 = \text{cnf}_0(\alpha) \), and since there is obviously a bijection between \( \alpha \) and \( \overline{\text{cnf}}(\alpha) \), there exists a bijection between \( \alpha \) and \( \text{cnf}_0(\alpha) \).

Now we are ready to show that for each infinite ordinal \( \alpha \), the cardinality of the set of all finite sequences which can be formed with elements of \( \alpha \) is the same as the cardinality of \( \alpha \). Moreover, we can show the following result:

Theorem 4.19. For each infinite ordinal \( \alpha \) we have

\[
|\alpha| = |\text{fin}(\alpha)| = |\text{seq}^+(\alpha)| = |\text{seq}(\alpha)|.
\]

Moreover, there exists a class function \( F \) such that for each infinite ordinal \( \alpha \geq \omega \), \( \{\alpha\} \times \text{seq}(\alpha) \subseteq \text{dom}(F) \) and \( F|\{\alpha\} \times \text{seq}(\alpha) \) induces an injection from \( \text{seq}(\alpha) \) into \( \alpha \).

Proof. Firstly notice that for every ordinal \( \alpha \), \( |\alpha| \leq |\text{fin}(\alpha)| \leq |\text{seq}^+(\alpha)| \leq |\text{seq}(\alpha)| \). In fact, there is a class function assigning to each ordinal \( \alpha \) some appropriate functions to witness these inequalities. Thus, it is enough to prove that for every infinite ordinal \( \alpha \), \( |\text{seq}(\alpha)| \leq |\alpha| \) uniformly; i.e., it is enough to show the existence of a class function \( F \) such that for every infinite ordinal \( \alpha \) and any distinct finite sequences \( s, t \in \text{seq}(\alpha) \) we have \( F((\alpha, s)) \in \alpha \) and \( F((\alpha, s)) \neq F((\alpha, t)) \). Let \( \alpha \) be an arbitrary but fixed infinite ordinal. In the following steps we will construct an injection \( F_\alpha : \text{seq}(\alpha) \mapsto \alpha \) such that the class function \( F \) defined by \( F((\alpha, s)) := F_\alpha(s) \) has the desired properties (notice that this requires that the function \( F_\alpha \) is fully determined by \( \alpha \).

First we give a detailed construction of an injection \( g_\alpha : \alpha \mapsto \omega^{\alpha_0} \), where \( \omega^{\alpha_0} \cdot k_0 = \text{cnf}_0(\alpha) \). By Lemma 4.18 there is a bijection between \( \alpha \) and \( \omega^{\alpha_0} \cdot k_0 \). Further, there is a bijection between the ordinal \( \omega^{\alpha_0} \cdot k_0 \) and the set \( \omega^{\alpha_0} \times k_0 \). Indeed, if \( \beta \in \omega^{\alpha_0} \cdot k_0 \), then there is a \( \beta' \in \omega^{\alpha_0} \) and an \( j \in k_0 \) such that \( \beta = \omega^{\alpha_0} \cdot j + \beta' \); let the image of \( \beta \) be \( (\beta', j) \). Similarly, there is a bijection between the set \( k_0 \times \omega^{\alpha_0} \) and the ordinal \( k_0 \cdot \omega^{\alpha_0} \), and since there is obviously a bijection between \( \omega^{\alpha_0} \times k_0 \) and \( k_0 \times \omega^{\alpha_0} \), there is a bijection between \( \alpha \) and
$k_0 \cdot \omega^{\alpha_0}$. Further, since $1 \leq k_0 < \omega$, there is an injection from $k_0 \cdot \omega^{\alpha_0}$ into $\omega \cdot \omega^{\alpha_0} = \omega^{1+\alpha_0}$, thus, there is an injection

$$g : \alpha \mapsto \omega^{1+\alpha_0}.$$ 

Notice that because $\alpha \geq \omega$, $\alpha_0 \geq 1$. Now we consider the following two cases:

If $\alpha_0 \geq \omega$, then $1 + \alpha_0 = \alpha_0$, thus, $g$ is an injection from $\alpha$ into $\omega^{\alpha_0}$; in this case let $g_\alpha := g$.

If $\alpha_0 < \omega$, then $1 + \alpha_0 = \alpha_0 + 1$ and there is a bijection between the ordinal $\omega^{\alpha_0 + 1}$ and the set of functions from $\alpha_0 + 1$ to $\omega$, denoted $\omega^{\alpha_0 + 1}$. Similar to the proof of Fact 4.2 let $p_0 < p_1 < \ldots < p_{\alpha_0}$ be the least $\alpha_0 + 1$ prime numbers and define $h : \alpha_0 + 1 \to \omega$ by stipulating $h(s) = \prod_{i \leq \alpha_0} p_i^{s(i) + 1}$. Then $h$ is injective and since $\alpha_0 \geq 1$ (notice that $\alpha \geq \omega$), there is an injection from $\alpha$ into $\omega^{\alpha_0}$; in this case let $g_\alpha$ be that injection.

Similarly, for each $n \in \omega$ we can construct an injection $f_{\alpha,n} : n \alpha \mapsto \alpha$. For $n = 0$ let $f_{\alpha,0}(\emptyset) := \emptyset$; and for $n > 0$ let $f_{\alpha,n}$ be defined by the following sequence of injections:

$$f_{\alpha,n} : n \alpha \overset{n(\omega^{\alpha_0})}{\longrightarrow} (\omega^{\alpha_0})^n \overset{\omega^{\alpha_0 n}}{\longrightarrow} \omega^{\alpha_0} \overset{\omega^{\alpha_0 \cdot \alpha_0}}{\longrightarrow} \omega^{n \alpha_0} \ldots$$

Now we can construct an injection from $\text{seq}(\alpha)$ into $\alpha$: Firstly notice that there is a natural bijection between $\text{seq}(\alpha)$ and $\bigcup_{n \in \omega} n \alpha$, thus, it is enough to construct an injection $F_\alpha$ from $\bigcup_{n \in \omega} n \alpha$ into $\alpha$. If $s \in \bigcup_{n \in \omega} n \alpha$, then $s$ is a finite set of ordered pairs (i.e., $|s| \in \omega$) and $f_{\alpha,|s|}$ is an injection from $|s|$ into $\alpha$, in particular, $f_{\alpha,|s|}(s) \in \alpha$. Finally let us define $F_\alpha : \bigcup_{n \in \omega} n \alpha \to \alpha$ by stipulating

$$F(s) := f_{\alpha,2}\left(\{0,|s|, (1,f_{\alpha,|s|}(s))\}\right).$$

Then, since $\alpha$ is infinite, $|s| \in \alpha$, and since $f_{\alpha,2}$ is an injection from $2 \alpha$ into $\alpha$, $F_\alpha$ is injective.

As an application of Theorem 4.19 let us prove that whenever we have an injection from $\mathcal{P}(A)$ into $A \times A$, then $A$ has at most four elements.
Theorem 4.20. \(2^m \leq m^2 \rightarrow m \leq 4\).

Proof. If \(m\) is finite, an easy calculation shows that \(2^m \leq m^2\) implies that \(m \in \{2, 3, 4\}\). Thus, let \(m\) be infinite and assume towards a contradiction that \(2^m \leq m^2\). Let \(A\) be a set of cardinality \(m\) and let \(f_0 : \mathcal{P}(A) \rightarrow A \times A\). With the function \(f_0\) we can construct an injective class function from \(\Omega\) into \(A\), which is obviously a contradiction to the Axiom Schema of Replacement — which implies that there is no injection from a proper class (like \(\Omega\)) into a set.

Firstly we construct an injection \(F_\omega : \omega \hookrightarrow A\). Let \(a_0, a_1, a_2, a_3, a_4\) be five distinct elements of \(A\) and define \(F_5 : 5 \rightarrow A\) by stipulating \(F_5(i) := a_i\) (for all \(i \in 5\)); further let \(S_5 := F_5[5]\) [i.e., \(S_5 = \{F_5(i) : i \in 5\}\)]. Assume that for some \(n \geq 5\) we have already constructed an injection \(F_n : n \hookrightarrow A\). For any distinct sets \(x, y \in \mathcal{P}(S_n)\), where \(S_n := F_n[n]\), let

\[x \prec y \iff |x| < |y| \lor \exists i \in n \left(F(i) \in (x \setminus y) \land \forall j \in i \left(F(j) \in x \leftrightarrow F(j) \in y\right)\right).\]

Since \(S_n\) is finite, the relation \(\prec\) is a well-ordering, and since \(n \geq 5\), \(|\mathcal{P}(S_n)| = 2^n > n^2 = |S_n \times S_n|\). Thus, there exists a \(\prec\)-minimal set \(x \subseteq S_n\) such that \(f_0(x) \notin S_n \times S_n\). Let \(f_0(x) = (b_0, b_1)\) and let

\[a_n = \begin{cases} b_0 & \text{if } b_0 \notin S_n, \\ b_1 & \text{otherwise}. \end{cases}\]

Define \(F_{n+1} := F_n \cup \{(n, a_n)\}\) and let \(S_{n+1} := S_n \cup \{F_{n+1}(n)\}\). Then \(F_{n+1}\) is an injection from \(n + 1\) into \(A\), and \(S_{n+1} = F_{n+1}[n+1]\). Proceeding this way we finally get an injection \(F_\omega : \omega \hookrightarrow A\) as well as a countably infinite set \(S_\omega = F_\omega[\omega] \subseteq A\).

Assume now that we have already constructed an injection \(F_\alpha : \alpha \rightarrow A\) for some infinite ordinal \(\alpha \geq \omega\) and let \(S_\alpha := F_\alpha[\alpha]\). By Theorem 4.19 there is a canonical bijection \(g : \alpha \rightarrow \alpha \times \alpha\). With \(g\) we can define a bijection \(\hat{g} : S_\alpha \rightarrow S_\alpha \times S_\alpha\) by stipulating

\[\hat{g}(F_\alpha(\beta)) = \langle F_\alpha(\beta_0), F_\alpha(\beta_1) \rangle \quad \text{where } \beta = g^{-1}(\langle \beta_0, \beta_1 \rangle).\]

Further, define a mapping \(\Gamma : S_\alpha \rightarrow \mathcal{P}(S_\alpha)\) by stipulating

\[\Gamma(a) = \begin{cases} x \subseteq S_\alpha & \text{if } f_0(x) = \hat{g}(a), \\ \emptyset & \text{otherwise.} \end{cases}\]

and let

\[M = \{a \in S_\alpha : a \notin \Gamma(a)\}.\]

Then \(M \in \mathcal{P}(S_\alpha)\) and let \(f_0(M) = \langle b_0, b_1 \rangle \in A \times A\). If \(\langle b_0, b_1 \rangle \in S_\alpha \times S_\alpha\), then \(f_0(M) = \hat{g}(a)\) for some \(a \in S_\alpha\), and hence \(\Gamma(a) = M\); but \(a \in \Gamma(a) \leftrightarrow a \in M \leftrightarrow a \notin \Gamma(a)\), which is obviously impossible. Thus, \(\langle b_0, b_1 \rangle \notin S_\alpha \times S_\alpha\) and we let
\[ a_\alpha = \begin{cases} b_0 & \text{if } b_0 \notin S_\alpha, \\ b_1 & \text{otherwise.} \end{cases} \]

Further, define \( F_{\alpha+1} := F_\alpha \cup \{(\alpha, a_\alpha)\} \) and let \( S_{\alpha+1} := S_\alpha \cup \{a_\alpha\} \). Then \( F_{\alpha+1} \) is an injection from \( \alpha + 1 \) into \( A \), and \( S_{\alpha+1} = F_{\alpha+1}[\alpha + 1] \). Finally, if \( \lambda \) is a limit ordinal and \( F_\beta \) is defined for each \( \beta \in \lambda \) we define \( F_\lambda := \bigcup_{\beta \in \lambda} F_\beta \).

Now, by the Transfinite Recursion Theorem 3.19, \( \bigcup_{\alpha \in \Omega} F_\alpha \) is an injective class function which maps \( \Omega \) into \( A \); a contradiction to Hartogs’ Theorem. \( \dashv \)

The idea of the previous proof — getting a contradiction by constructing an injective class function from \( \Omega \) into a given set — is used again in the proofs of Theorem 4.21, Proposition 4.22, and Lemma 4.23.

### More Cardinal Relations

**Theorem 4.21.** If \( m \) is an infinite cardinal, then \( \text{fin}(m) < 2^m \).

**Proof.** Let \( A \) be an arbitrary but fixed infinite set of cardinality \( m \). Obviously, the identity mapping is an injection from \( \text{fin}(A) \) into \( \mathcal{P}(A) \), hence, \( \text{fin}(m) \leq 2^m \). Now, assume towards a contradiction that \( |\mathcal{P}(A)| = |\text{fin}(A)| \) and let \( f_0 : \mathcal{P}(A) \rightarrow \text{fin}(A) \) be a bijection. The mapping will be used in order to construct an injective class function \( F : \Omega \leftrightarrow \text{fin}(A) \). First we define an injection \( F_\omega : \omega \leftrightarrow \text{fin}(A) \) by stipulating

\[ F_\omega(n) = f_0^{n+1}(A) \]

where \( f_0^0(A) := f_0(A) \) and for positive integers \( k \), \( f_0^{k+1}(A) := f_0(f_0^k(A)) \).

Then, since \( A \) is infinite, \( F_\omega \) is indeed an injection.

Assume that we have already constructed an injection \( F_\alpha : \alpha \leftrightarrow \text{fin}(A) \) for some infinite ordinal \( \alpha \geq \omega \) and for \( \iota \in \alpha \) let \( s_\iota := F(\iota) \). Notice that \( s_\iota \neq s_{\iota'} \) whenever \( \iota \neq \iota' \). Define an equivalence relation on \( A \) by

\[ x \sim y \iff \forall \iota \in \alpha (x \in s_\iota \leftrightarrow y \in s_\iota). \]

For \( x \in A \) and \( \mu \in \alpha \) define

\[ D_{x,\mu} = \bigcap \{s_\iota : \iota \in \mu \land x \in s_\iota \} \]

where we define for the moment \( \bigcap \emptyset := A \), and let

\[ g_x = \{\mu \in \alpha : x \in s_\mu \land (s_\mu \cap D_{x,\mu} \neq D_{x,\mu})\}. \]
More cardinal relations: \( \text{fin}(m) < \omega^m \)

We leave it as an exercise to the reader to show that for any \( x, y \in A \), \( g_x = g_y \) iff \( x \sim y \). Hence, there is a bijection between \( \{x^- : x \in A\} \) and \( \{g_x : x \in A\} \).

Further, for each \( x \in A \), \( g_x \in \text{fin}(\alpha) \). To see this, let \( \mu_0 < \mu_1 < \mu_2 < \ldots \) be the ordinals in \( g_x \) in increasing order. By definition we have:

1. \( x \notin s_\iota \), whenever \( \iota \in \mu_0 \)
2. \( x \in s_{\mu_0} \) and \( s_{\mu_0} = D_{x, \mu_0 + 1} \)
3. \( D_{x, \mu_0 + 1} \supseteq D_{x, \mu_1 + 1} \supseteq D_{x, \mu_2 + 1} \) \( \supseteq \ldots \)

By (2), \( D_{x, \mu_0 + 1} \) is finite, and therefore the decreasing sequence (3) must be finite too, which implies that also \( g_x \) is finite.

Since \( \{g_x : x \in A\} \subseteq \text{fin}(\alpha) \) we can apply Theorem 4.19 to obtain an injection \( h : \{g_x : x \in A\} \hookrightarrow \alpha \). The set \( h(\{g_x : x \in A\}) \), as a subset of \( \alpha \), is well-ordered by “\( \in \)”. Let \( \gamma \) be the order type of \( h(\{g_x : x \in A\}) \). Then \( \gamma \leq \alpha \) and for each \( g_x \) assign an ordinal number \( \eta(g_x) \) in \( \gamma \) such that the mapping \( \eta : \{g_x : x \in A\} \rightarrow \gamma \) is bijective. For each \( \iota \in \alpha \), \( s_\iota \) is the union of at most finitely many equivalence classes. Thus, we can construct an injection from \( \alpha \) into \( \text{fin}(\gamma) \) by stipulating

\[
\iota \mapsto \{\xi \in \gamma : \exists x \in s_\iota (\eta(g_x) = \xi)\}.
\]

Because by Theorem 4.19 we can construct a bijection between \( \text{fin}(\gamma) \) and \( \gamma \), we can also construct an injection from \( \alpha \) into \( \gamma \), and because \( \gamma \leq \alpha \), by the Cantor-Bernstein Theorem 3.17 we finally get a bijection \( H : \gamma \rightarrow \alpha \) between \( \gamma \) and \( \alpha \). Define the function \( \Gamma : A \rightarrow \mathcal{P}(A) \) by stipulating

\[
\Gamma(x) = f_0^{-1}(s_{H(\eta(g_x))})
\]

and consider the set

\[
M = \{x \in A : x \notin \Gamma(x)\}.
\]

We claim that the set \( M \) does not belong to \( \{f_0^{-1}(s_\iota) : \iota \in \alpha\} \). Indeed, if there would be a \( \beta \in \alpha \) such that \( f_0^{-1}(s_\beta) = M \), then there would also be an equivalence class \([x^-] \), which corresponds to \( g_x \), such that

\[
\beta = H(\eta(g_x)) = \Gamma(x).
\]

For each \( y \in [x^-] \) we have \( \Gamma(y) = M \), and \( y \in \Gamma(y) \equiv y \in M \equiv y \notin \Gamma(y) \), which is obviously impossible.

Now, let \( s_\alpha := f_0^{-1}(M) \) and define \( F_\alpha + 1 := F_\alpha \cup \{s_\alpha\} \). Then \( F_\alpha + 1 \) is an injection from \( \alpha + 1 \) into \( \text{fin}(A) \). Finally, if \( \lambda \) is a limit ordinal and \( F_\beta \) is defined for each \( \beta \in \lambda \), then define \( F_\lambda := \bigcup_{\beta \in \lambda} F_\beta \). Thus, by the Transfinite Recursion Theorem 3.19, \( \bigcup_{\alpha \in \Omega} F_\alpha \) is an injective class function which maps \( \Omega \) into \( \text{fin}(A) \); a contradiction to Hartogs’ Theorem.

Even though \( \text{fin}(m) < \omega^m \) (for all infinite cardinals \( m \)), it might be possible that for some natural number \( n \), \( n \cdot \text{fin}(m) = \omega^m \). The next result shows that in that case, \( n \) must be a power of 2.
Proposition 4.22. If $2^m = n \text{fin}(m)$ for some natural number $n$, then $n = 2^k$ for some $k \in \omega$.

Proof. If the cardinal $m$ is finite, then $2^m = \text{fin}(m) = 1 \cdot \text{fin}(m) = 2^0 \cdot \text{fin}(m)$. So, let $m$ be an infinite cardinal and let $A$ be an arbitrary but fixed set of cardinality $m$. Further, let $n$ be a natural number which is not a power of 2. Assume towards a contradiction that $|\mathcal{P}(A)| = |n \times \text{fin}(A)|$. Let

$$f_0 : n \times \text{fin}(A) \to \mathcal{P}(A)$$

be a bijection which will be used to construct an injective class function from $\Omega$ into $\text{fin}(A)$. Let $(m_0, x_0) := f_0^{-1}(A)$. Assume that for some $\ell \in \omega$, $x_0, x_1, \ldots, x_\ell$ are pairwise distinct finite subsets of $A$. For each $i \in n$ and $j \leq \ell$ let

$$X_{i,j} = f_0((i, x_j)) .$$

On $A$ define an equivalence relation by stipulating

$$a \sim b \iff \forall i \in n \forall j \leq \ell (a \in X_{i,j} \leftrightarrow b \in X_{i,j}) .$$

Further, let $\text{Eq} := \{[a] : a \in A\}$ be the set of all equivalence classes and let $k_0 := |\text{Eq}|$. Now define an ordering “$\prec$” on the set \{\(X_{i,j} : i \in n \wedge j \leq \ell\)\}, for example define

$$X_{i,j} \prec X_{i',j'} \iff j < j' \vee (j = j' \wedge i < i') .$$

The ordering “$\prec$” induces in a natural way an ordering on the set $\text{Eq}$, and consequently of the set $E = \{\bigcup Y : Y \subseteq \text{Eq}\}$. Since the equivalence classes in $\text{Eq}$ are pairwise disjoint, $|E| = 2^{k_0}$. Notice that $2^{k_0} \geq n \cdot (\ell + 1)$, and since $n$ is not a power of 2, there is a least set $\bigcup Y_0 \in E$ (least with respect to the ordering on $E$ induced by “$\prec$”) such that $f_0^{-1}(\bigcup Y_0) = (m_{\ell+1}, x_{\ell+1})$ and $x_{\ell+1} \notin \{x_j : j \leq \ell\}$. For $i \in n$ define $X_{i,\ell+1} := f_0((i, x_{\ell+1}))$ and proceed as before. Finally we get an infinite sequence $x_0, x_1, \ldots$ of pairwise distinct finite subsets of $A$ which shows that $\text{fin}(A)$ is transfinite, i.e., there exists an injection $F_\omega : \omega \to \text{fin}(A)$.

Assume that we have already constructed an injection $F_\alpha : \alpha \to \text{fin}(A)$ for some infinite ordinal $\alpha \geq \omega$. Using the fact that there is a bijection between $n \cdot \alpha$ and $\alpha$, by the same arguments as in the proof of Theorem 4.21 we can construct an injection $F_{\alpha+1} : \alpha + 1 \to \text{fin}(A)$ and finally obtain an injective class function from $\Omega$ into $\text{fin}(A)$; a contradiction to Hartogs’ Theorem. \hfill \Box

Even though Proposition 4.22 looks a little bland, one cannot do better in $ZF$, i.e., for all $k \in \omega$, the statement “$\exists m (2^m = 2^k \cdot \text{fin}(m))$” is consistent with $ZF$ (cf. Proposition 7.5).
More cardinal relations: $\text{seq}^{-1}(m) \neq 2^m \neq \text{seq}(m)$

First we prove that the inequality $\text{seq}^{-1}(m) \neq 2^m \neq \text{seq}(m)$ whenever $m \geq 2$.

**Lemma 4.23.** Let $m$ be a transfinite cardinal number. Then $2^m \not\leq \text{seq}(m)$ and consequently also $2^m \not\leq \text{seq}^{-1}(m)$.

**Proof.** Let $A$ be a set of cardinality $m$ and assume towards a contradiction that there exists an injection $f_0 : \mathcal{P}(A) \rightarrow \text{seq}(A)$. Since $A$ is transfinite there is an injection $F_\omega : \omega \rightarrow A$ and let $S_\omega := F_\omega[\omega]$. Assume that we have already constructed an injection $F_\alpha : \alpha \rightarrow A$ for some infinite ordinal $\alpha \geq \omega$ and let $S_\alpha := F_\alpha[\alpha]$. By Theorem 4.19 there is a bijection between $\alpha$ and $\text{seq}(\alpha)$, and consequently we can define a bijection $\bar{g} : S_\alpha \rightarrow \text{seq}(S_\alpha)$. Further, define $\Gamma : S_\alpha \rightarrow \mathcal{P}(S_\alpha)$ by stipulating

$$\Gamma(a) = \begin{cases} x \subseteq S_\alpha & \text{if } f_0(x) = \bar{g}(a), \\ \emptyset & \text{otherwise,} \end{cases}$$

and let

$$M = \{ a \in S_\alpha : a \notin \Gamma(a) \}.$$

Then $M \in \mathcal{P}(S_\alpha)$ and $f_0(M) = \langle b_0, b_1, \ldots, b_n \rangle \in \text{seq}(A) \setminus \text{seq}(S_\alpha)$. Now, let $a_i := b_i$, where $i \leq n$ is the least number such that $b_i \notin S_\alpha$ and define $F_{\alpha+1} := F_\alpha \cup \{ (\alpha, a_\alpha) \}$ and $S_{\alpha+1} := S_\alpha \cup \{ F_{\alpha+1}(\alpha) \}$. Then $F_{\alpha+1}$ is an injection from $\alpha + 1$ into $A$, and $S_{\alpha+1} = F_{\alpha+1}[\alpha + 1]$. Finally, if $\lambda$ is a limit ordinal and $F_\lambda$ is defined for each $\beta \in \lambda$ we define $F_\lambda := \bigcup_{\beta \in \lambda} F_\beta$ and finally get that $\bigcup_{\alpha \in \omega} F_\alpha$ is an injective class function; a contradiction to Hartogs’ Theorem.

To prove that $\text{seq}(m) \neq 2^m$ whenever $m \geq 1$ one could for example show that $\text{seq}(m) = 2^m$ implies that $m$ is transfinite by using similar ideas as above, but we get a slightly more elegant proof by showing that $\text{seq}(m) = 2^m$ implies that $\text{seq}(m + \aleph_0) = 2^{m+\aleph_0}$.

**Theorem 4.24.** For all cardinals $m \geq 1$, $\text{seq}(m) \neq 2^m$.

**Proof.** We will show that whenever $m \geq 1$ is a cardinal such that $2^m = \text{seq}(m)$, then $2^{m+\aleph_0} = \text{seq}(m + \aleph_0)$ which is a contradiction to Lemma 4.23. Let the set $A$ be such $|A| = m$ and $A \cap \omega = \emptyset$. Further, let $f_0 : \mathcal{P}(A) \rightarrow \text{seq}(A)$ be a bijection. For a fixed element $a_0 \in A$ and $n \in \omega$ let

$$s_n = \langle a_0, \ldots, a_0 \rangle_{n \text{-times}}.$$

With the sequences $s_n$ we can define an injection $g : \omega \rightarrow \mathcal{P}(A)$ by stipulating $g(n) := f_0^{-1}(s_n)$, which shows that $\mathcal{P}(A)$ is transfinite, i.e., $\aleph_0 \leq 2^m$. Thus,
by Proposition 4.4 we have $2^{\aleph_0} \leq 2^\text{m}$ which implies that there exists an injection $h : \mathcal{P}(\omega) \rightarrow \mathcal{P}(A)$. Finally let

$$F : \mathcal{P}(A) \times \mathcal{P}(\omega) \rightarrow \text{seq}(A \cup \omega)$$

$$(x, y) \mapsto f_0(x) \circ f_0(h(y))$$

where $\circ$ denotes the concatenation of the sequences $s$ and $t$. Then $F$ is injective and we consequently get $2^{\aleph_0 + \omega} = 2^\text{m} \cdot 2^{\aleph_0} = \text{seq}(\text{m} + \aleph_0)$. \\

In order to prove that $\text{seq}^{-1}(m) \neq 2^\text{m}$ whenever $m \geq 2$ we show that $\text{seq}^{-1}(m) = 2^\text{m}$ would imply that $m$ is transfinite, which is a contradiction to Lemma 4.23. However, before we have to introduce some notation concerning finite sequences of natural numbers.

For $n \in \omega$ let $n^* := \lfloor \text{seq}^{-1}(n) \rfloor$ be the number of non-repetitive sequences (i.e., sequences without repetitions) we can build with $n$ distinct objects (e.g., with $\{0, \ldots, n - 1\} = n$). It is not hard to verify that

$$n^* = \sum_{k=0}^{n} \binom{n}{k} k! = \sum_{j=0}^{n} \frac{n!}{j!},$$

and that for all positive integers $n$ we have $n^* = \lfloor en! \rfloor$, where $\lfloor x \rfloor$ denotes the integer part of a real number $x$ and $e$ is the Euler number. Obviously, $0^* = 1$ and $n^* = n \cdot (n - 1)^* + 1$, which implies that

$$n^* = e \int_1^\infty t^n e^{-t} dt.$$

The number $n^*$ is also the number of paths (without loops) in the complete graph on $n + 2$ vertices starting in one vertex and ending in another.

The first few numbers of the integer sequence $n^*$ are $0^* = 1, 1^* = 2, 2^* = 5, 3^* = 16, 4^* = 65, 5^* = 326$, and further we get e.g., $100^* \approx 2.53687 \cdot 10^{158}$ and $256^* \approx 2.33179 \cdot 10^{507}$.

For each positive integer $q$, an easy calculation modulo $q$ shows that for all $n \in \omega$ we have $n^* \equiv (n + q)^* \mod q$. In particular, if $q \mid n^*$, then $q \mid (n + q)^*$. Now we can ask whether there is a positive integer $t < q$ such that $q \mid (n + t)^*$ and $q \nmid n^*$. The following lemma shows that this is not the case whenever $q$ is a power of 2.

**Lemma 4.25.** If $2^k \mid n^*\text{ and } 2^k \mid (n + t)^*$ for some $t \in \omega$, then $2^k \mid t$.

**Proof.** For $k \leq 3$, an easy calculation modulo $2^k$ shows that for each $n$, if $2^k \mid n^*$, then $2^k \mid (n + t)^*$ whenever $0 < t < 2^k$.

Assume towards a contradiction that there is a smallest $k \geq 3$ such that $2^{k+1} \mid n^*$ and $2^{k+1} \mid (n + t)^*$ for some integer $t$ with $0 < t < 2^{k+1}$. Notice that since $k \geq 3$, $n \geq 3$. Then, because $2^k \mid 2^{k+1}$, we have $2^{k} \mid n^*$ and $2^{k} \mid (n + t)^*$,
and by the choice of $k$ we get $t = 2^k$. Let us now compute $(n + 2^k)^r$ by writing down $\sum_{i=0}^{n+2^k} \frac{(n+2^k)^r}{i!}$ explicitly:

$$
(n + 2^k)^r = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot 2^k \cdot (2^k + 1) \cdot \ldots \cdot (2^k + n) + [1]
2 \cdot 3 \cdot \ldots \cdot 2^k \cdot (2^k + 1) \cdot \ldots \cdot (2^k + n) + [2]
3 \cdot \ldots \cdot 2^k \cdot (2^k + 1) \cdot \ldots \cdot (2^k + n) + [3]
\vdots
\vdots
\vdots
2^k \cdot (2^k + 1) \cdot \ldots \cdot (2^k + n) + [2^k]
(2^k + 1) \cdot \ldots \cdot (2^k + n) + [2^k + 1]
\vdots
\vdots
(2^k + n) + [2^k + n]
1 [2^k + n + 1]
$$

Since $k \geq 4$ and $n \geq 3$, $2^{k+1}$ divides rows $[1] - [2^k]$. In order to calculate the products in rows $[2^k + 1] - [2^k + n + 1]$ (modulo $2^{k+1}$), we only have to consider products which are not obviously divisible by $2^{k+1}$. So, since $2^{k+1} \nmid (n + 2^k)^r$, for a suitable natural number $r$ we have

$$(n + 2^k)^r = 2^k \cdot \left( \sum_{j=0}^{n-1} \sum_{i>j} \frac{n!}{i \cdot j!} \right) + n^r + 2^{k+1} \cdot r.$$ 

We know that $2^{k+1} \mid n^r$ where $n \geq 3$ and $k \geq 4$, and because $n^r$ is even, $n$ has to be odd. If $j$ is equal to $n - 1$, $n - 2$, or $n - 3$, then $\sum_{i>j} \frac{n!}{i \cdot j!}$ is odd, and if $0 \leq j \leq (n - 4)$, then $\sum_{i>j} \frac{n!}{i \cdot j!}$ is even. So,

$$
\sum_{j=0}^{n-1} \sum_{i>j} \frac{n!}{i \cdot j!}
$$

is odd, and since $2^{k+1} \mid n^r$, $2^{k+1} \nmid (n + 2^k)^r$.

Now we are ready to prove the following result:

**Theorem 4.26.** For all cardinals $m \geq 2$, $\text{seq}^{-1}(m) \neq 2^m$.

**Proof.** By Lemma 4.23 it is enough to prove that for $m \geq 2$, $\text{seq}^{-1}(m) = 2^m \rightarrow \mathbb{N}_0 \leq m$. Let $A$ be an arbitrary set of cardinality $m$ and assume that

$$f_0 : \mathcal{P}(A) \rightarrow \text{seq}^{-1}(A)$$

is a bijection between $\mathcal{P}(A)$ and $\text{seq}^{-1}(A)$. We shall use this bijection to show that $A$ is transfinite. In fact it is enough to show that every finite sequence
\[ s_n = \langle a_0, \ldots, a_{n-1} \rangle \in \text{seq}^{-1}(A) \text{ of length } n \text{ can be extended canonically to a sequence } s_{n+1} = \langle a_n \rangle \in \text{seq}^{-1}(A) \text{ of length } n + 1. \]

Let \( a_0 \) and \( a_1 \) be two distinct elements of \( A \) and assume that for some \( n \geq 2 \) we already have constructed a sequence \( s_n = \langle a_0, a_1, \ldots, a_{n-1} \rangle \) of distinct elements of \( A \) and let \( S_n = \{ a_i : i \in n \} \). The sequence \( s_n \) induces in a natural way an ordering on the set \( \text{seq}^{-1}(S_n) \), e.g., order \( \text{seq}^{-1}(S_n) \) by length and lexicographically. Let us define an equivalence relation on \( A \) by stipulating

\[ a \sim b \iff \forall s \in \text{seq}^{-1}(S_n) \left( a \in f_0^{-1}(s) \leftrightarrow b \in f_0^{-1}(s) \right). \]

Let \( \text{Eq}(n) := \{ [a]^- : a \in A \} \) be the set of all equivalence classes. The ordering on \( \text{seq}^{-1}(S_n) \) induces an ordering on \( \text{Eq}(n) \). Let

\[ k_0 = |\text{Eq}(n)|. \]

Then \( 2^{k_0} \) is equal to the cardinality of \( \mathcal{P}(\text{Eq}(n)) \). Identify \( \bigcup Y : Y \subseteq \text{Eq}(n) \) with the set of all functions \( \bar{g} \in \text{Eq}(n) \). Now, the ordering on \( \text{Eq}(n) \) induces in a natural way an ordering on the set of functions \( \text{Eq}(n) \). By construction we have \( n^* = |\text{seq}^{-1}(S_n)| \leq 2^{k_0} \), i.e., we have either \( n^* < 2^{k_0} \) or \( n^* = 2^{k_0} \).

**Case 1:** If \( n^* < 2^{k_0} \), then there exists a least function \( \bar{g}_0 \in \text{Eq}(n) \) (least with respect to the ordering on \( \text{Eq}(n) \)) such that \( \bar{g}_0 \notin \{ x_n : s \in \text{seq}^{-1}(S_n) \} \), where \( x_n \) is the characteristic function of the set of all equivalence classes included in \( f_0^{-1}(s) \). In particular we get \( f_0(\bar{g}_0) \notin \text{seq}^{-1}(S_n) \). Let \( a_n \in A \) be the first element in the sequence \( f_0(\bar{g}_0) \) which does not belong to \( S_n \). Now, \( s_n \prec (a_n) \in \text{seq}^{-1}(A) \) is a sequence of length \( n + 1 \) and we are done.

**Case 2:** Suppose that \( n^* = 2^{k_0} \). For arbitrary elements \( a \in A \setminus S_n \) let us resume the construction with the sequence \( s_n \prec (a) \). By a parity argument one easily verifies that \( (n + 1)^* \) is not an integer power of 2, and thus, we are in Case 1. We proceed as long as we are in Case 1. If there is an element \( a \in A \setminus S_n \) such that we are always in Case 1, then we can construct an infinite non-repetitive sequence of elements of \( A \) and we are done.

Assume now that no matter which element \( a \in A \setminus S_n \) we resume our construction, we always get back to Case 2. We then have the following situation: Starting with any element \( a \in A \setminus S_n \) we get a non-repetitive sequence of elements of \( A \) of length \( n + \ell + 1 \) (for some positive integer \( \ell \)) where \( (n + \ell + 1)^* \) is an integer power of 2. Let \( s_{n+\ell}^0 = \langle a_0, a_1, \ldots, a_{n+\ell} \rangle \) be that sequence and let \( S_n^0 = \{ a_0, a_1, \ldots, a_{n+\ell} \} \). By construction we have \( a \in S_n^0 \), i.e., \( a \) belongs to the corresponding sequence \( s_{n+\ell}^0 \). However, \( S_n^0 \) is not necessarily the union of elements of \( \text{Eq}(n) \), which leads to the following definition:

A subset of \( A \) is called **good** if it is not the union of elements of \( \text{Eq}(n) \).
For every set $X \subseteq A$ which is good we have $f_0(X) \notin \text{seq}^{-1}(S_n)$, which implies that there is a first element in the sequence $f_0(X)$ which does not belong to the set $S_n$. Thus, it is enough to determine a good subset of $A$. For this, consider the set

$$T_{\text{min}} := \{ a \in A \setminus S_n : S_n^a \text{ is good and of least cardinality} \}.$$  

Notice that for every $a \in A \setminus S_n$, $S_n^a$ is finite and contains $a$, and since $A \setminus S_n$ is infinite, there is an $S_n^a$ (for some $a \in A \setminus S_n$) which is good, thus, $T_{\text{min}} \neq \emptyset$. If $T_{\text{min}}$ is good, use $f_0(T_{\text{min}})$ to construct a non-repetitive sequence in $A$ of length $(n+1)$, and we are done. Otherwise, let $m_T := |S_n^a|$ for some $a$ in $T_{\text{min}}$ (notice that by our assumptions, $m_T$ is a positive integer). For each $a \in T_{\text{min}}$ let us construct a non-repetitive sequence $\text{SEQ}^a$ of elements of $S_n^a$ of length $m_T$ in such a way that for all $a, b \in T_{\text{min}}$:

$$S_n^a = S_n^b \implies \text{SEQ}^a = \text{SEQ}^b$$

In order to do so, let $a \in T_{\text{min}}$ be arbitrary. Because $S_n^a \in T_{\text{min}}$, $S_n^a$ is good, thus

$$f_0(S_n^a) \notin \text{seq}^{-1}(S_n),$$

hence, there is a first element $a_n$ in the sequence $f_0(S_n^a)$ which does not belong to $S_n$. Repeat the construction starting with the sequence $s_{n+1} = s_n \setminus (a_n)$ and consider the set $S_n^{a_n}$. If $S_n^{a_n} = S_n^a$, then the corresponding sequence $s^{a_n} \in \text{seq}^{-1}(S_n^a)$ is of length $m_T$ and we define $\text{SEQ}^a := s^{a_n}$. On the other hand, if $S_n^{a_n} \subseteq S_n^a$, then $S_n^{a_n}$ is not good (since $S_n^a$ is a good set of least cardinality), i.e., $S_n^{a_n}$ is the union of elements of Eq(n). Let $S' = S_n^a \setminus S_n^{a_n}$ and let $s' \in \text{seq}^{-1}(S')$ be the corresponding sequence. Then $S'$ is good, which implies that $f_0(S') \notin \text{seq}^{-1}(S_n^a)$, and let $a'$ be the first element in the sequence $f_0(S')$ which does not belong to $S_n^{a_n}$. Now proceed building the sequence $\text{SEQ}^a$ by starting with the sequence $s'^{-1}(a')$. Notice that by construction the sequence $\text{SEQ}^a$ depends only on the set $S_n^a$, thus, for all $a, b \in T_{\text{min}}$, $\text{SEQ}^a = \text{SEQ}^b$ whenever $S_n^a = S_n^b$.  

So far, for each $a \in T_{\text{min}}$ with $|S_n^a| = m_T$ we can construct a non-repetitive sequence $\text{SEQ}^a \in \text{seq}^{-1}(S_n^a)$ of length $m_T > n$. On the other hand, we still have to determine in a constructive way a good subset of $A$ which contains $S_n$ — even though $S_n^a$ is good for each $a \in T_{\text{min}}$, it is not clear which set $S_n^a$ we should choose. Now, for $i < m_T$ define

$$Q_i := \{ b \in A : b \text{ is the } i \text{th element in } \text{SEQ}^a \text{ for some } a \in T_{\text{min}} \}.$$ 

**Claim.** There is a smallest $j_0 < m_T$ such that $Q_{j_0}$ is good. 

**Proof of Claim.** For any $a \in T_{\text{min}}$ let

$$a^a := \{ a' \in A : S_n^{a'} = S_n^a \}.$$
which are the elements of the finite set $S^a$ which are to some extent indistinguishable, and further let $t_0$ denote the least cardinality of the sets $a_i^*$, where $a_i \in T_{\min}$. Note that if for some $i \neq j_0$, $a_i \in Q_i \cap Q_j$, then $S^a_i$ cannot be good (otherwise, $SQ^a$ would not be unique). Consequently, for each $a_i \in T_{\min}$ there is exactly one $i_{a_i}$ such that $a_i \in Q_i$, and for all $b, b' \in a^*$ with $b \neq b'$ we have $i_b \neq i_{b'}$. Hence, if there are no good $Q_i$'s, then $t_0$ cannot exceed $k_0 = |Eq(n)|$. Let us now show that indeed, $t_0$ must exceed $k_0$: Recall that $n^* = 2^{k_0}$ and that $(n + \ell + 1)^*$ is an integer power of 2, where $\ell + 1 = m_T - n$. As a consequence of Lemma 4.25, for any positive integer $t$ we get:

$$\text{If } n^* = 2^k \text{ and } (n + t)^* = 2^k \text{ then } t \geq 2^k, \text{ in particular } t > k. \quad (*)$$

For every $a_i \in T_{\min}$ with $|a_i^*| = t_0$, and for any $b \in S_n \setminus S_n$, where $\hat{S}^b$ is not necessarily good, we have the following situation:

- $|S^b| = n + t$ where $(n + t)^* = 2^k$ for some $k > k_0$, and
- either $b \in a_i^*$ or $\hat{S}^b$ is not good.

Hence, for some integer $t' \geq 0$ we have

$$m_T = n + \ell + 1 = n + t' + t_0 = |S^a|,$$

where $(n + t')^*$ and $(n + t' + t_0)^*$ are both integer powers of 2. Say $(n + t')^* = 2^k$ and $(n + t' + t_0)^* = 2^{k'}$ where $k' > k \geq k_0$. Then, by $(*)$, $t_0 > k \geq k_0$ which completes the proof of the claim.

Since $f_0(Q_{j_0}) \notin \text{seq}^{-1}(S_n)$ there exists a first element $a_n$ in the sequence $f_0(Q_{j_0})$ which does not belong to $S_n$. Let $s_{n+1} = \hat{s_n}(a_n)$. Then $s_{n+1}$ is a non-repetitive sequence in $A$ of length $n + 1$, which is what we were aiming for.

To some extent, Theorem 4.24 and Theorem 4.26 are optimal, i.e., there are no other relations between $\text{seq}^{-1}(m)$, $\text{seq}(m)$, and $2^m$ which are provable in $ZF$ (see Chapter 7 | RELATED RESULT 49). It might be tempting to prove that for all cardinals $m$, $\text{seq}(m) \neq \text{fin}(m)$, however, such a proof cannot be carried out in $ZF$ (cf. Proposition 7.17).

$2^2 + 2^2 = 2^2$ whenever $m$ is infinite

The fact that $2^m + 2^m = 2^m$ whenever $m$ is infinite will turn out as a consequence of the following result:

**Lemma 4.27 (Läuchli’s Lemma).** If $m$ is an infinite cardinal, then

$$\left(2^{\text{fin}(m)}\right)^m = 2^{\text{fin}(m)}.$$  

**Proof.** Let $A$ be an arbitrary but fixed set of cardinality $m$. Recall that for $n \in \omega$, $[A]^n$ denotes the set of all $n$-element subsets of $A$. For natural numbers
More cardinal relations: $2^m + 2^n = 2^n$

$n, k \in \omega$, where $k \geq n$, we define two mappings $g_{n,k}$ and $d_{n,k}$ from $\mathcal{P}([A]^n)$ into itself as follows: For $X \subseteq [A]^n$ define

$$g_{n,k}(X) = \{y \in [A]^n : \forall z \in [A]^k \,(y \subseteq z \rightarrow \exists x \in X \,(x \subseteq z))\}$$

and let $d_{n,k}(X) := g_{n,k}(X) \setminus X$. To get familiar with the functions $g_{n,k}$ and $d_{n,k}$ respectively, consider the following example: Let $n = 2$, $k = 4$, take $\{a_0, a_1\} \in [A]^2$, and let $X_0 = \{x \in [A]^2 : x \cap \{a_0, a_1\} = \emptyset\}$. Then $g_{2,4}(X_0) = [A]^2$ and $Y := d_{2,4}(X_0) = \{y \in [A]^2 : y \cap \{a_0, a_1\} = \emptyset\}$. Further, $g_{2,4}(Y) = Y$ and $d_{2,4}(Y) = d_{2,4}(d_{2,4}(X_0)) = \emptyset$. We leave it as an exercise to the reader to show that the mapping $g_{n,k}$ has the following properties:

1. For all $X \subseteq [A]^n$, $X \subseteq g_{n,k}(X)$.
2. $g_{n,k} \circ g_{n,k} = g_{n,k}$, i.e., for all $X \subseteq [A]^n$, $g_{n,k}(g_{n,k}(X)) = g_{n,k}(X)$.
3. For all $X \subseteq [A]^n$, $g_{n,k}(X) \subseteq g_{n,k'}(X)$ whenever $k' \geq k$.

By induction on $j$ we define $d_{n,k}^{j+1} := d_{n,k} \circ d_{n,k}^j$, where $d_{n,k}^0$ denotes the identity. Then, we have $d_{n,k}^{j+1} = (g_{n,k} \circ d_{n,k}^j) \setminus d_{n,k}^j$, and therefore by (1) we get

$$d_{n,k}^{j+1} = (g_{n,k} \circ d_{n,k}^j) \setminus d_{n,k}^j.$$  

In order to show that $d_{n,k}^m = g_{n,k} \circ d_{n,k}^m$ we first prove a combinatorial result by applying the Finite Ramsey Theorem 23.

For any fixed integers $n, k \in \omega$ where $k \geq n$, for $U \subseteq A$ with $|U| \leq n$, and for any $X \subseteq [A]^n$, let $\psi(U, X, W)$ and $\varphi(U, X)$ be the following statements:

$$\psi(U, X, W) \equiv W \subseteq A \setminus U \land \forall V \in [W]^{n+|U|} \,(U \cup V \subseteq X)$$

and

$$\varphi(U, X) \equiv \forall m \in \omega \, \exists W \subseteq A \,(|W| \geq m \land \psi(U, X, W)).$$

Notice that if $U \subseteq X \subseteq [A]^n$, then we have $\psi(U, X, W)$ for every $W \subseteq A \setminus U$, and consequently we have $\varphi(U, X)$ for all $U \subseteq A$. To get familiar with the statements $\psi$ and $\varphi$ respectively consider again the example given above: Let $b \in A \setminus \{a_0, a_1\}$ and let $U = \{a_0, b\}$. Then we have $\varphi(U, d_{2,4}(X_0))$, since for any $m \in \omega$ we have $\psi(U, d_{2,4}(X_0), [A \setminus \{a_0, a_1, b\}]^m)$. Further, for $U' = \{b\} \subseteq U$ we have $\varphi(U', X_0)$, since for any positive $m \in \omega$ we have $\psi(U', X_0, [A \setminus \{a_0, a_1, b\}]^m)$.

**Claim 1.** If we have $\varphi(U, d_{n,k}(X))$, then there is a set $U'$ with $|U'| < |U|$ such that we have $\varphi(U', X)$. In particular we get that $\varphi(\emptyset, d_{n,k}(X))$ fails—a fact which can be easily verified directly.

**Proof of Claim 1.** Let us assume that $\varphi(U, d_{n,k}(X))$ holds for $U \subseteq A$ with $|U| \leq n$ and some set $X \subseteq [A]^n$. It is enough to show that for any integer $m \geq k$ there is a proper subset $U'$ of $U$ and a $W \subseteq [A]^m$ such that $\psi(U', X, W)$ holds. Indeed, since there are just finitely many proper subsets of $U$, there
must be a proper subset $U'$ of $U$ such that for arbitrarily large integers $m$
there is a set $W_m \in [A]^m$ such that $\psi(U', X, W_m)$ holds, we get that $\varphi(U', X)$ holds.

Recall that by the Finite Ramsey Theorem 2.3, for all $m, i, j \in \omega$, where $j \geq 1$ and $i \leq m$, there exists a smallest integer $N_{m, i, j} \geq m$ such that for each $j$-colouring of $[N]$ there is an $m$-element subset of $N$; all whose $i$-element subsets have the same colour. Let $m \geq k$, let $m'' = \max\{N_{m, i, 2} : 0 \leq i \leq n\}$, and let $m'' = N_{m', k - r + 2}$ where $r = |U|$. By $\varphi(U, d_{n, k}(X))$ there is a set $S$
with $|S| = m''$ such that $\psi(U, d_{n, k}(X), S)$. To each subset $U'$ of $U$ we assign the set $X(U')$ by stipulating

$$X(U') = \{Y \in [S]^{k-r} : \exists V' \subseteq Y (U' \cup V' \in X)\}.$$  

Now we show that $\bigcup_{U \subseteq U'} X(U') = [S]^{k-r}$. Let $V \in [S]^{k-r}$. By definition of $\psi(U, d_{n, k}(X), S)$, $S \subseteq A \setminus U$, and since $|U| = r$ we have $|U \cup V| = k$. Since $k - r \geq n - r$ there is a set $Q \in [V]^{n-r}$, and since $\psi(U, d_{n, k}(X), S)$ we get $U \cup Q \in d_{n, k}(X)$. Hence, by definition of $d_{n, k}$ and $g_{n, k}$ respectively, there is a set $x \in X$ such that $x \subseteq U \cup V$. If we let $U' = U \cap x$ and $V' = V \cap x$, then

$U' \cup V' \in X$ and consequently $V \in X(U')$.

Because $|S| = m'' = N_{m', k - r + 2}$, there is a set $T \in [S]^{m''}$ and a set $U' \subseteq U$ such that $[T]^{k-r} \subseteq X(U')$. Let $s = |U'|$, let

$$Z = \{V' \in [T]^{n-s} : U' \cup V' \in X\},$$

and let $Z' = [T]^{n-s} \setminus Z$. Since $|T| = m' \geq N_{m, n-s, 2}$, there exists a set $W \in [T]^{m'}$ such that either $[W]^{n-s} \subseteq Z$ or $[W]^{n-s} \subseteq Z'$. The latter case can be excluded. Indeed, since $m \geq k \geq k - r$, $[W]^{k-r} \neq \emptyset$. Now, each element $w$ of $[W]^{k-r}$ is a subset of $T$ and consequently an element of $X(U')$. Thus, there

is a $V' \subseteq w$ such that $U' \cup V' \in X$ which implies that $V' \in Z$, in particular,

$[W]^{n-s} \cap Z \neq \emptyset$. Hence, $[W]^{n-s} \subseteq Z$ and we finally have $\psi(U', X, W)$ where $|W| = m$.

It remains to show that $U' \neq U$: Since we have $\psi(U, d_{n, k}(X), S)$ and $W \subseteq S$, we also have $\psi(U, d_{n, k}(X), W)$. Now, if $U' = U$, then we would also have $\psi(U, X, W)$, but since $d_{n, k}(X) = g_{n, k}(X) \setminus X$, $d_{n, k}(X) \cap X = \emptyset$ which implies that the set $[W]^{n-r}$ is empty which is only the case when $|W| < n - r$; however, $|W| = m \geq k \geq n \geq n - r$.

Now we turn back to the sets $d_{n, k}(X)$ and show that $d_{n, k}^{m+1}(X) = \emptyset$. In fact we show a slightly stronger result:

**Claim 1.** If $d_{n, k}^l(X) \neq \emptyset$ for some set $X \subseteq [A]^n$, then $l \leq n$.

**Proof of Claim 1.** Take any $U \in d_{n, k}^l(X)$. Since $|U| = n$, for each set $W \subseteq A \setminus U$ we have $\psi(U, d_{n, k}(X), W)$, and since $A$ is not finite we have $\psi(U, d_{n, k}(X))$.

By applying **Claim 1** $l$ times we get a sequence $U = U_0, U_1, \ldots, U_l$ such that $|U_{j+1}| > |U_j|$ for all $j \leq l$, which implies that $|U_j| \geq j$ (for all $j$). In particular $|U| = |U_l| \geq l$, and since $|U| = n$ this implies that $l \leq n$. — **Claim 2**
More cardinal relations: $z^n + z^m = z^m$

As a consequence of Claim 2 we get

(5) $d_{n,k}^n = g_{n,k}^n \cdot d_{n,k}^n$.

Define now a mapping $f_{n,k}$ from $\mathcal{P}([A]^n)$ to $\mathcal{P}([A]^k)$ by stipulating

$$f_{n,k}(X) = \{z \in [A]^k : \exists x \in X (x \subseteq z)\}.$$ 

Further, let

$$I_{n,k}(X) = \{X \subseteq [A]^n : g_{n,k}(X) = X\}.$$ 

Then, by (1) and (3) we get

(6) $I_{n,k} \subseteq I_{n,k'}$ whenever $k' \geq k$.

Consider now $f_{n,k} := f_{n,k}|_{I_{n,k}}$. By definition of $g_{n,k}$ and $d_{n,k}$ respectively we have that $f_{n,k}$ is injective. Indeed, if $X, X' \in I_{n,k}$ (i.e., $g_{n,k}(X) = X$ and $g_{n,k}(X') = X'$) and $f_{n,k}(X) = f_{n,k}(X')$, then $X \subseteq g_{n,k}(X) = X$ and $X' \subseteq g_{n,k}(X) = X'$, and therefore $X = X'$. So, for sets in $\text{dom}(f_{n,k})$ we can define the inverse of $f_{n,k}$ by stipulating

$$f^{-1}_{n,k}(f_{n,k}(X)) = X.$$ 

Now we are ready to construct a one-to-one mapping $F$ from $\mathcal{P}(\text{fin}(A))$ into $\mathcal{P}(\text{fin}(A))$: Let $X \in \mathcal{P}(\text{fin}(A))^\omega$, i.e., $X = \{X_s : s \in \omega\}$ where for each $s \in \omega$, $X_s \in \mathcal{P}(\text{fin}(A))$. Define the function $F$ by stipulating

$$F(X) = \bigcup_{s \in \omega} \bigcup_{n \in \omega} \left( \bigcup_{0 \leq j \leq n} f_{n,k(s,n,j)} \circ d_{n,k(s,n,n)} \circ d_{n,k(s,n,n)}(X_s \cap [A]^n) \right)$$

where $k(s,n,j) := 2^s \cdot 3^n \cdot 5^j$. By definition we get that $F$ is a function from $\mathcal{P}(\text{fin}(A))^\omega$ to $\mathcal{P}(\text{fin}(A))$. So, it remains to show that $F$ is injective. To keep the notation short let

$$X_{s,n} = X_s \cap [A]^n,$$

$$X_{s,n,j} = g_{n,k(s,n,n)} \circ d_{n,k(s,n,n)}(X_{s,n}),$$

$$Y_{s,n,j} = f_{n,k(s,n,j)}(X_{s,n,j}).$$

Then

$$F(X) = \bigcup_{s \in \omega} \bigcup_{n \in \omega} \left( \bigcup_{0 \leq j \leq n} Y_{s,n,j} \right).$$

Since $Y_{s,n,j} \in \mathcal{P}([A]^{k(s,n,j)})$ and since the mapping $\langle s, n, j \rangle \mapsto k(s, n, j)$ is injective we get

$$Y_{s,n,j} = F(X) \cap [A]^{k(s,n,j)}.$$ 

By (2) we have $X_{s,n,j} \in I_{n,k(s,n,n)}$. Moreover, since $j \leq n$ we have $k(s,n,j) \leq k(s,n,n)$ and by (6) we get $X_{s,n,j} \in I_{n,k(s,n,j)}$. Thus, $Y_{s,n,j} = f_{n,k(s,n,j)}(X_{s,n,j})$ and therefore
\[ X_{s,n,j} = f_{n,k(s,n,j)}^{-1}(Y_{s,n,j}). \]

By (4) and (5) we get

\[ X_{s,n} = X_{s,n,0} \setminus \left( X_{s,n,1} \setminus \cdots \left( X_{s,n,n-1} \setminus X_{s,n,n} \right) \cdots \right), \]

and since

\[ X_s = \bigcup_{n \in \omega} X_{s,n} \]

we get that \( F \) is injective. This shows that \( (2^{\text{fin}(m)})^{\aleph_0} \leq 2^{\text{fin}(m)} \), and since we obviously have \( 2^{\text{fin}(m)} \leq (2^{\text{fin}(m)})^{\aleph_0} \), by the **CANTOR-BERNSTEIN THEOREM 3.17** we finally get \( (2^{\text{fin}(m)})^{\aleph_0} = 2^{\text{fin}(m)} \).

As a consequence of **LÄUCHLI'S LEMMA 4.27** we get the following equality:

**THEOREM 4.28.** If \( m \) is an infinite cardinal, then \( \aleph_0 \cdot 2^m = 2^m \), in particular we get \( 2^m + 2^m = 2^m \).

**Proof.** Let \( A \) be a set of cardinality \( m \). Further, let \( \inf(A) := \mathcal{P}(A) \setminus \text{fin}(A) \) and let \( \inf(m) := |\inf(A)| \). Then \( 2^m = \text{fin}(m) + \inf(m) \) and consequently

\[ 2^m = 2^{\text{fin}(m)} + \inf(m) = 2^{\text{fin}(m)} \cdot 2^{\text{inf}(m)}. \]

Since by **LÄUCHLI'S LEMMA 4.27**, \( 2^{\text{fin}(m)} = (2^{\text{fin}(m)})^2 \), and by **FACT 4.6**, \( 2^{\text{fin}(m)} \geq \aleph_0 \), we have

\[ 2^{\text{fin}(m)} \cdot 2^{\text{inf}(m)} = (2^{\text{fin}(m)})^2 \cdot 2^{\text{inf}(m)} = 2^{\text{fin}(m)} \cdot 2^m \geq \aleph_0 \cdot 2^m, \]

and since \( 2^m \leq \aleph_0 \cdot 2^m \), by the **CANTOR-BERNSTEIN THEOREM 3.17** we finally get \( \aleph_0 \cdot 2^m = 2^m \).

---

**Notes**

**D-finite and transfinite sets.** In [8, §5], Dedekind defined infinite and finite sets as follows: A set \( S \) is called infinite when it is similar to a proper subset of itself; otherwise, \( S \) is said to be finite. It is not hard to verify that Dedekind’s definition of finite and infinite sets correspond to our definition of D-finite and transfinite sets respectively. In the footnote to his definition Dedekind writes: In this form I communicated the definition of the infinite, which forms the core of my whole investigation, in September, 1882, to G. Cantor, and several years earlier to Schwarz and Weber. More historical background can be found in Fraenkel [12, Ch. I., §2, 5.].

\( \aleph_0 \leq 2^m \rightarrow \aleph_0 \leq 2^m \). The proof of **PROPOSITION 4.4** — which is Theorem 68 of Lindenbaum and Tarski [24] — is taken from Halbeisen [14, VIII] (see also Halbeisen and Shelah [17, Fact 8.1]); and for another proof see for example Sierpiński [34, VIII§2, Ex.9].
\(N_1 \leq^* 2^{\kappa_0}\). The relation symbol "\(\leq^*\)" was introduced by Tarski (cf. Lindenbaum and Tarski [24, p. 301]). The proof of Theorem 4.11 is essentially taken from Sierpinski [34, XV §2], and an alternative proof is given by Sierpinski [29]. Lemma 4.10 is due to Lebesgue [22, p. 233 ff.], and Church [7, Corollary 2, p. 383] showed that the set of all non-repetitive well-ordered sequences of natural numbers is of cardinality \(2^{\kappa_0}\).

If the reals are a countable union of countable sets. Proposition 4.12 is taken from Specker [36, III, §3], where one can find also some other implications like \(\aleph_1 < \kappa_0\), or that every subset of \(\mathbb{R}\) is either finite or transfinite. Corollary 4.13 (i.e., the paradoxical decomposition of \(\mathbb{R}\)) can also be found in Halbeisen and Shelah [18, Fact 8.6].

Cantor’s Normal Form Theorem. The proof of Cantor’s Normal Form Theorem 4.16 is taken from Cantor [4, §9, Satz B] (see also Cantor [6, p. 333 ff.]), but can also be found for example in Fraenkel [12, Ch. III, §11, Thm. 11]. For a slightly more general result see Bachmann [1, III, §12]. The proof of Theorem 4.19 is taken from Halbeisen [14, VII] (cf. Specker [35]).

Other cardinal relations. Theorem 4.20 — as well as the idea of getting a contradiction by constructing an injective class function from \(\Omega\) into a given set — is due to Specker [35, p. 334 ff.] (cf. Related Result 21). Theorem 4.21 and Proposition 4.22 are due to Halbeisen [14, IX] (see also Halbeisen and Shah [17, §2, Theorem 3 and p. 36]). Lemma 4.23 and Theorem 4.24 are due to Halbeisen [14, IX] (see also Halbeisen and Shah [17, §3, Theorem 3]). The proof of Theorem 4.26 is due to Shah (see Halbeisen and Shah [17, §3 Theorem 4]). Lemma 4.25 is due to Halbeisen, who proved that number-theoretic result when Theorem 4.26 was still a conjecture. For a generalisation of Theorem 4.26 see Related Result 20. Läuchli’s Lemma 4.27 as well as Theorem 4.28 is taken from Läuchli [21].

Related Results

13. Other definitions of finiteness. Among the many definitions of finiteness we would like to mention just one by von Neumann who defined in [25, p. 736] finite sets as follows: A set \(E\) is finite, if there is no non-empty set \(K \subseteq \mathcal{P}(E)\) such that for each \(x \in K\) there is a \(y \in K\) with \(|x| < |y|\). With respect to this definition of finiteness, a set \(I\) is infinite iff for each natural number \(n\) there exists an \(n\)-element subset of \(I\), or equivalently, a set \(E\) is finite iff there exists a bijection between \(E\) and a natural number \(n\). However, notice that von Neumann does not use the notion of natural numbers in his definition. In [25, VIII, 2], von Neumann investigated that notion of finiteness and showed for example that power sets of finite sets are finite. For some other definitions of finiteness and their dependencies we refer the reader to Kurepa [20], Lévy [23], Schröder [27], Sipka and Vogt [37], Tarski [38], and Truss [41].

14. The countability of the rationals. We have seen that the set of rational numbers is countable, but since we used the Cantor-Bernstein Theorem 3.17 to construct a bijection between \(\mathbb{Q}\) and \(\omega\), it is quite difficult to determine the image of a given rational number. However, there exists also a “computable” bijection
$f : \mathbb{Q} \to \omega$ due to Faber [10]: The image of a rational number \( q \), written in the form

$$q = \frac{a_1}{2!} + \frac{a_2}{3!} + \cdots + \frac{a_n}{(n+1)!},$$

where the \( a_i \)'s are computed by trigonometric series and for all \( 1 \leq i \leq n \) we have \( 0 \leq a_i < (i + 1)! \), is defined by

$$f(q) = a_1 \cdot 1! + a_2 \cdot 2! + a_3 \cdot 3! + \cdots + a_n \cdot n!.$$

15. **Goodstein sequences.** For positive integers \( m \) and \( n \), where \( n > 1 \), define the **hereditary base \( n \) representation of \( m \) as follows.** First write \( m \) as the sum of powers of \( n \), e.g., if \( m = 265 \) and \( n = 2 \) write \( 265 = 2^8 + 2^7 + 1 \). Then write each exponent as the sum of powers of \( n \) and repeat with exponents of exponents and so on until the representation stabilises, e.g., \( 265 \) stabilises at the representation \( 2^{2^2 + 1} + 2^{2^1 + 1} \). Now define the number \( G_n(m) \) as follows. If \( m = 0 \) let \( G_0(m) := 0 \); otherwise, let \( G_n(m) \) be the number produced by replacing every occurrence of \( n \) in the hereditarily base \( n \) representation of \( m \) by the number \( n + 1 \) and then subtracting 1, e.g., \( G_2(265) = 3^{3^3 + 1} + 3^{1^1 + 1} \). The Goodstein sequence \( m_0, m_1, \ldots \) for \( m \) starting at \( 2 \) is defined as follows: \( m_0 = m, m_1 = G_2(m_0), m_2 = G_3(m_1), m_3 = G_4(m_2), \) and so on; for example we get:

\[
\begin{align*}
265_0 &= 265 \\
265_1 &= 2^{2^2 + 1} + 2^{2^1 + 1} + 1 \\
265_2 &= 4^{4^{1 + 1}} + 4^{4^{1 + 1} - 1} \\
265_3 &= 5^{5^{1 + 1}} + 5^5 \cdot 3 + 5^5 \cdot 3 + 5^2 \cdot 3 + 5 \cdot 3 + 2 \\
265_4 &= 6^{6^{1 + 1}} + 6^6 \cdot 3 + 6^6 \cdot 3 + 6^2 \cdot 3 + 6 \cdot 3 + 1 \\
265_5 &= 7^{7^{1 + 1}} + 7^7 \cdot 3 + 7^7 \cdot 3 + 7^2 \cdot 3 + 7 \cdot 3 \\
265_6 &= 8^{8^{1 + 1}} + 8^8 \cdot 3 + 8^8 \cdot 3 + 8^2 \cdot 3 + 8 \cdot 3 + 2 + 7 \\
265_7 &= \cdots
\end{align*}
\]

Computing a few of the numbers \( 265_n \), one notices that the sequence \( 265_0, 265_1, 265_2, \ldots \) grows extremely fast and one would probably guess that it tends to infinity. Amazingly, Goodstein [13] showed that for every integer \( m \) there is a \( k \in \omega \) such that \( m_k = 0 \). Indeed, if we replace in the hereditarily base \( n \) representation of \( m_{n-2} \) each \( n \) by \( \omega \), we get an ordinal number, say \( \alpha_{n-2}(m) \); in fact we get \( \text{cnf}(\alpha_{n-2}(m)) \), e.g., \( \alpha_3(265) = \omega^{\omega^0 + 1} + \omega^{\omega^0 + 3} + \omega^3 \cdot 3 + \omega^3 \cdot 3 + \omega \cdot 3 + 2 \). We leave it as an exercise to the reader to show that the sequence of ordinal numbers \( \alpha_0(m), \alpha_1(m), \alpha_2(m), \ldots \) is strictly decreasing. In other words, \( \alpha_0(m) \nless \alpha_1(m) \nless \alpha_2(m) \nless \ldots \), thus, by the Axiom of Foundation, the sequence of ordinals must be finite which implies that the Goodstein sequence \( m_0, m_1, \ldots \) is eventually zero. However, Kirby and Paris [19] showed that Goodstein’s result is not provable in Peano Arithmetic (cf also Paris [20]).
16. **Ordinal arithmetic.** As we have seen, one can define various arithmetical operations on ordinals like addition, multiplication and exponentiation, and even subtraction. Moreover, one can also define division (cf. Fraenkel [12, Ch. III, §11, 4.], Bachmann [1, III, §17], or Sierpiński [31]): For any given ordinals $\alpha$ and $\delta$ ($\delta \neq 0$) there is a single pair of ordinals $\beta$, $\rho$ such that

$$\alpha = \delta \cdot \beta + \rho \quad \text{where } \rho < \delta.$$  

For the theory of ordinal arithmetic we refer the reader to Bachmann [1, III.] (cf. also Sierpiński [32, 33]).

17. **Cancellation laws.** Bernstein showed in his dissertation [2] (see [3, §2, Satz 3]) that for any finite cardinal $a \geq 1$ and arbitrary cardinals $m$ and $n$ we have

$$\alpha \cdot m = \alpha \cdot n \rightarrow m = n.$$  

In fact, Bernstein gave a quite involved proof for the case $a = 2$ ([3, §2, Satz 2]) and just outlined the proof for the general case. Later, Sierpiński [28] found a simpler proof for the case $a = 2$ and generalised the result in [30] to $(a \cdot m \leq a \cdot n) \rightarrow (m \leq n)$. Slightly later, Tarski showed in [39] that for any finite cardinal $a \geq 1$ and arbitrary cardinals $m$ and $n$ we have

$$a \cdot m \leq a \cdot n \rightarrow m \leq n.$$  

18. **On the cardinality of power sets of power sets.** As a consequence of Theorem 4.28 we get

$$2^{2^m} \times 2^{2^m} = 2^{2^m}.$$  

However, it is open if also $2^{m} \times 2^{m} = 2^{m}$ is provable in ZF.

19. **The hierarchy of $\aleph$’s.** By induction on $\Omega$ we define

$$\aleph_0 = |\omega|,$$

$$\aleph_{\alpha+1} = \aleph(\aleph_\alpha),$$

$$\aleph_\lambda = \bigcup_{\alpha \in \lambda} \aleph_\alpha \quad \text{for infinite limit ordinals } \lambda.$$  

For an ordinal $\alpha$, let $A$ be a set of cardinality $\aleph_\alpha$ and let $\gamma_0$ be the order type of a well-ordering of $A$. Then, since $|\gamma_0| = \aleph_\alpha$, $\gamma_0$ is an ordinal of cardinality $\aleph_\alpha$, and we define

$$\omega_\alpha = \bigcap \{ \gamma \in \gamma_0 + 1 : |\gamma| = \aleph_\alpha \}.$$  

20. **On the cardinality of the set of non-repetitive sequences.** Let $m$ be an infinite cardinal an let $S$ be a set of cardinality $m$. We defined $\mathcal{P}^m = |\mathcal{P}(S)|$, however, $\mathcal{P}^m$ can also be considered as the cardinality of the set of functions from $S$ to $\{0, 1\}$.

Similarly, for natural numbers $a \geq 2$ let $\mathcal{P}^a$ denote the cardinality of the set of functions from $S$ to $\{0, 1, \ldots , a - 1\}$. By Theorem 4.26 we have $\mathcal{P}^m \neq \text{seq}^{-1}(m)$ and it is natural to ask whether the following statement is provable in ZF:

For all finite cardinals $a$ and all infinite cardinals $m$, $\mathcal{P}^a \neq \text{seq}^{-1}(m)$. (⋆)

Obviously, if we would have a suitable generalisation of Lemma 4.25 at hand, then the proof of Theorem 4.26 would work for all natural numbers $a \geq 2.$
Halbeisen and Hungerbühler investigated in [16] the function $n^*$ and generalised Lemma 4.25 to numbers different from 2, and this generalisation was later used by Halbeisen [15] who showed that (*) holds for a large class of finite cardinals, e.g., for $a \in \{2, 3, 4, 6, 7, 8, 9, 11, 12, 14, 15, \ldots\}$; it is conjectured that (*) holds for all finite cardinals $a \geq 2$.

21. *On the cardinality of the set of ordered pairs.* By Cantor’s Theorem 3.25 we always have $2^n \leq n^*$. Furthermore, one can show that if there is a finite-to-one map from $2^n$ onto $m$, then $m$ is finite (see Forster [11]). Now, having Theorem 4.20 in mind, one could ask whether $2^{m} \leq m^2 \rightarrow m \leq 4$. This question is still open and is asked in Truss [40], where a dualisation of Theorem 4.20 is investigated.

REFERENCES

2. Felix Bernstein, *Untersuchungen aus der Mengenlehre*, Dissertation (1901), University of Göttingen (Germany).
8. Richard Dedekind, *Was sind und was sollen die Zahlen*, Friedrich Vieweg & Sohn, Braunschweig, 1888 (see also [9, pp. 335–390]).
14. Lorenz Halbeisen, *Vergleich zwischen unendlichen Kardinalzahlen in einer Mengenlehre ohne Auswahlaxiom*, Diplomarbeit (1990), University of Zürich (Switzerland).
4. Cardinal Relations in ZF only


The Axiom of Choice

Two terms occasionally used by musicians are “full” consonance and “pleasing” consonance. An interval is said to be “fuller” than another when it has greater power to satisfy the ear. Consonances are the more “pleasing” as they depart from simplicity, which does not delight our senses much.

Gioseppo Zarlino
Le istitutioni harmoniche, 1598

Zermelo’s Axiom of Choice and its Consistency with ZF

In 1904, Zermelo published his first proof that every set can be well-ordered. The proof is based on the so-called Axiom of Choice, denoted AC, which, in Zermelo’s words, states that the product of an infinite totality of sets, each containing at least one element, itself differs from zero (i.e., the empty set). The full theory $ZF + AC$, denoted $ZFC$, is called Set Theory.

In order to state the Axiom of Choice we first define the notion of a choice function: If $\mathcal{F}$ is a family of non-empty sets (i.e., $\emptyset \notin \mathcal{F}$), then a choice function for $\mathcal{F}$ is a function $f : \mathcal{F} \to \bigcup \mathcal{F}$ such that for each $x \in \mathcal{F}$, $f(x) \in x$.

The Axiom of Choice — which completes the axiom system of Set Theory and which is in our counting the ninth axiom of $ZFC$ — states as follows:

9. The Axiom of Choice

$$\forall \mathcal{F} \left( \emptyset \notin \mathcal{F} \to \exists f \left( f \in \mathcal{F} \cup \mathcal{F} \land \forall x \in \mathcal{F} \left( f(x) \in x \right) \right) \right)$$

Informally, every family of non-empty sets has a choice function, or equivalently, every Cartesian product of non-empty sets is non-empty.
Before we give some reformulations of the Axiom of Choice and show some of its consequences, we should address the question whether AC is consistent relative to the other axioms of Set Theory (i.e., relative to ZF), which is indeed the case.

Assume that ZF is consistent, then, by Proposition 3.5, ZF has a model, say V. To obtain the relative consistency of AC with ZF, we have to show that also ZF + AC has a model. In 1938, Gödel informed von Neumann at the Institute for Advanced Study in Princeton that he had found such a model. In fact he showed that there exists a smallest transitive subclass of V which contains all ordinals (i.e., contains \( \Omega \) as a subclass) in which AC as well as ZF holds. This unique submodel of V is called the constructible universe and is denoted by L, where “L” stands for the following “law” by which the constructible universe is built. Roughly speaking, the model L consists of all “mathematically constructible” sets, or in other words, all sets which are “constructible” or “describable”, but nothing else. To be more precise, let us give the following definitions:

Let \( M \) be a set and \( \varphi(x_0, \ldots, x_n) \) be a first-order formula in the language \{\( \in \}\}. Then \( \varphi^M \) denotes the formula we obtain by replacing all occurrences of “\( \exists x \)” and “\( \forall x \)” by “\( \exists x \in M \)” and “\( \forall x \in M \)” respectively. A subset \( y \subseteq M \) is definable over \( M \) if there is a first-order formula \( \varphi(x_0, \ldots, x_n) \) in the language \{\( \in \}\}, and parameters \( a_1, \ldots, a_n \) in \( M \), such that \( \{ z : \varphi^M(z, a_1, \ldots, a_n) \} = y \). Finally, for any set \( M \):

\[
def(M) = \{ y \subseteq M : y \text{ is definable over } M \}\]

Notice that for any set \( M \), \( \text{def}(M) \) is a set being itself a subset of \( \mathcal{P}(M) \). Now, by induction on \( \alpha \in \Omega \), define the following sets (compare with the cumulative hierarchy defined in Chapter 3):

\[
L_\alpha = \emptyset \\
L_\alpha = \bigcup_{\beta \prec \alpha} L_\beta \quad \text{if } \alpha \text{ is a limit ordinal} \\
L_{\alpha+1} = \text{def}(L_\alpha)
\]

and let

\[
L = \bigcup_{\alpha \in \Omega} L_\alpha.
\]

Like for the cumulative hierarchy one can show that for each \( \alpha \in \Omega \), \( L_\alpha \) is a transitive set, \( \alpha \subseteq L_\alpha \) and \( \alpha \in L_{\alpha+1} \), and that \( \alpha \in \beta \) implies \( L_\alpha \subseteq L_\beta \).

Moreover, Gödel showed that \( L \models ZF + AC \), and that \( L \) is the smallest transitive class containing \( \Omega \) as a subclass such that \( L \models ZFC \). Thus, by starting with any model \( V \) of ZF we find a subclass \( L \) of \( V \) such that \( L \models ZFC \). In other words we get that if ZF is consistent then so is ZFC (roughly speaking, if ZFC is inconsistent, then AC cannot be blamed for it).
Equivalent Forms of the Axiom of Choice

There are dozens of hypotheses which are equivalent to the Axiom of Choice, but among the best known and most popular ones are surely the Well-Ordering Principle, the Kuratowski-Zorn Lemma, Krein’s Principle, and Tichý’s Principle—sometimes called Tukey’s Lemma. Since the first three deal with orderings, we have to introduce first the corresponding definitions before we can state these—and some other—so-called choice principles.

A binary relation “≤” on a set $P$ is a partial ordering of $P$ if it is transitive (i.e., $p \leq q$ and $q \leq r$ implies $p \leq r$), reflexive (i.e., $p \leq p$ for every $p \in P$), and anti-symmetric (i.e., $p \leq q$ and $q \leq p$ implies $p = q$). If “≤” is a partial ordering on $P$, then $(P, \leq)$ is called a partially ordered set.

If $(P, \leq)$ is a partially ordered set, then we define

$$p < q \iff p \leq q \land p \neq q,$$

and call $(P, <)$ a partially ordered in the strict sense, (replacing reflexivity by $p \neq p$ for every $p \in P$).

Two distinct elements $p, q \in P$, where $(P, <)$ is a partially ordered set, are said to be comparable if either $p < q$ or $q < p$; otherwise, they are called incomparable. Notice that for $p, q \in P$ we could have $p \leq q$ as well as $p \nleq q$. However, if for any elements $p$ and $q$ of a partially ordered set $(P, <)$ we have $p < q$ or $p = q$ or $p > q$ (where these three cases are mutually exclusive), then $P$ is said to be linearly ordered by the linear ordering “<”. Two elements $p_1$ and $p_2$ of $P$ are called compatible if there exists a $q \in P$ such that $p_1 \leq q \geq p_2$; otherwise, they are called incompatible, denoted $p_1 \perp p_2$.

We would like to mention that in the context of forcing, elements of partially ordered sets are called conditions. Furthermore, it is worth mentioning that the definition of “compatible” given above incorporates a convention, namely the so-called Jerusalem convention for forcing—with respect to the American convention of forcing, $p_1$ and $p_2$ are compatible if there exists a $q$ such that $p_1 \leq q \leq p_2$.

Let $(P, <)$ be a partially ordered set. Then $p \in P$ is called maximal (or more precisely $<$-maximal) in $P$ if there is no $x \in P$ such that $p < x$. Similarly, $q \in P$ is called minimal (or more precisely $<$-minimal) in $P$ if there is no $x \in P$ such that $x < q$. Furthermore, for a non-empty subset $C \subseteq P$, an element $p' \in P$ is said to be an upper bound of $C$ if for all $x \in C$, $x \leq p'$. A non-empty set $C \subseteq P$, where $(P, <)$ is a partially ordered set, is a chain in $P$ if $C$ is linearly ordered by “≤” (i.e., for any distinct members $p, q \in C$ we have either $p < q$ or $p > q$). Conversely, if $A \subseteq P$ is such that any two distinct elements of $A$ are incomparable (i.e., neither $p < q$ nor $p > q$), then in Order Theory, $A$ is called an anti-chain. However, in the context of forcing we say that a subset $A \subseteq P$ is an anti-chain in $P$ if any two distinct elements of $A$ are incompatible. Furthermore, $A \subseteq P$ is a maximal anti-chain in $P$ if $A$ is an anti-chain in $P$ and $A$ is maximal with this property. Notice that if $A \subseteq P$
is a maximal anti-chain, then for every $p \in P \setminus A$ there is a $q \in A$ such $p$ and $q$ are compatible.

Recall that a binary relation $R$ on a set $P$ is a well-ordering on $P$, if there is an ordinal $\alpha \in \Omega$ and a bijection $f : P \to \alpha$ such that $R(x, y)$ iff $f(x) < f(y)$. This leads to the following equivalent definition of a well-ordering, where the equivalence follows from the proof of Theorem 5.1 (the details are left to the reader): Let $(P, <)$ be a linearly ordered set. Then “<” is a well-ordering on $P$ if every non-empty subset of $P$ has a $<$-minimal element. Furthermore, a set $P$ is said to be well-orderable (or equivalently, $P$ can be well-ordered) if there exists a well-ordering on $P$.

In general, it is not possible to define a well-ordering by a first-order formula on a given set (e.g., on $\mathbb{R}$). However, the existence of well-ordering is guaranteed by the following principle:

Well-Ordering Principle: Every set can be well-ordered.

To some extent, the Well-Ordering Principle (like the Axiom of Choice) postulates the existence of certain sets whose existence in general (i.e., without any further assumptions like $\mathbf{V} = \mathbf{L}$), cannot be proved within $\mathbf{ZF}$.

In particular, the Well-Ordering Principle postulates the existence of well-orderings of $\mathbb{Q}$ and of $\mathbb{R}$. Obviously, both sets are linearly ordered by “$<$”. However, since for any elements $x$ and $y$ with $x < y$ there exists a $z$ such that $x < z < y$, the ordering “$<$” is far away from being a well-ordering — consider for example the set of all positive elements. Even though $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ have similar properties (at least from an order-theoretical point of view), when we try to well-order these sets they behave very differently. Firstly, by Fact 4.1 we know that $\mathbb{Q}$ is countable and the bijection $f : \mathbb{Q} \to \omega$ allows us to define a well-ordering “$<$” on $\mathbb{Q}$ by stipulating $q < p \iff f(q) < f(p)$. Now, let us consider the set $\mathbb{R}$. For example we could first well-order the rational numbers, or even the algebraic numbers, and then try to extend this well-ordering to all real numbers. However, this attempt — as well as all other attempts — to construct explicitly a well-ordering of the reals will end in failure (the reader is invited to verify this claim by writing down explicitly some orderings of $\mathbb{R}$).

As mentioned above, Zermelo proved in 1904 that the Axiom of Choice implies the Well-Ordering Principle. In the proof of this result presented here we shall use the ideas of Zermelo’s original proof.

**Theorem 5.1.** The Well-Ordering Principle is equivalent to the Axiom of Choice.

**Proof.** ($\Leftarrow$) Let $M$ be a set. If $M = \emptyset$, then $M$ is already well-ordered. So, assume that $M \neq \emptyset$ and let $\mathcal{P}^*(M) := \mathcal{P}(M) \setminus \{\emptyset\}$. Further, let $f : \mathcal{P}^*(M) \to M$ be an arbitrary but fixed choice function for $\mathcal{P}^*(M)$ (which exists by AC).

A one-to-one function $w_\alpha : \alpha \mapsto M$, where $\alpha \in \Omega$, is an $f$-set if for all $\gamma \in \alpha$:
Equivalent forms of the Axiom of Choice

$$w_{\alpha}(\gamma) = f(M \setminus \{w_{\alpha}(\delta) : \delta \in \gamma\})$$

For example $w_1(0) = f(M)$ is an $f$-set, in fact, $w_1$ is the unique $f$-set with domain $\{0\}$. Further, by Hartogs’ Theorem 3.27, the collection of all $f$-sets is a set, say $S$. Define the ordering “$<$” on $S$ as follows: For two distinct $f$-sets $w_{\alpha}$ and $w_{\beta}$ let $w_{\alpha} \prec w_{\beta}$ if $\alpha \neq \beta$ and $w_{\beta}\mid_{\alpha} = w_{\alpha}$. Notice that $w_{\alpha} \prec w_{\beta}$ implies $\alpha \in \beta$.

**Claim.** The set $S$ of all $f$-sets is well-ordered by “$<$”.

**Proof of Claim.** Let $w_{\alpha}$ and $w_{\beta}$ be any two $f$-sets and let

$$\Gamma = \{\gamma \in (\alpha \cap \beta) : w_{\alpha}(\gamma) \neq w_{\beta}(\gamma)\}.$$ 

If $\Gamma \neq \emptyset$, then, for $\gamma_0 = \bigcap \Gamma$, we have $w_{\alpha}(\gamma_0) \neq w_{\beta}(\gamma_0)$. On the other hand, for all $\delta \in \gamma_0$ we have $w_{\alpha}(\delta) = w_{\beta}(\delta)$, thus, by the definition of $f$-sets, we get $w_{\alpha}(\gamma_0) = w_{\beta}(\gamma_0)$. Hence, $\Gamma = \emptyset$, and consequently we are in exactly one of the following three cases:

- $w_{\alpha} \prec w_{\beta}$ iff $\alpha \in \beta$.
- $w_{\alpha} = w_{\beta}$ iff $\alpha = \beta$.
- $w_{\beta} \prec w_{\alpha}$ iff $\beta \in \alpha$.

Thus, the ordering “$<$” on $S$ corresponds to the ordering of the ordinals by “$\in$”, and since the latter relation is a well-ordering on $\Omega$, the ordering “$<$” is a well-ordering, too.

Now, let $w := \bigcup S$ and let $M' := \{x \in M : \exists \gamma \in \text{dom}(w) \ (w(\gamma) = x)\}$. Then $w \in S$ and $M' = M$; otherwise, $w$ can be extended to the $f$-set $w \cup \{(\text{dom}(w), f(M \setminus M'))\}$.

Thus, the one-to-one function $w : \text{dom}(w) \to M$ is onto, or in other words, $M$ is well-orderable.

$(\Rightarrow)$ Let $\mathcal{F}$ be any family of non-empty sets and let “$<$” be any well-ordering on $\bigcup \mathcal{F}$. Define $f : \mathcal{F} \to \bigcup \mathcal{F}$ by stipulating $f(x)$ being the $<$-minimal element of $x$.

It turns out that in many cases, the Well-Ordering Principle — mostly in combination with transfinite induction — is easier to apply than the Axiom of Choice. For example in order to prove that every vector space has an algebraic basis, we would first well-order the set of vectors and then build a basis by transfinite induction (i.e., for every vector $v_{\alpha}$ we check whether it is in the linear span of the vectors $\{v_{\beta} : \beta \in \alpha\}$, and if it is not, we mark it as a vector of the basis). However, similarly to the well-ordering of $\mathbb{R}$, in many cases it is not possible to write down explicitly an algebraic basis of a vector space. For example consider the real vector space of all countably infinite sequences of real numbers, or any infinite dimensional Banach space.

The following three principles, which will be shown to be equivalent to the Axiom of Choice, are quite popular in Algebra and Topology. Even though these principles look rather different, all state that certain sets have maximal
elements or subsets (with respect to some partial ordering), and so they are usually called maximality principles. Let us first state the Kuratowski-Zorn Lemma and Kuripa’s Principle.

Kuratowski-Zorn Lemma: If \((P, \leq)\) is a non-empty partially ordered set such that every chain in \(P\) has an upper bound, then \(P\) has a maximal element.

Kuripa’s Principle: Each partially ordered set has a maximal subset of pairwise incomparable elements.

In order to state Teichmüller’s Principle we have to introduce one more notion: A family \(\mathcal{F}\) of sets is said to have finite character if for each set \(x, x \in \mathcal{F}\) iff \(\text{fin}(x) \subseteq \mathcal{F}\) (i.e., every finite subset of \(x\) belongs to \(\mathcal{F}\)).

Teichmüller’s Principle: Let \(\mathcal{F}\) be a non-empty family of sets. If \(\mathcal{F}\) has finite character, then \(\mathcal{F}\) has a maximal element (maximal with respect to inclusion “\(\subseteq\)”).

Below we shall see that the three maximality principles are all equivalent to the Axiom of Choice. However, in order to prove directly that the Axiom of Choice implies the Kuratowski-Zorn Lemma (i.e., without using the Well-Ordering Principle), we have to show first the following interesting lemma — which is the main reason why we do not want to derive the Kuratowski-Zorn Lemma from the Well-Ordering Principle, even though this would be much easier.

**Lemma 5.2.** Let \((P, \leq)\) be a non-empty partially ordered set. If there is a function \(b : \mathcal{P}(P) \to P\) which assigns to every chain \(C\) an upper bound \(b(C)\), and if \(f : P \to P\) is a function such that for all \(x \in P\) we have \(x \leq f(x)\), then there is a \(p_0 \in P\) such that \(p_0 = f(p_0)\).

**Proof.** Notice that because every well-ordered set is a chain, it is enough to require the existence of an upper bound \(b(W)\) just for every set \(W \subseteq P\) which is well-ordered by “\(<\)”. If \(W \subseteq P\) is a well-ordered subset of \(P\) and \(x \in W\), then \(W_x := \{y \in W : y < x\}\). A well-ordered set \(W \subseteq P\) is called an \(f\)-chain, if for all \(x \in W\) we have \(x = f(b(W_x))\). Notice that since \(\emptyset \subseteq P\) is well-ordered by “\(<\)”, the set \(\{f(b(\emptyset))\}\) is an \(f\)-chain.

We leave it as an exercise to the reader to verify that the set of \(f\)-chains is well-ordered by proper inclusion “\(\subseteq\)”. Hence, the set

\[
U = \bigcup \{W \subseteq P : W \text{ is an } f\text{-chain}\}
\]

is itself an \(f\)-chain. Consider \(p_0 := f(b(U))\) and notice that \(U \cup \{p_0\}\) is an \(f\)-chain. By the definition of \(U\) we get that \(p_0 \in U\), and consequently we have \(f(b(U_{p_0})) = p_0\). Now, since \(f(b(U_{p_0})) \geq b(U_{p_0}) \geq p_0\), we must have \(b(U_{p_0}) = p_0\), and therefore \(f(p_0) = p_0\).
Notice that the proof of Lemma 5.2 does not rely on any choice principles.

Now we are ready to prove that the Kuratowski-Zorn Lemma and Teichmüller’s Principle are both equivalent to the Axiom of Choice.

**Theorem 5.3.** The following statements are equivalent:

(a) Axiom of Choice.

(b) Kuratowski-Zorn Lemma.

(c) Teichmüller’s Principle.

**Proof.** (a)⇒(b) Let \((P, \leq)\) be a non-empty partially ordered set such that every chain in \(P\), in particular every well-ordered chain, has an upper bound. Then, for every non-empty well-ordered subset \(W \subseteq P\), the set of upper bounds \(B_W := \{ p \in P : \forall x \in W(x \leq p) \}\) is non-empty. Thus, the family

\[ F = \{ B_W : W \text{ is a well-ordered, non-empty subset of } P \} \]

is a family of non-empty sets and therefore, by the Axiom of Choice, for each \(W \in F\) we can pick an element \(b(W) \in B_W\). Now, for every \(x \in P\) let

\[ M_x = \begin{cases} \{x\} & \text{if } x \text{ is maximal in } P, \\ \{y \in P : y > x\} & \text{otherwise}. \end{cases} \]

Then \(\{M_x : x \in P\}\) is a family of non-empty sets and again by the Axiom of Choice, there is a function \(f : P \to P\) such that

\[ f(x) = \begin{cases} x & \text{if } x \text{ is maximal in } P, \\ y & \text{where } y > x. \end{cases} \]

Since \(f(x) \geq x\) (for all \(x \in P\)) and every non-empty well-ordered subset \(W \subseteq P\) has an upper bound \(b(W)\), we can apply Lemma 5.2 and get an element \(p_0 \in P\) such that \(f(p_0) = p_0\), hence, \(P\) has a maximal element.

(b)⇒(c) Let \(F\) be a non-empty family of sets and assume that \(F\) has finite character. Obviously, \(F\) is partially ordered by inclusion “\(\subseteq\)”. For every chain \(\mathcal{C}\) in \(F\) let \(U_{\mathcal{C}} = \bigcup \mathcal{C}\). Then every finite subset of \(U_{\mathcal{C}}\) belongs to \(F\). Thus, \(U_{\mathcal{C}}\) belongs to \(F\). On the other hand, \(U_{\mathcal{C}}\) is obviously an upper bound of \(\mathcal{C}\). Hence, every chain has an upper bound and we may apply the Kuratowski-Zorn Lemma and get a maximal element of the family \(F\).

(c)⇒(a) Given a family \(F\) of non-empty sets. We have to find a choice function for \(F\). Consider the family

\[ \mathcal{E} = \{ f : f \text{ is a choice function for some subfamily } F' \subseteq F \}. \]

Notice that \(f\) is a choice function if and only if every finite subfunction of \(f\) is a choice function. Hence, \(\mathcal{E}\) has finite character. Thus, by Teichmüller’s Principle, the family \(\mathcal{E}\) has a maximal element, say \(f_0\). Since \(f_0\) is maximal, \(\text{dom}(f_0) = F\), and therefore \(f_0\) is a choice function for \(F\). \(\Box\)
In order to prove that also Kurepa's Principle is equivalent to the Axiom of Choice, we have to change the setting a little bit: In the proof of Theorem 5.3, as well as in Zermelo's proof of Theorem 5.1, the Axiom of Foundation was not involved (in fact, the proofs can be carried out in Cantor's Set Theory). However, without the aid of the Axiom of Foundation it is not possible to prove that Kurepa's Principle implies the Axiom of Choice, whereas the converse implication is evident (compare the following theorem with Chapter 7 | Related Result 46).

**Theorem 5.4.** The following statements are equivalent in ZF:

(a) Axiom of Choice.

(b) Every vector space has an algebraic basis.

(c) Multiple Choice: For every family $\mathcal{F}$ of non-empty sets, there exists a function $f : \mathcal{F} \to 2^{\bigcup \mathcal{F}}$ such that for each $X \in \mathcal{F}$, $f(X)$ is a non-empty finite subset of $X$.

(d) Kurepa's Principle.

**Proof.** (a)⇒(b) Let $V$ be a vector space and let $\mathcal{F}$ be the family of all sets of linearly independent vectors of $V$. Obviously, $\mathcal{F}$ has finite character. So, by Teichmüller's Principle, which is, as we have seen in Theorem 5.3 equivalent to the Axiom of Choice, $\mathcal{F}$ has a maximal element. In other words, there is a maximal set of linearly independent vectors, which must be of course a basis of $V$.

(b)⇒(c) Let $\mathcal{F} = \{X_i : i \in I\}$ be a family of non-empty sets. We have to construct a function $f : \mathcal{F} \to 2^{\bigcup \mathcal{F}}$ such that for each $X_i \in \mathcal{F}$, $f(X_i)$ is a non-empty finite subset of $X_i$. Without loss of generality we may assume that the members of $\mathcal{F}$ are pairwise disjoint (if necessary, consider the family $\{X_i \times \{i\} : i \in I\}$ instead of $\mathcal{F}$). Adjoin all the elements of $X := \bigcup \mathcal{F}$ as indeterminates to some arbitrary but fixed field $F$ (e.g., $F = \mathbb{Q}$) and consider the field $F(X)$ consisting of all rational functions of the "variables" in $X$ with coefficients in $F$. For each $i \in I$, we define the $i$-degree of a monomial — i.e., a term of the form $ax_1^{k_1} \cdots x_i^{k_i}$ where $a \in F$ and $x_1, \ldots, x_i \in X$ — to be the sum of the exponents of members of $X_i$ in that monomial. A rational function $q \in F(X)$ is called $i$-homogeneous of degree $d$ if it is the quotient of two polynomials such that all monomials in the denominator have the same $i$-degree $n$, while all those in the numerator have $i$-degree $n + d$. The rational functions that are $i$-homogeneous of degree 0 for all $i \in I$ form a subfield $F_0$ of $F(X)$. Thus, $F(X)$ is a vector space over $F_0$, and we let $V$ be the subspace spanned by the set $X$.

By assumption, the $F_0$-vector space $V$ has an algebraic basis, say $B$. Below we use this basis $B$ to explicitly define the desired function $f : \mathcal{F} \to 2^{\bigcup \mathcal{F}}$. For each $i \in I$ and each $x \in X_i$, we can express $x$ as a finite linear combination of elements of $B$. Thus, every $x \in X_i$, can be written in the form
where \( B(x) \in \text{fin}(B) \) and for all \( b \in B(x) \), \( a^x_b \in F_0 \setminus \{0\} \). If \( y \) is another element of the same \( X \), as \( x \), then we have on the one hand

\[
y = \sum_{b' \in B(y)} a^{y}_{b'},
\]

and on the other hand, after multiplying the above representation of \( x \) by the element \( \frac{y}{x} \in F_0 \), we get

\[
y = \sum_{b \in B(x)} \left( \frac{y}{x} \cdot a^x_b \right) \cdot b.
\]

Comparing these two expressions for \( y \) and using the fact that \( B \) is a basis, i.e., that the representation of \( y \) is unique, we must have

\[
B(x) = B(y) \quad \text{and} \quad a^y_{b'} = \frac{y}{x} \cdot a^x_b \quad \text{for all } b \in B(x).
\]

Hence, the finite subset \( B(x) \) of \( B \) as well as the elements \( \frac{a^x_b}{y} \) of \( F(X) \) depend only on \( x \), not on the particular \( x \in X \), and we therefore call them \( B \) and \( a^x_b \) respectively. Notice that, since \( a^x_b \in F_0 \), \( a^x_b \) is \( \iota \)-homogeneous of degree \(-1\) (and \( \iota' \)-homogeneous of degree 0 for \( \iota' \neq \iota \)). So, when \( a^x_b \) is written as a quotient of polynomials in reduced form, some variables from \( X \) must occur in the denominator. Define \( f(X \iota) \) to be the set of all those members of \( X \iota \) that occur in the denominator of \( a^x_b \) (in reduced form) for some \( b \in B \). Then \( f(X \iota) \) is a non-empty finite subset of \( X \iota \), as required.

(c) \( \Rightarrow \) (d) Let \((P, <)\) be a partially ordered set. By Multiple Choice, there is a function \( f \) such that for each non-empty set \( X \subseteq P \), \( f(X) \) is a non-empty finite subset of \( X \). Let \( g : \mathcal{P}(P) \to \text{fin}(P) \) be such that \( g(\emptyset) := \emptyset \) and for each non-empty \( X \subseteq P \), \( g(X) := \{ y \in f(X) : y \text{ is } <\text{-minimal in } f(X) \} \). Obviously, for every non-empty \( X \subseteq P \), \( g(X) \) is a non-empty finite set of pairwise incomparable elements. Using the function \( g \) we construct by transfinite induction a maximal subset of pairwise incomparable elements: Let \( \mathcal{A}_\alpha := g(P) \), and for \( \alpha \in \Omega \) let \( \mathcal{A}_\alpha := g(X_\alpha) \), where

\[
X_\alpha := \{ x \in P : x \text{ is incomparable with each } a \in \bigcup \{ \mathcal{A}_\beta : \beta \in \alpha \} \}.
\]

By construction, the \( \mathcal{A}_\alpha \)'s are pairwise disjoint and any union of \( \mathcal{A}_\alpha \)'s is a set of pairwise incomparable elements. Again by construction there must be an \( \alpha_0 \in \Omega \) such that \( X_{\alpha_0} = \emptyset \). Thus, \( \bigcup \{ \mathcal{A}_\beta : \beta \in \alpha_0 \} \subseteq P \) is a maximal set of pairwise incomparable elements.

(d) \( \Rightarrow \) (a) By the Axiom of Foundation, for every set \( x \) there exists an ordinal \( \alpha \in \Omega \) such that \( x \subseteq V_\alpha \). Thus, since the Axiom of Choice is equivalent to the Well-Ordering Principle (see Theorem 5.1), it is enough to show that Kurepa’s Principle implies that for every \( \alpha \in \Omega \), \( V_\alpha \) can be well-ordered. The crucial
point in that proof is to show that power sets of well-orderable sets are well-orderable.

The first step is quite straightforward: Let $Q$ be a well-orderable set and let $\prec_Q$ be a well-ordering on $Q$. We define a linear ordering $\prec$ on $\mathcal{P}(Q)$ by stipulating $x \prec y$ iff the $\prec_Q$-minimal element of the symmetric difference $x \triangle y$ belongs to $x$. To see that $\prec$ is a linear ordering, notice that $\prec$ is just the lexicographic ordering on $\mathcal{P}(Q)$ induced by $\prec_Q$. The following claim is where Kuratowski’s Principle comes in.

**Claim.** Kuratowski’s Principle implies that every linearly orderable set is well-orderable.

**Proof of Claim.** Let $(P, \prec)$ be a linearly ordered set. Consider the set $W$ of all pairs $(X, x)$ where $X \subseteq P$ and $x \in X$. On $W$ we define a partial ordering $\prec$ by stipulating

$$(X, x) \prec (Y, y) \iff X = Y \land x \prec y.$$ 

By Kuratowski’s Principle, $(W, \prec)$ has a maximal set of pairwise incomparable elements, say $\mathcal{A} \subseteq W$. For every non-empty set $X \subseteq P$ let $f(X)$ be the unique element of $X$ such that $(X, f(X)) \in \mathcal{A}$. It is not hard to verify that $f$ is a choice function for $\mathcal{P}(P) \setminus \{\emptyset\}$, and consequently, $P$ can be well-ordered. \(\square\)

Now we are ready to show that Kuratowski’s Principle implies that every set $V_\alpha$ ($\alpha \in \Omega$) can be well-ordered. We consider the following two cases:

1. **$\alpha$ successor ordinal:** Let $\alpha = \beta_0 + 1$ and assume that $V_{\beta_0}$ is well-orderable. Then $V_\alpha = \mathcal{P}(V_{\beta_0})$, and as the power set of a well-orderable set, $V_\alpha$ is well-orderable.

2. **$\alpha$ limit ordinal:** Assume that for each $\beta \in \alpha$, $V_\beta$ is well-orderable, i.e., for each $\beta \in \alpha$ there exists a well-ordering $\prec_\beta$ on $V_\beta$. Let $\xi$ be the least ordinal such that there is no injection from $\xi$ into $V_\alpha$. The ordinal $\xi$ exists by Hartogs’ Theorem 3.27 and since every $V_\beta$ can be well-ordered. Since $\xi$ is well-ordered by $\in$, $\mathcal{P}(\xi)$ can be well-ordered; let us fix a well-ordering $\prec_\xi \subseteq (\mathcal{P}(\xi) \times \mathcal{P}(\xi))$. For every $\beta \in \alpha$ we choose a well-ordering $\prec_\beta$ on $V_\beta$ as follows:

   - If $\beta = 0$, then $\prec_0 = \emptyset$.
   - If $\beta = \bigcup_{\delta < \beta} \delta$ is a limit ordinal, then, for $x, y \in V_\beta$, let
     
     $$x \prec_\beta y \iff \rho(x) \in \rho(y) \lor (\rho(x) = \rho(y) \land x \prec_{\rho(x)} y),$$

     where for any $z$, $\rho(z) := \bigcap \{\gamma \in \Omega : z \in V_\gamma\}$.
   - If $\beta = \delta + 1$ is a successor ordinal, then, by the choice of $\xi$, there is an injection $f : V_\delta \hookrightarrow \xi$. Let $x = \operatorname{ran}(f)$; then $x \subseteq \xi$. Further, there exists a bijection between $\mathcal{P}(V_\delta) = V_\beta$ and $\mathcal{P}(x)$, and since $\mathcal{P}(x) \subseteq \mathcal{P}(\xi)$ and $\mathcal{P}(\xi)$ is well-ordered by $\prec_\xi$, the restriction of $\prec_\xi$ to $\mathcal{P}(x)$ induces a well-ordering on $V_\beta$.

Thus, for every $\beta \in \alpha$ we have a well-ordering $\prec_\beta$ on $V_\beta$. Now, for $x, y \in V_\alpha$ define
Equivalent forms of the Axiom of Choice

\[ x <_\alpha y \iff \rho(x) \in \rho(y) \lor (\rho(x) = \rho(y) \land x <_{\rho(x)} y). \]

Then, by construction, \( "<_\alpha" \) is a well-ordering on \( V_\alpha \).

We conclude this section on equivalent forms of AC by giving three cardinal relations which are equivalent to the Well-Ordering Principle.

**Theorem 5.5.** Each of the following statements is equivalent to the Well-Ordering Principle, and consequently to the Axiom of Choice:

(a) Every cardinal \( m \) is an aleph, i.e., contains a well-orderable set.

(b) Trichotomy of Cardinals: If \( n \) and \( m \) are any cardinals, then \( n < m \) or \( n = m \) or \( n > m \), where these three cases are mutually exclusive.

(c) If \( n \) and \( m \) are any cardinals, then \( n \leq \alpha m \) or \( m \leq \alpha n \).

(d) If \( m \) is any infinite cardinal, then \( m^2 = m \).

**Proof.** (a) If every set is well-orderable, then obviously every cardinal contains a well-orderable set and is therefore an aleph. On the other hand, for an arbitrary set \( x \) let \( m = |x| \) and let \( y_0 \in m \) be well-orderable. By definition of \( m \) there exists a bijection between \( y_0 \) and \( x \), and therefore, \( x \) is well-orderable as well.

(b) Firstly notice that any two alephs are comparable. Thus, by (a) we get that the Well-Ordering Principle implies the Trichotomy of Cardinals and consequently so does AC. On the other hand, by **Hartogs’ Theorem** 3.27 we know that for every cardinal \( m \) there is a smallest aleph, denoted \( \aleph(m) \), such that \( \aleph(m) \leq^* m \). Now, if any two cardinals are comparable we must have \( m < \aleph(m) \), which implies that \( m \) is an aleph.

(c) Notice that if every set can be well-ordered, then for any cardinals \( n \) and \( m \) we have \( n \leq^* m \) iff \( n \leq m \). For the other direction we first prove that for any cardinal \( m \) there exists an aleph \( \aleph'(m) \) such that \( \aleph'(m) \leq^* m \): Notice that if there exists a surjection from a set \( A \) onto a set \( B \), then there exist an injection from \( B \) into \( \mathcal{P}(A) \). So, by definition of \( \aleph(2^m) \) we have \( \aleph(2^m) \leq^* m \). Let now \( m \) be an arbitrary cardinal and let \( n = \aleph(2^m) \). If \( n \leq^* m \) or \( n \geq^* m \), then we must have \( n \geq^* m \) (since \( n \leq^* m \)), which implies that \( m \) is an aleph and completes the proof.

(d) Assume that for any infinite cardinal \( n \) we have \( n^2 = n \). Hence, we get

\[ m + \aleph(m) = (m + \aleph(m))^2 = m^2 + (m + m) \cdot \aleph(m) + \aleph(m)^2 = m + \aleph(m) + m \cdot \aleph(m), \]

and since \( m + \aleph(m) \leq m \cdot \aleph(m) \) we have

\[ m + \aleph(m) = m \cdot \aleph(m). \]

Now, let \( x \in m \) and let \( y_0 \in \aleph(m) \) be a set which is well-ordered by \( "<_{y_0}" \).

Without loss of generality we may assume that \( x \) and \( y_0 \) are disjoint. Since \( |x \cup y_0| = |x \times y_0| \), there exists a bijection \( f : x \cup y_0 \to x \times y_0 \). Using the bijection \( f \) we define \( \tilde{x} := \{ a \in x : \exists b \in y_0 ((a, b) \in f[y_0]) \} \subseteq x \). Firstly notice that \( \tilde{x} = x \). Indeed, if there would be an \( a_0 \in x \setminus \tilde{x} \), then for all \( b \in y_0 \)
we have \( f^{-1}(\langle a_0, b \rangle) \notin y_0 \), i.e., \( f^{-1}(\langle a_0, b \rangle) \in x \). Thus, since \( f \) is bijective, 
\[ f^{-1}\{\langle a_0 \rangle \times y_0 \} \subseteq x \] is a set of cardinality \( \aleph(m) \), contradicting the fact that
\( \aleph(m) \notin m \). So, for every \( a \in x \), the set
\[ u_a := \{ b \in y_0 : \exists b' \in y_0 (f(b) = \langle a, b' \rangle) \} \]
is a non-empty subset of \( y_0 \), and — since \( y_0 \) is well-ordered by “\(<_{y_0} \)" — has a 
\(<_{y_0}\)-minimal element, say \( \mu_a \). Finally, define an ordering “\(< \)" on \( x \) by stipulating \( a < a' \) iff \( \mu_a <_{y_0} \mu_{a'} \). It is easily checked that “\(< \)" is a well-ordering on x, and therefore, \( m \) is an aleph.

The converse implication — namely that the Well-Ordering Principle implies that \( m^2 = m \) for every infinite cardinal \( m \) — is postponed to the next section (see Theorem 5.7).

\[\text{Cardinal Arithmetic in the Presence of AC}\]

In the presence of AC we are able to define cardinal numbers as ordinals: For any set \( A \) we define
\[ |A| = \bigcap \{ \alpha \in \Omega : \text{there is a bijection between } \alpha \text{ and } A \} . \]

Recall that AC implies that every set \( A \) is well-orderable and that every well-ordering of \( A \) corresponds to exactly one ordinal (which is the order type of the well-ordering).

For example we have \( |n| = n \) for every \( n \in \omega \), and \( |\omega| = \omega \). However, for \( \alpha \in \Omega \) we have in general \( |\alpha| \neq \alpha \), e.g., \( |\omega + 1| = \omega \).

Ordinal numbers \( \kappa \in \Omega \) such that \( |\kappa| = \kappa \) are called cardinal numbers, or just cardinals, and are usually denoted by Greek letters like \( \kappa, \lambda, \mu, \) et cetera.

A cardinal \( \kappa \) is infinite if \( \kappa \notin \omega \), otherwise, it is finite. In other words, a cardinal is finite if and only if it is a natural number.

Since cardinal numbers are just a special kind of ordinals, they are well-ordered by “\(" \in \)". However, for cardinal numbers \( \kappa \) and \( \lambda \) we usually write \( \kappa < \lambda \) instead of \( \kappa \in \lambda \), thus,
\[ \kappa < \lambda \iff \kappa \in \lambda . \]

Let \( \kappa \) be a cardinal. The smallest cardinal number which is greater than \( \kappa \) is denoted by \( \kappa^+ \); thus,
\[ \kappa^+ = \bigcap \{ \alpha \in \Omega : \kappa < |\alpha| \} . \]

Notice that by Cantor’s Theorem 3.25, for every cardinal \( \kappa \) there is a cardinal \( \lambda > \kappa \), in particular, for every cardinal \( \kappa \), \( \bigcap \{ \alpha \in \Omega : \kappa < |\alpha| \} \) is non-empty and therefore \( \kappa^+ \) exists.
A cardinal $\mu$ is called a **successor cardinal** if there exists a cardinal $\kappa$ such that $\mu = \kappa^+$; otherwise, it is called a **limit cardinal**. In particular, every positive number $n \in \omega$ is a successor cardinal and $\omega$ is the smallest non-zero limit cardinal. By induction on $\alpha \in \Omega$ we define $\omega_\alpha := \omega^\alpha_\alpha$, where $\omega_0 := \omega$, and $\omega_\alpha := \cup_{\beta \in \alpha} \omega_\beta$ for limit cardinals $\alpha$; notice that $\cup_{\beta \geq \alpha} \omega_\beta$ is a cardinal. In particular, $\omega_\alpha$ is the smallest uncountable limit cardinal and $\omega_1 = \omega^\omega_\omega$ is the smallest uncountable cardinal. Further, the collection $\{\omega_\alpha : \alpha \in \Omega\}$ is the class of all infinite cardinals, i.e., for every infinite cardinal $\kappa$ there is an $\alpha \in \Omega$ such that $\kappa = \omega_\alpha$. Notice that the collection of cardinals is — like the collection of ordinals — a proper class and not a set.

Cardinal addition, multiplication, and exponentiation is defined as follows:

**Cardinal addition:** For cardinals $\kappa$ and $\mu$ let $\kappa + \mu := |(\kappa \times \{0\}) \cup (\mu \times \{1\})|$.

**Cardinal multiplication:** For cardinals $\kappa$ and $\mu$ let $\kappa \cdot \mu := |\kappa \times \mu|$.

**Cardinal exponentiation:** For cardinals $\kappa$ and $\mu$ let $\kappa^\mu := |\kappa^\mu|$.

Since for any set $A$, $|A|^2 = |\mathcal{P}(A)|$, the cardinality of the power set of a cardinal $\kappa$ is usually denoted by $2^\kappa$. However, because $2^\kappa$ is the cardinality of the so-called *continuum* $\mathbb{R}$, it is usually denoted by $\mathfrak{c}$. Notice that by Cantor’s Theorem 3.25 for all cardinals $\kappa$ we have $\kappa < 2^\kappa$.

As a consequence of the definition we get the following

**Fact 5.6.** Addition and multiplication of cardinals is associative and commutative and we have the distributive law for multiplication over addition, and for all cardinals $\kappa$, $\lambda$, $\mu$, we have

$$
\kappa^\lambda = \kappa^\lambda \cdot \kappa, \quad \kappa^\lambda \cdot \mu = (\kappa^\lambda)^\mu = \kappa^\mu \cdot \lambda^\mu.
$$

**Proof.** It is obvious that addition and multiplication is associative and commutative and that we have the distributive law for multiplication over addition. Now, let $\kappa$, $\lambda$, $\mu$, be any cardinal numbers. Firstly, for every function $f : (\lambda \times \{0\}) \cup (\mu \times \{1\}) \to \kappa$ let the functions $f_\lambda : (\lambda \times \{0\}) \to \kappa$ and $f_\mu : (\mu \times \{1\}) \to \kappa$ be such that for each $x \in (\lambda \times \{0\}) \cup (\mu \times \{1\})$,

$$
f(x) = \begin{cases} 
  f_\lambda(x) & \text{if } x \in \lambda \times \{0\}, \\
  f_\mu(x) & \text{if } x \in \mu \times \{1\}.
\end{cases}
$$

It is easy to see that each function $f : (\lambda \times \{0\}) \cup (\mu \times \{1\}) \to \kappa$ corresponds to a unique pair $(f_\lambda, f_\mu)$, and vice versa, each pair $(f_\lambda, f_\mu)$ defines uniquely a function $f : (\lambda \times \{0\}) \cup (\mu \times \{1\}) \to \kappa$. Thus, we have a bijection between $\kappa^{\lambda+\mu}$ and $\kappa^\lambda \cdot \kappa^\mu$.

Secondly, for every function $f : \mu \to \lambda \kappa$ let $\tilde{f} : \mu \cdot \lambda \to \kappa$ be such that for all $\alpha \in \mu$ and all $\beta \in \lambda$ we have

$$
\tilde{f}(\alpha, \beta) = f(\alpha)(\beta).
$$
We leave it as an exercise to the reader to verify that the mapping
\[ \mu(\lambda \kappa) \rightarrow \mu \times \lambda \kappa \]
\[ f \rightarrow \bar{f} \]
is bijective, and therefore we have \( \kappa^{\mu \times \lambda} = (\kappa^\lambda)^\mu \).

Thirdly, for every function \( f : \mu \rightarrow \kappa \times \lambda \) let the functions \( f_\kappa : \mu \rightarrow \kappa \) and \( f_\lambda : \mu \rightarrow \lambda \) be such that for each \( \alpha \in \mu \), \( f(\alpha) = (f_\kappa(\alpha), f_\lambda(\alpha)) \). We leave it again as an exercise to the reader to show that the mapping
\[ \mu(\kappa \times \lambda) \rightarrow \mu \kappa \times \mu \lambda \]
\[ f \rightarrow (f_\kappa, f_\lambda) \]
is a bijection.

The next result completes the proof of Theorem 5.5 (d):

**Theorem 5.7.** For any ordinal numbers \( \alpha, \beta \in \Omega \) we have
\[ \omega_\alpha + \omega_\beta = \omega_\alpha \cdot \omega_\beta = \omega_\alpha \cup \omega_\beta = \max\{\omega_\alpha, \omega_\beta\} \]
In particular, for every infinite cardinal \( \kappa \) we have \( \kappa^2 = \kappa \).

**Proof.** It is enough to show that for all \( \alpha \in \Omega \) we have \( \omega_\alpha \cdot \omega_\alpha = \omega_\alpha \). For \( \alpha = 0 \) we already know that \( |\omega \times \omega| = \omega \), thus, \( \omega_0 \cdot \omega_0 = \omega_0 \). Assume towards a contradiction that there exists a \( \beta_0 \in \Omega \) such that \( \omega_{\beta_0} \cdot \omega_{\beta_0} > \omega_{\beta_0} \). Let
\[ \alpha_0 = \bigcap \{ \alpha \in \beta_0 + 1 : \omega_\alpha \cdot \omega_\alpha > \omega_\alpha \} . \]
On \( \omega_{\alpha_0} \times \omega_{\alpha_0} \) we define an ordering “\(<\)” by stipulating
\[ \langle \gamma_1, \delta_1 \rangle < \langle \gamma_2, \delta_2 \rangle \iff \begin{cases} \gamma_1 \cup \delta_1 = \gamma_2 \cup \delta_2, & \text{or} \\ \gamma_1 \cup \delta_1 = \gamma_2 \cup \delta_2 \land \gamma_1 \in \gamma_2, & \text{or} \\ \gamma_1 \cup \delta_1 = \gamma_2 \cup \delta_2 \land \gamma_1 = \gamma_2 \land \delta_1 \in \delta_2. \end{cases} \]
This linear ordering can be visualised as follows:
Cardinal arithmetic in the presence of AC

It is easily verified that "\(<" is a well-ordering on \(\omega_\alpha \times \omega_\alpha\). Now, let \(\rho\) be the order type of the well-ordering "\(<" and let \(\Gamma : \omega_\alpha \times \omega_\alpha \to \rho\) be the unique order preserving bijection between \(\omega_\alpha \times \omega_\alpha\) and \(\rho\), i.e., \(\langle \gamma_1, \delta_1 \rangle < \langle \gamma_2, \delta_2 \rangle\) iff \(\Gamma((\gamma_1, \delta_1)) \in \Gamma((\gamma_2, \delta_2))\). Because \(\omega_\alpha \cdot \omega_\alpha > \omega_\alpha\), we have \(|\rho| > \omega_\alpha\).

Now, by the definition of the well-ordering "\(<"", there are \(\gamma_0, \delta_0 \in \omega_\alpha\) such that \(\Gamma((\gamma_0, \delta_0)) = \omega_\alpha\) and for \(\nu = \gamma_0 \cup \delta_0\) we have \(|\nu \times \nu| \geq \omega_\alpha\). Thus, for \(\omega_\beta = |\nu|\) we have \(\omega_\beta < \omega_\alpha\) (since \(\nu \in \omega_\alpha\)) and \(\omega_\beta \cdot \omega_\beta > \omega_\alpha\). In particular, \(\omega_\beta \cdot \omega_\beta > \omega_\beta\), which is a contradiction to the choice of \(\alpha_0\).

As a consequence of Theorem 5.7 we get the following

**Corollary 5.8.** If \(\kappa\) is an infinite cardinal, then \(\text{seq}(\kappa) = \kappa\) and \(\kappa^\kappa = 2^\kappa\).

**Proof.** Firstly we have \(\text{seq}(\kappa) = |\bigcup_{\beta \in \kappa} \kappa^\beta| = 1 + \kappa + \kappa^2 + \ldots = 1 + \kappa \cdot \omega = \kappa\).

Secondly, by definition we have \(\kappa^\kappa = |\kappa|^\kappa\). By identifying each function \(f \in \kappa^\kappa\) by its graph, which is a subset of \(\kappa \times \kappa\), we get \(|\kappa|^\kappa \leq |\kappa^{\kappa \times \kappa}|\), and since \(|\kappa \times \kappa| = \kappa\) we finally have \(\kappa^\kappa \leq |\kappa^{\kappa \times \kappa}| = 2^\kappa\).

Let \(\lambda\) be an infinite limit ordinal. A subset \(\mathcal{C}\) of \(\lambda\) is called **cofinal** in \(\lambda\) if \(\bigcup \mathcal{C} = \lambda\). The **cofinality** of \(\lambda\), denoted \(\text{cf}(\lambda)\), is the cardinality of a smallest cofinal set \(\mathcal{C} \subseteq \lambda\). In other words,

\[
\text{cf}(\lambda) = \min \{ |\mathcal{C}| : \mathcal{C} \text{ is cofinal in } \lambda \}.
\]

Notice that by definition, \(\text{cf}(\lambda)\) is always a cardinal number.

Let again \(\lambda\) be an infinite limit ordinal and let \(\mathcal{C} = \{ \beta_\xi : \xi \in \text{cf}(\lambda) \} \subseteq \lambda\) be cofinal in \(\lambda\). Now, for every \(\nu \in \text{cf}(\lambda)\) let \(\alpha_\nu := \bigcup \{ \beta_\xi : \xi \in \nu \}\) (notice that all the \(\alpha_\nu\)'s belong to \(\lambda\)). Then \(\langle \alpha_\nu : \nu \in \text{cf}(\lambda) \rangle\) is an increasing sequence (not necessarily in the strict sense) of length \(\text{cf}(\lambda)\) with \(\bigcup \{ \alpha_\nu : \nu \in \text{cf}(\lambda) \} = \lambda\). Thus, instead of cofinal **subsets** of \(\lambda\) we could equally well work with cofinal **sequences**.

Since every infinite cardinal is an infinite limit ordinal, \(\text{cf}(\kappa)\) is also defined for cardinals \(\kappa\). An **infinite cardinal** \(\kappa\) is called **regular** if \(\text{cf}(\kappa) = \kappa\); otherwise, \(\kappa\) is called **singular**. For example, \(\omega\) is regular and \(\omega_\omega\) is singular (since \(\{ \omega_n : n \in \omega \}\) is cofinal in \(\omega_\omega\)). In general, for non-zero limit ordinals \(\lambda\) we have \(\text{cf}(\omega_\lambda) = \text{cf}(\lambda)\). For example \(\text{cf}(\omega_\omega) = \text{cf}(\omega_{\omega + \omega}) = \text{cf}(\omega_{\omega \cdot \omega}) = \omega\).

**Fact 5.9.** For all infinite limit ordinals \(\lambda\), the cardinal \(\text{cf}(\lambda)\) is regular.

**Proof.** Let \(\kappa = \text{cf}(\lambda)\) and let \(\langle \alpha_\xi : \xi \in \kappa \rangle\) be an increasing, cofinal sequence of \(\lambda\). Further, let \(\mathcal{C} \subseteq \kappa\) be cofinal in \(\kappa\) with \(|\mathcal{C}| = \text{cf}(\kappa)\). Now, \(\langle \alpha_\nu : \nu \in \mathcal{C} \rangle\) is still a cofinal sequence of \(\lambda\), which implies that \(\text{cf}(\lambda) \leq \text{cf}(\kappa)\). On the other hand we have \(\text{cf}(\kappa) \leq \kappa = \text{cf}(\lambda)\). Hence, \(\text{cf}(\kappa) = \kappa = \text{cf}(\lambda)\), which shows that \(\text{cf}(\lambda)\) is regular.

The following result — which implicitly uses AC — shows that all infinite successor cardinals are regular.
Proposition 5.10. If \( \kappa \) is an infinite cardinal, then \( \kappa^+ \) is regular.

Proof. Assume towards a contradiction that there exists a subset \( C \subseteq \kappa^+ \) such that \( C \) is cofinal in \( \kappa^+ \) and \( |C| < \kappa^+ \), i.e., \( |C| \leq \kappa \). Since \( C \subseteq \kappa^+ \), for every \( \alpha \in C \) we have \( |(\alpha)\subseteq \kappa^+ \), which implies that \( C \) is not cofinal in \( \kappa^+ \).

For example, \( \omega_1, \omega_3, \) and \( \omega_{\omega+5} \) are regular, since \( \omega_1 = \omega_0^+, \omega_3 = \omega_1^+, \) and \( \omega_{\omega+5} = \omega_{\omega+4}^+ \).

We now consider arbitrary sums and products of cardinal numbers. For this, let \( I \) be a non-empty set and let \( \{ \kappa_i : i \in I \} \) be a family of cardinals. We define

\[
\sum_{i \in I} \kappa_i = \left| \bigcup_{i \in I} A_i \right|
\]

where \( \{ A_i : i \in I \} \) is a family of pairwise disjoint sets such that \( |A_i| = \kappa_i \) for each \( i \in I \), e.g., \( A_i = \kappa_i \times \{ i \} \) will do.

Similarly we define

\[
\prod_{i \in I} \kappa_i = \left| \prod_{i \in I} A_i \right|
\]

where \( \{ A_i : i \in I \} \) is a family of sets such that \( |A_i| = \kappa_i \) for each \( i \in I \), e.g., \( A_i = \kappa_i \) will do.

Theorem 5.11 (Inequality of König-Jourdain-Zermelo). Let \( I \) be a non-empty set and let \( \{ \kappa_i : i \in I \} \) and \( \{ \lambda_i : i \in I \} \) be families of cardinal numbers such that \( \kappa_i < \lambda_i \) for every \( i \in I \). Then

\[
\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i.
\]

Proof. Let \( \{ A_i : i \in I \} \) be a family of pairwise disjoint sets such that \( |A_i| = \kappa_i \) for each \( i \in I \). Firstly, for each \( i \in I \) choose a injection \( f_i : A_i \rightarrow \lambda_i \) and an element \( y_i \in \lambda_i \setminus f_i[A_i] \) (notice that since \( |A_i| < \lambda_i \), the set \( \lambda_i \setminus f_i[A_i] \) is non-empty).

As a first step we show that \( \sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i \). For this, define \( \bar{f} : \bigcup_{i \in I} A_i \rightarrow \prod_{i \in I} \lambda_i \) by stipulating \( \bar{f}(x) := (\bar{f}_i(x) : i \in I) \) where

\[
\bar{f}_i(x) = \begin{cases} f_i(x) & \text{if } x \in A_i, \\ y_i & \text{otherwise.} \end{cases}
\]

Then \( \bar{f} \) is obviously a one-to-one function from \( \bigcup_{i \in I} A_i \) into \( \prod_{i \in I} \lambda_i \), which shows that \( \sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i \).
To see that \(\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i\), take any function \(g : \bigcup_{i \in I} A_i \rightarrow \prod_{i \in I} \lambda_i\). For every \(i \in I\), let \(P_i(g[A_i])\) be the projection of \(g[A_i]\) on \(\kappa_i\). Then, for each \(i \in I\) we can choose an element \(z_i \in \lambda_i \setminus P_i(g[A_i])\). Evidently, the sequence \(\langle z_i : i \in I \rangle\) does not belong to \(g[\bigcup_{i \in I} A_i]\) which shows that \(g\) is not surjective, and consequently, \(g\) is not bijective.

As an immediate consequence we get the following

**Corollary 5.12.** For every infinite cardinal \(\kappa\) we have

\[\kappa < \kappa^{cf(\kappa)} \quad \text{and} \quad cf(2^\kappa) > \kappa.\]

In particular we get that \(cf(\kappa) \geq \omega\).

**Proof.** Let \(\langle \alpha_\nu : \nu \in cf(\kappa) \rangle\) be a cofinal sequence of \(\kappa\). On the one hand we have

\[\kappa = \left| \bigcup_{\nu \in cf(\kappa)} \alpha_\nu \right| \leq \sum_{\nu \in cf(\kappa)} |\alpha_\nu| \leq cf(\kappa) \cdot \kappa = \kappa,\]

and hence, \(\kappa = \sum_{\nu \in cf(\kappa)} |\alpha_\nu|\). On the other hand, for each \(\nu \in cf(\kappa)\) we have \(|\alpha_\nu| < \kappa\), and therefore, by **Theorem 5.11**, we have

\[\sum_{\nu \in cf(\kappa)} |\alpha_\nu| < \prod_{\nu \in cf(\kappa)} \kappa = \kappa^{cf(\kappa)}.\]

Thus, we have \(\kappa < \kappa^{cf(\kappa)}\).

In order to see that \(cf(2^\kappa) > \kappa\), notice that \(cf(2^\kappa) \leq \kappa\) would imply that \((2^\kappa)^{cf(2^\kappa)} \leq (2^\kappa)^\kappa = 2^{\kappa \cdot \kappa} = 2^\kappa\), which contradicts the fact that \(2^\kappa < (2^\kappa)^{cf(2^\kappa)}\).

---

**Some Weaker Forms of the Axiom of Choice**

**The Prime Ideal Theorem and Related Statements**

The following maximality principle — which is frequently used in areas like Algebra and Topology — is just slightly weaker than the **Axiom of Choice**. However, in order to formulate this choice principle we have to introduce the notion of **Boolean algebra** and **ideals**.

A **Boolean algebra** is an algebraic structure, say

\((B, +, \cdot, -, 0, 1)\)

where \(B\) is a non-empty set, “+” and “·” are two binary operations (called **Boolean sum** and **product**), “-” is an unary operation (called **complement**),...
and 0, 1 are two constants. For all \( u, v, w \in B \), the Boolean operations satisfy the following axioms:

\[
\begin{align*}
    u + v &= v + u & & \text{(commutativity)} \\
    u + (v + w) &= (u + v) + w & & \text{(associativity)} \\
    u \cdot (v + w) &= (u \cdot v) + (u \cdot w) & & \text{(distributivity)} \\
    u \cdot (u + v) &= u & & \text{(absorption)} \\
    u + (\neg u) &= 1 & & \text{(complementation)}
\end{align*}
\]

An **algebra of sets** is a collection \( \mathcal{S} \) of subsets of a given set \( S \) such that \( S \in \mathcal{S} \) and whenever \( X, Y \in \mathcal{S} \), then \( S \setminus (X \cap Y) \in \mathcal{S} \) (i.e., \( \mathcal{S} \) is closed under unions, intersections and complements). An algebra of sets \( \mathcal{S} \subseteq \mathcal{P}(S) \) is a Boolean algebra, with Boolean sum and product being \( \cup \) and \( \cap \) respectively, the complement \( \neg \) of a set \( X \in \mathcal{S} \) being \( S \setminus X \), and with \( \emptyset \) and \( S \) being the constants 0 and 1 respectively. In particular, for any set \( S \), \( (\mathcal{P}(S), \cup, \cap, \neg, \emptyset, S) \) is a Boolean algebra. The case when \( S = \omega \) plays an important role throughout this book and some combinatorial properties of the Boolean algebra \( (\mathcal{P}(\omega), \cup, \cap, \neg, \emptyset, \omega) \) will be investigated in Chapters 8–10.

From the axioms above one can derive additional Boolean algebraic rules that correspond to rules for the set operations \( \cup, \cap \) and \( \neg \). Among others we have

\[
    u + u = u \cdot u = \neg (\neg u) = u, \quad u + 0 = u, \quad u \cdot 0 = 0, \quad u + 1 = 1, \quad u \cdot 1 = u,
\]

as well as the two **De Morgan laws**

\[
    \neg (u + v) = \neg u \cdot \neg v \quad \text{and} \quad \neg (u \cdot v) = \neg u + \neg v.
\]

The De Morgan laws might be better recognised for example in set-theoretic notation as

\[
    S \setminus (X \cup Y) = (S \setminus X) \cap (S \setminus Y)
\]

where \( X, Y \in \mathcal{P}(S) \); or in Propositional Logic as

\[
    \neg (\varphi \lor \psi) \equiv \neg \varphi \land \neg \psi
\]

where \( \varphi \) and \( \psi \) are any sentences formulated in a certain language.

This last formulation in the language of Propositional Logic shows the relation between Boolean algebra and Logic and provides other examples of Boolean algebras:

Let \( \mathcal{L} \) be a first-order language and let \( S \) be the set of all \( \mathcal{L} \)-sentences. We define an equivalence relation “\( \sim \)” on \( S \) by stipulating

\[
    \varphi \sim \psi \iff \vdash \varphi \leftrightarrow \psi.
\]

The set \( B := S/\sim \) of all equivalence classes \( [\varphi] \) is a Boolean algebra under the operations \( [\varphi] + [\psi] := [\varphi \lor \psi], [\varphi] \cdot [\psi] := [\varphi \land \psi], -[\varphi] := [\neg \varphi] \), where
0 := \{ \varphi \land \neg \varphi \} and 1 := \{ \varphi \lor \neg \varphi \}. This algebra is called the **Lindenbaum algebra**.

Let us define
\[ u - v = u \cdot (-v) \]
and
\[ u \leq v \iff u - v = 0. \]

We leave it as an exercise to the reader to verify that “\( \leq \)” is a partial ordering on \( B \) and that
\[ u \leq v \iff u + v = v \iff u \cdot v = u. \]

Notice also that \([\varphi] \leq [\psi]\) is equivalent to \( \vdash \varphi \rightarrow \psi \).

With respect to that ordering, 1 is the greatest element of \( B \) and 0 is the least element. Also, for any \( u, v \in B \), \( u + v \) is the least upper bound of \( \{u, v\} \), and \( u \cdot v \) is the greatest lower bound of \( \{u, v\} \). Moreover, since \( -u \) is the unique element \( v \) of \( B \) such that \( u + v = 1 \) and \( u \cdot v = 0 \) we get that all Boolean-algebraic operations can be defined in terms of the partial ordering \( \leq \) (e.g., \( -u \) is the least element \( v \) of \( B \) with the property that \( u + v = 1 \)).

Now, let us define an additional operation “\( \oplus \)” on \( B \) by stipulating
\[ u \oplus v = (u - v) + (v - u). \]

Notice that for every \( u \in B \) we have \( u \oplus u = 0 \), thus, with respect to the operation “\( \oplus \)”, every element of \( B \) is its own (and unique) inverse. We leave it as an exercise to the reader to show that \( B \) with the two binary operations \( \oplus \) and \( \cdot \) is a **ring** with zero \( 0 \) and unit \( 1 \).

Before we give the definition of ideals in Boolean algebras, let us briefly recall the **algebraic notion** of ideals in commutative rings: Let \( R = (R, +, \cdot, 0) \) be an arbitrary commutative ring. An non-empty subset \( \mathcal{I} \subseteq R \) is an **ideal** in \( R \) if and only if for all \( x, y \in \mathcal{I} \) and all \( r \in R \) we have \( x - y \in \mathcal{I} \) and \( r \cdot x \in \mathcal{I} \).

The ideal \{0\} is called the **trivial ideal**. An ideal \( I \subseteq R \) of a ring is called **maximal** if \( I \neq R \) and the only ideals \( J \) in \( R \) for which \( I \subseteq J \) are \( J = I \) and \( J = R \). If \( R \) is a commutative ring and \( I \neq R \) is an ideal in \( R \), then \( I \) is called a **prime ideal** if given any \( r, s \in R \) with \( r \cdot s \in I \) we always have \( r \in I \) or \( s \in I \).

It is not hard to verify that in a commutative ring with \( 1 \), every maximal ideal is prime. Finally, if an ideal \( J \subseteq R \) is generated by a single element of \( R \), then \( J \) is so-called **principal ideal**.

With respect to the ring \((B, \oplus, \cdot, 0, 1)\), this leads to the following definition of ideals in Boolean algebras.

Let \((B, +, \cdot, 0, 1)\) be a Boolean algebra. An **ideal** \( I \) in \( B \) is a non-empty proper subset of \( B \) with the following properties:

- **0** \( \in I \) but **1** \( \notin I \).
- If \( u \in I \) and \( v \in I \), then \( u + v \in I \).
- For all \( w \in B \) and all \( u \in I \), \( w \cdot u \in I \) (or equivalently, if \( w \in B \), \( u \in I \) and \( w \leq u \), then \( w \in I \)).
Considering the Boolean algebra \( \mathcal{P}(\omega), \cup, \cap, -, \emptyset, \omega \), one easily verifies that the set of all finite subsets of \( \omega \) is an ideal over \( \omega \), i.e., an ideal on \( \mathcal{P}(\omega) \). This ideal is called the Fréchet ideal.

The dual notion of an ideal is a so-called filter. Thus, a filter \( F \) in \( B \) is a non-empty proper subset of \( B \) with the following properties:

1. \( 0 \notin F \) but \( 1 \in F \).
2. If \( u \in F \) and \( v \in F \), then \( u \cdot v \in F \).
3. For all \( w \in B \) and all \( u \in F \), \( w + u \in I \) (or equivalently, if \( w \in B \), \( u \in F \) and \( w \leq u \), then \( w \in F \)).

Moreover, if \( I \) is an ideal in \( B \), then \( I^* := \{ -u : u \in I \} \) is a filter, called dual filter. Similarly, if \( F \) is a filter in \( B \), then \( F^* := \{ -u : u \in F \} \) is an ideal, called dual ideal. The dual filter \( I^*_0 = \{ x \subseteq \omega : \omega \setminus x \text{ is finite} \} \) of the Fréchet ideal \( I_0 \) on \( \mathcal{P}(\omega) \) is called the Fréchet filter.

Let \( I \) be an ideal in \( B \), and let \( F \) be a filter in \( B \).

- \( I \) is called
  - **trivial** if \( I = \{0\} \);
  - **principal** if there is an \( u \in B \) such that \( I = \{ v : v \leq u \} \);
  - **prime** if for all \( u \in B \), either \( u \in I \) or \( -u \in I \);

- \( F \) is called
  - **trivial** if \( F = \{1\} \);
  - **principal** if there is an \( u \in B \) such that \( F = \{ v : v \geq u \} \);
  - **ultrafilter** if for all \( u \in B \), either \( u \in F \) or \( -u \in F \).

Let us consider a few ideals and filters over \( \omega \), i.e., ideals and filters in the Boolean algebra \( \mathcal{P}(\omega), \cup, \cap, -, \emptyset, \omega \). The trivial ideal is \( \{0\} \), and the trivial filter is \( \{ \omega \} \). For any non-empty subset \( x \subseteq \omega \), \( F_x := \{ y \in \mathcal{P}(\omega) : y \supseteq x \} \) is a principal filter, and the dual ideal \( I_{\omega \setminus x} := (F_x)^* = \{ z \in \mathcal{P}(\omega) : \omega \setminus z \in F_x \} = \{ z \in \mathcal{P}(\omega) : z \cap x = \emptyset \} \) is also principal. In particular, if \( x = \{a\} \) for some \( a \in \omega \), then \( F_x \) is a principal ultrafilter and \( I_{\omega \setminus \{a\}} \) is a principal prime ideal. We leave it as an exercise to the reader to show that every principal ultrafilter over \( \omega \) is of the form \( F_{\{a\}} \) for some \( a \in \omega \), and that every principal prime ideal is of the form \( I_{\omega \setminus \{a\}} \). Considering the Fréchet filter \( F \) on \( \mathcal{P}(\omega) \), one easily verifies that \( F \) is a non-principal filter, but not an ultrafilter (notice that neither \( x = \{2n : n \in \omega \} \) nor \( \omega \setminus x \) belongs to \( F \)). Similarly, the Fréchet ideal is not prime but non-principal.

Let us now summarize a few basic properties of ultrafilters over sets (the proofs are left to the reader):

**Fact 5.13.** Let \( U \) be an ultrafilter over a set \( S \).

1. If \( \{x_0, \ldots, x_{n-1}\} \subseteq \mathcal{P}(S) \) (for some \( n \in \omega \)) such that \( x_0 \cup \ldots \cup x_{n-1} \in U \) and for any distinct \( i, j \in n \) we have \( x_i \cap x_j \notin U \), then there is a unique \( k \in n \) such that \( x_k \in U \).
2. If \( x \in U \) and \( |x| \geq 2 \), then there exists a proper subset \( y \subsetneq x \) such that \( y \in U \).
3. If \( U \) contains a finite set, then \( U \) is principal.
On the one hand, prime ideals and ultrafilters in Boolean algebras are always maximal. On the other hand, one cannot prove in ZF that for example the Fréchet filter over $\omega$ can be extended to an ultrafilter. In particular, there are models of ZF in which every ultrafilter over $\omega$ is principal (cf. RELATED RESULT 38 and Chapter 17).

However, there is a choice principle which guarantees that every ideal in a Boolean algebra can be extended to a prime ideal, and consequently, that every filter can be extended to an ultrafilter.

**Prime Ideal Theorem:** If $I$ is an ideal in a Boolean algebra, then $I$ can be extended to a prime ideal.

In fact, the **Prime Ideal Theorem**, denoted PIT, is a choice principle which is just slightly weaker than the full Axiom of Choice. Below we shall present some equivalent formulations of the Prime Ideal Theorem, but first let us show that the Prime Ideal Theorem follows from the Axiom of Choice (for the fact that the converse implication does not hold see Theorem 7.16).

**Proposition 5.14. AC $\Rightarrow$ PIT.**

**Proof.** By Theorem 5.3 it is enough to show that the Prime Ideal Theorem follows from Teichmüller’s Principle. Let $(B, +, \cdot, -, 0, 1)$ be a Boolean algebra and let $I_0 \subseteq B$ be an ideal. Further, let $\mathcal{F}$ be the family of all sets $X \subseteq B \setminus I_0$ such that for every finite subset $\{u_0, \ldots, u_n\} \subseteq X \cup I_0$ we have

$$u_0 + \ldots + u_n \neq 1.$$ 

Obviously, $\mathcal{F}$ has finite character, and therefore, by Teichmüller’s Principle, $\mathcal{F}$ has a maximal element. In other words, there is a maximal subset $I_1$ of $B$ which has the property that whenever we pick finitely many elements $\{u_0, \ldots, u_n\}$ from $I := I_0 \cup I_1$ we have $u_0 + \ldots + u_n \neq 1$. Since $I_1$ is maximal we get that $I$ is an *ideal* in $B$ which extends $I_0$. Moreover, the ideal $I$ has the property that for any element $v \in B \setminus I$ there is a $u \in I$ such that $u + v = 1$, i.e., for any $v \in B$, $v \notin I$ implies $-v \in I$. Thus, $I$ is a prime ideal in $B$ which extends $I_0$.

A seemingly weaker version of PIT is the following statement.

**Ultrafilter Theorem:** If $F$ is a filter over a set $S$, then $F$ can be extended to an ultrafilter.

Notice that the Ultrafilter Theorem is the dual version of the Prime Ideal Theorem in the case when the Boolean algebra is an algebra of sets.

For the next version of the Prime Ideal Theorem we have to introduce first some terminology: Let $S$ be a set and let $\mathcal{B}$ be a set of binary functions (i.e., with values 0 or 1) defined on finite subsets of $S$. We say that $\mathcal{B}$ is a **binary mess** on $S$ if $\mathcal{B}$ satisfies the following properties:
• For each finite set \( P \subseteq S \), there is a function \( g \in B \) such that \( \text{dom}(g) = P \), i.e., \( g \) is defined on \( P \).
• For each \( g \in B \) and each finite set \( P \subseteq S \), the restriction \( g|_P \) belongs to \( B \).

Let \( f \) be a binary function on \( S \) and let \( B \) be a binary mess on \( S \). Then \( f \) is consistent with \( B \) if for every finite set \( P \subseteq S \), \( f|_P \in B \).

**Consistency Principle:** For every binary mess \( B \) on a set \( S \), there exists a binary function \( f \) on \( S \) which is consistent with \( B \).

In order to state the last version of the Prime Ideal Theorem we have to introduce first some terminology from **Propositional Logic:** The alphabet of Propositional Logic consists of an arbitrarily large but fixed set \( \mathcal{P} := \{ p_\lambda : \lambda \in A \} \) of so-called **propositional variables**, as well as all of the logical operators “\( \neg \)”, “\( \land \)”, and “\( \lor \)”. The formulae of Propositional Logic are defined recursively as follows:

• A single propositional variable \( p \in \mathcal{P} \) by itself is a formula.
• If \( \varphi \) and \( \psi \) are formulae, then so are \( \neg(\varphi) \), \( (\varphi \land \psi) \), and \( (\varphi \lor \psi) \); in Polish notation, the three composite formulae are \( \neg\varphi \), \( \land\varphi\psi \), and \( \lor\varphi\psi \), respectively.

A **realisation** of Propositional Logic is a map of \( \mathcal{P} \), the set of propositional variables, to the two element Boolean algebra \( (\{0, 1\}, +, \cdot, \neg, 0, 1) \). Given a realisation \( f \) of Propositional Logic. By induction on the complexity of formulae we extend \( f \) to all formulae of Propositional Logic (compare with the definition of Lindenbaum’s algebra): For any formulae \( \varphi \) and \( \psi \), if \( f(\varphi) \) and \( f(\psi) \) have already been defined, then

\[
f(\land\varphi\psi) = f(\varphi) \cdot f(\psi), \quad f(\lor\varphi\psi) = f(\varphi) + f(\psi),
\]

and

\[
f(\neg\varphi) = \neg f(\varphi).
\]

Let \( \varphi \) be any formula of Propositional Logic. If the realisation \( f \), extended in the way just described, maps the formula \( \varphi \) to 1, then we say that \( f \) satisfies \( \varphi \). Finally, a set \( \Sigma \) of formulae of Propositional Logic is **satisfiable** if there is a realisation which simultaneously satisfies all the formulae in \( \Sigma \).

**Compactness Theorem for Propositional Logic:** Let \( \Sigma \) be a set of formulae of Propositional Logic. If every finite subset of \( \Sigma \) is satisfiable, then also \( \Sigma \) is satisfiable.

Notice that the reverse implication of the Compactness Theorem for Propositional Logic is trivially satisfied.

Now we show that the above principles are all equivalent to the Prime Ideal Theorem.
The Prime Ideal Theorem and related statements

Theorem 5.15. The following statements are equivalent:

(a) Prime Ideal Theorem.
(b) Ultrafilter Theorem.
(c) Consistency Principle.
(d) Compactness Theorem for Propositional Logic.
(e) Every Boolean algebra has a prime ideal.

Proof. (a)⇒(b) The Ultrafilter Theorem is an immediate consequence of the Prime Ideal Theorem.

(b)⇒(c) Let $B$ be a binary mess on a non-empty set $S$. Assuming the Ultrafilter Theorem we show that there is a binary function $f$ on $S$ which is consistent with $B$. Let $\text{fin}(S)$ be the set of all finite subsets of $S$. For each $P \in \text{fin}(S)$, let

$$A_P = \{ g \in S^2 : |g|_P \in B \}.$$ 

Since $B$ is a binary mess, the intersection of finitely many sets $A_P$ is non-empty. Thus, the family $\mathcal{F}$ consisting of all supersets of intersections of finitely many sets $A_P$ is a filter over $S^2$. By the Ultrafilter Theorem, $\mathcal{F}$ can be extended to an ultrafilter $\mathcal{U} \subseteq \mathcal{P}(S^2)$. Since $\mathcal{U}$ is an ultrafilter, for each $s \in S$, either $\{ g \in S^2 : g(s) = 0 \}$ or $\{ g \in S^2 : g(s) = 1 \}$ belongs to $\mathcal{U}$, and we define the function $f \in S^2$ by stipulating that for each $s \in S$, the set $A_s = \{ g \in S^2 : g(s) = f(s) \}$ belongs to $\mathcal{U}$. Now, for any finite set $P = \{ s_0, \ldots, s_n \} \subseteq S$, $\bigcap_{s \in P} A_s \in \mathcal{U}$, which shows that $f|_P \in B$, i.e., $f$ is consistent with $B$.

(c)⇒(d) Let $\Sigma$ be a set of formulae of Propositional Logic and let $S \subseteq \mathcal{P}$ be the set of propositional variables which appear in formulae of $\Sigma$. Assume that every finite subset of $\Sigma$ is satisfiable, i.e., for every finite subset $\Sigma_0 \subseteq \Sigma$ there is a realisation $g_{\Sigma_0} : S_{\Sigma_0} \to \{ 0, 1 \}$ which satisfies $\Sigma_0$, where $S_{\Sigma_0}$ denotes the set of propositional variables which appear in formulae of $\Sigma_0$. Let

$$\mathcal{B}_\Sigma := \{ g_{\Sigma_0}|_P : \Sigma_0 \in \text{fin}(\Sigma) \land P \subseteq S_{\Sigma_0} \}.$$ 

Then $\mathcal{B}_\Sigma$ is obviously a binary mess and by Consistency Principle there exists a binary function $f$ on $S$ which is consistent with $\mathcal{B}_\Sigma$. Now, $f$ is a realisation of $\Sigma$ and therefore $\Sigma$ is satisfiable.

(d)⇒(e) Let $(B, +, \cdot, \neg, 0, 1)$ be a Boolean algebra and let $\mathcal{P} := \{ p_u : u \in B \}$ be a set of propositional variables. Further, let $\Sigma_B$ be the following set of formulae of Propositional Logic:

- $p_0$, $\neg p_1$;
- $p_u \lor \neg p_{-u}$ (for each $u \in B$);
- $\neg(p_{u_1} \land \ldots \land p_{u_n}) \lor p_{u_1 + \ldots + u_n}$ (for each finite set $\{ u_1, \ldots, u_n \} \subseteq B$);
- $\neg(p_{u_1} \lor \ldots \lor p_{u_n}) \lor p_{u_1 \ldots u_n}$ (for each finite set $\{ u_1, \ldots, u_n \} \subseteq B$).

Notice that every finite subset of $B$ generates a finite subalgebra of $B$ and that every finite Boolean algebra has a prime ideal. Now, since every finite
prime ideal in a finite subalgebra of \( B \) corresponds to a realisation of a finite subset of \( \Sigma_B \), and vice versa, every finite subset of \( \Sigma_B \) is satisfiable. Thus, by the Compactness Theorem for Propositional Logic, \( \Sigma_B \) is satisfiable. Let \( f \) be a realisation of \( \Sigma_B \) and let \( I = \{ u \in B : f(p_u) = 1 \} \). By definition of \( \Sigma_B \) and \( I \) respectively we get:

- \( f(p_0) = 1 \) and \( f(p_1) = 0 \); thus, \( 0 \in I \) but \( 1 \notin I \).
- \( f(p_u) = 1 - f(\neg p_u) \); thus, for all \( u \in B \), either \( u \in I \) or \( -u \in I \).
- If \( f(p_{u_1}) = f(p_{u_2}) = 1 \), then \( f(p_{u_1} \land p_{u_2}) = 1 \); thus, for all \( u_1, u_2 \in I \) we have \( u_1 + u_2 \in I \).
- If \( f(p_{u_1}) = 1 \), then \( f(p_{u_1} \lor p_{u_2}) = 1 \); thus, for all \( u_1 \in I \) and all \( u_2 \in B \) we have \( u_1 \cdot u_2 \in I \).

Thus, the set \( I = \{ u \in B : f(p_u) = 1 \} \) is a prime ideal in \( B \).

(e) \( \Rightarrow \) (a) Let \((B, +, \cdot, -, 0, 1)\) be a Boolean algebra and \( I \subseteq B \) an ideal in \( B \). Define the following equivalence relation on \( B \):

\[ u \sim v \iff (u - v) + (v - u) \in I \]

Let \( C \) be the set of all equivalence classes \([u]^-\) and define the operations \( +, -\), and \( \cdot\) on \( C \) as follows:


Now,

\[ (C, +, \cdot, -, [0]^-, [1]^-) \]

is a Boolean algebra, the so-called quotient of \( B \) modulo \( I \). By the Prime Ideal Theorem, \( C \) has a prime ideal \( J \). We leave it as an exercise to the reader to verify that the set

\[ \{ u \in B : [u]^- \in J \} \]

is a prime ideal in \( B \) which extends \( I \).

König’s Lemma and other Choice Principles

Let us begin by defining some choice principles:

- \( \mathcal{C}(\mathbb{N}_0, \infty) \): Every countable family of non-empty sets has a choice function (this choice principle is usually called Countable Axiom of Choice).
- \( \mathcal{C}(\mathbb{N}_0, \mathbb{N}_0) \): Every countable family of non-empty countable sets has a choice function.
- \( \mathcal{C}(\mathbb{N}_0, < \mathbb{N}_0) \): Every countable family of non-empty finite sets has a choice function.
- \( \mathcal{C}(\mathbb{N}_0, \omega) \): Every countable family of \( n \)-element sets, where \( n \in \omega \), has a choice function.
- \( \mathcal{C}(\infty, < \mathbb{N}_0) \): Every family of non-empty finite sets has a choice function (this choice principle is usually called Axiom of Choice for Finite Sets).
• \( C(\infty, n) \): Every family of \( n \)-element sets, where \( n \in \omega \), has a choice function. This choice principle is usually denoted \( C_n \).

Another — seemingly unrelated — choice principle is the Ramsey Partition Principle, denoted RPP.

• RPP: If \( X \) is an infinite set and \( |X|^2 \) is 2-coloured, then there is an infinite subset \( Y \) of \( X \) such that \( |Y|^2 \) is monochromatic.

Below we show how these choice principles are related to each other, but first let us show that \( C(\aleph_0, < \aleph_0) \) and König’s Lemma, denoted by KL, are equivalent.

**Proposition 5.16.** \( C(\aleph_0, < \aleph_0) \iff KL \).

**Proof.** (\( \Rightarrow \)) Let \( T = (V, E) \) be an infinite, finitely branching tree with vertex set \( V \), edge set \( E \), and root say \( v_0 \). The edge set \( E \) can be considered as a subset of \( V \times V \), i.e., as a set of ordered pairs of vertices indicating the direction from the root to the top of the tree. Let \( S_0 := \{v_0\} \), and for \( n \in \omega \) let

\[
S_{n+1} := \{v \in V : \exists u \in S_n((u, v) \in E)\}
\]

and let \( S := \bigcup_{n \in \omega} S_n \). Since \( T \) is infinite and finitely branching, \( S \) is infinite and for every \( n \in \omega \), \( S_n \) is a non-empty finite set. Further, for every \( v \in S \) let \( S(v) \) be the set of all vertices \( u \in S \) such that there exists a non-empty finite sequence \( s \in \text{seq}(S) \) of length \( k + 1 \) (for some \( k \in \omega \)) with \( s(0) = v \) and \( s(k) = u \), and for all \( i \leq k \) we have \( \langle s(i), s(i+1) \rangle \in E \). In other words, \( S(v) \) is the set of all vertices which can be reached from \( v \). Notice that \( (S(v), E|_{S(v)}) \) is a subtree of \( T \). Since \( S \) is infinite and for all \( n \in \omega \), \( \bigcup_{i \in n} S_i \) is finite, for each \( n \in \omega \) there exists a vertex \( v \in S_n \) such that \( S(v) \) is infinite.

We now proceed as follows: By \( C(\aleph_0, < \aleph_0) \), for each \( n \in \omega \) we can choose a well-ordering \( "<_n" \) on \( S_n \) and then construct a branch \( v_0, v_1, \ldots, v_n, \ldots \) through \( T \), where for all \( n \in \omega \), \( v_{n+1} \) is the \( <_{n+1} \)-minimal element of the non-empty set \( \{ v \in S_{n+1} : (v_n, v) \in E \land "S(v) is infinite" \} \).

(\( \Leftarrow \)) Let \( \mathcal{F} = \{F_n : n \in \omega\} \) be a countable family of non-empty finite sets. Further, let \( V = \bigcup_{k \in \omega} \left( \prod_{n \in k} F_n \right) \) and let \( E \subseteq V \times V \) be the set of all ordered pairs \( (s, t) \) of the form \( s = (x_0, \ldots, x_n) \) and \( t = (x_0, \ldots, x_n, x_{n+1}) \) respectively, where for each \( i \in n+2, x_i \in F_i \) (i.e., the sequence \( t \) is obtained by adding an element of \( F_{n+1} \) to \( s \)). Obviously, \( T = (V, E) \) is an infinite, finitely branching tree and therefore, by KL, has an infinite branch, say \( \langle a_n : n \in \omega \rangle \). Since, for all \( n \in \omega \), \( a_n \) belongs to \( F_n \), the function

\[
f : \mathcal{F} \to \bigcup \mathcal{F}
\]

\[F_n \mapsto a_n\]

is a choice function for \( \mathcal{F} \), and since the countable family of finite sets \( \mathcal{F} \) was arbitrary, we get \( C(\aleph_0, < \aleph_0) \). 

-1
Obviously, \( C(\aleph_0, \aleph_0) \Rightarrow C(\aleph_0, n) \) for all positive integers \( n \in \omega \). However, as a matter of fact we would like to mention that for each \( n \geq 2 \), \( C(\aleph_0, n) \) is a proper axiom, i.e., not provable within ZF (for \( n = 2 \) see for example Proposition 7.7).

The following result shows the strength of the choice principles RPP and KL compared to \( C(\aleph_0, \infty) \) and \( C(\aleph_0, n) \) respectively:

**Theorem 5.17.** \( C(\aleph_0, \infty) \Rightarrow \text{RPP} \Rightarrow \text{KL} \Rightarrow C(\aleph_0, n) \).

**Proof.** \( C(\aleph_0, \infty) \Rightarrow \text{RPP} \): Firstly we show that \( C(\aleph_0, \infty) \) implies that every infinite set \( X \) is transfinite, i.e., there is an infinite sequence of elements of \( X \) in which no element appears twice: Let \( X \) be an infinite set and for every \( n \in \omega \) let \( F_{n+1} \) be the set of all injections from \( n + 1 \) into \( X \). Consider the family \( \mathcal{F} = \{ F_{n+1} : n \in \omega \} \). Since \( X \) is infinite, \( \mathcal{F} \) is a countable family of non-empty sets. Thus, by \( C(\aleph_0, \infty) \), there is a choice function, say \( f \), on \( \mathcal{F} \). For every \( n \in \omega \) let \( g_n := f(F_{n+1}) \). With the countably many injections \( g_n \) we can easily construct an injection from \( \omega \) into \( X \). In particular, we get an infinite sequence \( \langle a_i : i \in \omega \rangle \) of elements of \( X \) in which no element appears twice.

For \( S := \{ a_i : i \in \omega \} \subseteq X \), every 2-colouring of \( \{X\}^2 \) induces a 2-colouring of \( [S]^2 \). Now, by Ramsey’s Theorem 2.1, there exists an infinite subset \( Y \) of \( S \) such that \( [Y]^2 \) is monochromatic (notice that no choice is needed to establish Ramsey’s Theorem for countable sets).

**RPP \Rightarrow KL:** Let \( T = (V, E) \) be an infinite, finitely branching tree and let the sets \( S_n \) (for \( n \in \omega \)) be as in the first part of the proof of Proposition 5.16. Define the colouring \( \pi : [V]^2 \to \{0, 1\} \) by stipulating \( \pi(\{u, v\}) = 0 \iff \{u, v\} \subseteq S_n \) for some \( n \in \omega \). By RPP there exists an infinite subset \( X \subseteq V \) such that \( [X]^2 \) is monochromatic. Now, since \( T \) is finitely branching, we get that if \( X \subseteq V \) is infinite and \( [X]^2 \) is monochromatic, then \( [X]^2 \) is of colour 1, i.e., no two distinct elements of \( X \) are in the same set \( S_n \). In order to construct an infinite branch through \( T \), just proceed as in the first part of the proof of Proposition 5.16.

**KL \Rightarrow C(\aleph_0, n):** Because \( C(\aleph_0, \aleph_0) \Rightarrow C(\aleph_0, n) \), this is an immediate consequence of Proposition 5.16.

The last result of this chapter deals with the relationship of the choice principles \( C_n \) (i.e., \( C(\infty, n) \)) for different natural numbers \( n \). Before we can state the theorem we have to introduce the following number-theoretical condition: Let \( m, n \) be two positive integers. Then we say that \( m, n \) satisfy condition \( (S) \) if the following condition holds:

There is no decomposition of \( n \) into a sum of primes, \( n = p_1 + \ldots + p_s \), such that \( p_i > m \) for all \( 1 \leq i \leq s \).

**Theorem 5.18.** If the positive integers \( m, n \) satisfy condition \( (S) \) and if \( C_k \) holds for every \( k \leq m \), then also \( C_n \) holds.
Proof. Firstly notice that $C_1$ is obviously true. Secondly notice that for $n \leq m$, the implication of the theorem is trivially true. So, without loss of generality we may assume that $n > m$.

The proof is now by induction on $n$: Let $m < n$ be a fixed positive integer such that $m, n$ satisfy condition (S) and assume that the implication of the theorem is true for every $l < n$. Since $n, m$ satisfy (S), $n$ is not a prime and consequently $n$ is divisible by some prime $p < n$. Necessarily, $p \leq m$, since otherwise we could write $n = p + \ldots + p$, contrary to (S). Let $\mathcal{F} = \{ A_\lambda : \lambda \in A \}$ be a family of $n$-element sets. We have to describe a way to choose an element from each set $A_\lambda (\lambda \in A)$. Take an arbitrary $A \in \mathcal{F}$ and consider $[A]^p$ (i.e., the set of all $p$-element subsets of $A$). Since $p \leq m$, by the premise of the theorem there is a choice function $g$ for $[A]^p$. In other words, for every $X \in [A]^p$, $g(X) \in X$, in particular, $g(X) \in A$. For every $a \in A$ let

$$q(a) = \{|X \in [A]^p : g(X) = a\}|$$

and let $q := \min \{ q(a) : a \in A \}$. Further, let $B := \{ a \in A : q(a) = q \}$. Obviously, the set $B$ is non-empty and the set $[A]^p$ has $\binom{n}{p}$ elements. In order to prove that $A \setminus B$ is non-empty, we have to show that $\binom{n}{p}$ is not divisible by $n$. Indeed, because $p$ divides $n$, there is a positive integer $k$ which is not divisible by $p$ such that $n = k \cdot p^{a+1}$ (for some $a \in \omega$). We have

$$\binom{n}{p} = \frac{k \cdot p^{a+1}}{p} \times \frac{(n-1) \cdots (n-p+1)}{1} = \frac{k \cdot p^{a+1}}{p} \times \frac{(n-1)}{p}$$

and since $p$ does obviously not divide $\binom{n-1}{p-1}$, we get that $\binom{n}{p}$ is divisible by $p^a$, but not by $p^{a+1}$; in particular, $\binom{n}{p}$ is not divisible by $n = k \cdot p^{a+1}$. Thus, the sets $B$ and $A \setminus B$ are both non-empty, and for $l_1 := |B|$ and $l_2 := |A \setminus B|$ we get that $l_1$ and $l_2$ are positive integers with $l_1 + l_2 = n$. Moreover, $m, l_1$ or $m, l_2$ satisfy condition (S), since otherwise we could write $l_1 = p_1 + \ldots + p_r$ and $l_2 = p_{r+1} + \ldots + p_s$, where $p_1, \ldots, p_s$ are primes bigger than $m$, which would imply that $n = p_1 + \ldots + p_s$, contrary to the assumption that $m, n$ satisfy (S). Thus, by the induction hypothesis, either $C_{l_1}$ holds and we choose an element in $B$, or, if $C_{l_1}$ fails, $C_{l_2}$ holds and we choose an element in $A \setminus B$. Finally, since $A \in \mathcal{F}$ was arbitrary, this completes the proof. −

Notes

The Axiom of Choice. Fraenkel writes in [26, p. 56 f.] that the Axiom of Choice is probably the most interesting and, in spite of its late appearance, the most discussed axiom of Mathematics, second only to Euclid’s axiom of parallels which was introduced more than two thousand years ago. We would also like to mention a different view to choice functions, namely the view of Peano. In 1890, Peano published a proof in which he was constrained to choose a single element from each set in a certain infinite sequence $A_1, A_2, \ldots$ of infinite subsets of $\mathbb{R}$. In that proof, he remarked carefully
But as one cannot apply infinitely many times an arbitrary rule by which one assigns to a class \( A \) an individual of this class, a determinate rule is stated here, by which, under suitable hypotheses, one assigns to each class \( A \) an individual of this class. To obtain his rule, he employed least upper bounds. According to Moore [66, p. 76], Peano was the first mathematician who—while accepting infinite collections—categorically rejected the use of infinitely many arbitrary choices.

The difficulty is well illustrated by a Russellian anecdote (cf. Sierpiński [82, p. 123]): A millionaire possesses an infinite number of pairs of shoes, and an infinite number of pairs of socks. One day, in a fit of eccentricity, he summons his valet and asks him to select one shoe from each pair. When the valet, accustomed to receiving precise instructions, asks for details as to how to perform the selection, the millionaire suggests that the left shoe be chosen from each pair. Next day the millionaire proposes to the valet that he select one sock from each pair. When asked as to how this operation is to be carried out, the millionaire is at a loss for a reply, since, unlike shoes, there is no intrinsic way of distinguishing one sock of a pair from the other. In other words, the selection of the socks cannot be carried out without the aid of some choice function.

As long as the implicit and unconscious use of the Axiom of Choice by Cantor and others involved only generalised arithmetical concepts and properties well-known from finite numbers, nobody took offence. However, the situation changed drastically after Zermelo [107] published his first proof that every set can be well-ordered—which was one of the earliest assertions of Cantor. It is worth mentioning that, according to Zermelo [107, p. 514] & [108, footnote p. 118], it was in fact the idea of Erhard Schmidt to use the Axiom of Choice in order to build the \( f \)-sets. Zermelo considered the Axiom of Choice as a logical principle, that cannot be reduced to a still simpler one, but is used everywhere in mathematical deductions without hesitation (see [107, p. 516]). Even though in Zermelo’s view the Axiom of Choice was “self-evident”, which is not the same as “obvious” (see Shaprio [81, §5] for a detailed discussion of the meaning of “self-evidence”), not all mathematicians at that time shared Zermelo’s opinion. Moreover, after the first proof of the Well-Ordering Principle was published in 1904, the mathematical journals (especially volume 60 of Mathematische Annalen) were flooded with critical notes rejecting the proof (see for example Moore [66, Chapter 2]), mostly arguing that the Axiom of Choice was either illegitimate or meaningless (cf. Fraenkel, Bar-Hillel, and Löwy [26, p. 82]). The reason for this was not only due to the non-constructive character of the Axiom of Choice, but also because it was not yet clear what a “set” should be. So, Zermelo decided to publish a more detailed proof, and at the same time taking the opportunity to reply to his critics. This resulted in [108], his second proof of the Well-Ordering Principle which was published in 1908, the same year as he presented his first axiomatisation of Set Theory in [108]. It seems that this was not a coincidence. Moore [66, p. 159] writes that Zermelo’s axiomatisation was primarily motivated by a desire to secure his demonstration of the Well-Ordering Principle and, in particular, to save his Axiom of Choice. Moreover, Hallett [32, p. xvi] goes even further by trying to show that the selection of the axioms themselves was guided by the demands of Zermelo’s reconstructed second proof. Hallett’s statement is motivated by a remark on page 124 in Zermelo [108], where he emphasises that the proof is just based on certain fixed principles to build initial sets and to derive new sets from given ones—exactly what we would require for principles to form an axiomatic system of Set Theory.
We would like to mention that because of its different character (cf. Bernays [3]) and since he considered the Axiom of Choice as a general logical principle, he did not include the Axiom of Choice in his second axiomatic system of Set Theory.

For a comprehensive survey of Zermelo's Axiom of Choice, its origins, development, and influence, we refer the reader to Moore [66] (see also Kanamori [46], Jech [41], and Fraenkel, Bar-Hillel, and Lévy [26, Chapter II, §4]); and for a biography of Zermelo (including the history of AC and axiomatic Set Theory) we refer the reader to Ebbinghaus [17].

Gödel's constructible universe. According to Kanamori [45, p. 28 ff.], in October of 1935 Gödel informed von Neumann at the Institute for Advanced Study in Princeton that he had established the relative consistency of the Axiom of Choice. This he did by devising his constructible (not constructive!) hierarchy L (for “law”) and verifying the Axiom of Choice and the rest of the ZF axioms there. Gödel conjectured that the Continuum Hypothesis would also hold in L, but he soon fell ill and only gave a proof of this and the Generalised Continuum Hypothesis (i.e., for all \(\alpha \in \Omega; 2^{\aleph_\alpha} = \omega_{\alpha+1}\)) two years later. The crucial idea apparently came to him during the night of June 14/15, 1937 (see also [31, pp. 1–8]).

Gödel's article [28] was the first announcement of these results, in which he describes the model \(L\) as the class of all “mathematically constructible” sets, where the term “constructible” is to be understood in the semi-intuitionistic sense which excludes impredicative procedures. This means “constructible” sets are defined to be those sets which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite ordinals. In the succeeding article [29], Gödel provided more details in the context of ZF, and in his monograph [30] — based on lectures given at the Institute for Advanced Study during the winter of 1938/39 — Gödel gave another presentation of \(L\). This time he used \(L\) to set off with a transfinite recursion in terms of eight elementary set generators, a sort of Gödel numbering into the transfinite (cf. Kanamori [45, p. 30], and for Gödel's work in Set Theory see Kanamori [47]).

Equivalent Forms of the Axiom of Choice. The literature gives numerous examples of theorems which are equivalent to the Axiom of Choice and a huge collection of such equivalent forms of the Axiom of Choice was accumulated by Rubin and Rubin [79, 80].

The most popular variants of the Axiom of Choice — and the most often used in mathematical proofs — are probably the Well-Ordering Principle (discussed above), the Kuratowski-Zorn Lemma, and Teichmüller’s Principle.

The Kuratowski-Zorn Lemma was proved independently by Kuratowski [53] and more than a decade later by Zorn [106] (see Moore [66, p. 223] and also Campbell [13]). Usually, the Kuratowski-Zorn Lemma is deduced quite easily from the Well-Ordering Principle. The direct deduction from the Axiom of Choice presented above (Theorem 5.3) is due to Kneser [51], who also proved lemma 5.2 which was stated without proof by Bourbaki [12, p. 37 (lemme fundamental)].

Teichmüller’s Principle was formulated independently by Tuley [103] and slightly earlier by Teichmuller in [97], where he provides also some equivalent forms of this very useful principle. Teichmüller himself was a member of the Nazi party and joined the army in 1939. Fighting first in Norway and then at the Eastern Front, he eventually died in 1943.
Kurepa's Principle was introduced by Kurepa in [54], where he showed that Kurepa's Principle together with the Linear-Ordering Principle — which states that every set can be linearly ordered — implies the Axiom of Choice. The proof that — in the presence of the Axiom of Foundation — Kurepa's Principle implies the Axiom of Choice is due to Felgner [38] (see also Felgner and Jech [20] or Jech [40, Theorem 9.1. (a)]).

The proof that "every vector space has an algebraic basis" implies Multiple Choice is taken from Blass [9], and the proof that Multiple Choice implies Kurepa's Principle is taken from Jech [40, Theorem 9.1. (a)] (compare with Chapter 7 | RELATED RESULT 44).

Among the dozens of cardinal relations which are equivalent to the Axiom of Choice (see for example Lindenbaum and Tarski [60], Bachmann [1, §31], or Moore [66, p. 330f.]), we just mentioned three.

In 1885, Cantor [14, §2] asserted the Trichotomy of Cardinals without proof, and in a letter of 28 July 1899 (cf. [16, pp. 443–447]) he wrote to Dedekind that the Trichotomy of Cardinals follows from the Well-Ordering Principle. However, their equivalence remained unproven until Hartogs [34] established it in 1915 (cf. also Moore [66, p. 10]). As a matter of fact we would like to mention that — according to Sierpiński [82, p. 99 f. ] — Lesniewski showed that Trichotomy of Cardinals is equivalent to the statement that for any two cardinals \( n \) and \( m \), where at least one of these cardinals is infinite, we always have \( n + m = n \) or \( n + m = m \).

**Theorem 5.5 (c) —** which is to some extent a dualisation of the Trichotomy of Cardinals — was stated without proof by Lindenbaum [60, p. 312 (A6)] and the proof given above is taken from Sierpiński [83, p. 426].

The fact that the cardinal equation \( m^2 = m \) implies the Axiom of Choice is due to Tarski [87] (see also Bachmann [1, V, p. 140 ff.]).

**Cardinal arithmetic in the presence of AC.** The definition of cardinals given above can also be found for example in von Neumann [72, VII.2, p. 731].

The first proof of **Theorem 5.7** appeared in Hessenberg [38, p. 593] (see also Jourdain [44]).

Regularity of cardinals was investigated by Hausdorff, who also raised the question of existence of regular limit cardinals (cf. [35, p. 131]).

The **Inequality of König-Jourdain-Zermelo 5.11** — also known as König's Theorem — was proven by König [52] (but only for countable sums and products), and independently by Jourdain [43] and by Zermelo [110] (for historical facts see Moore [66, p. 154] and Fraenkel [25, p. 98]). Obviously, the **Inequality of König-Jourdain-Zermelo** implies the Axiom of Choice (since it guarantees that every Cartesian product of non-empty sets is non-empty), and consequently we get that the **Inequality of König-Jourdain-Zermelo is equivalent to** the Axiom of Choice.

**Algebras.** Boolean algebra is named after George Boole who — according to Russell — discovered Pure Mathematics. Even though this might be an exaggeration, it is true that Boole was one of the first to view Mathematics as the study of abstract structures rather than as the science of magnitude, and he was the first who applied successfully mathematical techniques to Logic (cf. Boole [11, 10]) and his work evolved into the modern theory of Boolean algebras and algebraic Logic. In 1849, Boole was appointed at the newly founded Queen's College in Cork, where he died.
in 1864 as a result of pneumonia caused by walking to a lecture in a December
downpour and lecturing all day in wet clothes (see also MacHale [61]).

Lindenbaum’s algebra is named in memory of the Polish mathematician Adolf
Lindenbaum, who was killed by the Gestapo at Nova Wilejka in the summer of 1941. Lindenbaum and Tarski (see for example Tarski [90, 89, 91]) developed the idea of
viewing the set of formulae as an algebra (with operations induced by the logical
connectives) independently around 1935; however, Lindenbaum’s results were not
published (see Rasiowa and Sikorski [78, footnote to page 245]).

For the history of abstract algebraic Logic and Boolean algebras we refer the
reader to Font, Jansana, and Pigózzi [22].

Prime Ideals. Ideals and prime ideals on algebras of sets where investigated for
example by Tarski in [93].

The notion of Lindenbaum’s algebra and the Compactness Theorem for Propo-
sitional Logic is taken from Bell and Slomson [2, Chapter 2]. The equivalent forms
of the Prime Ideal Theorem are taken from Jech [40, Chapter 2, §3], and the corre-
sponding references can be found in [40, Chapter 2, §7]. We would like to mention
that the Ultrafilter Theorem, which is just the dual form of the Prime Ideal Theorem,
is due to Tarski [88].

Ramsey’s Theorem as a choice principle. Ramsey’s Original Theorem
(cf. Chapter 2) implies that every infinite set $X$ has the following property: For every
2-colouring of $[X]^2$ there is an infinite subset $Y$ of $X$ such that $[Y]^2$ is mono-
chromatic. As mentioned in Chapter 2, Ramsey [76] explicitly indicated that his proof of
this theorem used the Axiom of Choice. Later, Kleinberg [50] showed that every proof
of Ramsey’s Original Theorem must use the Axiom of Choice, although rather weak forms of the Axiom of Choice like $C(N, \infty)$ suffice (see Theorem 5.17). For
the position of Ramsey’s Original Theorem in the hierarchy of choice principles
we refer the reader to Blass [8] (see also Related Result 31).

For the fact that none of the implications in Theorem 5.17 is reversible we refer
the reader to Howard and Rubin [39].

From countable choice to choice for finite sets. The Countable Axiom of
Choice asserts that every countable family of non-empty sets has a choice function,
whereas the Axiom of Choice for Finite Sets asserts that every family of non-empty
finite sets has a choice function. Replacing the finite sets in the latter choice principle
by $n$-element sets (for natural numbers $n \geq 2$), we obtained the choice principles $C_n$
which assert that every family of $n$-element sets has a choice function. Combining
these two choice principles we get in fact versions of König’s Lemma, namely choice
principles like $C(N, \infty)$ and $C(N, n)$ (for positive integers $n \geq 2$).

The proof of Theorem 5.18 is taken from Jech [40, p. 111] and is optimal in the
following sense: If the positive integers $m, n$ do not satisfy condition (S), then there
is a model of Set Theory in which $C_k$ holds for every $k \leq m$ but $C_n$ fails (see the
proof of Theorem 7.16 in Jech [40]).
RELATED RESULTS

22. Hausdorff’s Principle. Among the numerous maximality principles which are equivalent to the Axiom of Choice, we like to mention the one known as Hausdorff’s Principle (cf. Hausdorff [35, VI, §1, p. 140]):

Hausdorff’s Principle: Every partially ordered set has a maximal chain (maximal with respect to inclusion “⊆”).

For the history of Hausdorff’s Principle see Moore [66, Section 3.4, p. 167 ff.] and a proof of the equivalence with the Axiom of Choice can be found for example in Bernays [5, p. 142 ff.].

23. Bases in vector spaces and the Axiom of Choice. Relations between the existence or non-existence of bases in vector spaces and some weaker forms of the Axiom of Choice are investigated for example in Keremedis [48, 49], Läuchli [55], and Halpern [33].

24. Cardinal relations which are equivalent to AC. Below we list a few of the dozens of cardinal relations which are equivalent to the Axiom of Choice (mainly taken from Tarski [87]):

(a) \( m \cdot n = m + n \) for all infinite cardinals \( m \) and \( n \).
(b) If \( m^2 = n^2 \), then \( m = n \).
(c) If \( m < n \) and \( p < q \), then \( m + p < n + q \).
(d) If \( m < n \) and \( p < q \), then \( m \cdot p < n \cdot q \).
(e) If \( m + p < n + p \), then \( m < n \).
(f) If \( m \cdot p < n \cdot p \), then \( m < n \).
(g) If \( 2m = m + n \), then \( m < n \).

For the proofs we refer the reader to Tarski [87] and Sierpiński [83, p. 421] (compare (g) with Chapter 4 | RELATED RESULT 17). More such cardinal relations can be found for example in Howard and Rubin [39, p. 82 ff.], Rubin and Rubin [90, p. 137 ff.], Moore [66, p. 330 ff.], and Bachmann [1, §31].

25. Successors of Cardinals. In [90] Tarski investigated the following three types of successor of a cardinal number:

\( S_1 \). For every cardinal \( m \) there is a cardinal \( n \) such that \( m < n \) and the formula \( m < p < n \) does not hold for any cardinal \( p \).

\( S_2 \). For every cardinal \( m \) there is a cardinal \( n \) such that \( m < n \) and for every cardinal \( p \) the formula \( m < p \) implies \( n \leq p \).

\( S_3 \). For every cardinal \( m \) there is a cardinal \( n \) such that \( m < n \) and for every cardinal \( p \) the formula \( p < n \) implies \( p \leq m \).

Tarski [90] showed that \( S_1 \) can be proved without the help of the Axiom of Choice, whereas \( S_2 \) is equivalent to this axiom. The relation of \( S_3 \) with the Axiom of Choice was further investigated by Sobociński [84] and Truss [100] (see also Bachmann [1, §31, p. 141]).

26. A formulation by Sudan. Sudan [85] showed that the following statement is equivalent to the Axiom of Choice: Let \( m \), \( n \), and \( p \) be arbitrary infinite cardinals. If \( m \) and \( n \) are either equal or \( n \) is a \( S_1 \)-successor (i.e., a successor in the in the sense of \( S_1 \)) of \( m \), then also \( p \cdot m \) and \( p \cdot n \) are either equal or \( p \cdot n \) is an \( S_1 \)-successor of \( p \cdot m \). For the influence of Tarski [87] on Sudan see Moore [66, p. 218].
27. A formulation by Tarski. There are also some equivalents of the Axiom of Choice which seemingly are far away of being choice principles. The following formulation by Tarski [92] is surely of this type: For every set \( N \) there is a set \( M \) such that \( X \in M \) if and only if \( X \subseteq M \) and for all \( Y \subseteq X \) we have \( |Y| \neq |N| \). Similar statements can be found in Tarski [94, 95] (see also Bachmann [1, §31.3]).

28. Singular Cardinal Hypothesis. The Singular Cardinal Hypothesis states that for every singular cardinal \( \kappa \), \( \beth^\kappa < \kappa \) implies \( \kappa^{\beth^\kappa} = \kappa^+ \). Obviously, the Singular Cardinal Hypothesis follows from the Generalised Continuum Hypothesis. On the other hand, the Singular Cardinal Hypothesis is not provable within ZFC and in fact, the failure of the Singular Cardinal Hypothesis is equiconsistent with the existence of a certain large cardinal (cf. Jech [12, p. 58f. & Chapter 24]).

29. Model Theory and the Prime Ideal Theorem. Using Lindenbaum's algebra, Rasiowa and Sikorski [77] gave an alternative proof of Gödel’s Completeness Theorem 3.4, and Henkin [36] proved that the Prime Ideal Theorem is equivalent to the Compactness Theorem 3.7. Notice that by Theorem 5.15 we just get that the Prime Ideal Theorem is equivalent to the Compactness Theorem for Propositional Logic, which is a seemingly weaker statement than the Compactness Theorem 3.7.

30. Colouring infinite graphs and the Prime Ideal Theorem. For \( n \) a positive integer consider the following statement:

\[ P_n. \text{ If } G \text{ is a graph such that every finite subgraph of } G \text{ is } n\text{-colourable,} \]

\[ \text{then } G \text{ itself is } n\text{-colourable.} \]

The following implications are provable in Set Theory without the Axiom of Choice (see Mycielski [69, 70]):

\[ \text{PIT } \Rightarrow P_{n+1} \Rightarrow P_n \Rightarrow C(\infty, n), \quad C(\infty, 2) \Rightarrow P_2 \]

On the other hand, Lévy [29] showed that for any \( n \), \( ZF \not\models C(\infty, n) \Rightarrow P_3 \).

Surprisingly, Läuchli showed in [57] that \( P_3 \) implies PIT, and consequently, for all \( n \geq 3 \), the equivalence \( P_n \Rightarrow \text{PIT} \) is provable in Set Theory without the Axiom of Choice. However, the question whether there is a “direct” proof of \( P_3 \Rightarrow \text{PIT} \) without involving PIT is still open.

31. Ramsey’s Theorem, König’s Lemma, and countable choice. Truss investigated in [102] versions of König’s Lemma, where restrictions are placed on the degree of branching of the finitely branching tree. In particular, he investigated \( C(n_0, n) \) for different \( n \in \omega \). Later in [24], Forster and Truss considered the relation between versions of Ramsey’s Original Theorem and these versions of König’s Lemma.

The choice principle \( C(n_0, n) \) was also investigated by Wiśniewski [105], where it is compared with \( C(\infty, n) \) and other weak forms of the Axiom of Choice.

32. Ramsey Choice. Related to \( C_n \) are the following two choice principles: \( C_n^c \) states that every infinite family \( X \) of \( n \)-element sets has an infinite subfamily \( Y \subseteq X \) with a choice function; and \( RC_n \) states that for every infinite set \( X \) there is an infinite subset \( Y \subseteq X \) such that \( |Y|^n \) has a choice function. These two choice principles are both strictly weaker than \( C_n \) (cf. Truss [99]). Montenegro
investigated in [65] the relation between $RC_n$ and $C_n$ for some small values of $n$.

It is not hard to see that $RC_2 \Rightarrow C_2$, $RC_3 \Rightarrow C_3$ (cf. [65, Lemma]). However, it is quite tricky to prove that $RC_4 \Rightarrow C_4$ (cf. [65, Theorem]) and it is still open whether $RC_5$ implies $C_5$.

33. Well-ordered and well-orderable subsets of a set. For a set $x$, $s(x)$ is the set of all subsets of $x$ which can be well-ordered, and $w(x)$ is the set of all well-orderings of subsets of $x$. Notice that $s(x) \subseteq \mathcal{P}(x)$, whereas $w(x) \subseteq \mathcal{P}(x \times x)$. Tarski [94] showed without the help of the Axiom of Choice that $|x| < |s(x)|$, for any set $x$, and his proof also yields $|x| < |w(x)|$. Later, Truss showed in [101] that for any infinite set $x$ and for any $n \in \omega$ we have $|s(x)| \geq |x^n|$ as well as $|x^n| < |w(x)|$.

Furthermore, he showed that if there is a choice function for the set of finite subsets of $x$, then $|x^n| < |s(x)|$. According to Howard and Rubin [39, p. 371] it is not known whether $|x^n| < |s(x)|$ (Form 283 of [39]) is provable in ZF. The cardinality of the set $w(x)$ was further investigated by Forster and Truss in [23].

34. Axiom of Choice for families of $n$-element sets. For different $n \in \omega$, $C_n$ has been extensively studied by Mostowski in [67], and most of the following results — which are all provable without the help of the Axiom of Choice — can be found in that paper (see also Truss [99], Gauntt [27], or Jech [40, Chapter 7, §4]):

(a) If $m, n$ satisfy condition ($S$), then $n < 8m^2$.

(b) $C_2 \Rightarrow C_n$ is provable if and only if $n \in \{1, 2, 4\}$.

(c) For a finite set $Z = \{m_1, \ldots, m_k\}$ of positive integers let $C_Z$ denote the statement $C_{m_1} \land \cdots \land C_{m_k}$. We say that $Z, n$ satisfy condition ($S$) if for every decomposition of $n$ into a sum of primes, $n = p_1 + \cdots + p_s$, at least one prime $p_i$ belongs to $Z$. Now, the following condition holds: If $Z, n$ satisfy condition ($S$), then $C_Z$ implies $C_n$.

(d) Let $S_n$ be the group of all permutation of $\{1, \ldots, n\}$. A subgroup $G$ of $S_n$ is said to be fixed point free if for every $i \in \{1, \ldots, n\}$ there is a $\pi \in S_n$ such that $\pi(i) \neq i$. Let $Z$ be again a finite set of positive integers. We say that $Z, n$ satisfy condition ($T$) if for every fixed point free subgroup $G$ of $S_n$ there is a subgroup $H$ of $G$ and a finite sequence $H_1, \ldots, H_k$ of proper subgroups of $H$ such that the sum of indices $|H : H_1| + \cdots + |H : H_k|$ is in $Z$. Now, the following condition holds: If $Z, n$ satisfy condition ($T$), then $C_Z$ implies $C_n$. Moreover we have: If $Z, n$ do not satisfy condition ($T$), then there is a model of ZF in which $C_Z$ holds and $C_n$ fails.

We would also like to mention that the Axiom of Choice for Finite Sets $C(\infty, < \aleph_0)$ is unprovable in ZF, even if we assume that $C_n$ is true for each $n \in \omega$ (cf. Jech [40, Chapter 7, §4], or Lévy [58] and Pincus [75]).

35. Ordering principles. Among the numerous choice principles which deal with ordering we mention just two:

Ordering Principle: Every set can be linearly ordered.

If $\prec$ and $\preccurlyeq$ are partial orderings of a set $P$, then we say that $\preccurlyeq$ extends $\prec$ if for any $p, q \in P$, $p \prec q$ implies $p \preccurlyeq q$.

Order-Extension Principle: Every partial ordering of a set $P$ can be extended to a linear ordering of $P$.
Obviously, the Order-Extension Principle implies the Ordering Principle, but the other direction fails (see Mathias [62]). Thus, the Ordering Principle is slightly weaker than the Order-Extension Principle. Furthermore, Siprašan (who changed his name from Siprašan to Marczewski while hiding from the Nazi persecution) showed in [86] that the Order-Extension Principle follows from the Axiom of Choice, where one can even replace the Axiom of Choice by the Prime Ideal Theorem (see for example Jech [40, 2.3.2]). We leave it as an exercise to the reader to show that the Ordering Principle implies \( C(\omega, < \aleph_0) \). Thus, we get the following sequence of implications:

\[
\text{PIT} \Rightarrow \text{Order-Extension Principle} \Rightarrow \text{Ordering Principle} \Rightarrow \text{C}(\omega, < \aleph_0)
\]

On the other hand, none of these implications is reversible (see Läuchli [56] and Pincus [74, §13.2], Felgner and Truss [21, Lemma 2.1], Mathias [62], or Jech [40, Chapter 7]; compare also with Chapter 7 | Related Result 48).

36. More ordering principles. Mathias showed in [62] that the following assertion does not imply the Order-Extension Principle:

If \( X \) is a set of well-orderable sets, then there is a function \( f \) such that for each \( x \in X \), \( f(x) \) is a well-ordering of \( x \).

On the other hand, Truss [98] showed that following assertion, apparently only slightly stronger than the ordering principle above, implies the Axiom of Choice:

If \( X \) is a set and \( f \) a function on \( X \) such that for each \( x \in X \), \( f(x) \) is a non-empty set of well-orderings of \( x \), then \( \{ f(x) : x \in X \} \) has a choice function.

37. Principle of Dependent Choices. Finally, let us mention a choice principle which is closely related to the Countable Axiom of Choice. Its meaning is that one is allowed to make a countable number of consecutive choices.

**Principle of Dependent Choices:** If \( R \) is a binary relation on a non-empty set \( S \), and if for every \( x \in S \) there exists \( y \in S \) with \( xRy \), then there is a sequence \( \langle x_n : n \in \omega \rangle \) of elements of \( S \) such that for all \( n \in \omega \) we have \( x_nRx_{n+1} \).

The Principle of Dependent Choices, usually denoted DC, was formulated by Bernays in [4] and for example investigated by Mostowski [68] (see also Jech [40, Chapter 8]). Even though DC is significantly weaker than AC, it is stronger than \( C(\aleph_0, \omega) \) and (thus) implies for example that every Dedekind-finite set is finite (i.e., every infinity set is transfinite). Thus, in the presence of DC, many — kind of natural — propositions are still provable. On the other hand, having just DC instead of full AC, most of the somewhat paradoxical constructions (e.g., making two tails from one) cannot be carried out anymore (see Herrlich [33] for some ‘disasters’ that happen with and without AC). In my opinion, DC reflects best our intuition, and consequently, ZF+DC would be a quite reasonable and smooth axiomatic system for Set Theory; however, it is not suitable for really exciting results.

38. An alternative to the Axiom of Choice. Let \( \omega \to (\omega)^\omega \) be the statement that whenever the set \( [\omega]^\omega \) is coloured with 2 colours, there exists an infinite subset
of $\omega$, all whose infinite subsets have the same colour (compare with the Ramsey property defined in Chapter 9). In Chapter 2 we have seen that $\omega \rightarrow (\omega)^{\omega}$ fails in the presence of the Axiom of Choice. On the other hand, Mathias proved that under the assumption of the existence of an inaccessible cardinal (defined on page 315), $\omega \rightarrow (\omega)^{\omega}$ is consistent with $ZF + DC$ (see Mathias [64, Theorem 5.1]). The combinatorial statement $\omega \rightarrow (\omega)^{\omega}$ has many interesting consequences: For example Mathias [63] gave an elementary proof of the fact that if $\omega \rightarrow (\omega)^{\omega}$ holds, then there are no so-called rare filters and every ultrafilter over $\omega$ is principal (see Mathias [64, p. 91ff] for similar results).

39. The Axiom of Determinacy. Another alternative to the Axiom of Choice is the Axiom of Determinacy, which asserts that all games of a certain type are determined. In order to be more precise we have to introduce first some terminology: With each subset $A$ of $\omega$ we associate the following game $G_A$, played by two players I and II. First I chooses a natural number $a_0$, then II chooses a natural number $b_0$, then I chooses $a_1$, then II chooses $b_1$, and so on. The game ends after $\omega$ steps; if the resulting sequence $(a_0, b_0, a_1, b_1, \ldots)$ is in $A$, then I wins, otherwise II wins. A strategy (for I or II) is a rule that tells the player what move to make depending on the previous moves of both players; and a strategy is a winning strategy if the player who follows it always wins (for a more formal definition see Chapter 10). The game $G_A$ is determined if one of the players has a winning strategy.

Axiom of Determinacy (AD): For every set $A \subseteq \omega$ the game $G_A$ is determined, i.e., either player I or player II has winning strategy.

An easy diagonal argument shows that AC is incompatible with AD, i.e., assuming the Axiom of Choice there exists a set $A \subseteq \omega$ such that the game $G_A$ is not determined (cf. Jech [42, Lemma 33.1]). In contrast we have that AD implies that every countable family of non-empty sets of reals has a choice function (cf. Jech [42, Lemma 33.2]). Moreover, one can show that $\text{Con}(ZF + AD)$ implies $\text{Con}(ZF + AD + DC)$, thus, even in the presence of AD we still can have DC. Furthermore, AD implies that sets of reals are well behaved, e.g., every set of reals is Lebesgue measurable, has the property of Baire, and every uncountable set of reals contains a perfect subset, i.e., a closed set without isolated points (cf. Jech [42, Theorem 33.3]); however, it also implies that every ultralayer over $\omega$ is principal (cf. Kanamori [45, Proposition 28.1]) and that $\aleph_1$ and $\aleph_2$ are both measurable cardinals (cf. Jech [42, Theorem 33.12]). Because of its nice consequences for sets of reals, AD is a reasonable alternative to AC, especially for the investigation of the real line (for the beauty of ZF + AD see for example Herrlich [37, Section 7.2]). In 1962, when Mycielski Steinhaus [73] introduced the Axiom of Determinacy, they did not claim this new axiom to be intuitively true, but stated that the purpose of their paper is only to prove another theory which seems very interesting although its consistency is problematic. Since AD implies the existence of large cardinals, the consistency of $ZF + AD$ cannot be derived from that of ZF. Moreover, using very sophisticated techniques — far beyond the scope of this book — Woodin proved that $ZF + AD$ is equiconsistent with ZFC + “There are infinitely many Woodin cardinals” (cf. Kanamori [45, Theorem 32.16] or Jech [42, Theorem 33.27]). Further results and the corresponding references can be found for example in Kanamori [45, Chapter 6] and Jech [42, Chapter 33].
REFERENCES


29. _______, *Consistency-proof for the generalized continuum-hypothesis*, *Proceedings of the National Academy of Sciences (U.S.A.)*, vol. 25 (1939), 220–224 (reprinted in [31]).


34. **Friedrich Hartogs**, *Über das Problem der Wohlordnung*, *Mathematische Annalen*, vol. 76 (1915), 438–443.

52. **Julius König**, *Zum Kontinuum-Problem*, *Mathematische Annalen*, vol. 60 (1905), 177-180.
53. **Casimir Kuratowski**, *Une méthode d’élimination des nombres transcendants des raisonnements mathématiques*, *Fundamenta Mathematicae*, vol. 3 (1922), 76-108.
56. **Läuchli**, *The independence of the ordering principle from a restricted axiom of choice*, *Fundamenta Mathematicae*, vol. 54 (1964), 31-43.


How to Make Two Balls from One

Rests, which are so convenient to the composer and singer, arose for two reasons: necessity and the desire for ornamentation. As for necessity, it would be impossible to sing an entire composition without pausing, for it would cause fatigue that might well prevent a singer from finishing. Rest were adopted also for the sake of ornament. With them parts could enter one after another in fatigue or consequence, procedures that give a composition an artful and pleasing effect.

Giuseppe Zarlino
Le Istitutioni Harmoniche, 1558

For two reasons we shall give the reader a rest: one reason is that the reader deserves a pause to reflect on the axioms of ZFC; the other reason is that we would like to show Robinson’s beautiful construction — relying on $\mathcal{AC}$ — of how to make two balls from one by dividing the ball into only five parts.

Equidecomposability

Two geometrical figures $A$ and $A'$ (i.e., two sets of points lying on the straight line $\mathbb{R}$, on the plane $\mathbb{R}^2$, or in the three-dimensional space $\mathbb{R}^3$) are said to be congruent, denoted $A \cong A'$, if $A$ can be obtained from $A'$ by translation and/or rotation, but we shall exclude reflections. Two geometrical figures $A$ and $A'$ are said to be equidecomposable, denoted $A \approx A'$, if there is a positive integer $n$ and partitions $A = A_1 \cup \ldots \cup A_n$ and $A' = A'_1 \cup \ldots \cup A'_n$
such that for all $1 \leq i \leq n$: $A_i \cong A'_i$. To indicate that $A$ and $A'$ are equidecomposable using at most $n$ pieces we shall write $A \cong_n A'$.

Below we shall present two somewhat paradoxical decompositions of the 2-dimensional unit sphere $S_2$ as well as of the 3-dimensional solid unit ball $B_1$:

Firstly we show that the unit sphere $S_2$ can be partitioned into four parts, say $S_2 = A \cup B \cup C \cup F$, such that $F$ is countable, $A \cong B \cong C$, and $A \cong B \cup C$. This result is known as Hausdorff’s Paradox, even though it is just a paradoxical partition of the sphere $S_2$ rather than a paradox.

Secondly we show how to make two balls from one, in fact we show that $B_1 \cong_3 B_1 \cup B_1$. This result is due to Robinson and is optimal with respect to the number of pieces needed, i.e., $B_1 \not\cong_4 B_1 \cup B_1$. We would like to mention that about two decades earlier, Banach and Tarski already showed that a unit ball and two unit balls are equidecomposable; however, their construction requires many more than five pieces.

Both decompositions, Hausdorff’s partition of the sphere as well as Robinson’s decomposition of the ball, rely on the Axiom of Choice. Moreover, it can be shown that in the absence of the Axiom of Choice neither decomposition is provable — but this is beyond the scope of this book (see Related Result 41). However, before we start the constructions, let us briefly discuss the measure-theoretical background of these somewhat paradoxical partitions, in particular of the decomposition of the ball: Firstly, why does Robinson’s decomposition of the ball seem paradoxical? Of course, it is because the volume is not preserved; but what are volumes? One could consider the notion of volume as a function $\mu$ which assigns to each set $X \subseteq \mathbb{R}^n$ a non-negative real number, called the volume of $X$. We require that the function $\mu$ has the following basic properties:

- $\mu(\emptyset) = 0$ and $\mu(B_1) > 0$ (e.g., $\mu(B_1) = 1$),
- $\mu(X \cup Y) = \mu(X) + \mu(Y)$ whenever $X$ and $Y$ are disjoint, and
- $\mu(X) = \mu(Y)$ whenever $X$ and $Y$ are congruent.

Now, by the fact that a unit ball and two unit balls are equidecomposable, and implicitly by Hausdorff’s result (see below), we see that there is no such measure on $\mathbb{R}^3$, i.e., $\mu$ is not defined for all subsets of $\mathbb{R}^3$. Roughly speaking, there are some dust-like subsets of $\mathbb{R}^3$ (like the sets we shall construct) to which we cannot assign a volume. Having this in mind, Robinson’s decomposition loses its paradoxical character — but certainly not its beauty.

**Hausdorff’s Paradox**

Before we show how to make two balls from one, we will present Hausdorff’s partition of the sphere. The itinerary is as follows: Firstly we define an infinite subgroup $H$ of $SO(3)$, where $SO(3)$ is the so-called *special orthogonal group*
consisting of all rotations in \( \mathbb{R}^3 \) leaving fixed the origin. Even though the group \( H \) is infinite, it is generated by just two elements. Since \( H \) is a subgroup of \( \text{SO}(3) \), there is a natural action of \( H \) on the unit sphere \( S_2 \) which induces an equivalence relation on \( S_2 \) by \( x \sim y \iff \exists g \in H(g(x) = y) \) (i.e., \( x \sim y \) iff \( y \) belongs to the orbit of \( x \)). Then we choose from each equivalence class a representative — this is where the Axiom of Choice comes in — and use the set of representatives to define Hausdorff’s partition of the sphere.

We begin the construction by defining the group \( H \). For this, consider the following two elements of \( \text{SO}(3) \), which will be the generators of \( H \):

\[
\varphi = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \psi = \frac{1}{4} \begin{pmatrix} -2 & -\sqrt{6} & \sqrt{6} \\ \sqrt{6} & 1 & 3 \\ -\sqrt{6} & 3 & 1 \end{pmatrix}
\]

The linear mapping \( \varphi \) is the rotation through \( \pi \) about the axis \((0,0,1)\), and \( \psi \) is the rotation through \( 2\pi/3 \) about the axis \((0,1,1)\). Thus, \( \varphi^2 = \psi^3 = \iota \) where \( \iota \) denotes the identity. We leave it as an exercise to the reader to show by induction on \( n \) that for all integers \( n \geq 1 \) and for all \( \varepsilon_k = \pm 1 \) (where \( 1 \leq k \leq n \)) we have:

\[
(\varphi \psi^{\varepsilon_n} \cdots \varphi \psi^{\varepsilon_1}) = \frac{1}{2^{n+1}} \begin{pmatrix} a_1 & a_2\sqrt{6} & a_3\sqrt{6} \\ b_1\sqrt{6} & b_2 & b_3 \\ b_1'\sqrt{6} & b_2' & b_3' \end{pmatrix}
\]

where all numbers \( a_1, a_2, \ldots, b_3' \) are integers with

- \( a_1 \equiv 2 \mod 4 \),
- \( a_2, a_3, b_1, \ldots, b_3' \) are odd, and
- \( b_1 \equiv b_1', b_2 \equiv b_2', b_3 \equiv b_3' \mod 4 \).

Hence, we conclude that for all \( n \geq 1 \): \( (\varphi \psi^{\varepsilon_n} \cdots \varphi \psi^{\varepsilon_1}) \notin \{ \iota, \varphi \} \). Consequently, for all \( n \geq 1 \), for all \( \varepsilon_k = \pm 1 \) (where \( 1 \leq k \leq n \)), and for \( \varepsilon_0 \in \{0,1\} \) and \( \varepsilon_{n+1} \in \{0,\pm 1\} \), we get:

\[
\psi^{\varepsilon_{n+1}} \cdot (\varphi \psi^{\varepsilon_n} \cdots \varphi \psi^{\varepsilon_1}) \cdot \varphi^{\varepsilon_0} \neq \iota
\]

(\#)

In other words, the only relations between \( \varphi \) and \( \psi \) are \( \varphi^2 = \psi^3 = \iota \). Let \( H \) be the group of linear transformations — in fact rotations — of \( \mathbb{R}^3 \) generated by the two rotations \( \varphi \) and \( \psi \). Then \( H \) is a subgroup of \( \text{SO}(3) \) and every element of \( H \) is a rotation which corresponds, by (\#), to a unique reduced “word” of the form

\[
\psi^{\varepsilon_{n+1}} \varphi \psi^{\varepsilon_n} \cdots \varphi \psi^{\varepsilon_1} \varphi^{\varepsilon_0}
\]

where \( n \geq 0 \), \( \varepsilon_k = \pm 1 \) (for all \( 1 \leq k \leq n \)), \( \varepsilon_0 \in \{0,1\} \), and \( \varepsilon_{n+1} \in \{0,\pm 1\} \).

We now consider the so-called Cayley graph of \( H \): The Cayley graph of \( H \) is a graph with vertex set \( H \), where for \( \rho_1, \rho_2 \in H \) there is a directed edge from \( \rho_1 \) to \( \rho_2 \) if either \( \rho_2 = \varphi \rho_1 \) or \( \rho_2 = \psi \rho_1 \). In the former case, the edge is labelled \( \varphi \), in the latter case it is labelled \( \psi \), e.g., \( \psi \rho \xrightarrow{\varphi} \varphi \psi \rho \) or \( \psi \varphi \xrightarrow{\psi} \varphi \).
To each vertex of the Cayley graph of $H$ (i.e., to each element of $H$) we assign a label, which is either $\mathbf{1}$, $\mathbf{2}$, or $\mathbf{3}$. The labelling is done according to the following rules:

- The identity $e$ gets the label $\mathbf{1}$.
- If $\rho \in H$ is labelled $\mathbf{2}$ or $\mathbf{3}$ and $\sigma = \varphi \rho$, then $\sigma$ is labelled $\mathbf{1}$.
- If $\rho \in H$ is labelled $\mathbf{1}$ and $\sigma = \varphi \rho$, then $\sigma$ is labelled either $\mathbf{2}$ or $\mathbf{3}$.
- If $\rho \in H$ is labelled $\mathbf{1}$ (or $\mathbf{2}$, or $\mathbf{3}$) and $\sigma = \psi \rho$, then $\sigma$ is labelled $\mathbf{2}$ (or $\mathbf{3}$, or $\mathbf{1}$, respectively).

These rules are illustrated by the following figures and diagrams:

The lightface label $\mathbf{3}$ indicates that if $\rho$ is a reduced word in $H$, labelled $\mathbf{1}$, of the form $\psi^e \rho'$ for $e = \pm 1$, then $\varphi \rho$ is always labelled $\mathbf{2}$ (not $\mathbf{3}$).

The following figure shows part of the labelled Cayley graph of $H$:

The group $H$ acts on the 2-dimensional unit sphere $S_2$ and we define the equivalence relation “$\sim$” on $S_2$ via $x \sim y$ iff there is a $\rho \in H$ such that
\( \rho(x) = y \). The equivalence classes of \("\sim"\) are usually called \( H \)-orbits, and the \( H \)-orbit containing \( x \in S_2 \) is written \([x]\). Let \( F \subseteq S_2 \) be the set of all fixed points (i.e., the set of all \( y \in S_2 \) such that there is a \( \rho \in H \setminus \{ x \} \) with \( \rho(y) = y \)) since \( H \) is countable and every rotation \( \rho \in H \) has two fixed points, \( F \) is countable. We notice first that any point equivalent to a fixed point is a fixed point (i.e., for every \( x \in S_2 \setminus F \) we have \([x] \subseteq S_2 \setminus F \)). Indeed, if \( \rho(y) = y \) for some \( \rho \in H \) and \( y \in S_2 \), then \( \sigma \rho \sigma^{-1}(\sigma(y)) = \sigma(y) \); that is, if \( y \) is fixed for \( \rho \), then \( \sigma(y) \) is fixed for \( \sigma \rho \sigma^{-1} \). Thus, a class of equivalent points consists either entirely of fixed points, or entirely of non-fixed points.

By the Axiom of Choice there is a choice function \( f \) for \( \mathcal{F} = \{ [x] : x \in S_2 \setminus F \} \) and let \( M = \{ f([x]) : x \in S_2 \setminus F \} \).

Now we define labels for all non-fixed points (i.e., points in \( S_2 \setminus F \)) as follows: Firstly, every element in \( M \) is labelled \( \bullet \). Secondly, if \( x \in S_2 \setminus F \), then there is a unique rotation \( \rho \in H \) such that \( \rho(y) = x \), where \( \{ y \} = M \cap [x] \).

We define the label of the point \( x \) by the label of \( \rho \) in the labelled Cayley graph of \( H \). This induces a partition of \( S_2 \setminus F \) into the following three parts:

\[
A = \{ x \in S_2 \setminus F : x \text{ is labelled } \bullet \}, \\
B = \{ x \in S_2 \setminus F : x \text{ is labelled } \mathfrak{0} \}, \\
C = \{ x \in S_2 \setminus F : x \text{ is labelled } \mathfrak{1} \}.
\]

Thus, \( S_2 = A \cup B \cup C \cup F \) and by the labelling of the vertices of the Cayley graph of \( H \) we get:

\[
B = \psi[A], \quad C = \psi^{-1}[A], \quad B \cup C = \varphi[A].
\]

Hence, we get that \( A \cong B, A \cong C \), and that \( A \cong B \cup C \). We leave it as an exercise to the reader to show that this implies \((S_2 \setminus F) \cong (S_2 \setminus F) \cup (S_2 \setminus F)\).

For each point \( x \in S_2 \) let \( l_x \) be the line joining the origin (i.e., the centre of the sphere) with \( x \), and for \( S \subseteq S_2 \) define \( \tilde{S} := \bigcup \{ l_x : x \in S \} \). Then the sets \( \tilde{A}, \tilde{B}, \) and \( \tilde{C} \), cannot be Lebesgue measurable (otherwise we would have \( 0 < \mu(B) = \mu(C) = \mu(B \cup C) \), a contradiction). In fact, Hausdorff’s decomposition shows that there is no non-vanishing measure on \( S_2 \) which is defined for all subsets of \( S_2 \) such that congruent sets have the same measure.

**Robinson’s Decomposition**

Robinson’s decomposition of the ball is similar to Hausdorff’s partition of the sphere. Firstly we define an infinite subgroup \( G \) of \( \text{SO}(3) \), where \( G \) is generated by four generators. The action of \( G \) on the unit ball \( B_1 \) (with centre the origin) induces an equivalence relation on \( B_1 \), and we choose from each equivalence class a representative. With the set of representatives and a sophisticated labelling we finally define a partition of \( B_1 \) into five parts \( A_1, \ldots, A_5 \), such that we can make a solid unit ball with either the two sets \( A_1 \) and \( A_3 \), or with the three sets \( A_2, A_4 \) and \( A_5 \).
Let the rotations \( \varphi \) and \( \psi \) be as above. Let \( \chi := \psi \varphi \psi \). Then, one easily verifies by induction on \( m \) that for all positive \( m \in \omega \) we have \( \chi^m = \psi (\varphi \psi^2)^{m-1} \varphi \psi \) and \( \chi^{-m} = \psi^2 \varphi (\psi \varphi^2)^{m-1} \psi^2 \). Now, by (\( * \)), we get that for every \( k \geq 1 \) and any non-zero integers \( p_1, p_2, \ldots, p_k \):

\[
\chi^{p_1} \varphi \chi^{p_2} \varphi \ldots \varphi \chi^{p_k} \neq \iota
\]

For \( 1 \leq m \leq 4 \) define

\[
\varphi_m = \chi^m \varphi \chi^m.
\]

We leave it again as an exercise to the reader to verify that for every \( k \geq 1 \), any non-zero integers \( p_1, p_2, \ldots, p_k \), and any \( i_1, \ldots, i_k \in \{1, 2, 3, 4\} \) where \( i_l \neq i_{l+1} \) for all \( 1 \leq l < k \):

\[
\varphi_{i_1}^{p_1} \varphi_{i_2}^{p_2} \ldots \varphi_{i_k}^{p_k} \neq \iota
\]

Let \( G \) be the subgroup of \( \text{SO}(3) \) generated by the four rotations \( \varphi_1, \ldots, \varphi_4 \).

We consider now the labelled Cayley graph of \( G \), where we allow again some freedom in the labelling process (indicated by lightface labels). The rules for labelling the vertices of the Cayley graph of \( G \) are illustrated by the following figure:
The following figure shows part of the labelled Cayley graph of $G$ in which just $\varphi_1$ and $\varphi_2$ are involved:

The group $G$ acts on the solid unit ball $B_1$ and we define the equivalence relation “$\sim$” on $B_1$ like above via $x \sim y \text{ iff }$ there is a $\rho \in G$ such that $\rho(x) = y$. The $G$-orbit containing $x \in B_1$ is again written $[x]^\sim$. Let $P$ be an arbitrary point on the unit sphere ($i.e.$, on the surface of $B_1$), which does not belong to any rotation axis, and finally let $E \subseteq B_1$ be the set of all points which belong to a rotation axis and which are distinct from the origin. It is easy to see that for every $x \in B_1 \setminus E$ we have $[x]^\sim \subseteq B_1 \setminus E$. By the Axiom of Choice there is a choice function $f$ for $\mathcal{F} = \{[x]^\sim : x \in B_1 \setminus E\}$ and let $M = \{f([x]^\sim) : x \in B_1 \setminus E\} \setminus \{0\}$, where 0 denotes the origin.

We first define labels for all points in $B_1 \setminus (E \cup [P]^\sim)$ as follows:

- Every element in $M$ is labelled 1.
- The origin is labelled 3.
- If $x \in B_1 \setminus E$ and $\rho(y) = x$, where $\{y\} = M \cap [x]^\sim$, then the label of the point $x$ is defined as the label of $\rho$ in the labelled Cayley graph of $G$.

Consider now the set $E$ and fix any class $[z]^\sim \subseteq E$. Choose a rotation $\theta \neq \iota$ having a fixed point in $[z]^\sim$ and which is as short as possible, or more precisely, which is expressible as a product of the smallest possible number of factors of the form $\varphi_m^{k1}$ with $m \in \{1, 2, 3, 4\}$. Fix an arbitrary point $x_0 \in [z]^\sim$ such that $\theta(x_0) = x_0$. 

Robinson's decomposition
Firstly we show that if \( \rho(x_0) = x_0 \), then \( \rho = \theta^n \) for some integer \( n \). If \( \rho = \iota \), then \( \rho = \theta^0 \) and we are done. Thus, we may assume that \( \rho \neq \iota \).

Notice first that the initial and final factors of \( \theta \) — where \( \theta \) and all other products of rotations are read from the right to the left — cannot be inverse, since otherwise, for some \( \sigma = \varphi_m \) where \( m \in \{1, 2, 3, 4\} \) and \( n \in \{ -1, 1 \} \), the rotation \( \sigma \theta \sigma^{-1} \) would be shorter than \( \theta \) and would have a fixed point in the same equivalence class \([z]^-\). Thus, the rotations \( \theta \) and \( \theta^{-1} \) neither begin nor end with the same factor. Now, if \( \rho \) has the same fixed point \( x_0 \) as \( \theta \), then \( \rho \theta = \theta \rho \). If \( \rho \theta \) does not simplify when \( \rho \) and \( \theta \) are written in terms of the \( \varphi_m \) where \( m \in \{1, 2, 3, 4\} \), then, by (i), \( \theta \rho \) must also not simplify. Hence, \( \rho \) must begin with the block \( \theta \). Inductively one finds that \( \rho \) is obtained by writing the block \( \theta \) \( n \)-times, that is, \( \rho = \theta^n \), where \( n \) is a positive integer. In case \( \rho \theta \) does simplify, then \( \rho \theta^{-1} \) does not (since \( \theta \) and \( \theta^{-1} \) end with different factors). Thus, we may apply the same argument as before to the equation \( \rho \theta^{-1} = \theta^{-1} \rho \) and find that \( \rho = \theta^{-n} \), where \( n \) is again a positive integer.

Secondly, notice that each point \( y \in [z]^- \) may be written in the form \( \sigma_y(x_0) \), where \( \sigma_y \in G \) is a rotation which starts neither with the block \( \theta \) (when written in terms of the \( \varphi_k \)), nor with the inverse of the last factor of \( \theta \) (where \( \theta \) is still read from the right to the left). The former property is obvious; and to achieve the latter property consider \( \sigma_\theta \theta^n \), where \( n \) is sufficiently large, and then simplify and remove any remaining blocks \( \theta \). Notice that this representation is unique. For suppose that \( \sigma(x_0) = \rho(x_0) \), where \( \sigma \) and \( \rho \) are again written in terms of the \( \varphi_k \). Then \( \rho^{-1} \sigma(x_0) = x_0 \), hence, \( \rho^{-1} \sigma = \theta^n \). If \( n > 0 \), this yields \( \sigma = \rho \theta^n \), which is impossible since \( \rho \theta^n \) does not simplify and \( \sigma \) does not begin with the block \( \theta \). If \( n < 0 \), we may interchange the roles of \( \sigma \) and \( \rho \) and again reach a contradiction. Hence we have \( n = 0 \), which is \( \sigma = \rho \).

Thirdly, assume that \( \theta \) is of the form

\[
\theta = \varphi_{i_k} \varphi_{i_{k-1}} \cdots \varphi_{i_1}
\]

where the \( i_l \)'s \( (1 \leq l \leq k) \) belong to \( \{1, 2, 3, 4\} \) and each exponent \( j_l \) is \( \pm 1 \).

So, starting with the point \( x_0 \), we obtain successively the \( k \) distinct points

\[
x_0, x_1 = \varphi_{i_k}^j(x_0), x_2 = \varphi_{i_{k-1}}^j(x_0), \ldots, x_k = \varphi_{i_1}^j(x_0) = x_0
\]

which form a closed cycle. As shown above, each point \( y \in [z]^- \) can be written uniquely in the form \( \sigma_y(x_0) \), where \( \sigma_y \) starts neither with the block \( \theta \) nor with the rotation \( \varphi_{i_k}^{-j_k} \).

Consider the following figure:
As a consequence of the preceding arguments we get that, starting with \( x_0 \), there are no other closed cycles in \( [z]^- \): Indeed, let \( y \in [z]^- \) and \( \rho \neq \iota \) be such that \( \rho(y) = y \). Now, \( y = \sigma_y(x_0) \) where \( \sigma_y \) is as above. Now, \( \rho \sigma_y(x_0) = \sigma_y(x_0) \) and therefore \( \sigma_y^{-1} \rho \sigma_y(x_0) = x_0 \). Consequently we have \( \sigma_y^{-1} \rho \sigma_y = \theta^n \) which implies \( y \in \{ x_0, \ldots, x_k \} \).

Now we are ready to assign a label to each point in \( E \): Firstly, for every \( [z]^- \), where \( z \in E \), we choose a rotation \( \theta_z \neq \iota \) having a fixed point in \( [z]^- \) and which is as short as possible, and then choose a point \( x_0 \in [z]^- \) such that \( \theta(x_0) = x_0 \). Assume that \( \theta_z \) is of the form \( \theta_z = \varphi_{i_1}^{j_1} \varphi_{i_2}^{j_1-1} \cdots \varphi_{i_l}^{j_l} \) where the \( i_l \)'s (for \( 1 \leq l \leq k \)) belong to \( \{ 1, 2, 3, 4 \} \) and each exponent \( j_l \) is \( \pm 1 \). Then from the point \( x_0 \) we obtain successively the points \( x_1, \ldots, x_{k-1} \), \( x_k = x_0 \). We know that every point \( y \in [z]^- \) can be written uniquely in the form \( \sigma_y(x_0) \), where \( \sigma_y \) starts neither with the block \( \theta_z \) nor with the rotation \( \varphi_{i_l}^{-j_l} \), and that, starting with \( x_0 \), there are no other closed cycles in \( [z]^- \). Thus, in order to label the points in \( [z]^- \) it is enough to assign a label to the \( k \) points of the cycle in a way which respects the labelling rules given above; the remaining points may be labelled like the non-fixed points, i.e., like the points in \( B_1 \setminus (E \cup \{ P \}) \).
For this, consider the following schemata which illustrate the labelling rules:

\[
\begin{align*}
\varphi_1 & & \varphi_2 & & \varphi_3 & & \varphi_4 \\
\begin{bmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1
\end{bmatrix} & = R_1 \\
\begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1
\end{bmatrix} & = R_2 \\
\begin{bmatrix}
1 & 1 & 0 & 1 \\
1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix} & = R_3 \\
\begin{bmatrix}
1 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix} & = R_4
\end{align*}
\]

For \(1 \leq m \leq 4\), the matrix \(R_m\), which corresponds to \(\varphi_m\), is such that \(a_{ij} = 0\) iff whenever \(\sigma\) has label \(\circ\), \(\varphi_m \sigma\) cannot get label \(\circ\). It is easy to see that for \(1 \leq m \leq 4\), the matrix \(R_m^T\) corresponds to \(\varphi_m^{-1}\). Consequently, the rotation \(\theta_i\) corresponds to a certain product of the matrices \(R_1, \ldots, R_4\) and their transposes. In particular, \(\theta_i\) corresponds to a \(4 \times 4\) matrix \(Q\). By considering the trace of \(Q\), \(\text{tr}(Q)\), and by applying the fact that for any matrices \(A\) and \(B\) we have \(\text{tr}(A^T) = \text{tr}(A)\) and \(\text{tr}(AB) = \text{tr}(BA)\), one can easily verify that \(\text{tr}(Q) \neq 0\). This implies that there exists a sequence of labels say \(\circ_0, \circ_1, \ldots, \circ_k\) with \(l_0 = l_k\) (here we use that \(\text{tr}(Q) \neq 0\)) such that labelling \(x_i^k\) with \(\varphi_i\) (for \(0 \leq i \leq k\)) respects the labelling rules.

So, we can assign a label to each of the \(k\) points \(x_0^k, \ldots, x_{k-1}^k\) of the cycle in a way which respects the labelling rules, and consequently, we can assign a label to every point in \(E\). Thus, the only points which are not labelled yet are the points in \([P]\). For the point \(P\), and only for this single point, we modify the labelling as illustrated by the following figure (the further labelling of the points in \([P]\) is done according to the labelling rules):
Finally, we have labelled all points of $B_1 \setminus \{ P \}$ with four labels, which induces a partition of $B_1$ into the following five parts:

\[
A_1 = \{ x \in B_1 : x \text{ is labelled } ① \}
\]
\[
A_2 = \{ x \in B_1 : x \text{ is labelled } ② \}
\]
\[
A_3 = \{ x \in B_1 : x \text{ is labelled } ③ \}
\]
\[
A_4 = \{ x \in B_1 : x \text{ is labelled } ④ \}
\]
\[
A_5 = \{ P \}
\]

Obviously, $B_1 = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$. We leave it as an exercise to the reader to check that by the labelling rules (and the labelling of $P$) we have:

- $\varphi_1[A_1] = A_1 \cup A_2 \cup A_5$.
- $\varphi_2[A_2] = A_1 \cup A_2 \cup A_3$.
- $\varphi_3[A_3] = A_3 \cup A_4$.
- $\varphi_4[A_4] = (A_3 \cup A_4) \setminus \{ 0 \}$, where $0$ denotes the origin.

Hence, we get that $A_1 \cong A_1 \cup A_2 \cup A_5 \cong A_2$, $A_3 \cong A_3 \cup A_4$, and $A_4 \cong (A_3 \cup A_4) \setminus \{ 0 \}$, and obviously we have $\{ P \} \cong \{ 0 \}$.

Now, with the two sets $A_1$ and $A_3$, as well as with the three sets $A_2$, $A_4$ and $A_5$, we can make a solid unit ball: Firstly, notice that $B_1 = \varphi_1[A_1] \cup \varphi_3[A_3]$. Secondly, let $\varphi$ be a translation which moves $P$ to the origin $0$. Then $B_1 = \varphi_2[A_2] \cup \varphi_4[A_4] \cup \varphi[A_5]$. Hence, we finally get

\[
B_1 \cong 5 B_1 \cup B_1.
\]

This result is optimal with respect to the number of pieces needed, in other words we have

\[
B_1 \not\cong 4 B_1 \cup B_1.
\]

To see this, assume towards a contradiction that there are distance-preserving (not necessarily orientation-preserving) transformations $\psi_1, \psi_2, \psi_3, \psi_4$ and a partition $B_1 = P_1 \cup P_2 \cup P_3 \cup P_4$ such that $B_1 = \psi_1[P_1] \cup \psi_2[P_2]$ and $B_1 = \psi_3[P_3] \cup \psi_4[P_4]$. Firstly notice that not all transformations $\psi_1, \psi_2, \psi_3, \psi_4$ could leave the origin fixed, for then one copy of $B_1$ would be without a centre. Now suppose for example that $\psi_4(0) \neq 0$. Then $S_2 \setminus \psi_4[B_1]$ (where $S_2$ denotes the surface of $B_1$) contains more than a hemisphere (i.e., more than half of $S_2$). In other words, $\psi_4[B_1] \cap S_2$, and in particular $\psi_4[P_4] \cap S_2$, is contained in less than a hemisphere. Since $\psi_4[P_4]$ must cover $S_2 \setminus \psi_4[P_4]$, it must cover more than a hemisphere, which is only possible if $\psi_4(0) = 0$ (otherwise, $\psi_4[P_4] \cup \psi_4[P_4]$ would not cover $S_2$). Thus, $P_3$ itself must cover more than a hemisphere, and consequently, $(P_1 \cup P_2) \cap S_2$ is contained in less than a hemisphere. Hence, $(\psi_1[P_1] \cup \psi_2[P_2]) \cap S_2$ is properly contained in $S_2$, and therefore $\psi_1[P_1] \cup \psi_2[P_2]$ cannot cover $S_2$. 

Robinson's decomposition
In 1924, Banach and Tarski proved in [2] that if $A$ and $A'$ are bounded subsets of Euclidean space of three or more dimensions and both sets have interior points, then $A$ and $A'$ are equidecomposable. In particular, for $A = B_1$ and $A' = B_1 \cup B_1$, $B_1 \cong B_1 \cup B_1$ (cf. [2, p. 262 (Lemme 22)]). However, no estimate was given for the number of pieces required to make two balls from one. Some years later, von Neumann [8, p. 77] stated without proof that nine pieces are sufficient, and about two decades later, Sierpiński improved von Neumann's result by showing that eight pieces are sufficient (cf. [13]). Finally, Robinson was able to show that in fact just five pieces are sufficient and that 5 is the smallest possible number of pieces, i.e., $B_1 \not\cong B_1 \cup B_1$.

The proof of $B_1 \cong B_1 \cup B_1$ given above is taken essentially from [10]. However, we have made a few modifications: For example, we have taken Sierpiński's construction given in [12] to obtain the four independent rotations $\varphi_1, \varphi_2, \varphi_3, \varphi_4$. Furthermore, we have replaced the parts in Robinson's proof which deal with products of relations with products of matrices, and introduced the trick with the trace in order to find fixed points in products of relations. Finally, we tried to visualise a few key steps in the proof by some figures.

The results of Banach and Tarski [2] — and indirectly also the other paradoxical decompositions of geometrical figures — were motivated by Hausdorff's decomposition of the sphere, given in [3] (see also [5, pp. 5–10] or [4, p. 469ff.]). The aim of Hausdorff's decomposition was to show that it is impossible to define a non-vanishing measure $\mu$ on $S_2$ which is defined for all subsets of $S_2$, is finitely additive (i.e., $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A$ and $B$ are disjoint), and has the property that congruent sets have the same measure.

Like Hartogs, also Hausdorff had to retire 1935 from his chair in Bonn and by October 1941 he was forced to wear the “yellow star”. Around the end of the year he was informed that he would be sent to Cologne — which he knew was just a preliminary to deportation to Poland — but managed to avoid being sent. Shortly later, in January 1942, he was informed again that he was to be interned now in Endenich, and together with his wife and his wife's sister, he committed suicide on 26 January.

**40. Further paradoxical decompositions.** In [8, p. 85f.] von Neumann introduced the following notion of decomposability: Let $A$ and $B$ be two subsets of a metric space $(X, d)$. $A$ is said to be metrically smaller than $B$ if there is a bijection $f : A \rightarrow B$ such that for any distinct points $x, y \in A$ we have $d(x, y) < d(f(x), f(y))$. Furthermore, $A$ is smaller by finite decomposition than $B$ if there is a positive integer $n$ and partitions $A = A_1 \cup \ldots \cup A_n$ and $B = B_1 \cup \ldots \cup B_n$ such that for all $1 \leq i \leq n$ we have that $A_i$ is metrically smaller than $B_i$. Now, von Neumann [8, p. 115f.] showed that every interval of the real line is smaller by finite decomposition than every other interval of the real line. About two decades later, Sierpiński [14] proved a 2-dimensional analogue by showing that every disc is smaller by finite decomposition than every other disc.

For the consequences of the paradoxical decompositions for Measure Theory and its connections with Group Theory, Geometry, and Logic, we refer the
reader to Wagon [18], and for some historical background see Wapner [19]. For other paradoxical decompositions see Laczkovich [7] or Sierpiński [15], and for a seemingly stronger notion of equidecomposability we refer the reader to Wilson [20].

41. Limits of decomposability. In 1923, Banach showed that there exists a finitely additive measure $m$ on $\mathbb{R}^2$, extending the Lebesgue measure $\mu$, such that $m$ is defined for all subsets of $\mathbb{R}^2$ and has the property that $m(A) = m(A')$ whenever $A \cong A'$ (see Banach [1, Théorème 1]). This implies that whenever $A$ and $A'$ are Lebesgue measurable subsets of $\mathbb{R}^2$ and $A \simeq A'$, then $\mu(A) = \mu(A')$ (see Banach and Tarski [2, Théorème 16]). In particular, the unit disc and two unit discs are not equidecomposable.

Neither Hausdorff’s partition of the sphere nor Robinson’s decomposition of the ball can be carried out without the aid of some form of the Axiom of Choice. The reason for this is that in the presence of inaccessible cardinals (cf. Chapter 15: Related Result 85), there exists a model of ZF in which every set of reals is Lebesgue measurable (see Solovay [17], and Shelah [11] or Raisonnier [9]).

42. Squaring the circle. As mentioned above, there is no 2-dimensional analogue of Robinson’s decomposition of the ball, i.e., there is no way of making two unit discs from one unit disc. However, Laczkovich [6] showed that a disc is equidecomposable — by translations only — with a square of the same area. The construction makes use of the Axiom of Choice and the figures are partitioned into about $10^{50}$ pieces.

References

2. Stefan Banach and Alfred Tarski, Sur la décomposition des ensembles de points en parties respectivement congruentes, *Fundamenta Mathematicae*, vol. 6 (1924), 244–277.


15. ______, *On the congruence of sets and their equivalence by finite decomposition*, [Lucknow University Studies, no. xx], The Lucknow University, Lucknow, 1954 (reprinted in [16]).


Models of Set Theory with Atoms

A musician regards consonances more highly than dissonances, so he composes principally with them. Nevertheless, it seems that he also values those sounds which are dissonant. Now intervals that are dissonant produce a sound that is disagreeable to the ear and render a composition harsh and without any sweetness. Therefore a musician must know them not only to avoid them where consonances are required, but to use them within the parts of a composition.

Gioseffo Zarlini
Le Istituzioni Harmoniche, 1558

In this chapter, we shall construct various models of Set Theory in which the Axiom of Choice fails. In particular, we shall construct a model in which $C(\aleph_0, 2)$ fails, and another one in which a cardinal $m$ exists such that $m^2 < |m|^2$. These somewhat strange models are constructed in a similar way to models of ZF (see the cumulative hierarchy introduced in Chapter 3). However, instead of starting with the empty set (in order to build the cumulative hierarchy) we start with a set of atoms and define a certain group $\mathcal{G}$ of permutations of these atoms. Roughly speaking, a set $x$ is in the model if $x$ is “stable” under certain subgroups $\mathcal{H} \subseteq \mathcal{G}$ (i.e., for all permutations $\pi \in \mathcal{H}$, $\pi x = x$). In this way we can make sure that some particular sets (e.g., choice functions for a given family in the model) do not belong to the model. Unfortunately, since we have to introduce atoms to construct these models, we do not get models of ZF; however, using the \textsc{Jech-Sochor Embedding Theorem 17.2}, we can embed arbitrarily large fragments of these models into models of ZF, which is sufficient for our purposes.
Permutation Models

In this section we shall give the definition of so-called permutation models, but first have to say a few words about Set Theory with atoms, denoted ZFA. Set theory with atoms is characterised by the fact that it admits so-called atoms or urelements.

Atoms are objects which do not have any elements but are distinct from the empty set. The collection of atoms — assumed to be a set — is usually denoted by $A$, and we add the constant symbol $A$ to the language of Set Theory. Thus, the language of Set Theory with atoms consists of the relation symbol “$\in$” and the constant symbol “$A$”, i.e., $\mathcal{L}_{ZFA} = \{\in, A\}$.

In ZFA we have two types of objects, namely sets and atoms, and since atoms behave slightly different than sets (e.g., they do not contain elements but are different from $\emptyset$), we have to add a new axiom for atoms (i.e., an axiom for the symbol $A$) and have to modify the Axiom of Empty Set as well as the Axiom of Extensionality.

**Axiom of Empty Set (for ZFA):**

$$\exists x (x \notin A \land \forall z (z \notin x))$$

**Axiom of Extensionality (for ZFA):**

$$\forall x \forall y ((x \notin A \land y \notin A) \rightarrow \forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

Roughly speaking, any two objects, which are not atoms but have the same elements, are equal. Notice that the Axiom of Extensionality implies that the empty set is unique, i.e., $\emptyset$ is the only object that has no elements but does not belong to $A$.

**Axiom of Atoms:**

$$\forall x (x \in A \leftrightarrow (x \neq \emptyset \land \exists \exists z (z \in x)))$$

In other words, an object is an atom if and only if it contains no elements but is different from the set $\emptyset$. For an alternative definition of atoms see Related Result 43.

It is time to mention that if $\forall z \neg \varphi(z)$, then we stipulate $\{z : \varphi(z)\} := \emptyset$ (not some atom, which would also be possible). For example, if $x$ and $y$ do not have any elements in common, i.e., $\forall z (z \in x \land z \notin y)$, then $x \cap y = \emptyset$. Notice that with this convention we do not have to modify the Axiom of Extensionality for ZFA.

The development of the theory ZFA is very much the same as that of ZF (except for the definition of ordinals, where we have to require that an ordinal does not have atoms among its elements). Let $S$ be a set. Then by transfinite recursion on $\alpha \in \Omega$ we can define $\mathcal{P}^\alpha(S)$ as follows: $\mathcal{P}^0(S) := S$, $\mathcal{P}^{\alpha+1}(S) := \mathcal{P}^\alpha(S) \cup \mathcal{P}(\mathcal{P}^\alpha(S))$ and $\mathcal{P}^\alpha(S) := \bigcup_{\beta \leq \alpha} \mathcal{P}^\beta(S)$ when $\alpha$ is
a limit ordinal. Furthermore, let \( \mathcal{P}^\infty(S) := \bigcup_{\alpha \in \Omega} \mathcal{P}^\alpha(S) \). If \( \mathcal{M} \) is a model of ZFA and \( A \) is the set of atoms of \( \mathcal{M} \), then \( \mathcal{M} = \mathcal{P}^\infty(A) \). The class \( \mathcal{V} := \mathcal{P}^\infty(\emptyset) \), which is a subclass of \( \mathcal{M} \), is a model of ZF and is called the kernel. Notice that all ordinals belong to the kernel.

Now, the underlying idea of permutation models, which are models of ZFA, is the fact that the axioms of ZFA do not distinguish between the atoms, and so a permutation of the set of atoms induces an automorphism of the universe.

Let \( A \) be a set of atoms and let \( \mathcal{M} = \mathcal{P}^\infty(A) \) be a model of ZFA. Furthermore, in \( \mathcal{M} \), let \( \mathcal{G} \) be a group of permutations (or automorphisms) of \( A \), where a permutation of \( A \) is a one-to-one mapping from \( A \) onto \( A \). We say that a set \( \mathcal{F} \) of subgroups of \( \mathcal{G} \) is a normal filter on \( \mathcal{G} \) if for all subgroups \( H, K \) of \( \mathcal{G} \) we have:

(A) \( \mathcal{G} \in \mathcal{F} \)
(B) if \( H \in \mathcal{F} \) and \( H \subseteq K \), then \( K \in \mathcal{F} \)
(C) if \( H \in \mathcal{F} \) and \( K \in \mathcal{F} \), then \( H \cap K \in \mathcal{F} \)
(D) if \( \pi \in \mathcal{G} \) and \( H \in \mathcal{F} \), then \( \pi H \pi^{-1} \in \mathcal{F} \)
(E) for each \( a \in A \), \( \{ \pi \in \mathcal{G} : \pi a = a \} \in \mathcal{F} \)

For every set \( x \in \mathcal{M} \) there is a least ordinal \( \alpha \) such that \( x \in \mathcal{P}^\alpha(A) \). So, by induction on the ordinals, for every \( \pi \in \mathcal{G} \) and for every set \( x \in \mathcal{M} \) we can define \( \pi x \) by stipulating

\[
\pi x = \begin{cases} 
\emptyset & \text{if } x = \emptyset, \\
\pi x & \text{if } x \in A, \\
\{ \pi y : y \in x \} & \text{otherwise.}
\end{cases}
\]

Notice that for all \( x, y \in \mathcal{M} \) and every \( \pi \in \mathcal{G} \) we have \( \pi x = y \iff x = \pi^{-1} y \) and \( x \in y \iff \pi x \in \pi y \), which leads to the following definition: A bijective class function \( F : \mathcal{M} \to \mathcal{M} \) is called an \( \mathcal{G} \)-automorphism of \( \mathcal{M} \) if for all \( x, y \in \mathcal{M} \) we have \( x \in y \iff F(x) \in F(y) \). In particular, \( \pi : \mathcal{M} \to \mathcal{M} \) is an \( \mathcal{G} \)-automorphism of \( \mathcal{M} \).

For \( x \in \mathcal{M} \), the symmetry group of \( x \), denoted \( \operatorname{sym}_G(x) \), is the group of all permutations in \( \mathcal{G} \) which map \( x \) to \( x \), in other words

\[
\operatorname{sym}_G(x) = \{ \pi \in \mathcal{G} : \pi x = x \}.
\]

A set \( x \) is said to be symmetric (with respect to a normal filter \( \mathcal{F} \)) if the symmetry group of \( x \) belongs to \( \mathcal{F} \), i.e., \( \operatorname{sym}_G(x) \in \mathcal{F} \). By (E) we have that every atom \( a \in A \) is symmetric. A set \( x \) is called hereditarily symmetric if \( x \) as well as each element of its transitive closure is symmetric. Notice that for all \( x \in \mathcal{M} \) and every \( \pi \in \mathcal{G} \), \( x \) is hereditarily symmetric iff \( \pi x \) is hereditarily symmetric.

Let \( \mathcal{V} \subseteq \mathcal{M} \) be the class of all hereditarily symmetric sets. Then \( \mathcal{V} \) is a transitive model of ZFA and we call \( \mathcal{V} \) a permutation model. Because \( A \),
as well as every \( a \in A \), is symmetric, we get that the set of atoms \( A \) belongs to \( \mathcal{V} \).

Because \( \emptyset \) is hereditarily symmetric and for all ordinals \( \alpha \) the set \( \mathcal{P}^\alpha(\emptyset) \) is hereditarily symmetric too, the kernel \( \mathcal{N} = \mathcal{P}^{\infty}(\emptyset) \) is a subclass of \( \mathcal{V} \). Notice that every \( \pi \in \mathcal{G} \) which is not the identity mapping is a non-trivial \( \in \)-automorphism of \( \mathcal{V} \). On the other hand, all \( \in \)-automorphisms of models of ZF are trivial. In particular, by induction on \( \alpha \) one easily verifies the following

**Fact 7.1.** For any set \( x \in \hat{\mathcal{V}} \) and any \( \pi \in \mathcal{G} \) we have \( \pi x = x \).

Since the atoms \( a \in A \) do not contain any elements, but are distinct from the empty set, the permutation models are not models of ZF. However, by the **Jech-Sochor Embedding Theorem 17.2** one can embed arbitrarily large fragments of a permutation model into a well-founded model of ZF.

Most of the well-known permutation models are of the following simple type: Let \( \mathcal{G} \) be a group of permutations of \( A \). A family \( I \) of subsets of \( A \), for example \( I = \text{fin}(A) \), is a **normal ideal** if for all subsets \( E, F \) of \( A \) we have:

(a) \( \emptyset \in I \)
(b) if \( E \in I \) and \( F \subseteq E \), then \( F \in I \)
(c) if \( E \in I \) and \( F \subseteq I \), then \( E \cup F \in I \)
(d) if \( \pi \in \mathcal{G} \) and \( E \in I \), then \( \pi E \in I \)
(e) for each \( a \in A \), \( \{a\} \in I \)

For each set \( S \subseteq A \), let

\[
\text{fix}_\mathcal{G}(S) = \{ \pi \in \mathcal{G} : \pi a = a \text{ for all } a \in S \}
\]

and let \( \mathcal{F} \) be the filter on \( \mathcal{G} \) generated by the subgroups \( \{\text{fix}_\mathcal{G}(E) : E \in I\} \).

Then \( \mathcal{F} \) is a normal filter. Furthermore, \( x \) is symmetric if and only if there exists a set of atoms \( E_x \in I \) such that

\[
\text{fix}_\mathcal{G}(E_x) \subseteq \text{sym}_\mathcal{G}(x)
\]

where \( E_x \) is called a **support** of \( x \). Notice that if \( E_x \) is a support of \( x \) and \( E_x \subseteq F_x \in I \), then \( F_x \) is a support of \( x \) as well.

Below, we give some relationships which are consistent with ZF between the cardinals defined in Chapter 4. We will do this by investigating the relations between certain sets in some permutation models. The general construction will be as follows: Let \( \mathcal{V} \) be a permutation model with a set of atoms \( A \) and let \( m \) be a set in \( \mathcal{V} \). Let \( \mathcal{C}(m) := \{ x \in \mathcal{V} : \mathcal{V} \models |x| = |m| \} \). Then \( \mathcal{C}(m) \) is in general a class in \( \mathcal{V} \). The cardinality of \( m \) in the model \( \mathcal{V} \) (denoted by \( m \)) is defined by \( m := \mathcal{C}(m) \cap \mathcal{P}^\alpha(A) \cap \mathcal{V} \), where \( \alpha \) is the smallest ordinal such that \( \mathcal{C}(m) \cap \mathcal{P}^\alpha(A) \cap \mathcal{V} \neq \emptyset \).

If \( m \) is a set in a permutation model \( \mathcal{V} \) and we have for example \( \mathcal{V} \models |\text{seq}(m)| < |\text{fin}(m)| \), and therefore \( \mathcal{V} \models \text{seq}(m) < \text{fin}(m) \), then, by the **Jech-Sochor Embedding Theorem 17.2**, there exist a well-founded model \( \mathcal{V} \).
of ZF and a set \( \hat{m} \) such that \( \hat{V} \models |\text{seq}(\hat{m})| < |\text{fin}(\hat{m})| \) and consequently \( \hat{V} \models \text{seq}(\hat{m}) < \text{fin}(\hat{m}) \), where \( \hat{m} \) and \( \hat{n} \) are the cardinalities of the sets \( \hat{m} \) and \( \hat{n} \) respectively. In fact, the Jech-Sochor Embedding Theorem 17.2 enables us to translate every relation between sets in a permutation model to a well-founded model. Hence, in order to prove that a relation between some cardinals is consistent with ZF, it is enough to find a permutation model in which the desired relation holds between the corresponding sets. Below we shall make use of this method without explicitly mentioning it.

The Basic Fraenkel Model

In this section we shall present a simple example of a permutation model in which the Axiom of Choice fails.

Let \( A \) be a countable infinite set (the atoms), let \( G \) be the group of all permutations of \( A \), and let \( I_{\text{fin}} \) be the set of all finite subsets of \( A \). Obviously, \( I_{\text{fin}} \) is a normal ideal and the filter derived from \( I_{\text{fin}} \) as described above is a normal filter.

Let \( \mathcal{V}_{F_0} \) (\( F \) for Fraenkel) be the corresponding permutation model, the so-called basic Fraenkel model. Note that a set \( x \) belongs to \( \mathcal{V}_{F_0} \), if and only if \( x \) is symmetric and each \( y \in x \) belongs to \( \mathcal{V}_{F_0} \), too.

Before we start with some results involving subsets of \( A \), let us recall that a set \( S \) is transfinite if \( \aleph_0 \leq |S| \); otherwise \( S \) is called D-finite.

**Lemma 7.2.** Let \( E \in I_{\text{fin}} \); then each \( S \subseteq A \) with support \( E \) is either finite or co-finite, i.e., \( A \setminus S \) is finite. Furthermore, if \( S \) is finite, then \( S \subseteq E \), and if \( S \) is co-finite, then \( (A \setminus S) \subseteq E \).

**Proof.** Let \( S \subseteq A \) with support \( E \). Because \( E \) is a support of \( S \), for all \( \pi \in \text{fix}(E) \) and every \( a \in A \) we have \( \pi a \in S \) iff \( a \in S \). If \( S \) contains an element \( a_0 \) of \( A \setminus E \), then it contains them all, since permutations in \( \text{fix}(E) \) can send \( a_0 \) to any other element of \( A \setminus E \). Thus, either \( S \subseteq E \) or \( (A \setminus S) \subseteq E \).

As a consequence we get the following result (cf. Chapter 4 \{Related Result 18\}: Let \( m \) denote the cardinality of the set of atoms of the basic Fraenkel model. Then

\[
\mathcal{V}_{F_0} \models (\beth^m)^{\aleph_0} = \beth^{\text{fin}(m)}.
\]

Indeed, every subset of \( A \) in \( \mathcal{V}_{F_0} \) is either finite or co-finite, and therefore, \( \beth^m = \beth^{\text{fin}(m)} \). Hence, \( (\beth^m)^{\aleph_0} = (\beth^{\text{fin}(m)})^{\aleph_0} \) and by LÄUCHLI’s Lemma 4.27 this is equal to \( \beth^{\text{fin}(m)} \).

**Proposition 7.3.** Let \( A \) be the set of atoms of the basic Fraenkel model and let \( m \) denote its cardinality. Then \( \mathcal{V}_{F_0} \not\models \aleph_0 \leq m \); in particular, in \( \mathcal{V}_{F_0} \) there are infinite D-finite sets. In particular, it is not provable in ZF that every D-finite set is finite.
Proof. If there is a one-to-one mapping \( f : \omega \to A \), then the set \( S = \{ f(2n) : n \in \omega \} \) would be an infinite, co-infinite set of atoms, which is a contradiction to Lemma 7.2.

We have seen in Chapter 4 that for every infinite cardinal \( m \), \( 2^{\aleph_0} \leq 2^{\text{fin}(m)} \). In contrast to this fact, the following result shows that in the model \( \mathcal{V}_{F_0} \), the power set of an infinite set can be \( D \)-finite, which shows that even for infinite cardinals \( m \), the statement \( \aleph_0 \leq 2^m \) is in general not provable in \( \mathcal{ZF} \).

**Proposition 7.4.** Let \( A \) be the set of atoms of the basic Fraenkel model and let \( m \) denote its cardinality. Then \( \mathcal{V}_{F_0} \models \aleph_0 \not\leq 2^m \). In particular, it is not provable in \( \mathcal{ZF} \) that the power set of an infinite set is transfinite.

Proof. Assume towards a contradiction that there exists a one-to-one function \( f : \omega \to \mathcal{P}(A) \) which belongs to \( \mathcal{V}_{F_0} \). Then, because \( f \) is symmetric, there is a finite set \( E_f \subseteq A \) (a support of \( f \)) such that \( \text{fix}_f(E_f) \subseteq \text{sym}_f(f) \). Now, let \( n \in \omega \) be such that \( \text{fix}_f(E_f) \not\subseteq \text{sym}_f(f(n)) \) (such an \( n \) exists because, by Lemma 7.2, \( E_f \) supports only finitely many subsets of \( A \)). Further, let \( \pi \in \text{fix}_f(E_f) \) be such that \( \pi f(n) \neq f(n) \). By Fact 7.1 we get that \( \pi n = n \), and therefore, \( f(\pi n) = f(n) \). So, \( E_f \) cannot be a support of \( f \) which contradicts the choice of \( E_f \) and shows that a one-to-one function from \( \omega \) into \( \mathcal{P}(A) \) cannot belong to the model \( \mathcal{V}_{F_0} \).

By Proposition 4.22 we know that if \( 2^m = n \cdot \text{fin}(m) \) for some \( n \in \omega \), then \( n = 2^k \) for some \( k \in \omega \). The next result shows that also a kind of converse is true:

**Proposition 7.5.** For every number \( n \) of the form \( n = 2^k \), where \( k \in \omega \), there is a set \( A_k \) in \( \mathcal{V}_{F_0} \) such that \( \mathcal{V}_{F_0} \models |\mathcal{P}(A_k)| = |n \times \text{fin}(A_k)| \).

Proof. If \( n = 2^k \), then the statement is true for every finite set \( A_k \) (in every model of Set Theory).

Let \( k \in \omega \setminus \{0\} \) and let \( n = 2^k \). Further, let \( A \) be the set of atoms of \( \mathcal{V}_{F_0} \) and let \( A_k = k \times A \). By Lemma 7.2 we know that every subset of \( A \) (in \( \mathcal{V}_{F_0} \)) is either finite or co-finite and therefore \( |\mathcal{P}(A)| = 2 \cdot |\text{fin}(A)| \). Thus, in \( \mathcal{V}_{F_0} \) we have \( |\mathcal{P}(A_k)| = |\mathcal{P}(k \times A)| = |\mathcal{P}(A)|^k = (2 \cdot |\text{fin}(A)|)^k = 2^k \times |\text{fin}(A)|^k \). And therefore \( \mathcal{V}_{F_0} \models |\mathcal{P}(A_k)| = |n \times \text{fin}(A_k)| \).

The Second Fraenkel Model

The set of atoms of the second Fraenkel model consists of countably many mutually disjoint 2-element sets:

\[
A = \bigcup_{n \in \omega} P_n, \quad \text{where } P_n = \{a_n, b_n\} \text{ (for } n \in \omega\)
\]
The second Fraenkel model

Let $\mathcal{G}$ be the group of those permutations of $A$ which preserve the pairs $P_n$, i.e., $\pi(\{a_n, b_n\}) = \{a_n, b_n\}$ (for each $\pi \in \mathcal{G}$ and every $n \in \omega$). Further, let $I_{\text{fin}}$ be the set of all finite subsets of $A$. Then $I_{\text{fin}}$ is a normal ideal and the filter generated by $I_{\text{fin}}$ is a normal filter.

Let $V_{F_2}$ be the corresponding permutation model, called the second Fraenkel model. The following theorem summarises the main features of this model.

**Theorem 7.6.** (a) For each $n \in \omega$ the set $P_n$ belongs to $V_{F_2}$.
(b) The sequence $\langle P_n : n \in \omega \rangle$ belongs to $V_{F_2}$. In particular, the set of pairs $\{P_n : n \in \omega\}$ is countable in $V_{F_2}$.
(c) There is no choice function on $\{P_n : n \in \omega\}$. In particular, $C(\aleph_0, 2)$ fails in $V_{F_2}$ which shows that $ZF \not\equiv C(\aleph_0, 2)$.

**Proof.** (a) For each $\pi \in \mathcal{G}$ and for every $n \in \omega$ we have $\pi P_n = P_n$, which implies that every $P_n$ is symmetric.
(b) For each $\pi \in \mathcal{G}$ we have $\pi(\langle P_n : n \in \omega \rangle) = \langle \pi P_n : n \in \omega \rangle = \langle P_n : n \in \omega \rangle$, and therefore by (a), $\langle P_n : n \in \omega \rangle$ is hereditarily symmetric.
(c) Assume that there is a choice function $f$ on $\{P_n : n \in \omega\}$ which belongs to $V_{F_2}$. The choice function $f$ would be a function from $\omega$ into $\bigcup \{P_n : n \in \omega\}$ such that $f(n) \in P_n$ (for every $n \in \omega$). Let $\{a_0, b_0, \ldots, a_k, b_k\}$ be a support of $f$ and let $\pi \in \text{fix}_\mathcal{G} \{\{a_0, b_0, \ldots, a_k, b_k\}\}$ be such that $\pi a_{k+1} = b_{k+1}$. Then $\pi(k + 1) = k + 1$, but $\pi(f(k + 1)) \neq f(k + 1)$, which implies that $\pi f \neq f$ and contradicts the fact that $\{a_0, b_0, \ldots, a_k, b_k\}$ is a support of $f$.

We leave it as an exercise to the reader to show that $C_2$, which is a more general choice principle than $C(\aleph_0, 2)$, already fails in $V_{F_0}$.

The following result shows that in $V_{F_2}$, König’s Lemma fails even for binary trees.

**Proposition 7.7.** In $V_{F_2}$ there exists an infinite binary tree which does not have an infinite branch.

**Proof.** We construct the binary tree $T = (V, E)$ with vertex set $V$ and edge set $E$ as follows: For $n \in \omega$ let $V_n = \{s \in {}^n A : \forall i \in n \exists j \in P_i\}$ and let $V = \bigcup_{n \in \omega} V_n$. Further, let $(s, t) \in E$ iff for some $n \in \omega$, $s \in V_n$, $t \in V_{n+1}$, and $t|_n = s$. It is easily verified that $T$ is an infinite tree and since every vertex $s \in V$ has exactly two successors, namely $\overline{s}a_n$ and $\overline{s}b_n$, where $s \in V_n$ and $\overline{s}x$ denotes the concatenation of the sequence $s$ and the element $x$, $T$ is even a binary tree. On the other hand, an infinite branch through $T$ would yield a choice function on $\{P_n : n \in \omega\}$, a contradiction to Theorem 7.6(c).

In a similar way one can show that Ramsey’s original theorem fails in $V_{F_2}$:

**Proposition 7.8.** In $V_{F_2}$ there exist an infinite set $S$ and a 2-colouring of $[S]^2$ such that no infinite subset of $S$ is homogeneous.
\textbf{Proof.} Let $S$ be the set of atoms of $\mathcal{V}_{F_2}$ and colour a 2-element set of atoms \{a, b\} red, if \{a, b\} = $P_n$ for some $n \in \omega$; otherwise, colour it blue. We leave it as an exercise to the reader to show that no infinite homogeneous set belongs to $\mathcal{V}_{F_2}$.

The last result of this section is a kind of infinite version of \textbf{Proposition 7.5}.

\textbf{Proposition 7.9.} In $\mathcal{V}_{F_2}$, let $m$ denote the cardinality of the set of atoms. Then $\mathcal{V}_{F_2} \models 2^m = 2^{\aleph_0} \cdot \text{fin}(m)$.

\textbf{Proof.} By the \textsc{Cantor-Bernstein Theorem} \ref{3.17} it is enough to find two one-to-one mappings $f : \mathcal{P}(A) \to 2 \times \text{fin}(A)$ and $g : 2 \times \text{fin}(A) \to \mathcal{P}(A)$. For every $n \in \omega$ let $U_n = \bigcup_{i \in n} P_i$.

For $S \subseteq A$ let $m = \bigcup \{n+1 : |P_n \cap S| = 1\}$. Then $F_S = S \cap U_m$ is finite and for every $n > m$ we have either $P_n \subseteq S$ or $P_n \cap S = \emptyset$. Now define $\chi_S : \omega \to 2$ by stipulating $\chi_S(n) = 0$ iff $P_{n+m+1} \cap S = \emptyset$, and define $f(S) := \langle \chi_S, F_S \rangle$. It is easily verified that the function $f$ is one-to-one.

Let $\langle \chi, F \rangle \in 2 \times \text{fin}(A)$ and define again $m = \bigcup \{n+1 : |P_n \cap F| = 1\}$. Then $F_0 = F \cap U_m$ and $F_1 = F \setminus F_0$ are finite. Further, let

$$S_{\chi,F} = F_0 \cup \bigcup \{P_{2n} : P_n \subseteq F_1\} \cup \bigcup \{P_{2n+m+1} : \chi(n) = 1\} \subseteq A$$

and define $g(\langle \chi, F \rangle) := S_{\chi,F}$. It is again easy to check that the function $g$ is one-to-one.

\textbf{The Ordered Mostowski Model}

The set of atoms $A$ of the ordered Mostowski model consists of an infinite countable set together with an ordering “$<^M$” such that $A$ is densely ordered and does not have a smallest or greatest element, \textit{i.e.}, $A$ is order-isomorphic to the rational numbers. Let $G$ be the group of all order-preserving permutations of $A$ and let $\text{fin}$ be the ideal of the finite subsets of $A$. Then again, $\text{fin}$ is a normal ideal and the filter generated by $\text{fin}$ is a normal filter.

Let $\mathcal{V}_M$ (M for Mostowski) be the corresponding permutation model, called the \textbf{ordered Mostowski model}.

First let us show that the binary relation “$<^M$” belongs to the model $\mathcal{V}_M$. In other words, for any two distinct atoms $a_1$ and $a_2$ we can decide in $\mathcal{V}_M$ whether we have $a_1 <^M a_2$ or $a_2 <^M a_1$.

\textbf{Lemma 7.10.} The set $R_\prec = \{\langle a_1, a_2 \rangle : a_1 <^M a_2\} \subseteq A \times A$ belongs to $\mathcal{V}_M$.

\textbf{Proof.} If $a_1 <^M a_2$, then $\pi a_1 <^M \pi a_2$ (for any $\pi \in G$), and therefore, $\langle a_1, a_2 \rangle \in R_\prec$ iff $\langle \pi a_1, \pi a_2 \rangle \in R_\prec$, which implies that $\text{sym}_G(R_\prec) = G$.

Because by definition all sets in the ordered Mostowski model must be symmetric, each set in $\mathcal{V}_M$ has a finite support. Moreover, each set in $\mathcal{V}_M$ has a unique least support:
Lemma 7.11. (a) If \( E_1 \) and \( E_2 \) are supports of \( x \), then also \( E = E_1 \cap E_2 \) is a support of \( x \).

(b) Every set \( x \in V_M \) has a least support.

(c) The class of all pairs \((x, E)\), where \( x \in V_M \) and \( E \) is the least support of \( x \), is symmetric.

Proof. (a) Let \( E_1 \) and \( E_2 \) be two finite supports of the set \( x \in V_M \) and let \( E = E_1 \cap E_2 \). Notice that for every \( \pi \in \text{fix}_G(E) \) there are infinitely many \( \rho_1, \ldots, \rho_n \in \text{fix}_G(E_1) \) and \( \sigma_1, \ldots, \sigma_n \in \text{fix}_G(E_2) \) such that \( \pi = \rho_1 \sigma_1 \cdots \rho_n \sigma_n \). To see this, it might be better to draw a picture than to prove it formally (e.g., let \( E_1 = \{a_0, a_1, a_2\} \) and \( E_2 = \{b_0, b_1, b_2\} \) be such that \( a_0 = b_0 <^G a_1 <^G b_1 <^G a_2 <^G b_2 \), and let \( \pi \in \text{fix}_G(\{a_0\}) \) be such that \( b_2 <^G \pi c \) for some \( a_0 <^G c <^G b_1 \). Since \( \rho_i x = \sigma_i x \) (for all \( 1 \leq i \leq n \)) we have

\[
\pi x = \rho_1 \sigma_1 \cdots \rho_n \sigma_n x = \rho_1 \sigma_1 \cdots \sigma_{n-1} \rho_n x = \cdots = \rho_1 x = x
\]

for all \( \pi \in \text{fix}_G(E) \), which shows that \( \pi \in \text{sym}_G(x) \). Hence, \( \text{fix}_G(E) \subseteq \text{sym}_G(x) \) which implies that \( E \) is a support of \( x \).

(b) Let \( E_0 \) be a support of \( x \). The least support of \( x \) is the intersection of all supports of \( x \), which are subsets of \( E_0 \). Since there are only finitely many of such supports, by (a), the intersection is a support of \( x \).

(c) Let \( x \in V_M \) and let \( E \) be the least support of \( x \). If \( \pi \in \mathcal{U} \), then \( \text{fix}_G(\pi E) = \pi \cdot \text{fix}_G(E) \cdot \pi^{-1} \) and \( \text{sym}_G(\pi x) = \pi \cdot \text{sym}_G(x) \cdot \pi^{-1} \), and thus, if \( E \) is a support of \( x \), then \( \pi E \) is a support of \( \pi x \).

\( \square \)

For every finite set \( E \subseteq A \), one can give a complete description of the subsets of \( A \) with support \( E \), which leads to the following

Lemma 7.12. If \( E \subseteq A \) is a finite set of cardinality \( n \), then there are \( 2^{2n+1} \) sets \( S \subseteq A \) in \( V_M \) such that \( E \) is a support of \( S \).

Proof. Let \( E = \{a_1, \ldots, a_n\} \) be such that \( a_1 <^M \cdots <^M a_n \). Assume that \( E \) is a support of the set \( S \subseteq A \). If there is an \( s_0 \in S \) such that \( a_i <^M s_0 <^M a_{i+1} \) (for some \( 1 \leq i < n \)), then \( \{s \in A : a_i <^M s <^M a_{i+1}\} \subseteq S \). To see this, notice that for every \( s \) with \( a_i <^M s <^M a_{i+1} \) there is a \( \pi \in \text{fix}_G(E) \) such that \( \pi s_0 = s \). Similarly, if there is an \( s \in S \) such that \( s <^M a_1 \) (or \( a_n <^M s \)), then \( \{s \in A : s <^M a_1\} \subseteq S \) (or \( \{s \in A : a_n <^M s\} \subseteq S \)). Now, there are \( n+1 \) such intervals and every interval is entirely contained in \( S \) or disjoint from \( S \). Further, for each \( 1 \leq i \leq n \), either \( a_i \in S \) or \( a_i \notin S \). Hence, there are \( 2^{2n+1} \) different subsets of \( A \) which have \( E \) as a support.

\( \square \)

Since the set of atoms in the ordered Mostowski model is infinite, the following result implies that the Axiom of Choice fails in \( V_M \) (compare this result with Proposition 7.4).

Lemma 7.13. Let \( A \) be the set of atoms of the ordered Mostowski model and let \( m \) denote its cardinality. Then \( V_M \models \aleph_0 \not\leq 2^m \).
Proof. We have to show that there is no one-to-one mapping \( f : \omega \to \mathcal{P}(A) \). Now, if a finite set \( E \subseteq A \) is a support of \( f \), then \( E \) supports each of the infinitely many distinct sets \( f(n) \) \( (n \in \omega) \), because all permutations fix each \( n \in \omega \). On the other hand, by Lemma 7.12, a finite set \( E \subseteq A \) can support just finitely many sets.

By Theorem 4.21, for every infinite cardinal \( m \) we have \( \text{fin}(m) < 2^m \). In contrast to this result we show now that \( \mathcal{V}_M \models 2^m \leq \text{fin}(m) \), where \( m \) denotes the cardinality of the set of atoms of \( \mathcal{V}_M \). As a consequence we get by Fact 4.8 that \( 2^{2^m} \leq \text{fin}(m) \), which implies by the Cantor-Bernstein Theorem 3.17 that \( \mathcal{V}_M \models 2^{2^m} = \text{fin}(m) \).

Proposition 7.14. Let \( A \) be the set of atoms of the ordered Mostowski model. Then in \( \mathcal{V}_M \) there is a surjection from \( \text{fin}(A) \) onto \( \mathcal{P}(A) \). Thus, it is consistent with \( \text{ZF} \) that there are infinite cardinals \( m \) such that \( 2^m < \text{fin}(m) \), even though \( \text{fin}(m) < 2^m \) is provable in \( \text{ZF} \) for every infinite cardinal \( m \).

Proof. The key idea in order to construct a surjective function \( g : \text{fin}(A) \to \mathcal{P}(A) \) is to define an ordering of the subsets of \( A \) sharing a given finite support. For \( E = \{a_1 <^m \ldots <^m a_n\} \in \text{fin}(A) \) let \( I_0 = \{a \in A : a <^m a_1\} \), \( I_n = \{a \in A : a_n <^m a\} \), and \( I_i = \{a \in A : a_i <^m a <^m a_{i+1}\} \) for \( 1 \leq i \leq n - 1 \). For every function \( \chi \in 2^{n+1} \) we assign a set \( S_\chi \in \mathcal{P}(A) \) by

\[
S_\chi = \bigcup_{\chi(2i)=1} I_i \cup \{a_i : \chi(2i-1) = 1\}.
\]

Then for every \( \chi \in 2^{n+1} \), \( E \) is a support of \( S_\chi \) and for every \( S_0 \subseteq A \) such that \( E \) is a support of \( S_0 \) there is a \( \chi_0 \in 2^{n+1} \) such that \( S_0 = S_{\chi_0} \) (this follows from Lemma 7.12).

We now consider for a moment the set \( 2^{n+2} \). Let \( "<" \) be the lexicographic ordering on \( 2^{n+2} \), i.e., \( \xi < \xi' \) if there is a \( j \in 2n + 2 \) such that \( \xi(j) < \xi'(j) \), but for all \( i < j \) we have \( \xi(i) = \xi'(i) \). For \( \xi \in 2^{n+2} \) let \( \xi \in 2^{n+2} \) be such that for all \( i \in 2n + 2 \), \( \xi(i) := 1 - \xi(i) \). We define the function \( \mu : 2^{n+2} \to 2^{n+2} \) by stipulating

\[
\mu(\xi) = \begin{cases} 
\xi & \text{if } \xi < \xi', \\
\xi & \text{otherwise},
\end{cases}
\]

in other words, \( \mu(\xi) \) is \( \xi \) or \( \xi' \), whichever begins with 0.

Let us turn back to the set \( 2^{n+1} \). For \( \chi \in 2^{n+1} \) let \( \chi^+ := \chi \cup \{0\} \). Notice that \( \chi^+ \in 2^{n+2} \). We define the ordering \( "<_n" \) on \( 2^{n+1} \) by stipulating

\[
\chi_0 <_n \chi_1 \iff \mu(\chi_0) < \mu(\chi_1).
\]

Now, we are ready to define a surjection from \( \text{fin}(A) \) onto \( \mathcal{P}(A) \). For this, consider the following function:

\[
g : \text{fin}(A) \to \mathcal{P}(A) \\
E \mapsto S_{\chi_{\mu(\xi)}}
\]
The Prime Ideal Theorem revisited

where for $|E| = n$, $\chi_n^k$ denotes the $n^{th}$ function of $2^{n+1}2$ with respect to the ordering $\prec_n$.

By construction, for every set $S_0 \in \mathcal{P}(A)$ there is a finite set $E$ such that $E$ is a support of $S_0$ and $S_0 = S_{\chi_n^k}$. Indeed, let $E_0$ be the least support of $S_0$. Then there is an $n \in \omega$ such that $S_0 = S_{\chi_n^k}$. By the properties of the ordering $\prec_n$, $n \geq |E_0|$ and we leave it as an exercise to show that $E_0$ can be extended to a finite set $E$ such that $|E| = n$ and $S_{\chi_n^k} = S_0$. Hence, the mapping $g$ is surjective as required.

Proposition 7.15. Let $m$ denote the cardinality of the set of atoms of the ordered Mostowski model. Then

$$\forall n \in \omega. n \cdot \text{fin}(m) < 2^m < \aleph_0 \cdot \text{fin}(m)$$

for every $n \in \omega$.

Proof (Sketch). $2^m \leq \aleph_0 \cdot \text{fin}(m)$: For $S \subseteq A$ let $E$ be the least support of $S$, let $n = |E|$, and let $k \in \omega$ be such that $S = S_{\chi_k}$, where $\chi_k$ denotes the $k^{th}$ function of $2^{n+1}2$ with respect to the ordering $\prec_n$ defined above. Then the mapping $S \mapsto (k, S_{\chi_k})$ is an injective function from $\mathcal{P}(A)$ into $\omega \times \text{fin}(A)$. $2^m \neq \aleph_0 \cdot \text{fin}(m)$: This is an immediate consequence of Lemma 7.13.

$n \cdot \text{fin}(m) \leq 2^m$: For $j \in n$ and $E \in \text{fin}(A)$ large enough we can define $S_j,E$ as the $j^{th}$ set which has $E$ as its least support. For $E \in \text{fin}(A)$ which are not large enough to allow such an encoding, we have to work with a large enough auxiliary set $E_0$ and then do some encoding for example on $E \cup E_0$.

$n \cdot \text{fin}(m) \neq 2^m$: Assume towards a contradiction that there is an injective function $f : \mathcal{P}(A) \mapsto \omega \times \text{fin}(A)$. Let $k \in \omega$ be such that $2^{2k+1} > n \cdot 2^k$ and let $E_0 \subseteq A$ be a finite set of size $k$. By Lemma 7.12 there are $2^{2k+1}$ subsets of $A$, say $S_1, S_2, \ldots$, which have $E_0$ as their support. Since there are only $2^k$ subsets of $E_0$, by the choice of $k$ there is a first $S_i (1 \leq i \leq 2^{2k+1})$ such that $f(S_i) \notin n \times \text{fin}(E_0)$. Now, $f(S_i) = (m, F_0)$ for some $m \in k$ and $F_0 \in \text{fin}(A)$. Since $F_0 \not
subseteq E_0$ we have $|E_0 \cup F_0| > |E_0|$ and can proceed with $E_1 = E_0 \cup F_0$. Finally, with the sets $E_0, E_1, \ldots$ we get $\aleph_0 \leq 2^m$, which contradicts Lemma 7.13. \hfill \Box

The Prime Ideal Theorem Revisited

In this section we show that the Prime Ideal Theorem holds in the ordered Mostowski model. In other words, the Axiom of Choice is not provable in ZFA from the Prime Ideal Theorem.

Theorem 7.16. The Prime Ideal Theorem holds in the ordered Mostowski model.
Proof. By Theorem 5.15 it is enough to show that in $\mathcal{V}_M$, for every binary mess $B$ there is a function $f$ which is consistent with $B$.

Let $B \in \mathcal{V}_M$ be a binary mess on a set $S$, and let $E_B$ be the least support of $B$. On $S$ define an equivalence relation by stipulating $x \sim y$ iff there is a $\pi \in \text{fix}_{\mathcal{V}}(E_B)$ such that $y = \pi x$. For every $x \in S$ let

$$\{x\}^\pi = \{\pi x : \pi \in \text{fix}_{\mathcal{V}}(E_B)\}$$

and let $\tilde{S} = \{\{x\}^\pi : x \in S\}$. Notice that $x \sim y$ iff $\{x\}^\pi = \{y\}^\pi$.

The goal—which will become clear later—is to lift some functions $t$ of the binary mess on $S$ to functions $h$ defined on finite subsets of $\tilde{S}$ in order to get a binary mess $\tilde{B}$ on $\tilde{S}$ so that every function $g$ on $\tilde{S}$ which is consistent with $\tilde{B}$ induces a function $f \in \mathcal{V}_M$ which is consistent with $B$. Let $B$ consist of all binary functions $h$ defined on finite subsets $Q$ of $\tilde{S}$ that satisfy the following condition: For every finite set $P \subseteq \bigcup \{\{x\}^\pi : [x]^\pi \in Q\}$ there is a $t \in B$ such that $t$ is defined on $P$ and

$$t(x) = h([x]^\pi) \text{ for every } x \in P.$$ 

If this is the case, we say that the set $P$ admits the function $h$. In other words, $P$ admits $h$ if and only if there is a binary function $t \in B$ which is defined on $P$ such that whenever $x, y \in P$ and $x \sim y$, then $t(x) = t(y) = h([x]^\pi)$. In order to show that $\tilde{B}$ is a binary mess, we have to verify that for every finite set $\tilde{Q} \subseteq \tilde{S}$ there is a binary function $h \in \tilde{B}$ which is defined on $\tilde{Q}$.

Once we know that $\tilde{B}$ is a binary mess, we can take any $g$ on $\tilde{S}$ consistent with $\tilde{B}$ and define

$$f(x) = g([x]^\pi)$$

for every $x \in S$. The function $f$ is obviously symmetric, hence $f \in \mathcal{V}_M$, and we are done. So, all that we have to do is to prove the following claim:

For every finite set $\tilde{Q} \subseteq \tilde{S}$ there is a binary function $h \in \tilde{B}$ defined on $\tilde{Q}$, such that for every finite set $P \subseteq \bigcup \{\{x\}^\pi : [x]^\pi \in \tilde{Q}\}$, $P$ admits $h$.

For simplicity we distinguish two cases:

$E_B$ is empty: Let $\tilde{Q}$ be a finite subset of $\tilde{S} = \{\{x\}^\pi : x \in S\}$ and let $Q = \{x \in S : [x]^\pi \in \tilde{Q}\}$. We are looking for a binary function $h$ on $Q$ such that every finite subset of $Q$ admits $h$. Notice that we have $r = 2^q$ binary functions $h$ on $Q$ to choose from, where $q = |Q|$. In $\mathcal{M}_1$, fix some $P_0 \subseteq Q$ which has exactly one element in each equivalence class $[x]^\pi \in \tilde{Q}$ and notice that by definition of $\tilde{S}$, $Q = \bigcup \{\pi P_0 : \pi \in \mathcal{G}\}$. Let us say that $P \subseteq Q$ is a $k$-set if there are $k$ permutations $\pi_1, \ldots, \pi_k \in \mathcal{G}$ such that $P = \pi_1 P_0 \cup \ldots \cup \pi_k P_0$. Since every finite subset of $Q$ is included in a $k$-set for some $k$, it is sufficient to show that for every $k$ and every $k$-set $P$ there is a binary function $h$ on $Q$ such that $P$ admits $h$.
Let $k$ be arbitrary but fixed. We say that two $k$-sets $P_1$ and $P_2$ are isomorphic if $P_2 = \pi_P(P_1)$ for some $\pi \in \mathcal{G}$. Notice that being isomorphic is an equivalence relation. If $P_1$ and $P_2$ are isomorphic and $P_1$ admits $h$ (where $h$ is some binary function on $\hat{Q}$), then also $P_2$ admits $h$. To see this, first notice that since $E_B = \emptyset$, a binary function $t$ belongs to $B$ iff $\pi t$ belongs to $B$ (for any $\pi \in \mathcal{G}$). If $P_1$ admits $h$, then there is a $t \in B$ such that $t$ is defined on $P_1$ and for all $x \in P_1$, we have $t(x) = h(x^\pi)$. Let $P_2 = \pi(P_1)$ and consider the binary function $\pi t \in B$. Since $t(x) \in \{0, 1\}$, $\pi(t(x)) = h(x^\pi)$. Further, for each $y \in P_2$ there is an $x \in P_1$ such that $y = \pi x$, which implies that the binary function $\pi t$ is defined on $P_2$. Hence, for any $y \in P_2$ and $x = \pi^{-1}y \in P_1$ we have $(\pi t)(y) = (\pi t)(\pi x) = t(x) = h(x^\pi) = h(y^\pi)$, which shows that $P_2$ admits $h$. Thus, if a $k$-set $P$ admits the binary function $h$, then all $k$-sets belonging to the same isomorphism class as $P$ also admit $h$.

Now we show that there are only finitely many isomorphism classes of $k$-sets: Let $E_0$ be the least support of $P_0$, and let $n = |E_0|$. Let $\{E_1, \ldots, E_k\}$ and $\{E_1', \ldots, E_k'\}$ be two sets of $n$-element subsets of $A$ (where $A$ is the set of atoms of $\mathcal{V}_M$). We say that these two so-called $(k, n)$-sets are isomorphic if there is a $\pi \in \mathcal{G}$ which transforms the set $\{E_1, \ldots, E_k\}$ into the set $\{E_1', \ldots, E_k'\}$. Notice that there are only finitely many isomorphism classes of $(k, n)$-sets. To see this, let us just consider the case when $n = k = 2$. Let $E_1 = \{a, b\}$ and $E_2 = \{c, d\}$, and without loss of generality let us assume that $a < b$, $c < d$, and that $a = \min\{a, b, c, d\}$. Then the seven different types we can have are represented by $a < b < c < d$, $a < b < c < d$, $a < c < b < d$, $a < c < b < d$, $a = c < b < d$, $a = c < b = d$, and $a < c < d < b$.

For each $E = \pi E_0$ let $P_E := \pi P_0$. Notice that for every $E = \pi E_0$ there is a function $h$ defined on $\hat{Q}$ such that for any function defined on $P_E$ it fits. Further, for each $(k, n)$-set $E = \{E_1, \ldots, E_k\} = \{\pi_1 E_0, \ldots, \pi_k E_0\}$ let $P_E := \pi_1 P_0 \cup \ldots \cup \pi_k P_0$. If $E$ and $E'$ are isomorphic, then so are the two $k$-sets $P_E$ and $P_{E'}$. On the other hand, for every $k$-set $P$ there are $k$ permutations $\pi_1, \ldots, \pi_k \in \mathcal{G}$ such that $P = \pi_1 P_0 \cup \ldots \cup \pi_k P_0$, which implies that $P = P_E$ where $E = \{\pi_1 E_0, \ldots, \pi_k E_0\}$, and consequently we get that $P_E$ and $P_{E'}$ are isomorphic iff $E$ and $E'$ are isomorphic. Hence, since there are only finitely many isomorphism classes of $(k, n)$-sets, there are only finitely many isomorphism classes of $k$-sets.

Thus it suffices to find a binary function $h$ such that for any set of representatives $\{E_1, \ldots, E_p\}$, where $p$ is the number of isomorphism classes of $(k, n)$-sets, we have that each $k$-set $P_{E_i}$ (1 $\leq i \leq p$) admits $h$.

Now we apply the Finite Ramsey Theorem 23 which tells us that for all $m, n, r \in \omega$ there exists an $N \in \omega$ such that for every colouring of $[N]^n$ with $r$ colours, there exists a set $H \in [N]^m$, all whose $n$-element subsets have the same colour. Let $m = k \cdot n$ and $r = 2^k$, and let $F \in [A]^N$ be a set of $N$ atoms. Further, let $P = \bigcup \{P_E : E \in [F]^n\}$ and take any $t \in B$ which is defined on $P$. Then each $t|_{P_{E_i}}$ corresponds to one of the $r$ possible binary functions $h_1, \ldots, h_r$ defined on $\hat{Q}$, which induces a colouring on $[F]^n$ with $r$ colours. By the Finite Ramsey Theorem 23 we find a set $H \in [F]^m$ such
that for every $E \in [H]^n$, $t|_{p_E}$ is the same function and therefore induces a unique function on $Q$, say $h$. Finally, by the choice of $m$, the set $H$ contains members from each isomorphism class, which implies that each $k$-set $P \subseteq Q$ admits $h$.

$E_B$ is non-empty: Assume $E_B = \{a_1, \ldots, a_l\}$ where $a_1 < \ldots < a_l$. Instead of $\mathcal{G}$ we have to work with $\text{fix}_\alpha(E_B)$. Let $I_1 = \{a \in A : a < a_1\}$, $I_j = \{a \in A : a_{j-1} < a < a_j\}$ (for $1 < j < l$), and $I_l = \{a \in A : a_l < a\}$. Let $P_0$ and $E_0$ be as above and for $1 \leq j \leq l$ let $n_j := |E_0 \cap I_j|$. Instead of $(k, n)$-sets consider sets of the form $(\mathcal{E}_1, \ldots, \mathcal{E}_n)$, where for $1 \leq i \leq n$, $\mathcal{E}_i = (E_{i,1}, \ldots, E_{i,l})$ and for each $1 \leq j \leq l$, $E_{i,j} \subseteq I_j$ and $|E_{i,j}| = n_j$. Now we can proceed as above until we reach the point where the Finite Ramsey Theorem comes in. Here, the combinatorics gets slightly more involved and instead of the Finite Ramsey Theorem we need Rado’s generalisation, which is Theorem 2.7 given in Chapter 2: It says that for all $r, l, m_1, \ldots, m_l \in \omega$ there is some $N \in \omega$ such that whenever $[N]^{m_1} \times \ldots \times [N]^{m_l}$ is coloured with $r$ colours, then there are $M_1, \ldots, M_l \in [N]^m$ such that $[M_1]^{m_1} \times \ldots \times [M_l]^{m_l}$ is monochromatic. Let $m = \max\{k \cdot n_i : 1 \leq i \leq l\}$ and $r = 2^m$, and let $F_1, \ldots, F_l \in [A]^N$ be $N$-element sets of atoms such that for every $1 \leq j \leq l$, $F_{i,j} \subseteq I_j$. Then we find $l$ sets $M_j \in [F_j]^m$ such that $[M_1]^{m_1} \times \ldots \times [M_l]^{m_l}$ is monochromatic, which implies again that each $k$-set $P \subseteq Q$ admits the same function $h$.}

### Custom-Built Permutation Models

Below we shall construct two permutation models. The first one is designed in order to show that the existence of infinite cardinals $m$ for which $\text{seq}(m) < \text{fin}(m)$ is consistent with $\text{ZF}$. By modifying the first custom-built permutation model, this somewhat counter intuitive result can even be pushed a little bit further by showing that also the existence of infinite cardinals $m$ for which $m^2 < |m|^2$ is consistent with $\text{ZF}$.

#### The first custom-built permutation model

The set of atoms of the first custom-built permutation model is built by induction, where every atom encodes a finite sequence of atoms on a lower level and every finite sequence of atoms appears in finitely many atoms.

By induction on $n \in \omega$ we construct sets $A_n$, functions $\text{Seq}_n$ from $A_n$ to $\text{seq}(A_{n-1})$, and groups $G_n$ which are subgroups of the group of permutations of $A_n$ as follows:

1. $A_0 := \{a_0\}$, where $a_0$ is an atom, $\text{Seq}_0(a_0) = \langle \rangle$, and $G_0 = \{\iota\}$ is the group of all permutations of $A_0$.

For $n \in \omega$ let $k_n = |G_n|$, and let $\mathcal{S}_n$ be the set of sequences of $A_n$ of length less than or equal to $n + 1$ which do not belong to the range of $\text{Seq}_n$. Then
(β) $A_{n+1} := A_n \cup \{(n+1, \zeta, i) : \zeta \in \mathcal{F}_n \land i < k_n + k_n\}$.

(γ) $\text{Seq}_{n+1}$ is a function from $A_{n+1}$ to $\text{seq}(A_n)$ defined as follows:

$$\text{Seq}_{n+1}(x) = \begin{cases} 
\text{Seq}_n(x) & \text{if } x \in A_n, \\
\zeta & \text{if } x = (n+1, \zeta, i) \in A_{n+1} \setminus A_n.
\end{cases}$$

(δ) $G_{n+1}$ is the subgroup of the group of permutations of $A_{n+1}$ containing all permutations $h$ such that for some $g_h \in G_n$ and $j_h < k_n + k_n$ we have

$$h(x) = \begin{cases} 
g_h(x) & \text{if } x \in A_n, \\
(n+1, g_h(\zeta), i + \_a j_h) & \text{if } x = (n+1, \zeta, i) \in A_{n+1} \setminus A_n,
\end{cases}$$

where $g_h(\zeta)(m) := g_h(\zeta(m))$ and $\_a$ is addition modulo $(k_n + k_n)$.

Let $A := \bigcup\{A_n : n \in \omega\}$. For each triple $(n, \zeta, i) \in \mathcal{A}$ we assign an atom $\alpha_{(n, \zeta, i)}$ and define the set of atoms by stipulating $\mathcal{A} := A_0 \cup \{\alpha_{(n, \zeta, i)} : (n, \zeta, i) \in A\}$. However, for the sake of simplicity we shall work with $A$ as the set of atoms rather than with $\mathcal{A}$. Let $\text{Seq} := \bigcup\{\text{Seq}_n : n \in \omega\}$; then $\text{Seq}$ is a function from $A$ onto $\text{seq}(A)$. Furthermore, let $\text{Aut}(A)$ be the group of all permutations of $A$. Then $\mathcal{G} := \{H \in \text{Aut}(A) : \forall n \in \omega(H|_{A_n} \in G_n)\}$ is a group of permutations of $A$. Finally, let $\mathcal{F}$ be the filter on $\mathcal{G}$ generated by $\{\text{fix}_\mathcal{G}(E) : E \in \text{fin}(A)\}$ (which happens to be normal) and let $\mathcal{V}_s$ ($s$ for sequences) be the class of all hereditarily symmetric objects. Now we are ready to prove the following result.

**Proposition 7.17.** Let $m$ denote the cardinality of the set of atoms $A$ of $\mathcal{V}_s$. Then $\mathcal{V}_s \models \text{seq}(m) \leq \text{fin}(m)$.

**Proof.** Firstly we prove that $\mathcal{V}_s \models \text{seq}(m) \leq \text{fin}(m)$ by constructing a one-to-one function $f$ in $\mathcal{V}_s$ which maps $\text{seq}(A)$ into $\text{fin}(A)$. For any sequence $\zeta \in \text{seq}(A)$ there is a least $n_\zeta \in \omega$ such that $\zeta \in \mathcal{S}_{n_\zeta}$. Define $f : \text{seq}(A) \to \text{fin}(A)$ by stipulating

$$f(\zeta) = \{a \in A : \exists i(a = \alpha_{(n_\zeta + 1, \zeta, i)})\}.$$  

Obviously, $f$ is injective and it remains to show that $f$ belongs to $\mathcal{V}_s$. Take an arbitrary permutation $\pi \in \mathcal{G}$ and let $\zeta \in \text{seq}(A)$ be an arbitrary sequence. Notice first that by the definition of $\mathcal{G}$, $n_\zeta = n_{\pi\zeta}$. Thus, for each $i < k_{n_\zeta} + k_{n_\zeta}$ there is a $j < k_{n_\zeta} + k_{n_\zeta}$ such that $\pi(n_\zeta + 1, \zeta, i) = (n_{\pi\zeta} + 1, \pi\zeta, j)$, which shows that $\pi(\zeta, f(\zeta)) = (\pi\zeta, f(\pi\zeta))$, and since $\zeta$ was arbitrary we get $\pi f = f$.

In order to prove that $\mathcal{V}_s \models \text{seq}(m) \neq \text{fin}(m)$ assume towards a contradiction that there is a one-to-one function $g \in \mathcal{V}_s$ from $\text{fin}(A)$ into $\text{seq}(A)$.

Notice first that for every $E \in \text{fin}(A)$ there are $C, F \in \text{fin}(A)$ such that $E \subseteq C$, and for all $x \in A \setminus C$ we have $|\{\pi x : \pi \in \text{fix}_\mathcal{G}(C)\}| > 2$, and $|\{\pi F : \pi \in \text{fix}_\mathcal{G}(C)\}| = 2$. Indeed, choose $n \geq 1$ such that $E \subseteq A_n$, and let
$C := A_n$ and $F := \{(n + 1, \zeta, i) \in A_{n+1} : i \text{ is even}\}$. Then $F$ has exactly two images under the permutations of $\text{fix}_g(C)$, and for all $x \in A \setminus C$ we have $|\{\pi x : \pi \in \text{fix}_g(C)\}| \geq (k_n + k_n) > 2$.

Let $E$ be a support of $g$ and let $C$ and $F$ be as above. If the sequence $g(F)$ belongs to $\text{seq}(C)$, then for some $\pi \in \text{fix}_g(C)$, $\pi F \neq F$, hence, $g(\pi F) \neq g(F)$. But this contradicts that $C$ is a support of $g$ and that $\pi \in \text{fix}_g(C)$. Otherwise, if the sequence $g(F)$ does not belong to $\text{seq}(C)$, there is an $m \in \omega$ such that $x_0 := g(F)(m) \notin C$. Hence, by the choice of $C$ and $F$ we have $|\{\pi x_0 : \pi \in \text{fix}_g(C)\}| > 2$, and $|\{\pi F : \pi \in \text{fix}_g(C)\}| = 2$. Since every $\pi \in \text{fix}_g(C)$ maps $g$ to itself, in particular $(F, g(F))$ to $(\pi F, \pi g(F))$, and since $|\{\pi F : \pi \in \text{fix}_g(C)\}| < |\{\pi x_0 : \pi \in \text{fix}_g(C)\}|$, the image under $g$ of a 2-element set has strictly more than two elements, which is obviously a contradiction.

The second custom-built permutation model

The set of atoms of the second custom-built permutation model is also built by induction, and every atom encodes an ordered pair of atoms on a lower level. The model we finally get will be a model in which there exists a cardinal $m$ such that $m^2 < [m]^2$, which is to some extent just a finite version of Proposition 7.17. The atoms are constructed as follows:

(a) $A_0$ is an arbitrary countable infinite set of atoms.

(b) $\mathcal{G}_0$ is the group of all permutations of $A_0$.

(c) $A_{n+1} := A_n \cup \{ (n + 1, p, \varepsilon) : p \in A_n \times A_n, \varepsilon \in \{0, 1\}\}$.

(d) $\mathcal{G}_{n+1}$ is the subgroup of the permutation group of $A_{n+1}$ containing all permutations $h$ for which there are $g_h \in \mathcal{G}_n$ and $\varepsilon_h \in \{0, 1\}$ such that

$$h(x) = \begin{cases} g_h(x) & \text{if } x \in A_n, \\ (n + 1, g_h(p), \varepsilon_h + 2 \varepsilon) & \text{if } x = (n + 1, p, \varepsilon), \end{cases}$$

where for $p = (p_1, p_2) \in A_n$, $g_h(p) := (g_h(p_1), g_h(p_2))$ and $+2$ denotes addition modulo 2.

Let $A := \bigcup\{ A_n : n \in \omega \}$ and let $\text{Aut}(A)$ be the group of all permutations of $A$. Then

$$\mathcal{G} := \{ H \in \text{Aut}(A) : \forall n \in \omega \ (H|_{A_n} \in \mathcal{G}_n) \}$$

is a group of permutations of $A$. Let $\mathcal{F}$ be the filter on $\mathcal{G}$ generated by $\{ \text{fix}_E : E \in \text{fin}(A) \}$ (which happens to be normal) and let $\mathcal{V}_p$ ($p$ for pairs) be the class of all hereditarily symmetric objects. Now we are ready to prove the following

**Proposition 7.18.** Let $m$ denote the cardinality of the set of atoms $A$ of $\mathcal{V}_p$. Then $\mathcal{V}_p \models m^2 < [m]^2$. 
\textbf{Proof.} First we show that $\mathcal{V}_p \models m^2 \leq [m]^2$. For this it is sufficient to find a one-to-one function $f \in \mathcal{V}_p$ from $A^2$ into $[A]^2$. We define such a function as follows. For $x, y \in A$ let

\begin{equation*}
  f((x, y)) := \{ (n + m + 1, (x, y), 0), (n + m + 1, (x, y), 1) \},
\end{equation*}

where $n$ and $m$ are the smallest numbers such that $x \in A_n$ and $y \in A_m$, respectively. For any $\pi \in \mathcal{G}$ and $x, y \in A$ we have $\pi f((x, y)) = f((\pi x, \pi y))$ and therefore, the function $f$ is as desired and belongs to $\mathcal{V}_p$.

Now assume towards a contradiction that there exists a one-to-one function $g \in \mathcal{V}_p$ from $[A]^2$ into $A^2$ and let $E_g$ be a finite support of $g$. Without loss of generality we may assume that if $(n+1, (x, y), z) \in E_g$ then also $x, y \in E_g$ (this will be needed later). Let $k := |E_g|$ and for $x, y \in A$ let $g((x, y)) = (t_0(x, y), t_1(x, y))$. Let $r := k + 4$ and let $N \in \omega$ be such that for every colouring $\tau : [N]^2 \to r^2$ we find a 3-element set $H \in [N]^3$ such that $\tau |H|_2$ is constant. Such a number $N$ exists by the \textsc{finite Ramsey theorem} 2.3. Choose $N$ distinct elements $x_0, \ldots, x_{N-1} \in A_0 \setminus E_g$ let $X = \{ x_0, \ldots, x_{N-1} \}$ and let $\{ c_h : h < k \}$ be an enumeration of $E_g$ (recall that $k = |E_g|$). We define a colouring $\tau : [X]^2 \to r \times r$ as follows. For $\{ x_i, x_j \} \in [X]^2$, where $i < j$, let $\tau(\{ x_i, x_j \}) = (\tau_0(\{ x_i, x_j \}), \tau_1(\{ x_i, x_j \}))$ where for $l \in \{ 0, 1 \}$ we define

\begin{equation*}
  \tau_l(\{ x_i, x_j \}) := \begin{cases} 
h & \text{if } t_l(x_i, x_j) = c_h, \\
k & \text{if } t_l(x_i, x_j) = x_i, \\
k+1 & \text{if } t_l(x_i, x_j) = x_j, \\
k+2 & \text{if } t_l(x_i, x_j) \in A_0 \setminus (\{ x_i, x_j \} \cup E_g), \\
k+3 & \text{if } t_l(x_i, x_j) \in A \setminus (A_0 \cup E_g). \end{cases}
\end{equation*}

By the definition of $N$ we find 3 elements $x_{i_0}, x_{i_1}, x_{i_2} \in X$ with $i_0 < i_1 < i_2$ such that for both $l \in \{ 0, 1 \}$, $\tau_l$ is constant on $[\{ x_{i_0}, x_{i_1}, x_{i_2} \}]^2$. So, for $\{ x_i, x_j \} \in [\{ x_{i_0}, x_{i_1}, x_{i_2} \}]^2$ with $i < j$ and for some $l \in \{ 0, 1 \}$, we are in at least one of the following cases:

\begin{enumerate}
  \item $t_l(x_i, x_j) = c_h$ and $t^{1-l}_{x_i, x_j} = c_h$,
  \item $t_l(x_i, x_j) = c_h$ and $t^{1-l}_{x_i, x_j} = x_i$,
  \item $t_l(x_i, x_j) = c_h$ and $t^{1-l}_{x_i, x_j} = x_j$,
  \item $t_l(x_i, x_j) = t^{1-l}_{x_i, x_j}$ and $t^{1-l}_{x_i, x_j} \in \{ x_i, x_j \}$
  \item $t_l(x_i, x_j) = x_i$ and $t^{1-l}_{x_i, x_j} = x_j$,
  \item $t_l(x_i, x_j) \in A_0 \setminus (E_g \cup \{ x_i, x_j \})$
  \item $t_l(x_i, x_j) \in A \setminus (E_g \cup A_0)$
\end{enumerate}
If we are in case (1) or (2), then \( g(x_{i_0}, x_{i_1}) = g(x_{i_0}, x_{i_2}) \), and therefore \( g \) is not a one-to-one function. If we are in case (3), then \( g \) is also not a one-to-one function because \( g(x_{i_0}, x_{i_2}) = g(x_{i_1}, x_{i_2}) \), and the same is true for \( g \) if we are in case (4), e.g., \( g(x_{i_0}, x_{i_1}) = (x_{i_0}, x_{i_1}) = g(x_{i_0}, x_{i_2}) \).

If we are in case (5), then let \( \pi \in \text{fix}(E_g) \) be such that \( \pi x_{i_0} = x_{i_1} \) and \( \pi x_{i_1} = x_{i_2} \). Assume that \( g(x_{i_0}, x_{i_1}) = (x_{i_0}, x_{i_1}) \) (the case when \( g(x_{i_0}, x_{i_1}) = (x_{i_1}, x_{i_2}) \) is similar). Then we have \( \pi x_{i_0}, x_{i_1} = \{ x_{i_0}, x_{i_1} \} \), but \( \pi g(x_{i_0}, x_{i_1}) = (x_{i_1}, x_{i_2}) \), and therefore \( E_g \) is not a support of \( g \) which contradicts the choice of \( E_g \);—which, by our assumption, has the property that whenever \( (n+1, (x, y), z) \in E_g \) also \( x, y \in E_g \).

If we are in case (6), then let \( l \in \{ 0, 1 \} \) be such that \( l^l(x_{i_0}, x_{i_1}) \in A_0 \setminus (E_g \cup \{ x_{i_0}, x_{i_1} \}) \) and let \( a := l^l(x_{i_0}, x_{i_1}) \). Without loss of generality we may assume \( l = 0 \), thus, \( a = l^l(x_{i_0}, x_{i_1}) \). Take an arbitrary \( a' \in A_0 \setminus (E_g \cup \{ a, x_{i_1}, x_{i_1} \}) \) and let \( \pi \in \text{fix}(E_g \cup \{ x_{i_0}, x_{i_1} \}) \) be such that \( \pi a = a' \) and \( \pi a' = a \). Then we get \( \pi x_{i_0}, x_{i_1} = \{ x_{i_0}, x_{i_1} \} \) but \( g(\pi x_{i_0}, x_{i_1}) = g(\{ x_{i_0}, x_{i_1} \}) = (a, x) \neq (a', x') = \pi (a, x) = \pi g(\{ x_{i_0}, x_{i_1} \}) \).

Hence, \( E_g \) is not a support of \( g \) which contradicts the choice of \( E_g \).

If we are in case (7), then let \( l \in \{ 0, 1 \} \) be such that \( l^l(x_{i_0}, x_{i_1}) \in A \setminus (E_g \cup A_0) \), thus \( l^l(x_{i_0}, x_{i_1}) = (n+1, p, \epsilon) \) for some \( (n+1, p, \epsilon) \in A \). Further, let \( \pi \in \text{fix}(E_g \cup \{ x_{i_0}, x_{i_1} \}) \) be such that \( \pi (n+1, p, \epsilon) = (n+1, p, 1-\epsilon) \). Then we have \( \pi x_{i_0}, x_{i_1} = \{ x_{i_0}, x_{i_1} \} \) but \( \pi g(\{ x_{i_0}, x_{i_1} \}) = g(\{ x_{i_0}, x_{i_1} \}) \), and therefore \( E_g \) is not a support of \( g \) which contradicts the choice of \( E_g \).

So, in all the cases, either \( g \) is not one-to-one or \( E_g \) is not a support of \( g \), which contradicts our assumption and completes the proof.

Notes

Permutation models. The method of permutation models was introduced by Fraenkel [2, 4, 3, 5, 6], and, in a precise version with supports, by Lindenbaum and Mostowski [18] and by Mostowski [20, 21, 22]. The present version with filters is due to Specker [23]. In particular, the second Fraenkel model can be found for example in Fraenkel [2], where he proved that the Axiom of Choice for countable families of pairs is unprovable in ZFA (for a proof in a more general setting see Mendelson [39]), and the ordered Mostowski model is introduced in [21, §4, p. 236] in order to show that the Axiom of Choice is independent from the Ordering Principle. (Some more background can be found for example in Lévy [17].)

The Prime Ideal Theorem. The independence of the Axiom of Choice from the Prime Ideal Theorem in ZFA was proved first by Halpern [10] (but the proof presented above is taken from Jech [13, Chapter 7, §1]). A few years later, the same result in ZF was proved by Halpern and Lévy [12], using the Halpern–Láuchli Theorem.

The custom-built models. The first custom-built permutation model as well as Proposition 7.17 is due to Shelah and can be found in [8, Theorem 2]. The second
custom-built permutation model, which is just a modification of the first one, is
due to Halbeisen, but the crucial part of Proposition 7.18 is again due to Shelah
(cf. Halbeisen and Shelah [9, Proposition 7.3.1]).

43. Alternative definition of atoms. Atoms could also be defined by stipulating
\( a \in A \iff a = \{a\} \). This approach has the advantage that we do not need to
modify the Axiom of Extensionality; however, it has the disadvantage that
models of ZFA would not be well-founded—except in the case when \( A = \emptyset \).

44. The Axiom of Choice in Algebra. Läuchli shows in [14] that many classical
results in Algebra cannot be proved without the aid of the Axiom of Choice.
For example he shows that it is consistent with ZFA that there exists vector
spaces without algebraic bases, or in which there exist two algebraic bases with
different cardinalities.

45. More cardinal relations. Let \( m \) denote the cardinality of the set of atoms of
the basic Fraenkel model \( \mathcal{V}_{F_0} \). Then the following statements hold in
\( \mathcal{V}_{F_0} \) (cf. Halbeisen and Shelah [9, Proposition 7.13]):

- (a) \( \text{fin}(m) \perp \text{seq}^{-1}(m) \) and \( \text{fin}(m) \perp \text{seq}(m) \).
- (b) \( \text{seq}^{-1}(m) \perp z^m \) and \( \text{seq}(m) \perp z^m \).
- (c) \( \text{seq}^{-1}(m) < \text{seq}(m) \).

Unlike in the basic Fraenkel model, the cardinalities \( \text{fin}(m) \), \( z^m \), \( \text{seq}^{-1}(m) \), and
\( \text{seq}(m) \) are all comparable in the ordered Mostowski model. Let \( m \) denote the
cardinality of the set of atoms of \( \mathcal{V}_M \). Then the following sequence of inequalities
holds in \( \mathcal{V}_M \):

\[
m < [m]^2 < m^2 < \text{fin}(m) < z^m < \text{seq}^{-1}(m) < \text{fin}^2(m) < \text{seq}^{-1}(\text{fin}(m)) < \\
< \text{fin}(z^2) < \text{fin}^3(m) < \text{fin}^4(m) < \ldots < \text{fin}^n(m) < \text{seq}(m) < z^{\text{fin}(m)} = z^n
\]

(See for example Halbeisen and Shelah [9, p. 249] or Halbeisen [7], or just use
the ideas of the proof of Proposition 7.15.) Furthermore we have that

\[
\mathcal{V}_M \models \left(z^m \right)^{\text{fin}(m)} = z^m
\]

which follows for example from the fact that \( \mathcal{V}_M \models z^m = z^{\text{fin}(m)} \) and LÄUCHLI'S
Lemma 4.27.

Finally, let \( m \) denote the cardinality of the set of atoms of the second Fraenkel
model. Then, by Proposition 7.9 and LÄUCHLI'S Lemma 4.27 we have

\[
\mathcal{V}_{F_2} \models \left(z^m \right)^{\text{fin}(m)} = z^m.
\]

46. Multiple Choice and Kurepa’s Principle in Fraenkel’s models. In Chapter 5 we
have seen that Multiple Choice and Kurepa’s Principle are both equivalent in ZF
to the Axiom of Choice. On the other hand, one can show that Multiple Choice
holds in the model \( \mathcal{V}_{F_0} \) and that Kurepa’s Principle holds in the model \( \mathcal{V}_{F_2} \) (see
Levy [16] and Halpern [13] respectively, or Jech [13, Theorem 9.2]). This shows that
these two choice principles—which imply AC in ZF—are weaker than AC
in ZFA.
47. **Countable unions of countable sets.** In order to show that a union of countably many countable sets is not necessarily countable, one can work for example in the permutation model given by Fraenkel [9]: The set of atoms consists of countably many mutually disjoint countable sets. So, \( A = \bigcup_{n<\omega} C_n \) where each \( C_n \) is countable. For each \( n \in \omega \), the group \( G_n \) consists of all permutations of \( C_n \) and \( \mathcal{G} = \prod_{n<\omega} G_n \). The normal filter \( \mathcal{F} \) on \( \mathcal{G} \) is generated by products of the form \( \prod_{n<\omega} H_n \), where \( H_n \) is either equal to \( G_n \) or the trivial group, and the former is the case for all but finitely many \( n \)’s.

48. **Ordering principles in Mostowski’s model.** Mostowski showed in [21] that in ZFA, the Axiom of Choice is not provable from the Ordering Principle (see also Jech [23, Theorem 4.7]). In fact he showed that the Ordering Principle holds in the ordered Mostowski model \( \mathcal{V}_M \), whereas the Axiom of Choice obviously fails in that model. Notice also that even the Prime Ideal Theorem, which implies the Ordering Principle, holds in \( \mathcal{V}_M \). In [1], Felgner and Truss gave a direct proof — not referring to the Prime Ideal Theorem — of the fact that the Order-Extension Principle holds in \( \mathcal{V}_M \) and, then, by modifying \( \mathcal{V}_M \), they were able to show that in ZFA, the Prime Ideal Theorem is not provable from the Order-Extension Principle.

Läuchli showed in [15] (see also Jech [13, p. 53]) that the following form of the Axiom of Choice holds in \( \mathcal{V}_M \): For every family of non-empty well-orderable sets there is a choice function. Notice that this implies that in \( \mathcal{V}_M \), the union of a countable set of countable sets is always countable.

49. **Another custom-built permutation model.** Let \( m \) denote the cardinality of the set of atoms of the first custom-built permutation model \( \mathcal{V}_s \). Then one can show that \( \mathcal{V}_s \models \text{seq}^{\omega \omega}(m) < \text{seq}(m) < \omega^m \) (see Halbeisen and Shelah [9, Proposition 7.4.1]), or use Proposition 7.17 and show that \( m \) is \( D \)-finite.) So, for an infinite cardinals \( m \) we can have \( \text{seq}^{\omega \omega}(m) < \text{seq}(m) < \omega^m \) (which holds in \( \mathcal{V}_s \) as well as \( \omega^m < \text{seq}^{\omega \omega}(m) < \text{seq}(m) \) (which holds in \( \mathcal{V}_M \)), and therefore both statements are consistent with ZF. It is now natural to ask whether it is also possible to put \( \omega^m \) between the cardinals \( \text{seq}^{\omega \omega}(m) \) and \( \text{seq}(m) \) (recall that by Theorem 4.24, for all infinite cardinals \( m \) we have \( \omega^m \neq \text{seq}^{\omega \omega}(m) \neq \text{seq}(m) \)). Indeed, the existence of an infinite cardinal \( m \) for which

\[
\text{seq}^{\omega \omega}(m) < \omega^m < \text{seq}(m)
\]

is also consistent with ZF and the permutation model in which this holds — given in Halbeisen and Shelah [9, Section 7.4] — is due to Shelah.

**References**


References

7. Lorenz Halbeisen, Vergleichung zwischen unendlichen Kardinalzahlen in einer Mengenlehre ohne Auswahlaxiom, Diplomarbeit (1990), University of Zürich (Switzerland).
22. [Author], *Axiom of choice for finite sets*, *Fundamenta Mathematicae*, vol. 33 (1945), 137–168.


Twelve Cardinals and their Relations

The consonances are those intervals which are formed from the natural steps.
An interval may be diminished when one of its steps is replaced by a smaller one.
Or it may be augmented when one of its steps is replaced by a larger one.

GIOSEFFO ZARLINO
Le Institutioni Harmoniche, 1558

In this chapter we investigate twelve cardinal characteristics and their relations to one another. A cardinal characteristic of the continuum is an uncountable cardinal number which is less than or equal to $c$ that describes a combinatorial or analytical property of the continuum. Like the power of the continuum itself, the size of a cardinal characteristic is often independent from ZFC. However, some restrictions on possible sizes follow from ZFC, and we shall give a complete list of what is known to be provable in ZFC about their relation. Later in Part II, but mainly in Part III, we shall see how one can diminish or augment some of these twelve cardinals without changing certain other cardinals. In fact, these cardinal characteristics are also used to investigate combinatorial properties of the various forcing notions introduced in Part III.

We shall encounter some of these cardinal characteristics (e.g., $p$) more often than others (e.g., $i$). However, we shall encounter each of these twelve cardinals again, and like the twelve notes of the chromatic scale, these twelve cardinals will build the framework of our investigation of the combinatorial properties of forcing notions that is carried out in Part III.

On the one hand, it would be good to have the definition of a cardinal characteristic at hand when it is needed; but on the other hand, it is also convenient to have all the definitions together (especially when a cardinal characteristic is used several times), rather than scattered over the entire
book. Defining all twelve cardinals at once also gives us the opportunity to show what is known to be provable in ZFC about the relationship between these twelve cardinals. Thus, one might first skip this chapter and go back to it later and take bits and pieces when necessary.

The Cardinals $\omega_1$ and $c$

We have already met both cardinals, $c$ and $\omega_1$: $c$ is the cardinality of the continuum $\mathbb{R}$, and $\omega_1$ is the smallest uncountable cardinal. According to FACT 4.3, $c = 2^\omega$ is also the cardinality of the sets $[0,1]$, $\omega^2$, $\omega$, and $[0,1] \setminus Q$; and by LEMMA 4.10, $\omega_1$ can also be considered as the set of order types of well-orderings of $\mathbb{Q}$.

The Continuum Hypothesis, denoted CH, states that $c$ is the least uncountable cardinal, i.e., $c = \omega_1$ (cf. Chapter 4), which is equivalent to saying that every subset of $\mathbb{R}$ is either countable or of the same cardinality as $\mathbb{R}$. Furthermore, the Generalised Continuum Hypothesis, denoted GCH, states that for every ordinal $\alpha \in \Omega$, $2^{\omega_\alpha} = \omega_{\alpha+1}$. Gödel showed that $L \models \text{GCH}$, where $L$ is the constructible universe (see the corresponding note in Chapter 5), thus, GCH is consistent with ZFC.

Each of the following ten combinatorial cardinal characteristics of the continuum is uncountable and less than or equal to $c$. Thus, if we assume CH, then these cardinals are all equal to $c$. However, as we shall see in Part II, CH is not provable in ZFC. In other words, if ZFC is consistent then there are models of ZFC in which CH fails, i.e., models in which $\omega_1 < c$. In those models, possible (i.e., consistent) relations between the following cardinal characteristics will be provided in Part II and Part III.

The Cardinal $p$

For two sets $x, y \subseteq \omega$ we say that $x$ is almost contained in $y$, denoted $x \subseteq^* y$, if $x \setminus y$ is finite, i.e., all but finitely many elements of $x$ belong to $y$. For example a finite subset of $\omega$ is almost contained in $\emptyset$, and $\omega$ is almost contained in every co-finite subset of $\omega$ (i.e., in every $y \subseteq \omega$ such that $\omega \setminus y$ is finite). A pseudo-intersection of a family $\mathcal{F} \subseteq [\omega]^{\omega}$ of infinite subsets of $\omega$ is an infinite subset of $\omega$ that is almost contained in every member of $\mathcal{F}$. For example $\omega$ is a pseudo-intersection of the family of co-finite sets. Furthermore, a family $\mathcal{F} \subseteq [\omega]^{\omega}$ has the strong finite intersection property (sfip) if every finite subfamily has infinite intersection. Notice that every family with a pseudo-intersection necessarily has the sfip, but not vice versa. For example any filter $\mathcal{F} \subseteq [\omega]^{\omega}$ has the sfip, but no ultrafilter on $[\omega]^{\omega}$ has a pseudo-intersection.
The Cardinals $\mathfrak{b}$ and $\mathfrak{d}$

**Definition of $\mathfrak{p}$.** The *pseudo-intersection number* $\mathfrak{p}$ is the smallest cardinality of any family $\mathcal{F} \subseteq [\omega]^\omega$ which has the $\text{slip}$ but which does not have a pseudo-intersection; more formally

$$\mathfrak{p} = \min \{ |\mathcal{F}| : \mathcal{F} \subseteq [\omega]^\omega \text{ has the slip but no pseudo-intersection} \}.$$ 

Since ultrafilters on $[\omega]^\omega$ are families which have the slip but do not have a pseudo-intersection, and since every ultrafilter on $[\omega]^\omega$ is of cardinality $\mathfrak{c}$, the cardinal $\mathfrak{p}$ is well-defined and $\mathfrak{p} \leq \mathfrak{c}$. It is natural to ask whether $\mathfrak{p}$ can be smaller than $\mathfrak{c}$; however, the following result shows that $\mathfrak{p}$ cannot be too small.

**Theorem 8.1.** $\omega_1 \leq \mathfrak{p}$.

**Proof.** Let $\mathcal{E} = \{ X_n \in [\omega]^\omega : n \in \omega \}$ be a countable family which has the slip. We construct a pseudo-intersection of $\mathcal{E}$ as follows: Let $a_0 := \bigcap X_0$ and for positive integers $n$ let

$$a_n = \bigcap \left( \bigcap \{ X_i : i \in n \} \setminus \{ a_i : i \in n \} \right).$$

Further, let $Y = \{ a_n : n \in \omega \}$; then for every $n \in \omega$, $Y \setminus \{ a_i : i \in n \} \subseteq X_n$ which shows that $Y \subseteq^* X_n$, hence, $Y$ is a pseudo-intersection of $\mathcal{E}$. \hfill \Box

The Cardinals $\mathfrak{b}$ and $\mathfrak{d}$

For two functions $f, g \in {}^\omega \omega$ we say that $g$ *dominates* $f$, denoted $f <^* g$, if for all but finitely many integers $k \in \omega$, $f(k) < g(k)$, i.e., if there is an $n_0 \in \omega$ such that for all $k \geq n_0$, $f(k) < g(k)$. Notice that ordering $<^*$ is transitive, however, $<^*$ it is not a linear ordering (we leave it as an exercise to the reader to find functions $f, g \in {}^\omega \omega$ such that neither $f <^* g$ nor $g <^* f$).

A family $\mathcal{D} \subseteq {}^\omega \omega$ is *dominating* if for each $f \in {}^\omega \omega$ there is a function $g \in \mathcal{D}$ such that $f <^* g$.

**Definition of $\mathfrak{d}$.** The *dominating number* $\mathfrak{d}$ is the smallest cardinality of any dominating family; more formally

$$\mathfrak{d} = \min \{ |\mathcal{D}| : \mathcal{D} \subseteq {}^\omega \omega \text{ is dominating} \}.$$ 

A family $\mathcal{B} \subseteq {}^\omega \omega$ is *unbounded* if there is no single function $f \in {}^\omega \omega$ which dominates all functions of $\mathcal{B}$, i.e., for every $f \in {}^\omega \omega$ there is a $g \in \mathcal{B}$ such that $g \not<^* f$. Since $<^*$ is not a linear ordering, an unbounded family is not necessarily dominating — but vice versa (see Fact 8.2).

**Definition of $\mathfrak{b}$.** The *bounding number* $\mathfrak{b}$ is the smallest cardinality of any unbounded family; more formally

$$\mathfrak{b} = \min \{ |\mathcal{B}| : \mathcal{B} \subseteq {}^\omega \omega \text{ is unbounded} \}.$$
Obviously, the family \( \omega \) itself is dominating and therefore unbounded, which shows that \( d \) and \( b \) are well-defined and \( b, d \leq c \). Moreover, we have the following

**Fact 8.2.** \( b \leq d \).

**Proof.** It is enough to show that every dominating family is unbounded. So, let \( D \subseteq \omega \) be a dominating family and let \( f \in \omega \) be an arbitrary function. Since \( D \) is dominating, there is a \( g \in D \) such that \( f <^\ast g \), i.e., there is an \( n_0 \in \omega \) such that for all \( k \geq n_0 \), \( f(k) < g(k) \). Hence we get \( g \not<^\ast f \), and since \( f \) was arbitrary this implies that \( D \) is unbounded.

It is natural to ask whether \( b \) can be smaller than \( d \), or at least smaller than \( c \), however, the following result shows that \( b \) cannot be too small.

**Theorem 8.3.** \( \omega_1 \leq b \).

**Proof.** Let \( \mathcal{E} = \{g_n \in \omega : n \in \omega \} \) be a countable family. We construct a function \( f \in \omega \) which dominates all functions of \( \mathcal{E} \): For each \( k \in \omega \) let

\[
  f(k) = \bigcup \{g_i(k) : i \in k\}.
\]

Then for every \( k \in \omega \) and each \( i \in k \) we have \( f(k) \geq g_i(k) \) which shows that for all \( n \in \omega \), \( g_n <^\ast f \), hence, \( f \) dominates all functions of \( \mathcal{E} \).

One could also define dominating and unbounded families with respect to the ordering \( <^\ast \) defined by stipulating \( f < g \iff \forall k \in \omega (f(k) < g(k)) \). Then the corresponding dominating number would be the same as \( d \), as any dominating family can be made dominating in the new sense by adding all finite modifications of its members; but the corresponding bounding number would drop to \( \omega \), as the family of all constant functions is unbounded (we leave the details to the reader).

**The Cardinals \( s \) and \( \tau \)**

A set \( x \subseteq \omega \) splits an infinite set \( y \in [\omega]^\omega \) if both \( y \cap x \) and \( y \setminus x \) are infinite (i.e., \( |y \cap x| = |y \setminus x| = \omega \)). Notice that any \( x \subseteq \omega \) which splits a set \( y \in [\omega]^\omega \) must be infinite. A splitting family is a family \( \mathcal{F} \subseteq [\omega]^\omega \) such that each \( y \in [\omega]^\omega \) is split by at least one \( x \in \mathcal{F} \).

**Definition of \( s \).** The splitting number \( s \) is the smallest cardinality of any splitting family; more formally

\[
  s = \min \{|\mathcal{F}| : \mathcal{F} \subseteq [\omega]^\omega \text{ is splitting}\}.
\]
The cardinals $s$ and $\tau$

By Theorem 8.1 and later results we get $\omega_1 \leq s$—we leave it as an exercise to the reader to find a direct proof of the uncountability of $s$.

In the proof of the following result we will see how to construct a splitting family from a dominating family.

**Theorem 8.4.** $s \leq \delta$.

**Proof.** For each strictly increasing function $f \in \omega^\omega$ with $f(0) > 0$ let

\[
\sigma_f = \bigcup \left\{ \left[ f^{2n}(0), f^{2n+1}(0) \right] : n \in \omega \right\},
\]

where for $a, b \in \omega$, $[a, b) := \{ k \in \omega : a \leq k < b \}$ and $f^{n+1}(0) = f(f^n(0))$ with $f^0(0) := 0$. Let $\mathcal{D} \subseteq \omega^\omega$ be a dominating family. Without loss of generality we may assume that every $f \in \mathcal{D}$ is strictly increasing and $f(0) > 0$, and let

\[
\mathcal{S} = \{ \sigma_f : f \in \mathcal{D} \}.
\]

We show that $\mathcal{S}$ is a splitting family. So, fix an arbitrary $x \in [\omega]^\omega$ and let $f_x \in \omega^\omega$ be the (unique) strictly increasing bijection between $\omega$ and $x$. More formally, define $f_x : \omega \to x$ by stipulating

\[
f_x(k) = \bigcap \left( x \setminus \{ f_x(i) : i \in k \} \right).
\]

Notice that for all $k \in \omega$, $f_x(k) \geq k$. Since $\mathcal{D}$ is dominating there is an $f \in \mathcal{D}$ such that $f_x <^* f$, which implies that there is an $n_0 \in \omega$ such that for all $k \geq n_0$ we have $f_x(k) < f(k)$. For each $k \in \omega$ we have $k \leq f^k(0)$ as well as $k \leq f_x(k)$. Moreover, for $k \geq n_0$ we have

\[
f^k(0) \leq f_x(f^k(0)) < f(f^k(0)) = f^{k+1}(0)
\]

and therefore $f_x(f^k(0)) \in \left[ f^k(0), f^{k+1}(0) \right)$. Thus, for all $k \geq n_0$ we have $f_x(f^k(0)) \in \sigma_f$ iff $k$ is even, which shows that both $x \cap \sigma_f \cap x$ and $x \setminus \sigma_f$ are infinite. Hence, $\sigma_f$ splits $x$, and since $x$ was arbitrary, $\mathcal{S}$ is a splitting family.

A **reaping family**—also known as a *refining* or *unsplittable family*—is a family $\mathcal{R} \subseteq [\omega]^\omega$ such that there is no single set $x \in [\omega]^\omega$ which splits all elements of $\mathcal{R}$, i.e., for every $x \in [\omega]^\omega$ there is a $y \in \mathcal{R}$ such that $y \cap x$ or $y \setminus x$ is finite. In other words, a family $\mathcal{R}$ is reaping if for every $x \in [\omega]^\omega$ there is a $y \in \mathcal{R}$ such that $y \subseteq^* (\omega \setminus x)$ or $y \subseteq^* x$. The origin of "reaping" in this context is that $A$ reaps $B$ iff $A$ splits $B$, by analogy with a scythe cutting the stalks of grain when one reaps the grain. So, a **reaping family** would be a splitting family. However, the more logical approach, where "reaps" means "is unsplittable by", seems to have no connection with the everyday meaning of the word "reap".
The reaping number \( \tau \) is the smallest cardinality of any reaping family; more formally

\[
\tau = \min \{ |\mathcal{R}| : \mathcal{R} \subseteq [\omega]^\omega \text{ is reaping} \}.
\]

Since the family \([\omega]^\omega\) is obviously reaping, \( \tau \) is well-defined and \( \tau \leq \mathfrak{c} \). Furthermore, by Theorem 8.3, the following result implies that every reaping family is uncountable:

**Theorem 8.5.** \( \mathfrak{b} \leq \tau \).

**Proof.** Let \( \mathcal{E} = \{ x_\xi : \xi < \kappa < b \} \) be an arbitrary family of infinite subsets of \( \omega \) of cardinality strictly less than \( b \). We show that \( \mathcal{E} \) is not a reaping family. For each \( x_\xi \in \mathcal{E} \), let \( g_\xi \in \omega^\omega \) be the unique strictly increasing bijection between \( \omega \) and \( x_\xi \setminus \{ 0 \} \). Further, let \( \tilde{g}_\xi(k) := g_\xi(k+1) \), where \( g_\xi(0) = g_{\xi}(0) \) and \( g_\xi(0) := 0 \). Consider \( \tilde{\mathcal{E}} = \{ \tilde{g}_\xi : \xi \in \kappa \} \). Since \( \kappa < b \), the family \( \tilde{\mathcal{E}} \) is bounded, i.e., there exists a function \( f \in \omega^\omega \) such that for all \( \xi \in \kappa \), \( \tilde{g}_\xi(k) < f(k+1) \). Let \( x = \bigcup_{k \in \omega} \{ f(k), f(k+1) \} \). Then for each \( \xi \in \kappa \), there is an \( n_\xi \in \omega \) such that for all \( k \geq n_\xi \), \( f(k) \leq \tilde{g}_\xi(f(k)) < f(f(k)) \). This implies that neither \( x_\xi \subseteq^* x \) nor \( x_\xi \subseteq^* (\omega \setminus x) \), and hence, \( \mathcal{E} \) is not a reaping family. \( \dashv \)

**The Cardinals** \( a \) and \( i \)

Two sets \( x, y \subseteq [\omega]^\omega \) are **almost disjoint** if \( x \cap y \) is finite. A family \( \mathcal{A} \subseteq [\omega]^\omega \) of pairwise almost disjoint sets is called an **almost disjoint family**; and a **maximal almost disjoint** (mad) family is an infinite almost disjoint family \( \mathcal{A} \subseteq [\omega]^\omega \) which is maximal with respect to inclusion, i.e., \( \mathcal{A} \) is not properly contained in any almost disjoint family \( \mathcal{A}' \subseteq [\omega]^\omega \).

**Definition of** \( a \). The **almost disjoint number** \( a \) is the smallest cardinality of any maximal almost disjoint family; more formally

\[
a = \min \{ |\mathcal{A}| : \mathcal{A} \subseteq [\omega]^\omega \text{ is mad} \}.
\]

Before we show that \( \mathfrak{b} \leq a \) (which implies that \( a \) is uncountable), let us show first that there is a mad family of cardinality \( \mathfrak{c} \).

**Proposition 8.6.** There exists a maximal almost disjoint family of cardinality \( \mathfrak{c} \).

**Proof.** Notice that by Teichmüller’s Principle, every almost disjoint family can be extended to a mad family. So, it is enough to construct an almost disjoint family \( \mathcal{A}_0 \) of cardinality \( \mathfrak{c} \). Let \( \{ s_i : i \in \omega \} \) be an enumeration of \( \bigcup_{n \in \omega} n \omega \), i.e., for each \( t : n \to \omega \) there is a unique \( i \in \omega \) such that \( t = s_i \). For \( f \in \omega^\omega \) let

\[
x_f = \{ i \in \omega : \exists n \in \omega (f | n = s_i) \}.
\]
The theorems in the previous section show that for any $a < \mathfrak{c}$, there is a mad family. It is enough to construct an uncountable mad family from a $\mathfrak{m}$-mad family. Let $\mathcal{F} = \{ x_\xi : \xi < \kappa \}$ be a mad family. Let $\mathcal{A} = \{\{x_\xi : \xi < \kappa \} \cup \{ y \} : y \in \omega \}$ be a mad family. Then $\mathcal{A}$ is an almost disjoint family of cardinality $\mid \omega \mid$.

**Theorem 8.7.** $b \leq a$.

**Proof.** Let $\mathcal{A} = \{ x_\xi : \xi < \kappa \}$ be a mad family. It is enough to construct an uncountable family of cardinality $\mid \mathcal{A} \mid$. Let $z = \omega \setminus \bigcup_{\xi \in \kappa} x_\xi$; then $z$ is finite (otherwise, $\mathcal{A} \cup \{ z \}$ would be an almost disjoint family which properly contains $\mathcal{A}$). Let $z_0 := z \cup \{ n \in \omega \setminus \{ 0 \} \}$ and for positive integers $n \in \omega$ let $x'_n := \{ x_n \} \setminus \bigcup_{k \in \omega} x_k$. Then, since $\mathcal{A}$ is an almost disjoint family, $\{ x'_n : n \in \omega \}$ is a family of pairwise disjoint subsets of $\omega$ and for $n \in \omega$ let $g_n \in \omega$ be the unique strictly increasing bijection from $x'_n$ to $\omega$. Let $h : \omega \rightarrow \omega \times \omega$ defined by stipulating

$$h(m) = (n, k) \text{ where } m \in x'_n \text{ and } k = g_n(m).$$

By definition, for each $n \in \omega$, $h[x'_n] = \{ (n, k) : k \in \omega \}$, and for all $\xi < \kappa$, $h[x_\omega+\xi] \cap x'_n$ is finite. Further, for each $\xi < \kappa$ define $f_\xi \in \omega$ by stipulating

$$f_\xi(k) = \bigcup \{ h(x_\omega+\xi) \cap x'_k \}$$

and let $\mathcal{B} = \{ f_\xi \in \omega : \xi < \kappa \}$. Then by definition $\mid \mathcal{B} \mid = \mid \mathcal{A} \mid$; moreover, $\mathcal{B}$ is unbounded. Indeed, if there would be a function $f \in \omega$ which dominates all functions of $\mathcal{B}$, then the infinite set $\{ h^{-1}(\{ n, f(n) \}) : n \in \omega \}$ would have finite intersection which each element of $\mathcal{A}$ contrary to maximality of $\mathcal{A}$. $\dashv$

A family $\mathcal{F} \subseteq [\omega]^\omega$ is called independent if the intersection of any finitely many members of $\mathcal{F}$ and the complements of any finitely many other members of $\mathcal{F}$ is infinite. More formally, $\mathcal{F} \subseteq [\omega]^\omega$ is independent if for any $n, m \in \omega$ and disjoint sets $\{ x_i : i \in n \}, \{ y_j : j \in m \} \subseteq \mathcal{F}$,

$$\bigcap_{i \in n} x_i \cap \bigcap_{j \in m} (\omega \setminus y_j)$$

is infinite,

where we stipulate $\bigcap \emptyset = \omega$. Equivalently, $\mathcal{F} \subseteq [\omega]^\omega$ is independent if for any $I, J \in \text{fin}(\mathcal{F})$ with $I \cap J = \emptyset$ we have

$$\bigcap I \setminus \bigcup J$$

is infinite.

We leave it as an exercise to the reader to show that if $\mathcal{F}$ is infinite, then $\mathcal{F}$ is independent iff for any disjoint sets $I, J \in \text{fin}(\mathcal{F})$, $\bigcap I \setminus J \neq \emptyset$. The following result implies that $a$ is uncountable and in the proof we will show how one can construct an uncountable family from a mad family.
A **maximal independent** family is an independent family $\mathcal{I} \subseteq [\omega]^\omega$ which is maximal with respect to inclusion, i.e., $\mathcal{I}$ is not properly contained in any independent family $\mathcal{I}' \subseteq [\omega]^\omega$.

**Definition of $i$.** The **independence number** $i$ is the smallest cardinality of any maximal independent family; more formally

$$i = \min \{|\mathcal{I}| : \mathcal{I} \subseteq [\omega]^\omega \text{ is independent} \}.$$  

We shall see that $\max\{r, \mathfrak{b}\} \leq i$ (which implies that $i$ is uncountable), but first let us show that there is a maximal independent family of cardinality $\mathfrak{c}$.

**Proposition 8.8.** There is a maximal independent family of cardinality $\mathfrak{c}$.  

**Proof.** It is enough to construct an independent family of cardinality $\mathfrak{c}$ on some countably infinite set. So, let us construct an independent family of cardinality $\mathfrak{c}$ on the countably infinite set

$$C = \{(s, A) : s \in \operatorname{fin}(\omega) \land A \subseteq \mathcal{P}(s)\}.$$  

Further, for each $x \subseteq [\omega]^\omega$ define

$$P_x := \{(s, A) \in C : x \cap s \in A\}.$$  

Notice that for any distinct $x, y \in [\omega]^\omega$ there is a finite set $s \in \operatorname{fin}(\omega)$ such that $x \cap s \neq y \cap s$, and consequently we get $P_x \neq P_y$ which implies that the set $\mathcal{I}_0 = \{P_x : x \in [\omega]^\omega\} \subseteq [C]^\omega$ is of cardinality $\mathfrak{c}$. Moreover, $\mathcal{I}_0$ is an independent family on $C$. Indeed, for any finitely many distinct finite subsets of $\omega$, say $x_0, \ldots, x_m, \ldots, x_m+n$ where $m, n \in \omega$, there is a finite set $s \subseteq \omega$ such that for all $i, j$ with $0 \leq i < j \leq m + n$ we have $x_i \cap s \neq x_j \cap s$. Let $A = \{s \cap x_i : 0 \leq i \leq m\} \subseteq \mathcal{P}(s)$, and for every $k \in \omega \setminus s$ let $s_k := s \cup \{k\}$ and $A_k := A \cup \{t \cup \{k\} : t \in A\}$. Then

$$\{(s_k, A_k) : k \in \omega \setminus s\} \subseteq \bigcap_{0 \leq i \leq m} P_{x_i} \setminus \bigcup_{1 \leq j \leq n} P_{x_{m+j}},$$  

which shows that $\bigcap\{P_x : 0 \leq i \leq m\} \setminus \bigcup\{P_{x_{m+j}} : 1 \leq j \leq n\}$ is infinite, and therefore, $\mathcal{I}_0$ is an independent family on $C$ of cardinality $\mathfrak{c}$.  

The following result implies that $i$ is uncountable.

**Theorem 8.9.** $\max\{r, \mathfrak{b}\} \leq i$.  

**Proof.** $r \leq i$: The idea is to show that every maximal independent family yields a reaping family of the same cardinality. For this, let $\mathcal{I} \subseteq [\omega]^\omega$ be a maximal independent family of cardinality $i$ and let

$$\mathcal{R} = \left\{ \bigcap I \setminus \bigcup J : I, J \in \operatorname{fin}(\mathcal{I}) \land I \cap J = \emptyset \right\}.$$  


Then \( \mathcal{R} \) is a family of cardinality \( \kappa \). Furthermore, since \( \mathcal{I} \) is a maximal independent family, for every \( x \in [\omega]^\omega \) we find a \( y \in \mathcal{R} \) (i.e., \( y = \bigcap I \setminus \bigcup J \)) such that either \( x \cap y \) or \( (\omega \setminus x) \cap y \) is finite, and because \( (\omega \setminus x) \cap y = y \setminus x \), this shows that \( x \) does not split all elements of \( \mathcal{R} \). Thus, \( \mathcal{R} \) is a reaping family of cardinality \( \kappa \), and therefore \( \tau \leq \kappa \).

\( \vartheta \leq \kappa \): The idea is to show that an independent family of cardinality strictly less than \( \vartheta \) cannot be maximal. For this, suppose \( \mathcal{I} = \{X_\xi : \xi < \kappa \} \subseteq [\omega]^\omega \) is an infinite independent family of cardinality \( \kappa < \vartheta \). We shall construct a set \( Z \in [\omega]^\omega \) such that \( \mathcal{I} \cup \{Z\} \) is still independent, which implies that the independent family \( \mathcal{I} \) is not maximal. For this it is enough to show that for any finite, disjoint subfamilies of \( \mathcal{I} \), say \( I \) and \( J \), the infinite set \( \bigcap I \setminus \bigcup J \) meets both \( Z \) and \( \omega \setminus Z \) in an infinite set.

Let \( \mathcal{I}_\omega := \{X_\omega : n \in \omega\} \subseteq \mathcal{I} \) be a countably infinite subfamily of \( \mathcal{I} \) and for each \( n \in \omega \) let \( X_0^n := X_n \) and \( X_1^n := \omega \setminus X_n \). Further, for each \( g \in \omega^2 \) let

\[
C_{n,g} = \bigcap_{k \in n} X_{g(k)}^k
\]

and for \( \mathcal{I}' := \mathcal{I} \setminus \mathcal{I}_\omega \) define

\[
\mathcal{F} = \left\{ \bigcap I' \setminus \bigcup J' : I' \text{ and } J' \text{ are finite, disjoint subfamilies of } \mathcal{I}' \right\}.
\]

**Claim.** The family \( \mathcal{C} = \{C_{n,g} : n \in \omega\} \) has a pseudo-intersection that has infinite intersection with every set in \( \mathcal{F} \).

**Proof of Claim.** Since \( \mathcal{I} \) is an infinite independent family of cardinality \( \kappa < \vartheta \), \( \mathcal{C} \subseteq [\omega]^\omega \) is a family of cardinality \( \kappa \) such that each set in \( \mathcal{F} \) has infinite intersection with every member of \( \mathcal{C} \). For any \( h \in \omega^\omega \) define

\[
Y^h_g = \bigcup_{n \in \omega} \left( C_{n,g} \cap h(n) \right).
\]

Since \( \{C_{n,g} : n \in \omega\} \) is decreasing (i.e., \( C_{n,g} \supseteq C_{m,g} \) whenever \( n \leq m \)), \( Y^h_g \) is almost contained in each member of \( \mathcal{C} \) — however, \( Y^h_g \) is not necessarily infinite. It remains to choose the function \( h \in \omega^\omega \) so that \( Y^h_g \) is infinite (i.e., \( Y^h_g \) is a pseudo-intersection of \( \mathcal{C} \)) and has infinite intersection with every set in \( \mathcal{F} \). Notice first that for every \( A \in \mathcal{F} \) and for every \( n \in \omega \), \( A \cap C_{n,g} \) is infinite; thus, for every \( A \in \mathcal{F} \) we can define a function \( f_A(n) \in \omega^\omega \) by stipulating

\[
f_A(n) = \text{the } n^{th} \text{ element (in increasing order) of } A \cap C_{n,g}.
\]

Since \( |\mathcal{F}| < \vartheta \), the family \( \{f_A : A \in \mathcal{F}\} \) is not dominating. In particular, there is a function \( h_0 \in \omega^\omega \) with the property that for each \( A \in \mathcal{F} \) the set

\[
D_A = \{ n \in \omega : h_0(n) > f_A(n) \}
\]

is infinite. Now, for each \( A \in \mathcal{F} \) and every \( n \in D_A \) we have \( h_0(n) \geq f_A(n) + 1 \) which implies that \( |A \cap h_0(n)| \geq |A \cap f_A(n) + 1| = n \), and since \( D_A \) is infinite,
also \( A \cap Y^h \) is infinite. Finally, by construction \( Y^h \) is a pseudo-intersection of \( \mathcal{G} \) that has infinite intersection with every set in \( \mathcal{F} \).

By the CLAIM, for every \( g \in \omega^2 \) there is a set, say \( Y_g \in [\omega]^\omega \), which has the following two properties:

1. For all \( n \in \omega \), \( Y_g \subseteq \bigcap_{k \in n} X^g(k) \).
2. \( Y_g \cap (\bigcap I \setminus \bigcup J') \) is infinite whenever \( I' \) and \( J' \) are finite, disjoint subfamilies of \( \mathcal{F}' \).

It follows from (1) that for any distinct \( g, g' \in \omega^\omega \), \( Y_g \) and \( Y_{g'} \) are almost disjoint. Let now

\[
Q_0 = \{ g \in \omega^\omega : \exists n_0 \in \omega \forall k \geq n_0 \ (g(k) = 0) \} \\
Q_1 = \{ g \in \omega^\omega : \exists n_1 \in \omega \forall k \geq n_1 \ (g(k) = 1) \} .
\]

Then \( Q_0 \cup Q_1 \) is a countably infinite subset of \( \omega^\omega \). Let \( \{ g_n : n \in \omega \} \) be an enumeration of \( Q_0 \cup Q_1 \) and for each \( n \in \omega \) let \( Y_{g_n} := Y_{g_n} \setminus \bigcup \{ Y_{g_k} : k \in n \} \). Then \( \{ Y_{g_n} : n \in \omega \} \) is a countable family of pairwise disjoint infinite subsets of \( \omega \). Finally let

\[
Z = \bigcup_{g \in Q_0} Y_g \quad \text{and} \quad Z' = \bigcup_{g \in Q_1} Y_g .
\]

Then \( Z \) and \( Z' \) are disjoint. Now we show that \( Z \) has infinite intersection with every \( \bigcap I \setminus J \), where \( I \) and \( J \) are arbitrary finite subfamilies of \( \mathcal{F} \); and since the same also holds for \( Z' \subseteq \omega \setminus Z \), \( \mathcal{F} \cup \{ Z \} \) is an independent family, i.e., the independent family \( \mathcal{F} \) of cardinality \( \omega \) is not maximal.

Given any finite, disjoint subfamilies \( I, J \subseteq \mathcal{F} \), and let \( I_0 = I \setminus \mathcal{F}_\omega \), \( J_0 = J \setminus \mathcal{F}_\omega \), \( I' = I \setminus I_0 \), \( J' = J \setminus J_0 \), where \( \mathcal{F}_\omega = \{ X_n : n \in \omega \} \). Further, let \( m \in \omega \) be such that \( I_0 \cup J_0 \subseteq \{ X_n : n \in m \} \subseteq \mathcal{F}_\omega \) and fix \( g \in Q_0 \) such that for all \( n \in m \),

\[
(X_n \in (I_0 \cup J_0) \land g(n) = 0) \leftrightarrow X_n \in I_0 .
\]

We get the following inclusions:

\[
\bigcap I \setminus J \supseteq \left( \bigcap I' \setminus J' \right) \cap \bigcap_{n \in m} X_n^{g(n)} \supseteq \left( \bigcap I' \setminus J' \right) \cap Y_g
\]

The intersection on the very right is infinite (by property (2) of \( Y_g \)) and is contained in \( Z \) (because \( g \in Q_0 \)). Hence, we have found an infinite set which is almost contained in \( Z \cap (\bigcap I \cup J) \), and therefore \( Z \) is infinite.

\[\] 

The Cardinals \( \text{par} \) and \( \text{hom} \)

By RAMSEY’S THEOREM 2.1, for every colouring \( \pi : [\omega]^2 \rightarrow 2 \) there is an \( x \in [\omega]^\omega \) which is homogeneous for \( \pi \), i.e., \( \pi|_{[x]^2} \) is constant. This leads to the following cardinal characteristic:

\[\]
The cardinals \textit{par} and \textit{hom} are defined.  The **homogeneity number** \textit{hom} is the smallest cardinality of any family \( \mathcal{F} \subseteq [\omega]^\omega \) with the property that for every colouring \( \pi : [\omega]^2 \to 2 \) there is an \( x \in \mathcal{F} \) which is homogeneous for \( \pi \).

The following result implies that \( \textit{hom} \) is uncountable. In fact we will show that each family which contains a homogeneous set for every 2-colouring of \( [\omega]^2 \) is reaping and that each such family yields a dominating family of the same cardinality.

**Theorem 8.10.** \( \max \{ r, \delta \} \leq \textit{hom} \).

**Proof.** Let \( \mathbb{F} \subseteq [\omega]^\omega \) be a family such that for every colouring \( \pi : [\omega]^2 \to 2 \) there is an \( x \in \mathbb{F} \) which is homogeneous for \( \pi \). We shall show that \( \mathbb{F} \) is reaping and that \( \mathbb{F}' = \{ f_x \in [\omega] : x \in \mathbb{F} \} \) is dominating, where \( f_x \) is the strictly increasing bijection between \( \omega \) and \( x \).

\[ \delta \leq \textit{hom} \]:  Firstly we show that \( \mathbb{F} \) is a dominating family. For any strictly increasing function \( f \in [\omega]^\omega \) with \( f(0) = 0 \) define \( \pi_f : [\omega]^2 \to 2 \) by stipulating

\[ \pi_f(\{ n, m \}) = 0 \iff \exists k \in \omega (f(2k) \leq n, m < f(2k + 2)) . \]

Then, for every \( x \in \mathbb{F} \) which is homogeneous for \( \pi_f \) we have \( f <^* f_x \) which implies that \( \mathbb{F}' \) is dominating.

\( r \leq \textit{hom} \):  Now we show that \( \mathbb{F} \) is a reaping family. Take any \( y \in [\omega]^\omega \) and define \( \pi_y : [\omega]^2 \to 2 \) by stipulating

\[ \pi_y(\{ n, m \}) = 0 \iff \{ n, m \} \subseteq y \lor \{ n, m \} \cap y = \emptyset . \]

Now, for every \( x \in \mathbb{F} \) which is homogeneous for \( \pi_y \) we have either \( x \subseteq y \) or \( x \cap y = \emptyset \), and since \( y \) was arbitrary, \( \mathbb{F} \) is reaping.

Recall that a set \( H \in [\omega]^\omega \) is called **almost homogeneous** for a colouring \( \pi : [\omega]^2 \to 2 \) if there is a finite set \( K \subseteq H \) such that \( H \setminus K \) is homogeneous for \( \pi \). This leads to the following cardinal characteristic:

**Definition of par.** The **partition number** \textit{par} is the smallest cardinality of any family \( \mathcal{P} \) of 2-colourings of \( [\omega]^2 \) such that no single \( H \in [\omega]^\omega \) is almost homogeneous for all \( \pi \in \mathcal{P} \).

By Proposition 28 we get that \textit{par} is uncountable, and the following result gives an upper bound for \textit{par}.

**Theorem 8.11.** \( \textit{par} = \min \{ s, b \} \).

**Proof.** First we show that \( \textit{par} \leq \min \{ s, b \} \) and then we show that \( \textit{par} \geq \min \{ s, b \} \). \( \textit{par} \leq s \):  Let \( \mathbb{F} \subseteq [\omega]^\omega \) be a splitting family and for each \( x \in \mathbb{F} \) define the colouring \( \pi_x : [\omega]^2 \to 2 \) by stipulating

\[ \pi_x(\{ n, m \}) = 0 \iff \{ n, m \} \subseteq x \lor \{ n, m \} \cap x = \emptyset . \]
and let \( P = \{ \pi_x : x \in P \} \). Then, since \( P \) is splitting, no infinite set is almost homogeneous for all \( \pi \in P \).

\[ \par \leq b: \text{Let } B \subseteq \omega \omega \text{ be an unbounded family. Without loss of generality we may assume that each } g \in B \text{ is strictly increasing. For each } g \in B \text{ define the colouring } \pi_g : [\omega]^2 \to 2 \text{ by stipulating} \]

\[ \pi_g([n, m]) = 0 \iff g(n) < m \text{ where } n < m. \]

Assume towards a contradiction that some infinite set \( H \subseteq [\omega]^\omega \) is almost homogeneous for all colourings in \( P = \{ \pi_g : g \in B \} \). We shall show that \( H \) yields a function which dominates the unbounded family \( B \), which is obviously a contradiction. Consider the function \( h \in \omega^\omega \) which maps each natural number \( n \) to the second member of \( H \) above \( n \); more formally, \( h(n) := \min \{ m \in H : \exists k \in H(n < k < m) \} \). For each \( n \in \omega \) we have \( n < k < h(n) \) with both \( k \) and \( h(n) \) in \( H \). By almost homogeneity of \( H \), for each \( g \in B \) there is a finite set \( K \subseteq \omega \) such that \( H \setminus K \) is homogeneous for \( \pi_g \), i.e., for all \( [n, m] \in [H \setminus K]^2 \) with \( n < m \) we have either \( g(n) < m \) or \( g(n) \geq m \). Since \( H \) is infinite, the latter case is impossible. On the other hand, the former case implies that for all \( n \in H \setminus K, g(n) < h(n) \), hence, \( h \) dominates \( g \) and consequently \( h \) dominates each function of \( B \).

\[ \par \geq \min\{s, b\}: \text{Suppose } P = \{ \pi_\xi : \xi \in \kappa < \min\{s, b\} \} \text{ is a family of 2-colouring of } [\omega]^2. \text{ We shall construct a set } H \subseteq [\omega]^\omega \text{ which is almost homogeneous for all colourings } \pi \in P \text{. For each } \xi \in \kappa \text{ and all } n \in \omega \text{ define the function } f_{\xi, n} \in [\omega]^2 \text{ by stipulating} \]

\[ f_{\xi, n}(m) = \begin{cases} \pi_\xi([n, m]) & \text{for } m \neq n, \\ 0 & \text{otherwise.} \end{cases} \]

Since \( |\{f_{\xi, n} : \xi \in \kappa \land n \in \omega\}| = \kappa \cdot \omega = \kappa < s \), there is an infinite set \( A \subseteq \omega \) on which all functions \( f_{\xi, n} \) are almost constant; more formally, for each \( \xi \in \kappa \) and each \( n \in \omega \) there are \( g_\xi(n) \in \omega \) and \( j_\xi(n) \in \{0, 1\} \) such that for all \( m \geq g_\xi(n) \), \( f_{\xi, n}(m) = j_\xi(n) \). Moreover, since \( \kappa < s \) there is an infinite set \( B \subseteq A \) on which each function \( j_\xi \in [\omega]^2 \) is almost constant, say \( j_\xi(n) = i_\xi \) for all \( n \in B \) with \( n \geq b_\xi \). Further, since \( \kappa < b \) there is a strictly increasing function \( h \in [\omega]^\omega \) which dominates each \( g_\xi \), i.e., for each \( \xi \in \kappa \) there is an integer \( c_\xi \) such that for all \( n \geq c_\xi, g_\xi(n) < h(n) \). Let \( H = \{ x_\xi : \xi \in \kappa \} \subseteq B \) be such that for all \( k \in \omega, h(x_k) < x_{k+1} \). Then \( H \) is almost homogeneous for each \( \pi_\xi \in P \). Indeed, if \( n, m \in H \) are such that \( \max\{b_\xi, c_\xi\} \leq n < m \), then \( g_\xi(n) < h(n) < m \) and therefore \( \pi_\xi([n, m]) = f_{\xi, n}(m) = j_\xi(n) = i_\xi \), i.e., \( H \setminus \max\{b_\xi, c_\xi\} \) is homogeneous for \( \pi_\xi \).

The Cardinal \( \theta \)

A family \( \mathcal{H} = \{ \omega_\xi : \xi \in \kappa \} \subseteq P([\omega]^\omega) \) of mad families of cardinality \( \kappa \) is called shattering if for each \( x \in [\omega]^\omega \) there is a \( \xi \in \kappa \) such that \( x \) has infinite
intersection with at least two distinct members of \( \mathcal{A}_x \), i.e., at least two sets of \( \mathcal{A}_x \) split \( x \). We leave it as an exercise to the reader to show that there are shattering families of cardinality \( c \) (for each \( x \in [\omega]^\omega \) take two disjoint sets \( y, y' \subseteq x \) such that \( \omega \setminus (y \cup y') \) is infinite and extend \( \{y, y'\} \) to a mad family of cardinality \( c \).

**Definition of \( \mathbf{h} \).** The **shattering number** \( \mathbf{h} \) is the smallest cardinality of a shattering family; more formally

\[
\mathbf{h} = \min \{ |\mathcal{H}| : \mathcal{H} \text{ is shattering} \}.
\]

If one tries to visualise a shattering family, one would probably draw a kind of matrix with \( c \) columns, where the rows correspond to the elements of the family (i.e., to the mad families). Having this picture in mind, the size of the shattering family would then be the **height** of the matrix, and this where the letter “h” comes from.

In order to prove that \( \mathbf{h} \leq \text{par} \) we shall show how to construct a shattering family from any family \( \mathcal{P} \) of 2-colourings of \( [\omega]^2 \) such that no single set is almost homogeneous for all \( \pi \in \mathcal{P} \); the following lemma is the key idea in that construction:

**Lemma 8.12.** For every colouring \( \pi : [\omega]^2 \to 2 \) there is a mad family \( \mathcal{A}_x \) of cardinality \( c \) such that each \( A \in \mathcal{A}_x \) is homogeneous for \( \pi \).

**Proof.** Let \( \mathcal{A} \subseteq [\omega]^\omega \) be an arbitrary almost disjoint family of cardinality \( c \) and let \( \pi \) be a 2-colouring of \( [\omega]^2 \). By **Ramsey’s Theorem 2.1**, for each \( A \in \mathcal{A} \) we find an infinite set \( A' \subseteq A \) such that \( A' \) is homogeneous for \( \pi \). Let \( \mathcal{A}' = \{A' : A \in \mathcal{A}\} \); then \( \mathcal{A}' \) is an almost disjoint family of cardinality \( c \) where each member of \( \mathcal{A}' \) is homogeneous for \( \pi \). Let \( \{x_\xi : \xi \in \kappa \leq c\} \) be an enumeration of \( [\omega]^\omega \setminus \mathcal{A}' \). By transfinite induction define \( \mathcal{A}_x = \mathcal{A}' \) and for each \( \xi \in \kappa \) let

\[
\mathcal{A}_{\xi+1} = \begin{cases} 
\mathcal{A}_\xi \cup \{x_\xi\} & \text{if } x_\xi \text{ is homogeneous for } \pi \text{ and for each } A \in \mathcal{A}_\xi, x_\xi \cap A \text{ is finite}, \\
\mathcal{A}_\xi & \text{otherwise.}
\end{cases}
\]

By construction, \( \mathcal{A}_x = \bigcup_{\xi \in \kappa} \mathcal{A}_\xi \) is an almost disjoint family of cardinality \( c \), all whose members are homogeneous for \( \pi \). Moreover, \( \mathcal{A}_x \) is a mad family. Indeed, if there would be an \( x' \in [\omega]^\omega \) such that for all \( A \in \mathcal{A}_x \), \( x' \cap A \) is finite, then, by **Ramsey’s Theorem 2.1**, there would be an \( x_{\xi_0} \in [x]'^\omega \) (for some \( \xi_0 \in \kappa \)) which is homogeneous for \( \pi \). In particular, \( x_{\xi_0} \) would belong to \( \mathcal{A}_{\xi_0+1} \). Hence, \( x \cap x_{\xi_0} \) is infinite, where \( x_{\xi_0} \in \mathcal{A}_x \), which is a contradiction to the choice of \( x \).

**Theorem 8.13.** \( \mathbf{h} \leq \text{par} \).
Proof. Let \( \mathcal{P} \) be a family of 2-colourings of \( [\omega]^2 \) such that no single set is almost homogeneous for all \( \pi \in \mathcal{P} \) and let \( \mathcal{H}_\mathcal{P} = \{ \mathcal{A}_\pi : \pi \in \mathcal{P} \} \), where \( \mathcal{A}_\pi \) is like in Lemma 8.12. We claim that \( \mathcal{H}_\mathcal{P} \) is shattering. Indeed, let \( H \subseteq \omega \) be an arbitrary infinite subset of \( \omega \). By the property of \( \mathcal{P} \), there is a \( \pi \in \mathcal{P} \) such that \( H \) is not almost homogeneous for \( \pi \). Consider \( \mathcal{A}_\pi \in \mathcal{H}_\mathcal{P} \): Since \( \mathcal{A}_\pi \) is mad, there is an \( A \in \mathcal{A}_\pi \) such that \( H \cap A \) is infinite, and since \( A \) is homogeneous for \( \pi \), \( H \setminus A \) is infinite too; and again, since \( \mathcal{A}_\pi \) is mad, there is an \( A' \in \mathcal{A}_\pi \) (distinct from \( A \)) such that \( (H \setminus A) \cap A' \) is infinite. This shows that \( H \) has infinite intersection with two distinct members of \( \mathcal{A}_\pi \). Hence, \( \mathcal{H}_\mathcal{P} \) is shattering.

In order to prove that \( p \leq h \) we have to introduce some notions: If \( \mathcal{A} \) and \( \mathcal{A}' \) are mad families (of cardinality \( c \)), then \( \mathcal{A}' \) refines \( \mathcal{A} \), denoted \( \mathcal{A}' \succ \mathcal{A} \), if for each \( A' \in \mathcal{A}' \) there is an \( A \in \mathcal{A} \) such that \( A' \subseteq^+ A \) a shattering family \( \{ A_\xi : \xi \in \kappa \} \) is called refining if \( \mathcal{A}_\xi \succ \mathcal{A}_\xi \) whenever \( \xi' > \xi \).

The next result is the key lemma in the proof that every shattering family of size \( h \) induces a refining shattering family of the same cardinality.

**Lemma 8.14.** For every family \( \mathcal{E} = \{ A_\xi : \xi < \kappa < h \} \) of cardinality \( \kappa < h \) of mad families of cardinality \( c \) there exists a mad family \( \mathcal{A}' \) which refines each \( A_\xi \in \mathcal{E} \). Furthermore \( \mathcal{A}' \) is of cardinality \( c \).

**Proof.** Let \( \mathcal{E} = \{ A_\xi : \xi < h \} \) be a family of less than \( h \) mad families of cardinality \( c \). For every \( x \in [\omega]^\omega \) we find an \( x' \in [x^\omega]^\omega \) with the property that for each \( A_\xi \in \mathcal{H} \) there is an \( A \in A_\xi \) such that \( x' \subseteq^+ A \). Indeed, if there is no such \( x' \) (for some given \( x \in [\omega]^\omega \)), then a bijection between \( x \) and \( \omega \) would yield a shattering family of cardinality \( \kappa < h \), contrary to the definition of \( h \).

Now, if \( \mathcal{A}' \subseteq \{ x' : x \in [\omega]^\omega \} \) is a mad family, then \( \mathcal{A}' \) is of cardinality \( c \) (since \( \mathcal{A}_0 \) is of cardinality \( c \)) and refines each \( A_\xi \in \mathcal{E} \) (since \( \mathcal{A}_0 \subseteq \{ x' : x \in [\omega]^\omega \} \)). It remains to show that mad families \( \mathcal{A}' \subseteq \{ x' : x \in [\omega]^\omega \} \) exist. Indeed, if \( \mathcal{A} \subseteq \{ x' : x \in [\omega]^\omega \} \) is an almost disjoint family which is not maximal, then there exists an \( x \in [\omega]^\omega \) such that for all \( A \in \mathcal{A} \), \( x \cap A \) is finite. Notice that \( \mathcal{A} \cup \{ x' \} \) is still an almost disjoint family, hence, by Teichmüller’s Principle, every almost disjoint family \( \mathcal{A} \subseteq \{ x' : x \in [\omega]^\omega \} \) can be extended to a mad family \( \mathcal{A}' \subseteq \{ x' : x \in [\omega]^\omega \} \).

**Proposition 8.15.** If \( \mathcal{H} = \{ A_\xi : \xi < h \} \) is a shattering family of cardinality \( h \), then there exists a refining shattering family \( \mathcal{H}' = \{ A_\xi : \xi < h \} \) such that for each \( \xi \in h \) we have \( A_\xi \succ A_\xi \).

**Proof.** The proof is by transfinite induction: Let \( A_\eta := A_0 \) and assume we have already defined \( A_\xi \) for all \( \xi \in \eta \) where \( \eta \in h \). Apply Lemma 8.14 to the family \( \{ A_\xi : \xi \in \eta \} \) to obtain \( A_\eta' \) and let \( \mathcal{H}' = \{ A_\xi' : \xi \in h \} \).

Now, the proof of \( p \leq h \) is straightforward.

**Theorem 8.16.** \( p \leq h \).
Proof. By Proposition 8.15 there exists a refining shattering family $\mathcal{H} = \{\mathcal{A}_\xi : \xi \in h\}$ of cardinality $h$. With $\mathcal{H}$ we shall build a family $\mathcal{F} \subseteq [\omega]^\omega$ of cardinality $h$ which has the sftp but which does not have a pseudo-intersection: Choose any $x_0 \in \mathcal{A}_0$ and assume we have already chosen $x_\xi \in \mathcal{A}_\xi$ for all $\xi \in \eta$ where $\eta \in h$. Since $\mathcal{H}$ is refining we can choose a $x_\eta \in \mathcal{A}_\eta$ such that $x_\eta$ is a pseudo-intersection of $\{x_\xi : \xi \in \eta\}$. Finally let $\mathcal{F} = \{x_\xi : \xi \in h\}$. Then $\mathcal{F}$ is a family of cardinality $\leq h$ which has the sftp, but since $\mathcal{H}$ is shattering, no infinite set is almost contained in every member of $\mathcal{F}$, i.e., $\mathcal{F}$ does not have a pseudo-intersection. \[\blacksquare\]

Summary

The diagram below shows the relations between the twelve cardinals. A line connecting two cardinals indicates that the cardinal lower on the diagram is less than or equal to the cardinal higher on the diagram (provably in ZFC).

\[\begin{array}{c}
\omega_1 \\
| \\
\uparrow \\
h \\
| \\
p \\
| \\
\otimes \\
c \\
| \\
i \\
| \\
hom \\
| \\
a \\
\end{array}\]

Later we shall see that each of following relations is consistent with ZFC:

- $a < c$ (Proposition 18.5)
- $i < c$ (Proposition 18.11)
- $\omega_1 < p = c$ (Proposition 19.1)
• \( a < \delta = \tau \) (Corollary 21.11)
• \( s \equiv b < \delta \) (Proposition 21.13)
• \( \delta < \tau \) (Proposition 22.4)
• \( \delta > \tau \) (Proposition 23.7)
• \( p < h \) (Proposition 24.12)

Notes

Most of the classical cardinal characteristics and their relations presented here can be found for example in van Douwen [42] and Vaughan [43], where one finds also a few historical notes (for \( \delta \) see also Kanamori [27, p. 179 f]). Proposition 8.8 is due to Fichtenholz and Kantorovitch [22], but the proof we gave is Hausdorff’s, who generalised in [26] the result to arbitrary infinite cardinals (see also Exercise (A6) on p. 288 of Kunen [29]). Theorem 8.9 is due to Shelah [33], however, the proof is taken from Blass [5] (see also [4, Theorem 21]), where the claim in the proof is due to Ketenen [28, Proposition 1.3]. Theorem 8.10 and Theorem 8.11 are due to Blass and the proofs are taken from Blass [5] (see also [4, Section 6]). The shattering cardinal \( h \) was introduced and investigated by Balcar, Pelant, and Simon in [2] (cf. Related Result 51).

Related Results

50. The Continuum Hypothesis. There are numerous statements from areas like Algebra, Combinatorics, or Topology, which are equivalent to CH. For example Erdős and Kakutani showed that CH is equivalent to the statement that \( R \) is the union of countably many sets of rationaly independent numbers (cf. [20, Theorem 2]). Many more equivalents to CH can be found in Sierpiński [39]. For the historical background of CH we refer the reader to Felgner [21].

51. On the shattering number \( h \). Balcar, Pelant, and Simon showed that \( h \leq \text{cf}(c) \) (see [2, Theorem 4.2]), gave a direct prove for \( h \leq b \) (see [2, Theorem 4.5]) and for \( h \leq s \) (follows from [2, Lemma 2.11(c)]), and showed that \( h \) is regular (see [2, Lemma 2.11(b)]). Furthermore, Lemma 2.11(c) of Balcar, Pelant, and Simon [2] states that there are shattering families of size \( h \) which have a very strong combinatorial property:

Base Matrix Lemma. There exists a shattering family \( \mathcal{H} = \{ \mathcal{A}_\xi : \xi \in h \} \) which has the property that for each \( X \in [\omega]^\omega \) there is a \( \xi \in h \) and an \( A \in \mathcal{A}_\xi \) such that \( A \subseteq^* X \).

Proof. Let \( \mathcal{F} = \{ \mathcal{A}_\xi : \xi \in h \} \) be an arbitrary but fixed refining shattering family of cardinality \( h \). We first prove the following

Claim. For every infinite set \( X \in [\omega]^\omega \) there exists an ordinal \( \xi \in h \) such that \( |\{ C \in \mathcal{A}_\xi : |C \cap X| = \omega \}| = c \).

Proof of Claim. Let \( X \in [\omega]^\omega \) be an arbitrary infinite subset of \( \omega \). Firstly we show that there exists a strictly increasing sequence \( (\xi_n : n \in \omega) \) in \( h \), such that for each \( n \in \omega \) and \( f \in {}^n \omega \) we find a set \( C_f \in \mathcal{A}_{\xi_n} \) with the following properties:
Related Results

- $|C_f \cap X| = \omega$,
- if $f, f' \in \omega^2$ are distinct, then $C_f \neq C_{f'}$, and
- for all $f \in \omega^2$ and $m \in n$, $C_f \subseteq C_{f^m}$.

The sequence $(\xi_n : n \in \omega)$ is constructed by induction on $n$: First we choose an arbitrary $\xi_0 \in \mathfrak{b}$. Now, suppose we have already found $\xi_n \in \mathfrak{b}$ for some $n \in \omega$. Since $\mathcal{F}$ is a shattering family, for every $h \in \omega^2$ there exists a $\xi_n > \xi_n$ such that the infinite set $C_h \cap X$ has infinite intersection with at least two members of $\mathcal{A}_{\mathfrak{b}_n}$. Let $\xi_{n+1} = \bigcup \{ \xi_h : h \in \omega^2 \}$ Then, since $\mathcal{F}$ is refining, we find a family $\{C_f : f \in \omega^{n+2}\} \subseteq \mathcal{A}_{\mathfrak{b}_{n+2}}$ with the desired properties.

Let $\xi^* := \bigcup_{\xi \in \mathfrak{b}} \xi_n$; then the ordinal $\xi^*$ is smaller than $\mathfrak{b}$. Otherwise, since $\mathcal{F}$ is refining, the family $\{\mathcal{A}_{\xi_n} : n \in \omega\}$ would be a shattering family of cardinality $\omega$, contradicting the fact that $\mathfrak{b} \geq \omega$.

By construction, for each $f \in \omega^2$ we find a $C_f \in \mathcal{A}_\mathfrak{b}$ such that $C_f \cap X$ is infinite (notice that for each $n \in \omega$, $|C_{f^m} \cap X| = \omega$), and since $\mathcal{F}$ is refining we have $C_f \neq C_{f'}$ whenever $f, f' \in \omega^2$ are distinct. Thus, $\{|C_f : f \in \omega^2\} = \mathfrak{c}$ and for each $f \in \omega^2$ we have $|C_f \cap X| = \omega$.

Now we construct the shattering family $\mathcal{H} = \{\mathcal{A}_\xi : \xi \in \mathfrak{b}\}$ as follows: For each $\xi \in \mathfrak{b}$, let $\mathcal{A}_\xi$ be the family of all $X \in [\omega]^{\omega^2}$ such that

$$\{|C \in \mathcal{A}_\xi : |C \cap X| = \omega\} = \mathfrak{c}.$$ 

If $\mathcal{A}_\xi = \emptyset$, then let $\mathcal{A}_\xi = \mathcal{A}_{\mathfrak{b}}$. Otherwise, define (e.g., by transfinite induction) an injection $g_\xi : \mathcal{A}_\xi \rightarrow \mathcal{A}_\mathfrak{b}$ such that for each $X \in \mathcal{A}_\xi$, $|X \cap g_\xi(X)| = \omega$.

Now, for each $C \in \mathcal{A}_\mathfrak{b}$, let $\mathcal{C}_C \subseteq [\omega]^{\omega}$ be an almost disjoint family such that $\bigcup \mathcal{C}_C = C$, and whenever $C = g_\xi(X)$ for some $X \in \mathcal{A}_\xi$ (i.e., $|X \cap C| = \omega$), then there exists an $A \in \mathcal{C}_C$ with $A \subseteq X$. Let $\mathcal{A}_\xi := \{A \in \mathcal{C}_C : C \in \mathcal{A}_\xi\}$ and let $\mathcal{H} := \{\mathcal{A}_\xi : \xi \in \mathfrak{b}\}$. Then, by construction, for every $X \in [\omega]^{\omega^2}$ we find an ordinal $\xi \in \mathfrak{b}$ and an infinite set $A \in \mathcal{A}_\xi$ such that $A \subseteq X$.

52. The tower number $\mathfrak{t}$. A family $\mathcal{F} = \{T_\alpha : \alpha < \kappa\} \subseteq [\omega]^{\omega^2}$ is called a tower if $\mathcal{F}$ is well-ordered by $\supseteq$ (i.e., $T_\alpha \subseteq T_\beta \Rightarrow \alpha < \beta$) and does not have a pseudo-intersection. The tower number $\mathfrak{t}$ is the smallest cardinality (or height) of a tower. Obviously we have $p \leq \mathfrak{t}$ and the proof of Theorem 8.16 shows that $\mathfrak{t} \leq \mathfrak{b}$. However, it is open whether $p < \mathfrak{t}$ is consistent with ZFC (for partial results see for example van Douwen [42], Blass [5], or Shelah [35]).

53. A linearly ordered subset of $[\omega]^{\omega^2}$ of size $\mathfrak{c}$. Let $\{q_n \in \mathbb{Q} : n \in \omega\}$ be an enumeration of the rational numbers $\mathbb{Q}$ and for every real number $r \in \mathbb{R}$ let $C_r := \{n \in \omega : q_n \leq r\}$. Then, for any real numbers $r_0 < r_1$ we have $C_{r_0} \subseteq C_{r_1}$ and $|C_{r_1} \setminus C_{r_0}| = \omega$. Thus, with respect to the ordering “$\subseteq$”, $\{C_r : r \in \mathbb{R}\} \subseteq [\omega]^{\omega^2}$ is a linearly ordered set of size $\mathfrak{c}$. In general one can show that whenever $M$ is infinite, the partially ordered set $(\mathcal{P}(M), \subseteq)$ contains a linearly ordered subset of size strictly greater than $|M|$.

54. The $\sigma$-reaping number $\tau_{\sigma}$. A family $\mathcal{R} \subseteq [\omega]^{\omega^2}$ is called $\sigma$-reaping if no countably many sets suffice to split all members of $\mathcal{R}$. The $\sigma$-reaping number $\tau_{\sigma}$ is the smallest cardinality of any $\sigma$-reaping family (for a definition of $\tau_{\sigma}$ in terms of bounded sequences see Vojtěš [44]). Obviously we have $\tau \leq \tau_{\sigma}$, but it is not known whether $\tau = \tau_{\sigma}$ is provable in ZFC, i.e., it is not known whether $\tau < \tau_{\sigma}$ is consistent with ZFC (see also Vojtěš [44] and Brendle [8]).
55. On \( i \) and \( \text{hom}^* \). We have seen that \( \max \{ \tau, \delta \} \leq \text{hom}^* \) (see Theorem 8.10) and that \( \max \{ \tau, \delta \} \leq i \) (see Theorem 8.9). Moreover, Blass [4, Section 6] showed that \( \text{hom}^* = \max \{ \tau, \delta \} \) (see also Blass [5]). Thus, in every model in which \( \tau = \tau_i \) we have \( \text{hom}^* \leq i \). Furthermore, one can show that \( \text{hom}^* < i \) is consistent with ZFC. In Balcar, Hernández-Hernández, and Hrušák [1] it is shown that \( \max \{ \tau, \text{cof} (\mathcal{M}) \} \leq i \), where \( \text{cof} (\mathcal{M}) \) is the cofinality of the ideal of meagre sets. On the other hand, it is possible to construct models in which \( \delta = \tau_i = \omega_1 \) and \( \text{cof} (\mathcal{M}) = \omega_0 = \omega \) (see for example Shelah and Zapletal [36] or Brendle and Khomskii [15]). Thus, in such models we have \( \omega_1 = \text{hom}^* < i = \omega_2 \). However, it is open whether \( i < \text{hom}^* \) (which would imply \( \tau < \tau_i \)) is consistent with ZFC.

56. The ultrafilter number \( u \). A family \( \mathcal{F} \subseteq [\omega]^\omega \) is a base for an ultrafilter \( \mathcal{U} \subseteq [\omega]^\omega \) if \( \mathcal{U} = \{ y \in [\omega]^\omega : \exists x \in \mathcal{F} (x \subseteq y) \} \). The ultrafilter number \( u \) is the smallest cardinality of any ultrafilter base. We leave it as an exercise to the reader to show that \( \tau \leq u \).

57. Consistency results. The following statements are consistent with ZFC:

- \( \tau < u \) (cf. Goldstern and Shelah [23])
- \( u < \delta \) (cf. Blass and Shelah [6] or see Chapter 23 | Related Result 130)
- \( u < a \) (cf. Shelah [34], see also Brendle [13])
- \( b < \text{pr} \) (cf. Shelah [32, Theorem 5.2] or Dow [19, Proposition 2.7])
- \( \text{hom} < \epsilon \) (see Chapter 23 | Related Result 138)
- \( \delta < a \) (cf. Shelah [34], see also Brendle [10])
- \( \omega_1 = b < a = s = \delta = \omega_2 \) (cf. Shelah [32, Sections 1 & 2])
- \( \gamma = b = a < s = \lambda \) for any regular uncountable cardinals \( \kappa < \lambda \) (cf. Brendle and Fischer [14])
- \( b = \kappa < \kappa^+ = a = c \) for \( \kappa > \omega_1 \) (cf. Brendle [7])
- \( \omega_1 = s < b = \delta = \tau = a = \omega_2 \) (cf. Shelah [32, Section 4])
- \( \text{cf} (\alpha) = \omega \) (cf. Brendle [11])
- \( b = \omega_1 + \text{there are no towers of height} \omega_2 \) (cf. Dordal [17]).

Some more results can be found for example in Blass [3], Brendle [9, 12], van Douwen [42], Dow [19], and Dordal [18].

58. Combinatorial properties of maximal almost disjoint families. An uncountable set of reals is a \( \sigma \)-set if every relative Borel subset is a relative \( G_\delta \) set. Brendle and Piper showed in [16] that CH implies the existence of a MAD family which is also a \( \sigma \)-set (in that paper, they also discuss related results assuming Martin’s Axiom).

59. Applications to Banach space theory. Let \( \ell_p (\kappa) \) denote the Banach space of bounded functions \( f : \kappa \to \mathbb{R} \) with finite \( \ell_p \)-norm, where for \( 1 \leq p < \infty \),

\[
\|f\| = \left( \sum_{\alpha < \kappa} |f(\alpha)|^p \right)^{\frac{1}{p}},
\]

and for \( p = \infty \),

\[
\|f\| = \sup \{ |f(\alpha)| : \alpha \in \kappa \}.
\]

As mentioned above, Hausdorff generalised Proposition 8.8 to arbitrary infinite cardinals \( \kappa \), i.e., if \( \kappa \) is an infinite cardinal then there are independent families on \( \kappa \) of cardinality \( 2^\kappa \). Now, using independent families on \( \kappa \) of cardinality \( 2^\kappa \) it is quite straightforward to show that \( \ell_\infty (\kappa) \) contains an isomorphic
copy of $\ell_1(\kappa)$ (the details are left to the reader), and Halbeisen [24] showed that the dual of $\ell_\infty(\kappa)$ contains an isomorphic copy of $\ell_2(\kappa^*)$ (for an analytic proof in the case $\kappa = \omega$ see Rosenthal [31, Proposition 3.4]).

We have seen that there are almost disjoint families on $\omega$ of cardinality $\kappa = \kappa^\omega$.

Unlike for independent families, this result cannot be generalised to arbitrary cardinals $\kappa$, i.e., it is consistent with ZFC that for some infinite $\kappa$, there no almost disjoint family on $\kappa$ of cardinality $\kappa^\omega$ (see Baumgartner[3, Theorem 5.6 (b)]).

However, one can prove that for all infinite cardinals $\kappa$ there is an almost disjoint family on $\kappa$ of cardinality $> \kappa$ (cf. Tarski [41], Sierpiński [37, 38] or [40, p. 448 f.], or Baumgartner [3, Theorem 2.8]). Using an almost disjoint family of cardinality $> \kappa$ it is not hard to show that every infinite dimensional Banach space of cardinality $\kappa$ has more than $\kappa$ pairwise almost disjoint normalised Hamel bases (cf. Halbeisen [25]), and Pelczyński and Sudyłov [30] showed that $c_0(\kappa)$, which is a subspace of $\ell_\infty(\kappa)$, is not complemented in $\ell_\infty(\kappa)$.

References


11. ______. The almost-disjointness number may have countable cofinality, Transactions of the American Mathematical Society, vol. 355 (2003), 2633–2649.

The Shattering Number revisited

As variety brings pleasure and delight, so excessive repetition generates boredom and annoyance. Besides, the composer would be thought by connoisseurs of the art to have a meagre store of ideas. But it is not only permitted but admirable to duplicate a passage or melody as many times as one wishes if the counterpoint is always different and varied. For such repetitions strike us as being somehow ingenious, and we should try to write them wherever they seem suitable.

Giuseppe Zarlino

*Le istituzioni harmoniche*, 1568

In this chapter we shall have a closer look at the shattering number $\mathfrak{h}$. In the preceding chapter, $\mathfrak{h}$ was introduced as the minimum height of a shattering matrix. However, like other cardinal characteristics, $\mathfrak{h}$ has different facets. In this chapter we shall see that $\mathfrak{h}$ is closely related to the Ramsey property, a combinatorial property of subsets of $\omega$ (discussed at the end of Chapter 2) which can be regarded as a generalisation of Ramsey’s Theorem.

The Ramsey Property

By Ramsey’s Theorem 2.1, for every 2-colouring of $[\omega]^2$ there is a homogeneous set; on the other hand we have seen that there are 2-colourings of $[\omega]^\omega$ without a homogeneous set (see the example given in Chapter 2). Obviously, every colouring $\pi : [\omega]^\omega \to \{0, 1\}$ induces a set $C_\pi \subseteq [\omega]^\omega$ by stipulating $C_\pi = \{ x \in [\omega]^\omega : \pi(x) = 1 \}$.

By identifying 2-colourings of $[\omega]^\omega$ with subsets of $[\omega]^\omega$, the existence of a 2-colouring of $[\omega]^\omega$ without a homogeneous set is equivalent to the existence
of a set \( C \subseteq \omega^\omega \) such that for all \( x \in \omega^\omega \) there are \( y_0, y_1 \in [x]^\omega \) such that \( y_0 \notin A \) and \( y_1 \in A \).

Now, a set \( C \subseteq \omega^\omega \) has the **Ramsey property**, if there exists a set \( x \in \omega^\omega \) such that either \([x]^\omega \subseteq C\) or \([x]^\omega \cap C = \emptyset\). Notice that the finite as well as the co-finite subsets of \( \omega^\omega \) have the Ramsey property, but notice also that not all subsets of \( \omega^\omega \) have the Ramsey property (cf. Chapter 5 | Related Result 38).

Below, we investigate a property of subsets of \( \omega^\omega \) which is slightly stronger than the Ramsey property, but first we have to introduce the following notation.

For a finite set \( s \in \text{fin}(\omega) \) and an infinite set \( x \in \omega^\omega \) such that \( \max(s) < \min(x) \) (i.e., \( \bigcup s \prec \bigcap x \)), let

\[
[s, x]^\omega = \{ z \in \omega^\omega : s \subseteq z \subseteq s \cup x \}.
\]

Now, a set \( C \subseteq \omega^\omega \) is called **completely Ramsey** if for every set \( [s, x]^\omega \) there is a \( y \in [x]^\omega \) such that either \([s, y]^\omega \subseteq C\) or \([s, y]^\omega \cap C = \emptyset\). If we are always in the latter case (i.e., for each \( [s, x]^\omega \) there is a \( y \in [x]^\omega \) such that \([s, y]^\omega \cap C = \emptyset\)), then \( C \) is called **completely Ramsey-null**. In particular, for \( s = \emptyset \) and \( x = \omega \) we conclude that any completely Ramsey set has the Ramsey property. On the other hand, not every set which has the Ramsey property is completely Ramsey (we leave it as an exercise to the reader to find a counterexample).

The proof of the following result uses a so-called **fusion argument**, a technique which we will meet again in Part III (Lemma 9.1 itself is used in the proof of Theorem 9.2).

**Lemma 9.1.** If \( C \subseteq \omega^\omega \) is completely Ramsey-null, then for each \( x \in \omega^\omega \) there is a \( y \in [x]^\omega \) such that \( C \) contains no infinite set \( z \subseteq^* y \).

**Proof.** Let \( C \) be completely Ramsey-null and \( x \in \omega^\omega \) be arbitrary. By definition of completely Ramsey-null there is a \( y_0 \in [x]^\omega \) such that \([0, y_0]^\omega \cap C = \emptyset\) and let \( a_0 = \min(y_0) \). Assume we have already constructed a sequence \( x \supseteq y_0 \supseteq \ldots \supseteq y_n \) of infinite subsets of \( \omega \) as well as a sequence \( a_0 < \ldots < a_n \) of natural numbers such that for all \( s \in \mathcal{P}(a_{n-1} + 1) \),

\[
[s, y_n]^\omega \cap C = \emptyset.
\]

For \( h = 2^{a_{n+1}} + 1 \) let \( \{ s_i : i \in h \} \) be an enumeration of \( \mathcal{P}(a_n + 1) \) where \( s_0 = \emptyset \). Further let \( z_0 = y_n \setminus (a_n + 1) \) and for each \( i \in h \) choose an infinite set \( z_{i+1} \subseteq z_i \) such that \([s_{i+1}, z_{i+1}]^\omega \cap C = \emptyset\) (notice that we can do this because \( C \) is completely Ramsey-null). Finally let \( y_{n+1} = z_{n-1} \); then for all \( s \in \mathcal{P}(a_n + 1) \) we have

\[
[s, y_{n+1}]^\omega \cap C = \emptyset.
\]

Let now \( a_{n+1} = \min(y_{n+1}) \) and start the process again with the sequences \( x \supseteq y_0 \supseteq \ldots \supseteq y_{n+1} \) and \( a_0 < \ldots < a_{n+1} \). At the end we get an infinite
sequence $a_0 < a_1 < \ldots < a_n < \ldots$ and by construction the set $y = \{a_i : i \in \omega\}$ has the property that for each $s \in \text{fin}(\omega)$ with $\max(s) \in y$, 

$$[s, y \setminus (\max(s) + 1)]^\omega \cap C = \emptyset,$$

which implies that for each infinite set $z \subseteq^* y$ we have $[\emptyset, z]^\omega \cap C = \emptyset$, i.e., $C$ contains no infinite set $z \subseteq^* y$.

### The Ideal of Ramsey-Null Sets

Below, we consider the set of completely Ramsey-null sets. So, let 

$$\mathcal{R}_0 = \{C \subseteq [\omega]^\omega : C \text{ is completely Ramsey-null}\}$$

be the collection of all subsets of $[\omega]^\omega$ which are completely Ramsey-null. Since $\mathcal{R}_0$ is closed under subsets (i.e., $C \in \mathcal{R}_0$ and $C' \subseteq C$ implies $C' \in \mathcal{R}_0$) and finite unions (i.e., $C_0, \ldots, C_n \in \mathcal{R}_0$ implies $C_0 \cup \ldots \cup C_n \in \mathcal{R}_0$), $\mathcal{R}_0$ is an ideal on $\mathcal{P}([\omega]^\omega)$.

Obviously, $[\omega]^\omega \notin \mathcal{R}_0$ but for every $x \in [\omega]^\omega$ we have $\{x\} \in \mathcal{R}_0$. Thus, the set $[\omega]^\omega$ can be covered by $\mathfrak{c}$ completely Ramsey-null sets which implies that the union of $\mathfrak{c}$ sets from $\mathcal{R}_0$ can be a set which does not belong to $\mathcal{R}_0$. These observations lead to the following two cardinal numbers.

**Definition.** The *additivity* of $\mathcal{R}_0$, denoted $\text{add}(\mathcal{R}_0)$, is the smallest number of sets in $\mathcal{R}_0$ with union not in $\mathcal{R}_0$; more formally

$$\text{add}(\mathcal{R}_0) = \min \{|\mathcal{C}| : \mathcal{C} \subseteq \mathcal{R}_0 \land \bigcup\mathcal{C} \notin \mathcal{R}_0\}.$$ 

**Definition.** The *covering number* of $\mathcal{R}_0$, denoted $\text{cov}(\mathcal{R}_0)$, is the smallest number of sets in $\mathcal{R}_0$ with union $[\omega]^\omega$; more formally

$$\text{cov}(\mathcal{R}_0) = \min \{|\mathcal{C}| : \mathcal{C} \subseteq \mathcal{R}_0 \land \bigcup\mathcal{C} = [\omega]^\omega\}.$$ 

We leave it as an exercise to the reader to show (using a fusion argument) that any countable union of completely Ramsey-null sets is completely Ramsey-null. Hence, $\omega_1 \leq \text{add}(\mathcal{R}_0)$, and consequently we get $\omega_1 \leq \text{add}(\mathcal{R}_0) \leq \text{cov}(\mathcal{R}_0) \leq \mathfrak{c}$. Moreover, we even have the following result.

**Theorem 9.2.** $\text{add}(\mathcal{R}_0) = \text{cov}(\mathcal{R}_0) = \mathfrak{h}.$

**Proof.** Because $\text{add}(\mathcal{R}_0) \leq \text{cov}(\mathcal{R}_0)$ it is enough to show that $\text{cov}(\mathcal{R}_0) \leq \mathfrak{h}$ and that $\mathfrak{h} \leq \text{add}(\mathcal{R}_0)$.

$\text{cov}(\mathcal{R}_0) \leq \mathfrak{h}$: Let $\{\mathcal{A}_\xi : \xi \in \mathfrak{h}\}$ be a shattering family of cardinality $\mathfrak{h}$. For each $\xi \in \mathfrak{h}$ let $D_\xi = \{y \in [\omega]^\omega : \exists x \in \mathcal{A}_\xi (y \subseteq^* x)\}$ and let $C_\xi = [\omega]^\omega \setminus D_\xi$. Firstly notice that for each $\xi \in \mathfrak{h}$, $C_\xi \in \mathcal{R}_0$. Indeed, take any $[s, y]^\omega$, then, since $\mathcal{A}_\xi$ is mad, there is an $x \in \mathcal{A}_\xi$ such that $y \cap x$ is infinite; thus, $[s, y \cap x]^\omega \subseteq D_\xi$, or
equivalently \([s, y \cap x]^{\omega} \cap C_{\xi} = \emptyset\). Secondly notice that \(\bigcup_{\xi \in h} C_{\xi} = [\omega]^{\omega}\). Indeed, take any \(y \in [\omega]^{\omega}\), then, since \(\{\mathcal{A}_\xi : \xi \in h\}\) is shattering, there is a \(\xi \in h\) and two distinct elements \(x, x' \in \mathcal{A}_\xi\) such that \(y \cap x\) as well as \(y \cap x'\) is infinite; hence, \(y \notin D_{\xi}\), or equivalently \(y \in C_{\xi}\).

\(h \leq \text{add}(\mathcal{R}_0)\): Let \(\{C_{\xi} \subseteq [\omega]^{\omega} : \xi \in \kappa < h\} \subseteq \mathcal{R}_0\) be a family of completely Ramsey-null sets of cardinality \(\kappa < h\). We will show that \(\bigcup_{\xi \in \kappa} C_{\xi} \in \mathcal{R}_0\). For each \(\xi \in \kappa\) let

\[
D_{\xi} = \{y \in [\omega]^{\omega} : \forall z \in [\omega]^{\omega} (z \subseteq^* y \rightarrow [0, z]^{\omega} \cap C = \emptyset)\}.
\]

Now we choose for each \(\xi \in \kappa\) an almost disjoint family \(\mathcal{A}_\xi \subseteq D_{\xi}\) of cardinality \(c\) which is maximal with respect to inclusion. Notice that by Lemma 9.1, for each \(x \in [\omega]^{\omega}\) there is a \(y \in \mathcal{A}_\xi\) such that \(x \cap y\) is infinite, i.e., \(\mathcal{A}_\xi \subseteq D_{\xi}\) is a mad family (on \([\omega]^{\omega}\)) of cardinality \(c\). Indeed, if there would be an \(x \in [\omega]^{\omega}\) \(\mathcal{A}_\xi\) which has finite intersection with each member of \(\mathcal{A}_\xi\), then, by Lemma 9.1, there is a \(y \in [x]^{\omega}\) such that \(y \in D_{\xi} \setminus \mathcal{A}_\xi\) which would imply that \(\mathcal{A}_\xi\) is not maximal. Because \(\kappa < h\) we can apply Lemma 8.14 and get a mad family \(\mathcal{A}'\) which refines each \(\mathcal{A}_\xi\). Take any set \([s, x]^{\omega}\). Since \(\mathcal{A}'\) is mad, there is a \(y' \in \mathcal{A}'\) such that \(x \cap y'\) is infinite; let \(z = x \cap y'\). Because \(\mathcal{A}'\) refines all \(\mathcal{A}_\xi\)'s, for each \(\xi \in \kappa\) there is a \(y \in \mathcal{A}_\xi\) such that \(z \subseteq^* y\), and since \(\mathcal{A}_\xi \subseteq D_{\xi}\), by definition of \(D_{\xi}\) we get \([0, s \cup z]^{\omega} \cap C_{\xi} = \emptyset\), in particular, \([s, z]^{\omega} \cap C_{\xi} = \emptyset\). Thus, for every set \([s, x]^{\omega}\) there exists a \(z \in [x]^{\omega}\) such that for all \(\xi \in \kappa\), \([s, z]^{\omega} \cap C_{\xi} = \emptyset\), i.e., \([s, z]^{\omega} \cap \bigcup_{\xi \in \kappa} C_{\xi} = \emptyset\), hence \(\bigcup_{\xi \in \kappa} C_{\xi} \in \mathcal{R}_0\).

The Ellentuck Topology

Below, we give a topological characterisation of completely Ramsey sets, but before we have to introduce the basic notions of General Topology:

A topological space is a pair \((X, \mathcal{O})\) consisting of a set \(X\) and a family \(\mathcal{O}\) of subsets of \(X\) satisfying the following conditions:

(O1) \(\emptyset \in \mathcal{O}\) and \(X \in \mathcal{O}\).

(O2) If \(O_1 \in \mathcal{O}\) and \(O_2 \in \mathcal{O}\), then \(O_1 \cap O_2 \in \mathcal{O}\).

(O3) If \(\mathcal{F} \subseteq \mathcal{O}\), then \(\bigcup \mathcal{F} \in \mathcal{O}\).

The set \(X\) is called a space, the elements of \(X\) are called points of the space, and the subsets of \(X\) belonging to \(\mathcal{O}\) are called open and the complements of open sets are called closed. The family \(\mathcal{O}\) of open subsets of \(X\) is also called a topology on \(X\).

Let us consider for example the real line \(\mathbb{R}\). For \(r_1, r_2 \in \mathbb{R}\) define \((r_1, r_2) := \{r \in \mathbb{R} : r_1 < r < r_2\}\). Now, a set \(O \subseteq \mathbb{R}\) is called open if for every \(r \in O\) there exists a real \(\varepsilon > 0\) such that \((r - \varepsilon, r + \varepsilon) \subseteq O\) (i.e., every \(r \in O\) is contained in an open interval contained in \(O\)). We leave it as an exercise to the reader to show that the family of open sets satisfies conditions (O1)-(O3).

From (O2) it follows that the intersection of any finite family of open sets is an open set, and from (O3) it follows that the union of any family of open
The Ellentuck topology

sets is open. Notice that arbitrary intersections of closed sets as well as finite unions of closed sets are closed sets. For an arbitrary set \( A \subseteq X \) let

\[ A^* = \bigcup \{ O \in \mathcal{O} : O \subseteq A \} \]

be the interior of \( A \); and let

\[ \bar{A} = \bigcap \{ C : C \text{ is closed and } A \subseteq C \} \]

be the closure of \( A \). Notice that \( A^* \) is the largest open set contained in \( A \) and that \( A \) is the smallest closed set containing \( A \).

A family \( \mathcal{B} \subseteq \mathcal{O} \) is called a base for a topological space \( (X, \mathcal{O}) \) if every non-empty open subset of \( X \) can be represented as the union of a subfamily of \( \mathcal{B} \). The sets in a basis \( \mathcal{B} \) are also called basic open sets. If a family \( \mathcal{B} \) of subsets of \( X \) is such that \( X \in \mathcal{B} \) and every non-empty finite intersection of sets in \( \mathcal{B} \) belongs to \( \mathcal{B} \), then \( (X, \mathcal{O}) \), where

\[ \mathcal{O} = \left\{ \bigcup \mathcal{F} : \mathcal{F} \subseteq \mathcal{B} \right\}, \]

is a topological space with base \( \mathcal{B} \) (notice that \( \bigcup \emptyset = \emptyset \)). In this case we say that the topology on \( X \) is generated by the basic open sets \( O \in \mathcal{B} \).

For example the topology on \( \mathbb{R} \) introduced above is generated by the countably many basic open intervals \( (q_1, q_2) \), where \( q_1, q_2 \in \mathbb{Q} \).

Let \( (X, \mathcal{O}) \) be a topological space and let \( A \subseteq X \) be a subset of \( X \).

- \( A \) is called dense if for every open set \( O \in \mathcal{O} \), \( A \cap O \neq \emptyset \).
- \( A \) is called nowhere dense if \( X \setminus A \) contains an open dense set.
- \( A \) is called meagre if \( A \) is the union of countably many nowhere dense sets.
- \( A \) has the Baire property if there is an open set \( O \in \mathcal{O} \) such that \( O \Delta A \) is meagre, where \( O \Delta A = (O \setminus A) \cup (A \setminus O) \) (i.e., \( x \in A \cap O \) or \( x \notin A \cup O \)).

Obviously, meagre sets and open sets have the Baire property and countable unions of meagre sets are meagre. Moreover, the following result shows that the Baire property is closed under complementation and countable unions and intersections.

**Fact 9.3.** (a) Every closed set has the Baire property.
(b) The complement of a set with the Baire property has the Baire property.
(c) Unions and intersections of countably many sets with the Baire property have the Baire property.

**Proof.** (a) Let \( A \subseteq X \) be a closed subset of \( X \). We shall show that \( A \setminus A^* \) is nowhere dense. Firstly, \( A \setminus A^* = A \cap (X \setminus A^*) \), thus, \( A \setminus A^* \) is closed and \( X \setminus (A \setminus A^*) \) is open. Secondly, no open set \( O \in \mathcal{O} \) is contained in \( A \setminus A^* \), and
therefore \( O \cap (X \setminus (A \setminus A^c)) \) is a non-empty open set. Thus, \( X \setminus (A \setminus A^c) \) is open dense, or equivalently, \( A \setminus A^c \) is nowhere dense. In particular, \( A^c \Delta A \) is meagre which shows that \( A \) has the Baire property.

(b) Assume that \( A \subseteq X \) has the Baire property and let \( O \in \mathcal{B} \) be such that \( O \setminus A \) is meagre. Let \( \tilde{O} := X \setminus (X \setminus O)^c \) be the closure of \( O \). By (a), \( \tilde{O} \setminus O \) is nowhere dense. Thus, \( A \setminus \tilde{O} \) is meagre and therefore \( (X \setminus A) \Delta (X \setminus \tilde{O}) \) is also meagre, which shows that \( X \setminus A \) has the Baire property.

(c) By (b) it is enough to prove (c) for unions. So, let \( \{ A_n \subseteq X : n \in \omega \} \) be a family of sets which have the Baire property. For each \( n \in \omega \) let \( O_n \in \mathcal{B} \) be an open set such that \( O_n \setminus A_n \) is meagre. Then

\[
M = \bigcup_{n \in \omega} O_n \setminus \bigcup_{n \in \omega} A_n \subseteq \bigcup_{n \in \omega} (O_n \setminus A_n)
\]

is a subset of a countable union of meagre sets. Hence, \( M \) is meagre which shows that \( \bigcup_{n \in \omega} A_n \) has the Baire property.

Consider now the set \([\omega]^\omega\). The aim is to define a topology on \([\omega]^\omega\) such that a set \( A \subseteq [\omega]^\omega \) has the Baire property (with respect to that topology) if and only if \( A \) is completely Ramsey. For this let

\[
\mathcal{B} = \{ [s, x]^\omega : s \in \text{fin}(\omega) \wedge x \in [\omega]^\omega \wedge \max(s) < \min(x) \}
\]

where we defined \([s, x]^\omega := \{ z \in [\omega]^\omega : s \subseteq z \subseteq s \cup x \}\). Obviously, \([\omega]^\omega = [\emptyset, \omega]^\omega \in \mathcal{B} \) and we leave it as an exercise to the reader to show that every non-empty finite intersection of sets in \( \mathcal{B} \) belongs to \( \mathcal{B} \) — notice that \([s, x]^\omega \cap [t, y]^\omega \) is either empty or it is \([s \cup t, x \cap y]^\omega \). Thus, \( \mathcal{B} = \{ \bigcup \mathcal{F} : \mathcal{F} \subseteq \mathcal{B} \} \) is a topology on \([\omega]^\omega\), called the Ellentuck topology.

In Chapter 21 we shall introduce a topology on \( ^\omega \omega \) which corresponds to the topology on \([\omega]^\omega \) generated by the basic open sets \([s, \omega \setminus \max(s) + 1]^\omega \).

Notice that with respect to the Ellentuck topology, each singleton set \( \{ x \} \subseteq [\omega]^\omega \) is nowhere dense and all countable sets are meagre. Furthermore, by definition, subsets of meagre sets as well as countable unions of meagre sets are meagre. Thus, the collection of all meagre subsets of \([\omega]^\omega \) is an ideal on \( \mathcal{P}([\omega]^\omega) \). The following theorem shows that the ideal of meagre sets coincide with the ideal of completely Ramsey-null sets, and that a set is completely Ramsey if it has the Baire property; for the latter result we have to prove first the following lemma, whose proof uses twice a fusion argument.

**Lemma 9.4.** Every open set is completely Ramsey.

**Proof.** Firstly we introduce some terminology: Let \( \mathcal{O} \subseteq [\omega]^\omega \) be an arbitrary but fixed open set. A basic open set \([s, x]^\omega \) is called **good** (with respect to \( \mathcal{O} \)), if there is a set \( y \in [x]^\omega \) such that \([s, y]^\omega \subseteq \mathcal{O} \); otherwise it is called **bad**. Further, \([s, x]^\omega \) is called **ugly** if \([s \cup \{a\}, x \setminus a^+\]^\omega \) is bad for all \( a \in x \), where \( a^+ := a + 1 \). Notice that if \([s, x]^\omega \) is ugly, then \([s, x]^\omega \) is bad, too.
Finally, \([s, x]^{\omega}\) is called \textbf{completely ugly} if \([s \cup \{a_0, \ldots, a_n\}, x \setminus a_{n+1}^{+}]^{\omega}\) is bad for all \(\{a_0, \ldots, a_n\} \subseteq x\) with \(a_0 < \ldots < a_n\). If \([s, x]^{\omega}\) is completely ugly, then \([s, x]^{\omega} \cap O = \emptyset\) (notice that \([s, x]^{\omega} \cap O\) is open, and therefore is either empty or contains a basic open set \([t, y]^{\omega} \subseteq [s, x]^{\omega}\).

Now, in order to show that the open set \(O\) is completely Ramsey it is enough to prove that every basic open set \([s, x]^{\omega}\) is either good or there exists a \(z \in [x]^{\omega}\) such that \([s, z]^{\omega}\) is completely ugly. This is done in two steps: Firstly we show that if \([s, x]^{\omega}\) is bad, then there exists a \(y \in [x]^{\omega}\) such that \([s, y]^{\omega}\) is ugly, and secondly we show that if \([s, y]^{\omega}\) is ugly, then there exists a \(z \in [y]^{\omega}\) such that \([s, z]^{\omega}\) is completely ugly.

**Claim 1.** \textbf{If the basic open set} \([s, x]^{\omega}\) \textbf{is bad, then there exists a set} \(y \in [x]^{\omega}\) \textbf{such that} \([s, y]^{\omega}\) \textbf{is ugly.}

**Proof of Claim 1.** Let \(x_0 := x\) and \(a_0 := \min(x_0)\), and for \(i \in \omega\) let \(x_{i+1} \subseteq (x_i \setminus a_i^+)\) such that \([s \cup \{a_i\}, x_{i+1}]^{\omega} \subseteq O\) if possible, and \(x_{i+1} = x_i \setminus a_i^+\) otherwise. Further, let \(a_{i+1} := \min(x_{i+1})\). Strictly speaking we assume that \([\omega]^{\omega}\) is well-ordered and that \(x_{i+1}\) is the first element of \([\omega]^{\omega}\) with the required properties. Now, let \(y = \{a_i : s \cup \{a_i\}, x_{i+1}^{\omega} \subseteq O\}\). Because \([s, x]^{\omega}\) is bad, \(y \in [\omega]^{\omega}\), which implies that \([s, y]^{\omega}\) is ugly.

**Claim 2.** \textbf{If the basic open set} \([s, y]^{\omega}\) \textbf{is ugly, then there exists a set} \(z \in [y]^{\omega}\) \textbf{such that} \([s, z]^{\omega}\) \textbf{is completely ugly.}

**Proof of Claim 2.** This follows by an iterative application of Claim 1. Let \(y_0 := y\) and let \(a_0 := \min(y_0)\). For every \(i \in \omega\) we can choose a set \(y_{i+1} \subseteq (y_i \setminus a_i^+)\), where \(a_i := \min(y_i)\), such that for each \(i \in \omega\), \(y_{i+1} \subseteq (y_i \setminus a_i^+)\) and we have either \([s \cup t, y_{i+1}]^{\omega}\) is ugly or \([s \cup t, y_{i+1}]^{\omega} \subseteq O\). Let \(z := \{a_i : i \in \omega\}\) and assume towards a contradiction that there exists a finite set \(t \subseteq z\) such that \([s \cup t, z \setminus \max(t)^{+}]^{\omega}\) is good. Notice that since \([s, y]^{\omega}\) was assumed to be ugly, \(t \neq \emptyset\). Now, let \(t_0\) be a smallest finite subset of \(z\) such that \([s \cup t_0, z \setminus \max(t_0)^{+}]^{\omega}\) is good and let \(t_0^+ = t_0 \setminus \{\max(t_0)\}\). By definition of \(t_0\), \([s \cup t_0, z \setminus \max(t_0)^{+}]^{\omega}\) cannot be good (i.e., it is bad), and therefore, by construction of \(z\), it must be ugly. On the other hand, if \([s \cup t_0, z \setminus \max(t_0)^{+}]^{\omega}\) is ugly, then \([s \cup t_0, z \setminus \max(t_0)^{+}]^{\omega}\) is bad, which is a contradiction to our assumption that \([s \cup t_0, z \setminus \max(t_0)^{+}]^{\omega}\) is good. Thus, for all finite subsets \(t \subseteq z\), \([s \cup t, z \setminus \max(t)^{+}]^{\omega}\) is ugly, and therefore \([s, z]^{\omega}\) is completely ugly.

Let \([s, x]^{\omega}\) be an arbitrary basic open set. If \([s, x]^{\omega}\) is good, then there exists a \(y \in [x]^{\omega}\) such that \([s, y]^{\omega}\) \(\subseteq O\). Otherwise, \([s, x]^{\omega}\) is bad and we find a \(z \in [x]^{\omega}\) such that \([s, z]^{\omega}\) is completely ugly, i.e., \([s, z]^{\omega} \cap O = \emptyset\). Hence, the arbitrary open set \(O\) is completely Ramsey.

We shall use the very same fusion arguments again in Chapter 24 in order to prove that Mathias forcing has pure decision (see proof of Theorem 24.3).

**Theorem 9.5 (Ellentuck).** \textbf{For every} \(A \subseteq [\omega]^{\omega}\) \textbf{we have:}

(a) \(A\) \textbf{is nowhere dense if and only if} \(A\) \textbf{is completely Ramsey-null.}
(b) $A$ is meagre if and only if $A$ is nowhere dense.

(c) $A$ has the Baire property if and only if $A$ is completely Ramsey.

Proof. (a) A set $A \subseteq [\omega]^\omega$ is nowhere dense iff for each basic open set $[s, x]^\omega$ there exists a basic open set $[t, y]^\omega \subseteq [s, x]^\omega$ such that $[t, y]^\omega \cap A = \emptyset$. Hence, we obviously have that every completely Ramsey-null set is nowhere dense. For the other direction assume that $A \subseteq [\omega]^\omega$ is not completely Ramsey-null, i.e., there is a basic open set $[s, x]^\omega$ such that for all basic open sets $[s, y]^\omega \subseteq [s, x]^\omega$ we have $[s, y]^\omega \cap A \neq \emptyset$. By a fusion argument we can construct a set $z_0 \in [x]^\omega$ such that for all $[t, y]^\omega \subseteq [s, z_0]^\omega$ we have $[t, y]^\omega \cap A \neq \emptyset$, i.e., $A$ is not nowhere dense.

(b) On the one hand, nowhere dense sets are meagre. On the other hand, by Theorem 9.2 we have $\text{add}(\mathcal{R}_0) = \mathfrak{b}$ and since $\mathfrak{b}$ is uncountable we get that countable unions of completely Ramsey-null sets (i.e., of nowhere dense sets) are completely Ramsey-null. Thus, meagre sets are completely Ramsey-null and therefore nowhere dense.

(c) On the one hand, if $A \subseteq [\omega]^\omega$ is completely Ramsey, then $\mathcal{O} = \bigcup \{[s, y]^\omega : [s, y]^\omega \subseteq A\}$ is an open subset of $A$ and for each basic open set $[s, x]^\omega$ there is a $y \in [x]^\omega$ such that either $[s, y]^\omega \subseteq A$ (i.e., $[s, y]^\omega \subseteq (A \cap \mathcal{O})$ and in particular $[s, y]^\omega \cap (\mathcal{O} \triangle A) = \emptyset$) or $[s, y]^\omega \cap A = \emptyset$ (i.e., $[s, y]^\omega \cap (A \cup \mathcal{O}) = \emptyset$ and in particular $[s, y]^\omega \cap (\mathcal{O} \triangle A) = \emptyset$). In both cases we have $[s, y]^\omega \cap (\mathcal{O} \triangle A) = \emptyset$ which implies that $\mathcal{O} \triangle A$ is meagre and shows that $A$ has the Baire property.

On the other hand, if $A \subseteq [\omega]^\omega$ has the Baire property then there is an open set $\mathcal{O} \subseteq [\omega]^\omega$ such that $\mathcal{O} \triangle A$ is meagre, thus by (b), $\mathcal{O} \triangle A$ is completely Ramsey-null. Now, $\mathcal{O} \triangle A \in \mathcal{R}_0$ iff for each basic open set $[s, y]^\omega$ there is a $z \in [y]^\omega$ such that $[s, z]^\omega \cap (\mathcal{O} \triangle A) = \emptyset$. Because $\mathcal{O}$ is completely Ramsey (by Lemma 9.4), for every basic open set $[s, x]^\omega$ there is a set $y \in [x]^\omega$ such that either $[s, y]^\omega \subseteq \mathcal{O}$ or $[s, y]^\omega \cap \mathcal{O} = \emptyset$, and in both cases there is a $z \in [y]^\omega$ such that $[s, z]^\omega \cap (\mathcal{O} \triangle A) = \emptyset$. Thus, we have either $[s, z]^\omega \subseteq A$ or $[s, z]^\omega \cap A = \emptyset$, which shows that $A$ is completely Ramsey.

As a consequence we get the following

Corollary 9.6. The union of less than $\mathfrak{b}$ completely Ramsey sets is completely Ramsey.

Proof. Let $\kappa < \mathfrak{b}$ and let $\{C_\xi \subseteq [\omega]^\omega : \xi \in \kappa\}$ be a family of completely Ramsey sets. For each $\xi \in \kappa$ let $O_\xi \subseteq [\omega]^\omega$ be an open set such that $O_\xi \triangle C_\xi$ is meagre. Then

$$D = \bigcup_{\xi \in \kappa} O_\xi \triangle \bigcup_{\xi \in \kappa} C_\xi \subseteq \bigcup_{\xi \in \kappa} (O_\xi \triangle C_\xi)$$

is a subset of a union of $\kappa$ meagre sets, and since $\kappa < \mathfrak{b}$, $D$ is meagre and therefore $\bigcup_{\xi \in \kappa} C_\xi$ is completely Ramsey.


A generalised Suslin operation

First we introduce an operation on certain families of sets and then we show that the collection of completely Ramsey sets is closed under that operation.

Recall that for arbitrary cardinals $\kappa$, $\text{seq}(\kappa)$ denotes the set of all finite sequences which can be formed with elements of $\kappa$. As usual we identify the set $\text{seq}(\kappa)$ with the set $\bigcup_{n \in \omega} n^\kappa$. Let $\{ Q_s : s \in \text{seq}(\kappa) \}$ be a family of sets indexed by elements of $\text{seq}(\kappa)$ and define

$$A_\kappa \{ Q_s : s \in \text{seq}(\kappa) \} = \bigcup_{f \in \text{seq}(\kappa) \cap n \in \omega} Q_{f|n}.$$ 

The operation $A_\kappa$ is called the Suslin operation.

Now we will show that the collection of completely Ramsey sets (i.e., the collection of sets having the Baire property) is closed under the generalised Suslin operation $A_\kappa$ whenever $\omega \leq \kappa < \aleph_1$, i.e., for every family $\{ Q_s : s \in \text{seq}(\kappa) \}$ of completely Ramsey sets, $A_\kappa \{ Q_s : s \in \text{seq}(\kappa) \}$ is completely Ramsey.

A set $A \subseteq [\omega]^\omega$ is meagre in the basic open set $[s, x]^\omega$ if the intersection $A \cap [s, x]^\omega$ is meagre. Thus, by (a) & (b) of Theorem 9.5, $A$ is meagre in $[s, x]^\omega$ if for every $[t, y]^\omega \subseteq [s, x]^\omega$ there is a $y' \in [y]^\omega$ such that $A \cap [t, y']^\omega = \emptyset$. Now, for an arbitrary but fixed set $A \subseteq [\omega]^\omega$ let

$$M = \bigcup \{ [s, x]^\omega : A \text{ is meagre in } [s, x]^\omega \}.$$ 

The main part of the following lemma is that $A \cup ([\omega]^\omega \setminus M)$ has the Baire property.

**Lemma 9.7.** For $A$ and $M$ as above we have:

(a) $A$ is meagre in each basic open set $[s, x]^\omega \subseteq M$.

(b) $M \cap A$ is meagre.

(c) $A \cup ([\omega]^\omega \setminus M)$ has the Baire property.

**Proof.** (a) Let $[s, x]^\omega \subseteq M$ be an arbitrary basic open subset of $M$ and let

$$N = \{ [t, y]^\omega \subseteq [s, x]^\omega : A \text{ is meagre in } [t, y]^\omega \}.$$ 

Then, by definition of $M$ and since the basic open sets of the Ellentuck topology are closed under finite intersections, $\bigcup N = [s, x]^\omega$. So, for each basic open set $[u, z]^\omega \subseteq [s, x]^\omega$ there is a $[t, y]^\omega \subseteq [u, z]^\omega$ which belongs to $N$ and we find a $y' \in [y]^\omega$ such that $[t, y']^\omega \cap A = \emptyset$. Since $[u, z]^\omega \subseteq [s, x]^\omega$ was arbitrary and $[t, y']^\omega \subseteq [u, z]^\omega$, this shows that $A$ is meagre in $[s, x]^\omega$.

(b) We have to show that $[\omega]^\omega \setminus (M \cap A)$ contains an open dense set, i.e., for every basic open set $[s, x]^\omega$ there is a $[t, y]^\omega \subseteq [s, x]^\omega$ such that $[t, y]^\omega \cap M \cap A = \emptyset$. Let $[s, x]^\omega$ be an arbitrary basic open set. If $[s, x]^\omega \cap M = \emptyset$, then we are done.
Otherwise, since $M$ is open, $[s, x]^\omega \cap M \supseteq [t, y]^\omega$ for some basic open set $[t, y]^\omega$, and since $[t, y]^\omega \subseteq M$, by (a), $A$ is meagre in $[t, y]^\omega$. Hence, there is a $[t, y']^\omega \subseteq [t, y]^\omega$ such that $[t, y']^\omega \cap A = \emptyset$ which shows that $[t, y']^\omega \cap (M \cap A) = \emptyset$.

(c) Notice that $A \cup ([\omega]^{\omega} \setminus M) = ([\omega]^{\omega} \setminus M) \cup (M \cap A)$. Now, by (b), $M \cap A$ is meagre, and because $M$ is open, $[\omega]^{\omega} \setminus M$ is closed. Thus, $A \cup ([\omega]^{\omega} \setminus M)$ is the union of a meagre set and a closed set and therefore has the Baire property.

The following result is used in the proof of Theorem 9.9.

**Proposition 9.8.** For every $A \subseteq [\omega]^{\omega}$ there is a set $C \supseteq A$ which has the Baire property and whenever $Z \subseteq C \setminus A$ has Baire property, then $Z$ is meagre.

**Proof.** Let $C = A \cup ([\omega]^{\omega} \setminus M)$ where $M = \bigcup \{[s, x]^{\omega} : A$ is meagre in $[s, x]^{\omega}\}$. By Lemma 9.7 (c) we know that $C$ has the Baire property. Now let $Z \subseteq C \setminus A$ be such that $Z$ has the Baire property. If $Z$ is not meagre, then there exists a basic open set $[t, y]^{\omega}$ such that $[t, y]^{\omega} \setminus Z$ is meagre. In particular, $A$ is meagre in $[t, y]^{\omega}$ and therefore $[t, y]^{\omega} \subseteq M$. On the other hand, since $[t, y]^{\omega} \cap Z \neq \emptyset$ and $Z \cap M = \emptyset$ we get that $[t, y]^{\omega} \not\subseteq M$, a contradiction.

Now we are ready to prove that the collection of completely Ramsey sets (i.e., the Baire property) is closed under the generalised Suslin operation $A_\kappa$ whenever $\kappa < \aleph_1$.

**Theorem 9.9.** Let $\kappa < \aleph_1$ be an infinite cardinal and for each $s \in \text{seq}(\kappa)$ let $Q_s \subseteq [\omega]^{\omega}$. If all sets $Q_s$ are completely Ramsey, then

$A_\kappa \{Q_s : s \in \text{seq}(\kappa)\}$

is completely Ramsey too.

**Proof.** Let $\{Q_s : s \in \text{seq}(\kappa)\}$ be a family of completely Ramsey sets. We have to show that the set $A = A_\kappa \{Q_s : s \in \text{seq}(\kappa)\}$ is completely Ramsey. Without loss of generality we may assume that $Q_s \supseteq Q_t$ whenever $s \subseteq t$. For every $s \in \text{seq}(\kappa)$ let

$A_s := \bigcup_{f \in \text{seq}(\kappa)} \bigcap_{n \in \omega} Q_{f|n}^{s|n}$.

We leave it as an exercise to the reader to verify that $A = A_\emptyset$ and that for every $s \in \text{seq}(\kappa)$ we have $A_s \subseteq Q_s$ and $A_s \supseteq \bigcup_{\alpha < \kappa} A_{s^{\langle \alpha \rangle}}$. Further, notice that

$A = A_\kappa \{A_s : s \in \text{seq}(\kappa)\}$.

By Proposition 9.8, for each $s \in \text{seq}(\kappa)$ we find a set $C_s \supseteq A_s$ which is completely Ramsey and whenever $Z \subseteq C_s \setminus A_s$ completely Ramsey, then $Z$ is completely Ramsey-null. Because $Q_s \supseteq A_s$ and $Q_s$ is completely Ramsey, we may assume that $C_s \subseteq Q_s$, and thus,
Related Results

\[ A = \mathcal{A}_\kappa \{ C_s : s \in \text{seq}(\kappa) \} . \]

Let \( C := C_\emptyset \) and notice that \( A = \bigcup_{\alpha \in \kappa} A_{(\alpha)} \subseteq \bigcup_{\alpha \in \kappa} C_{(\alpha)} \), in particular, \( C \subseteq \bigcup_{\alpha \in \kappa} C_{(\alpha)} \). Now we show that

\[ C \setminus A \subseteq \bigcup_{\alpha \in \kappa} C_{(\alpha)} \subseteq \bigcup_{f \in \kappa^n \, n \in \omega} C_{f[n]} \subseteq \bigcup_{s \in \text{seq}(\kappa)} \left( C_s \setminus \bigcup_{\alpha \in \kappa} C_{s^{-\langle \alpha \rangle}} \right). \]

Let \( x \in [\omega]^{\omega} \) be such that

\[ x \notin \bigcup_{s \in \text{seq}(\kappa)} \left( C_s \setminus \bigcup_{\alpha \in \kappa} C_{s^{-\langle \alpha \rangle}} \right) . \]  

(\( \xi \))

If for all \( \alpha \in \kappa \), \( x \notin C_{(\alpha)} \), then \( x \notin C \). On the other hand, if there exists an \( \alpha_0 \in \kappa \) such that \( x \in C_{(\alpha_0)} \), then by (\( \xi \)) we find an \( \alpha_1 \) such that \( x \in C_{(\alpha_0, \alpha_1)} \), and again by (\( \xi \)) we find an \( \alpha_2 \) such that \( x \in C_{(\alpha_0, \alpha_1, \alpha_2)} \), and so on, and finally we find an \( f \in \kappa^n \) such that for all \( n \in \omega \), \( x \in C_{f[n]} \), which implies that \( x \in A \). Further, \( C_s \setminus \bigcup_{\alpha \in \kappa} C_{s^{-\langle \alpha \rangle}} \subseteq C_s \setminus \bigcup_{\alpha \in \kappa} A_{s^{-\langle \alpha \rangle}} = C_s \setminus A_s \), and since \( \bigcup_{\alpha \in \kappa} C_{s^{-\langle \alpha \rangle}} \) is the union of less than \( \xi \) completely Ramsey sets, \( C_s \setminus \bigcup_{\alpha \in \kappa} C_{s^{-\langle \alpha \rangle}} \) is completely Ramsey-and, as a subset of \( C_s \setminus A_s \), it is completely Ramsey-null. Thus, \( C \setminus A \), as a subset of a union of less than \( \xi \) completely Ramsey-null sets, is completely Ramsey-null, and because \( C \) is completely Ramsey, \( A \) is completely Ramsey too. \( \blacksquare \)

Notes

Lemma 9.1 and Theorem 9.2 are due to Plewik [18]. The Ellentuck topology on \( [\omega]^\omega \) was introduced by Ellentuck in [6] (for a comprehensive exposition of General Topology we refer the reader to Engelking [7]). The main result of that paper is Theorem 9, which is now known as Ellentuck’s Theorem 9.5 (see also Matet [16]). However, the aim of Ellentuck’s paper was to give a simpler proof for the fact that every analytic set is completely Ramsey—a fact which also follows from Theorem 9.9 (cf. Galvin and Prikry [8] and Silver [19]). The proof of Theorem 9.9 is similar to the proof of Jech [12, Theorem 11.18] and is essentially taken from Halbeisen [9, Section 3] (see also Matet [15, Proposition 9.8]).

Related Results

60. *The ideal of completely doughnut null sets*. In Chapter 2, the doughnut property was introduced. Now, similarly as we defined the ideal \( R_0 \) of completely Ramsey-null sets one can define the ideal \( v_0 \) of completely doughnut null sets. By Theorem 9.2 we know that \( \text{add}(R_0) = \text{cov}(R_0) \), however, it is not known whether we also have \( \text{add}(v_0) = \text{cov}(v_0) \) (see Halbeisen [10, Question 4]). A partial answer to this problem can be found in Kalmbach, Plewik, and Wojciechowska [13], where it is shown that \( t = \min\{ \text{cf}(c), t \} \) implies \( \text{add}(v_0) = \text{cov}(v_0) \).
61. \( \mathcal{R}_\omega \) and other \( \sigma \)-ideals on \( [\omega]^{\omega} \). In [3], Corazza compares the ideal of completely Ramsey-null sets with other \( \sigma \)-ideals like the ideal of Lebesgue measure zero, meagre, and Marczewski measure zero sets of reals (see also Louveau [14], Aniszczyk, Franklin, Plewik [1], and Brown [3]).

62. Ellentuck type theorems. In [4], Carlson and Simpson survey the interplay between topology and Ramsey Theory. In particular, an abstract version of Ellentuck’s Theorem 9.5 is introduced and discussed. For a further development of this theory see for example Mijares [17].

Let \( \beta \omega \setminus \omega \) denote the set of all non-principal ultrafilters over \( \omega \). For \( A \subseteq \omega \) define

\[
A^* = \{ \mathcal{U} \in \beta \omega \setminus \omega : A \in \mathcal{U} \},
\]

and let \( \mathcal{B}^* = \{ A^* : A \subseteq \omega \} \). Notice that \( \omega^* = \beta \omega \setminus \omega \) and that \( A^* = \emptyset \) iff \( A \) is finite. Furthermore, for all \( A^*, B^* \in \mathcal{B}^* \) we have

\[
A^* \cap B^* = (A \cap B)^* \quad \text{and} \quad A^* \cup B^* = (A \cup B)^*.
\]

In particular, \( \mathcal{B}^* \) has the property that intersections of sets in \( \mathcal{B}^* \) belong to \( \mathcal{B}^* \), thus, \( \mathcal{B}^* \) is a base for a topology on \( \beta \omega \setminus \omega \). The set \( \beta \omega \setminus \omega \) with the topology generated by the basic open sets \( A^* \in \mathcal{B}^* \) is a topological space which has many interesting properties; the following results can be found for example in Todorcević [20, Section 14].

- \( \beta \omega \setminus \omega \) is Hausdorff ([20, Lemma 1]).
- \( \beta \omega \setminus \omega \) is compact ([20, Lemma 2]).
- \( \beta \omega \setminus \omega \) contains no non-trivial converging sequences ([20, Theorem 2]).

For an introduction to \( \beta \omega \setminus \omega \) see van Mill [21], and for combinatorial properties of \( \beta \omega \setminus \omega \) we refer the reader to Hindman and Strauss [11].

63. The minimum height of a tree \( \pi \)-base of \( \beta \omega \setminus \omega \). A family \( \mathcal{P} \subseteq \mathcal{B}^* \) of basic open sets is a \( \pi \)-base for \( \beta \omega \setminus \omega \) if every non-empty element of \( \mathcal{B}^* \) contains a member of \( \mathcal{P} \). If a \( \pi \)-base \( \mathcal{P} \subseteq \mathcal{B}^* \) is a tree when considered as a partially ordered set under reverse inclusion (i.e., for every \( A^* \in \mathcal{P} \), \( A^*_\geq := \{ B^* \in \mathcal{P} : A^* \subseteq B^* \} \) is well-ordered by \( \supseteq \)), then \( \mathcal{P} \) is called a tree \( \pi \)-base of \( \beta \omega \setminus \omega \). If \( \mathcal{P} \subseteq \mathcal{B}^* \) is a tree \( \pi \)-base of \( \beta \omega \setminus \omega \), then the height of an element \( A^* \in \mathcal{P} \), denoted \( h(A^*) \), is the order type of \( A^*_\geq \) (well-ordered by \( \supseteq \)), and the height of \( \mathcal{P} \) is defined by

\[
h(\mathcal{P}) := \bigcup \{ h(A^*) : A^* \in \mathcal{P} \}.
\]

Now, the Base Matrix Lemma 2.11 of Balcar, Pelant, and Simon [2] (see Chapter 8 [Related Result 51]) shows that \( h \) is the minimum height of a tree \( \pi \)-base of \( \beta \omega \setminus \omega \), i.e.,

\[
h = \min \{ h(\mathcal{P}) : \mathcal{P} \subseteq \mathcal{B}^* \text{ is a tree } \pi \text{-base of } \beta \omega \setminus \omega \}.
\]

References

References

Happy Families and their Relatives

A cadence is a certain simultaneous progression of all the voices in a composition accompanying a re-
pose in the harmony or the completion of a meaning-
ful segment of the text.

Gioseffo Zarlino
Le Istituzioni Harmoniche, 1558

In this chapter we shall investigate combinatorial properties of certain families of infinite subsets of \( \omega \). In order to do so, we shall use many of the combinatorial tools developed in the preceding chapters. The families we investigate—particularly \( P \)-families and Ramsey families—will play a key role in understanding the combinatorial properties of Silver and Mathias forcing notions (see Chapter 22 and Chapter 24 respectively).

Happy Families

The \( P \)-families and Ramsey families mentioned above are relatives to the so-called happy families. The name “happy families” comes from a children’s card game, where the idea of the game is to collect the members of fictional families. The connection to families in Set Theory is that a family \( \mathcal{E} \subseteq [\omega]^{\omega} \) is happy if for every countable decreasing sequence \( y_0 \supseteq y_1 \supseteq \cdots \) of elements of \( \mathcal{E} \) there is a member of \( \mathcal{E} \) which selects certain elements from the sets \( y_i \) (cf. Proposition 10.6(b)). This explains why happy families are also called selective co-ideals—which is more sober but less amusing.

Firstly recall that a family \( \mathcal{F} \subseteq [\omega]^\omega \) is a filter if it is closed under supersets and finite intersections, and that the Fréchet filter is the filter consisting of all co-finite subsets of \( \omega \) (i.e., all \( x \in [\omega]^\omega \) such that \( \omega \setminus x \) is finite). To keep the notation short, for \( x \subseteq \omega \) define \( x^c := \omega \setminus x \). For a filter \( \mathcal{F} \subseteq [\omega]^\omega \), \( \mathcal{F}^+ \)
denotes the collection of all sets \( x \subseteq \omega \) such that \( \omega \setminus x \) does not belong to \( \mathcal{F} \), i.e.,

\[
\mathcal{F}^+ = \{ x \subseteq \omega : x^c \notin \mathcal{F} \}.
\]

An equivalent definition of \( \mathcal{F}^+ \) is given by the following

**Fact 10.1.** For any filter \( \mathcal{F} \subseteq [\omega]^\omega \), \( x \in \mathcal{F}^+ \) if and only if \( x \cap z \) is non-empty whenever \( z \in \mathcal{F} \).

**Proof.** On the one hand, if, for some \( z \in \mathcal{F} \), \( x \cap z = \emptyset \), then \( x^c \supseteq z \), which implies that \( x^c \in \mathcal{F} \) and therefore \( x \notin \mathcal{F}^+ \). On the other hand, if, for some \( x \subseteq \omega \), \( x^c \in \mathcal{F} \), then we obviously have \( x \cap x^c = \emptyset \), thus, \( x \) does not meet every member of \( \mathcal{F} \).

If \( \mathcal{U} \) is an ultrafilter and \( x \cup y \in \mathcal{U} \), then at least one of \( x \) and \( y \) belongs to \( \mathcal{U} \). In general, this is not the case for filters \( \mathcal{F} \), but it holds for \( \mathcal{F}^+ \).

**Lemma 10.2.** Let \( \mathcal{F} \subseteq [\omega]^\omega \) be a filter. If \( \mathcal{F}^+ \) contains \( x \cup y \), then it contains at least one of \( x \) and \( y \).

**Proof.** If neither \( x \) nor \( y \) belongs to \( \mathcal{F}^+ \), then \( x^c, y^c \in \mathcal{F} \). Hence, \( (x \cup y)^c = x^c \cap y^c \in \mathcal{F} \), and therefore \( x \cup y \notin \mathcal{F}^+ \).

Now, a filter \( \mathcal{F} \subseteq [\omega]^\omega \) is called a free filter if it contains the Fréchet filter. In particular, every ultrafilter on \( [\omega]^\omega \) is free. Notice that for a free filter \( \mathcal{F} \), \( \mathcal{F}^+ = \{ x \subseteq \omega : \forall z \in \mathcal{F} (|x \cap z| = \omega) \} \), and that a filter \( \mathcal{U} \subseteq [\omega]^\omega \) is an ultrafilter iff \( \mathcal{U} = \mathcal{U}^+ \). Finally, a family \( \mathcal{E} \) of subsets of \( \omega \) is called a free family if there is a free filter \( \mathcal{F} \subseteq [\omega]^\omega \) such that \( \mathcal{E} = \mathcal{F}^+ \). In particular, \( [\omega]^\omega \) and all ultrafilters on \( [\omega]^\omega \) are free families. Notice that a free family does not contain any finite sets and is closed under superset unions. Moreover, a free family \( \mathcal{E} \) is closed under finite intersections iff \( \mathcal{E} \) is an ultrafilter on \( [\omega]^\omega \).

Recall that \( \text{fin}(\omega) \) denotes the set of all finite subsets of \( \omega \). To keep the notation short, for \( s \in \text{fin}(\omega) \) let \( \bar{s} := \bigcup s \), and for \( n \in \omega \) let \( n^+ := n + 1 \) (in other words, \( n^+ \) is the successor cardinal of \( n \)). In particular, for non-empty sets \( s \in \text{fin}(\omega) \) we have \( \bar{s} = \max(s) \) and \( s^+ = \max(s) + 1 \).

A set \( x \subseteq \omega \) is said to diagonalise the set \( \{ x_s : s \in \text{fin}(\omega) \} \subseteq [\omega]^\omega \) if the following conditions are satisfied:

- \( x \subseteq x_{\bar{s}} \);
- for all \( s \in \text{fin}(\omega) \), if \( \bar{s} \in x \) then \( x \setminus s^+ \subseteq x_s \).

For \( \mathcal{A} \subseteq [\omega]^\omega \) we write \( \text{fil}(\mathcal{A}) \) for the filter generated by the members of \( \mathcal{A} \), i.e., \( \text{fil}(\mathcal{A}) \) consists of all subsets of \( \omega \) which are supersets of intersections of finitely many members of \( \mathcal{A} \).

Now, a set \( \mathcal{E} \subseteq [\omega]^\omega \) is a happy family if \( \mathcal{E} \) is a free family and whenever \( \text{fin}(\{ x_s : s \in \text{fin}(\omega) \}) \subseteq \mathcal{E} \), there is an \( x \in \mathcal{E} \) which diagonalises the set \( \{ x_s : s \in \text{fin}(\omega) \} \).

Below, we give two examples of happy families: in the first the family is as large as possible, and in the second the family is of medium size — in the next section we shall see examples of happy families which are as small as possible.
**Fact 10.3.** The family $[\omega]^\omega$ is happy.

**Proof.** Let $\{x_s : s \in \text{fin}(\omega)\} \subseteq [\omega]^\omega$ be a subfamily of $[\omega]^\omega$ and assume that $$\text{fil}\left(\{x_s : s \in \text{fin}(\omega)\}\right) \subseteq [\omega]^\omega,$$ i.e., the intersection of finitely many elements of $\{x_s : s \in \text{fin}(\omega)\}$ is infinite. Let $n_0 := \bigcap x_0$ and for $k \in \omega$ choose $n_{k+1} > n_k$ such that $$n_{k+1} \in \bigcap \{x_s : \hat{s}^+ \leq n_k + 1\}.$$ By our assumption, those choices are possible. Let $x = \{n_k : k \in \omega\};$ then $x \subseteq x_0,$ and whenever $\hat{s} = n_k$ (i.e., $\hat{s}^+ \leq n_k + 1$), we get $$x \setminus \hat{s}^+ \subseteq \bigcap \{x_s : \hat{s}^+ \leq n_k + 1\}.$$ In particular, $x \setminus \hat{s}^+ \subseteq x_s,$ as required.

In order to construct non-trivial examples of happy families, we have to introduce first the following notion: For a mad family $\mathcal{A} \subseteq [\omega]^\omega,$ let $\mathcal{F}_{\mathcal{A}}$ be the collection of all subsets of $\omega$ which are almost contained in supersets of complements of finite unions of members of $\mathcal{A}.$

The goal is to show that $\mathcal{F}_{\mathcal{A}}$ is a happy family whenever $\mathcal{A} \subseteq [\omega]^\omega$ is a mad family, but for this we have to prove first that $\mathcal{F}_{\mathcal{A}}$ is a free filter.

**Proposition 10.4.** If $\mathcal{A} \subseteq [\omega]^\omega$ is a mad family, then $\mathcal{F}_{\mathcal{A}}$ is a free filter but not an ultrafilter.

**Proof.** Let $\mathcal{A} \subseteq [\omega]^\omega$ be a mad family and let $$\mathcal{F}_{\mathcal{A}} = \left\{y \in [\omega]^\omega : \exists x_0 \ldots x_n \in \mathcal{A}\left((x_0 \cup \ldots \cup x_n)^c \subseteq^* y\right)\right\}.$$ Firstly, $\mathcal{F}_{\mathcal{A}}$ is a free filter: By definition, $\mathcal{F}_{\mathcal{A}}$ is closed under superset and contains all co-finite sets, and since $\mathcal{A}$ is mad, no co-finite set is the union of finitely many members of $\mathcal{A};$ hence, $\mathcal{F}_{\mathcal{A}}$ does not contain any finite set. Further, for any $y, y' \in \mathcal{F}_{\mathcal{A}}$ there are $x_0, \ldots, x_n$ and $x_0', \ldots, x_m'$ in $\mathcal{A}$ such that $$(\bigcup_{i \in n} x_i)^c \subseteq^* y \quad \text{and} \quad (\bigcup_{j \in m} x_j')^c \subseteq^* y',$$ which shows that $$\left(\bigcup_{i \in n} x_i \cup \bigcup_{j \in m} x_j'\right)^c \subseteq^* y \cap y' \in \mathcal{F}_{\mathcal{A}}.$$ Secondly, $\mathcal{F}_{\mathcal{A}}$ is not an ultrafilter: We have to find a set $z \in [\omega]^\omega$ such that neither $z$ nor $z^c$ belongs to $\mathcal{F}_{\mathcal{A}}.$ Let $\{x_i : i \in \omega\}$ be distinct elements of $\mathcal{A}.$ Notice that it is enough to construct a set $z \in [\omega]^\omega$ such that both $z$ and $z^c$ have infinite intersection with each $x_i.$ To construct such a set $z,$ take a strictly increasing sequence $n_0 < \ldots < n_k < \ldots$ of natural numbers such that for each $k \in \omega,$ if $k = 2^i(2m + 1),$ then both $n_{2k}$ and $n_{2k+1}$ are in $x_m$ and put $z = \{n_{2k} : k \in \omega\}.$
Now we are ready to give non-trivial examples of happy families. Even though the proof of the following proposition becomes considerably easier by the characterisation of happy families given by Proposition 10.6(b), we think it makes sense to have some non-trivial examples of happy families — and to work with the original definition — before giving an equivalent definition of happy families.

**Proposition 10.5.** Let $\mathcal{A} \subseteq [\omega]^\omega$ be a mad family. Then $\mathcal{F}_\mathcal{A}$ is a happy family.

*Proof.* Given any family $\{y_t : t \in \text{fin}(\omega)\}$ with $\text{fil}(\{x_s : s \in \text{fin}(\omega)\}) \subseteq \mathcal{F}_\mathcal{A}^+$. For $s \in \text{fin}(\omega)$, let $x_s = \bigcap \{y_t : t \leq s\}$. Then for any $n \in \omega$, $x_{(n)} = x_s$ whenever $n = s$. We shall construct an $x \in \mathcal{F}_\mathcal{A}$ which diagonalises $\{y_t : t \in \text{fin}(\omega)\}$ by showing that for all $n \in \omega$, $x \setminus n^+ \subseteq x_{(n)}$. For this, let $x^0$ — constructed as in the proof of Fact 10.3 — diagonalise $\{x_s : s \in \text{fin}(\omega)\}$.

We may not assume that $x^0$ belongs to $\mathcal{F}_\mathcal{A}$, i.e., there might be a $z \in \mathcal{F}$ such that $x^0 \cap z$ is finite. However, since $\mathcal{A}$ is mad, there is a $y^0 \in \mathcal{A}$ such that $x^0 \cap y^0$ is infinite. For each $s \in \text{fin}(\omega)$ define $x^1_s := x_s \setminus y^0$. Notice that all $x^1_s$ are infinite and that $\text{fil}(\{x^1_s : s \in \text{fin}(\omega)\}) \subseteq \mathcal{F}_\mathcal{A}^+$, as $y^0 \in \mathcal{A}$. Let $x^1$ diagonalise $\{x^1_s : s \in \text{fin}(\omega)\}$ and let $y^1 \in \mathcal{A}$ be such that $x^1 \cap y^1$ is infinite. Since $x^1 \subseteq x^0 \subseteq \omega \setminus y^0$ we get $y^1 \neq y^0$. Further, notice that $x^1$ also diagonalises $\{x_s : s \in \text{fin}(\omega)\}$. Now, for each $s \in \text{fin}(\omega)$ define $x^2_s := x_s \setminus (y^0 \cup y^1)$ and proceed as before. After countably many steps we have constructed two sequences of infinite sets, $(x^i : i \in \omega)$ and $(y^i : i \in \omega)$, such that each $y^i$ belongs to $\mathcal{A}$, $y^i \neq y^j$ whenever $i \neq j$, $x^i \cap y^i$ is infinite (for all $i \in \omega$), and $x^i$ diagonalises $\{x_s : s \in \text{fin}(\omega)\}$. Construct a strictly increasing sequence $n_0 < \ldots < n_k < \ldots$ of natural numbers such that $n_0 \in x^2_0$ and for each $k \in \omega$, if $k = 2(2m + 1)$, then

$$n_k \in y^i \cap x^i \cap x_{(n_{k-1})}.$$ 

Such a sequence of natural numbers exists because all sufficiently large numbers in $x^i$ belong to $x_{(n_{k-1})}$ and since $y^i \cap x^i$ is infinite. Finally, let $x = \{n_k : k \in \omega\}$. Then $x$ diagonalises $\{x_s : s \in \text{fin}(\omega)\}$ and it remains to show that $x \in \mathcal{F}_\mathcal{A}$, i.e., $x$ has infinite intersection with each member of $\mathcal{F}_\mathcal{A}$. By construction, for each $i \in \omega$, $x \cap y^i$ is infinite, and since $\mathcal{A}$ is mad, $x \setminus y^i$ is infinite as well. Thus, $x$ has infinite intersection with the complement of any finite union of elements in $\mathcal{A}$, hence $x \in \mathcal{F}_\mathcal{A}$.

After having seen that there are non-trivial happy families, let us give now another characterisation of happy families which will be used later in this chapter.

**Proposition 10.6.** For a free family $\mathcal{A}$, the following statements are equivalent:

(a) $\mathcal{A}$ is happy.
(b) If \( y_0 \geq y_1 \geq \cdots \geq y_n \geq \cdots \) is a countable decreasing sequence of elements of \( \mathcal{U} \), then there is a function \( f \in {}^\omega \omega \) such that \( f[\omega] \in \mathcal{U} \), \( f(0) \in y_0 \), and for all \( n \in \omega \) we have \( f(n + 1) \in y_{f(n)} \).

**Proof.** (a)⇒(b) Assume that \( \mathcal{U} \) is happy and let \( \{ y_i : i \in \omega \} \subseteq \mathcal{U} \) be such that for all \( i \in \omega \), \( y_{i+1} \subseteq y_i \). For each \( s \in \text{fin}(\omega) \) define

\[
x_s = \bigcap \{ y_i : i \leq s \}.
\]

Notice that \( \text{fil}(\{ x_s : s \in \text{fin}(\omega) \}) \subseteq \mathcal{U} \). Since \( \mathcal{U} \) is assumed to be happy there is an \( x \) which diagonalises the family \( \{ x_s : s \in \text{fin}(\omega) \} \). Let \( f = f_\mathcal{U} \) — recall that \( f_\mathcal{U} \in {}^\omega \omega \) is the unique strictly increasing bijection between \( \omega \) and \( x \) (defined in Chapter 8). For an arbitrary \( n \in \omega \) let \( s := x \cap (f(n) + 1) \). Then \( \hat{s} = f(n) + 1 \) and \( \hat{s} \in x \). As \( (f(n) + 1) \in x \setminus \hat{s}^+ \) and \( x \setminus \hat{s}^+ \subseteq x_s \subseteq y_{f(n)} \), we have \( f(n + 1) \in y_{f(n)} \), and since \( n \) was arbitrary, \( f \) has the required properties.

(b)⇒(a) Assume now that \( \mathcal{U} \) has property (b) and let \( \{ x_s : s \in \text{fin}(\omega) \} \subseteq \mathcal{U} \) be such that \( \text{fil}(\{ x_s : s \in \text{fin}(\omega) \}) \subseteq \mathcal{U} \). We have to find an \( x \in \mathcal{U} \) which diagonalises \( \{ x_s : s \in \text{fin}(\omega) \} \). For each \( i \in \omega \) define

\[
y_i = \bigcap \{ x_s : \hat{s} \leq i \}.
\]

Obviously, for each \( i \in \omega \) we have \( y_i \in \mathcal{U} \) and \( y_{i+1} \subseteq y_i \). By (b) there is a function \( f \in {}^\omega \omega \) such that \( f[\omega] \in \mathcal{U} \) and for all \( n \in \omega \) we have \( f(n + 1) \in y_{f(n)} \). Let \( x := f[\omega] \) and let \( s \in \text{fin}(\omega) \) be such that \( \hat{s} \in x \). Then there exists an \( n \in \omega \) such that \( f(n) = \hat{s} \), and for every \( k \in x \setminus \hat{s}^+ \) we have \( k = f(m) \) for some \( m > n \), hence, \( k \in y_{f(n)} \). Now, \( \hat{s}^+ = f(n) + 1 \), and since \( y_{f(n)} \subseteq x_s \) we get \( k \in x_s \). Hence, for all \( s \in \text{fin}(\omega) \) with \( \hat{s} \in x \) we have \( x \setminus \hat{s}^+ \subseteq x_s \), which shows that \( x \) diagonalises \( \{ x_s : s \in \text{fin}(\omega) \} \).

We leave it as an exercise to the reader to find an easier proof of Proposition 10.5 by using the characterisation of happy families given by Proposition 10.6.(b).

**Ramsey Ultrafilters**

So far we have seen two examples of happy families. In the first example (Fact 10.3), the happy family was as large as possible, and in the second example (Proposition 10.5), the happy families were of medium size. Below, we consider happy families which are as small as possible, i.e., happy families which are ultrafilters.

A free ultrafilter \( \mathcal{U} \subseteq [\omega]^\omega \) is a **Ramsey ultrafilter** if for every colouring \( \pi : [\omega]^2 \to 2 \) there exists an \( x \in \mathcal{U} \) which is homogeneous for \( \pi \), i.e., \( \pi|_{[x]^2} \) is constant.
The following result gives two alternative characterisations of Ramsey ultrafilter. The first characterisation of Ramsey ultrafilters is related to P-points and Q-points (introduced below), and the second characterisation show that a Ramsey ultrafilter is an ultrafilter that is also a happy family.

**Proposition 10.7.** For every free ultrafilter \( \mathcal{U} \), the following conditions are equivalent:

(a) \( \mathcal{U} \) is a Ramsey ultrafilter.

(b) Let \( \{ u_i : i \in \omega \} \) be a partial partition of \( \omega \), i.e., \( \bigcup \{ u_i : i \in \omega \} \subseteq \omega \) and for any distinct \( i, j \in \omega \) we have \( u_i \cap u_j = \emptyset \). Then either \( u_i \in \mathcal{U} \) for a (unique) \( i \in \omega \), or there exists an \( x \in \mathcal{U} \) such that for each \( i \in \omega \), \( |x \cap u_i| \leq 1 \).

(c) \( \mathcal{U} \) is happy.

**Proof.** (a)⇒(b) Let \( \{ u_i : i \in \omega \} \) be a partition of \( \omega \). With respect to \( \{ u_i : i \in \omega \} \) define the colouring \( \pi : [\omega]^2 \to 2 \) as follows:

\[
\pi(\{n,m\}) = \begin{cases} 
0 & \text{if there is an } i \in \omega \text{ such that } \{n,m\} \subseteq u_i, \\
1 & \text{otherwise.}
\end{cases}
\]

By (a) there is an \( x \in \mathcal{U} \) such that \( \pi|_{[\omega]^2} \) is constant. Now, if \( \pi|_{[\omega]^2} \) is constantly zero, then there exists an \( i \in \omega \) such that \( x \subseteq u_i \), hence \( u_i \in \mathcal{U} \). On the other hand, if \( \pi|_{[\omega]^2} \) is constantly one, then for any distinct \( n, m \in x \) and any \( i \in \omega \) we get that \( \{n, m\} \cap u_i \) has at most one element, hence, for each \( i \in \omega \), \( x \cap u_i \) has at most one element.

(b)⇒(c) By Proposition 10.6 it is enough to show that for every countable decreasing sequence \( y_0 \supseteq y_1 \supseteq \ldots \supseteq y_n \supseteq \ldots \) of elements of \( \mathcal{U} \) there is a function \( f \in \omega^\omega \) such that \( f[w] \in \mathcal{U} \), \( f(0) \in y_0 \), and for all \( k \in \omega \) we have \( f(k+1) \in y_f(k) \). If \( y = \bigcap_{n \in \omega} y_n \in \mathcal{U} \), then the function \( f_y \in \mathcal{U}^\omega \) has the required properties. So, let us assume that \( \bigcap_{n \in \omega} y_n \notin \mathcal{U} \) and without loss of generality let us further assume that for all \( n \in \omega \), \( y_n \setminus y_{n+1} \neq \emptyset \). Consider the partition \( \{y_n^0 \cup \bigcap_{n \in \omega} y_n \} \cup \{y_n \setminus y_{n+1} : n \in \omega \} \) and notice that none of the pieces are in \( \mathcal{U} \). By (b), there exists a set \( x = \{a_n : n \in \omega \} \in \mathcal{U} \) such that for all \( n \in \omega \), \( x \cap (y_n \setminus y_{n+1}) = \{a_n\} \), in particular, \( x \cap \bigcap_{n \in \omega} y_n = \emptyset \).

Let \( g \in \omega^\omega \) be a strictly increasing function such that \( g(0) > 0 \), \( g|\omega \subseteq x \), and for all \( n \in \omega \), \( x \setminus g(n) \subseteq y_n \). For \( k \in \omega \) let \( g^{k+1}(0) := g(g^k(0)) \), where \( g^0(0) := 0 \). Further, for \( k \in \omega \) let \( x_k := x \cap [g^{2k}(0), g^{2k+1}(0)) \) —recall that \( [a,b) = \{i \in \omega : a \leq i < b\} \). Now, by (b) and since \( \mathcal{U} \) is an ultrafilter, there exists a set \( z = \{c_k : k \in \omega \} \subseteq x \) such that \( z \in \mathcal{U} \) and for all \( k \in \omega \), \( z \cap x_k = \{c_k\} \). Notice that by construction, for each \( k \in \omega \) we have \( c_{k+2} > g(c_k) \) and \( c_{k+2} \in y_{c_k} \). Finally, since \( \mathcal{U} \) is an ultrafilter and \( \{c_k : k \in \omega \} \notin \mathcal{U} \), either \( \{c_{2k} : k \in \omega \} \) or \( \{c_{2k+1} : k \in \omega \} \) belongs to \( \mathcal{U} \). In the former case define \( f \in \mathcal{U}^\omega \) by stipulating \( f(k) := c_{2k} \), otherwise define \( f(k) := c_{2k+1} \). Then \( f \) has the required properties.

(c)⇒(a) Let \( \mathcal{U} \) be an ultrafilter that is also a happy family, and further let
\[\pi : [\omega]^2 \to 2\] be an arbitrary but fixed colouring. We have to find a \(y \in \mathcal{U}\) such that \(\pi|_{[y]^2}\) is constant. The proof is similar to the proof of Proposition 2.2.

First we construct a family \(\{x_s : s \in \text{fin}(\omega)\} \subseteq \mathcal{U}\). Let \(x_0 = \omega\), and let \(x_{\{0\}} \in \mathcal{U}\) be such that \(x_{\{0\}} \subseteq \omega \setminus \{0\}\) and for all \(k, k' \in x_{\{0\}}\) we have \(\pi(\{0, k\}) = \pi(\{0, k'\})\). Notice that since \(\mathcal{U}\) is an ultrafilter, \(x_{\{0\}}\) exists. In general, if \(x_s\) is defined and \(s > \emptyset\), then let \(x_{s \cup \{n\}} \in \mathcal{U}\) be such that \(x_{s \cup \{n\}} \subseteq x_s \setminus n^+\) and for all \(k, k' \in x_{s \cup \{n\}}\) we have \(\pi(\{n, k\}) = \pi(\{n, k'\})\). Since \(\mathcal{U}\) is happy, there is a \(y \in \mathcal{U}\) which diagonalises the family \(\{x_s : s \in \text{fin}(\omega)\}\). By construction, there is a \(y \in \mathcal{U}\) and for all \(k, k' \in y \setminus n^+\) we have \(\pi(\{n, k\}) = \pi(\{n, k'\})\) and we can define the colouring \(\tau : x \to 2\) by stipulating

\[\tau(n) = \begin{cases} 0 & \text{if there is a } k \in x \setminus n^+ \text{ such that } \pi(\{n, k\}) = 0, \\ 1 & \text{otherwise.} \end{cases}\]

Since \(\mathcal{U}\) is an ultrafilter, there exists an \(x \in \mathcal{U}\) such that \(x \subseteq y\) and \(\tau|_x\) is constant, hence, \(\pi|_{[y]^2}\) is constant.

At first glance, condition (a) is just related to Proposition 2.2 and not to Ramsey's Theorem. However, the following fact shows that this is not the case. Moreover, even Proposition 2.8 is related to Ramsey ultrafilters (the proofs are left to the reader).

**Fact 10.8.** For every free ultrafilter \(\mathcal{U}\), the following conditions are equivalent:

(a) \(\mathcal{U}\) is a Ramsey ultrafilter, i.e., for every colouring \(\pi : [\omega]^2 \to 2\) there exists an \(x \in \mathcal{U}\) which is homogeneous for \(\pi\).

(b) For any \(n \in \omega\), for any positive integer \(r \in \omega\), and for every colouring \(\pi : [\omega]^n \to r\), there exists an \(x \in \mathcal{U}\) which is homogeneous for \(\pi\).

(c) Let \(\{r_k : k \in \omega\}\) and \(\{n_k : k \in \omega\}\) be two (possibly finite) sets of positive integers, and for each \(k \in \omega\) let \(\pi_k : [\omega]^n_k \to r_k\) be a colouring. Then there exists an \(x \in \mathcal{U}\) which is almost homogeneous for each \(\pi_k\).

It is time now to address the problem of the existence of Ramsey ultrafilters. On the one hand, it can be shown that there are models of ZFC in which no Ramsey ultrafilters exist (see Proposition 25.11). Thus, the existence of Ramsey ultrafilters is not provable in ZFC. On the other hand, if we assume for example CH (or just \(p = \text{c}\)), then we can easily construct a Ramsey ultrafilter.

**Proposition 10.9.** If \(p = \text{c}\), then there exists a Ramsey ultrafilter.

**Proof.** Let \(\{\pi_\alpha : \alpha \in \text{c}\}\) be an enumeration of the set of all 2-colourings of \([\omega]^2\), i.e., for every colouring \(\pi : [\omega]^2 \to 2\) there exists an \(\alpha \in \text{c}\) such that \(\pi = \pi_\alpha\). By transfinite induction we first construct a sequence \(\langle x_\alpha : \alpha \in \text{c} \rangle \subseteq [\omega]^\omega\) such that \(\{x_\alpha : \alpha \in \text{c}\}\) has the finite intersection property and for all \(\alpha \in \text{c}\), \(\pi_\alpha|_{[x_{\alpha+1}]^2}\) is constant. Let \(x_0 := \omega\) and assume that for some \(\alpha \in \text{c}\) we
have already constructed $x_\beta$ ($\beta \in \alpha$) such that \( \{x_\beta : \beta \in \alpha\} \) has the finite intersection property and for all $\gamma + 1 \in \alpha$ we have $\pi_\gamma |_{x_{\alpha+1}}^2$ is constant. If $\alpha$ is a successor ordinal, say $\alpha = \beta_0 + 1$, then let $x_\alpha \in [x_{\beta_0}]^\omega$ be such that $\pi_{\beta_0} |_{x_\alpha}^2$ is constant (notice that by \textsc{Ramsey’s Theorem} 2.1, $x_{\alpha+1}$ exists). If $\alpha$ is a limit ordinal, then let $x_\alpha$ be a pseudo-intersection of $\{x_\beta : \beta \in \alpha\}$ (notice that since $|\alpha| < p$, $x_{\alpha+1}$ exists). In either case, the family $\{x_\beta : \beta \in \alpha\}$ has the required properties. In particular, the family $\mathcal{E} = \{x_\alpha : \alpha \in \mathcal{E}\}$ has the finite intersection property and for each colouring $\pi : [\omega]^2 \to 2$ there is an $x \in \mathcal{E}$ such that $\pi |_{x^2}$ is constant. Finally, extend the family $\mathcal{E}$ to an ultrafilter $\mathcal{U}$. Then $\mathcal{U}$ is a \textsc{Ramsey} ultrafilter.

\section*{P-points and Q-points}

Below, we consider ultrafilters which are weaker than \textsc{Ramsey} ultrafilters, but which share with them some combinatorial properties.

A free ultrafilter $\mathcal{U}$ is a \textbf{P-point} if for each partition $\{u_n \subseteq \omega : n \in \omega\}$ of $\omega$, either $u_n \in \mathcal{U}$ for a (unique) $n \in \omega$, or there exists an $x \in \mathcal{U}$ such that for each $n \in \omega$, $x \cap u_n$ is finite.

Furthermore, a free ultrafilter $\mathcal{U}$ is a \textbf{Q-point} if for each partition of $\omega$ into finite pieces $\{I_n \subseteq \omega : n \in \omega\}$, (i.e., for each $n \in \omega$, $I_n$ is finite), there exists an $x \in \mathcal{U}$ such that for each $n \in \omega$, $x \cap I_n$ has at most one element.

Comparing these definitions of P-points and Q-points with \textsc{Proposition} 10.7 (b), it is evident that a \textsc{Ramsey} ultrafilter is both a P-point as well as a Q-point; but also the converse is true.

\textbf{Fact 10.10.} $\mathcal{U}$ is a \textsc{Ramsey} ultrafilter if and only if $\mathcal{U}$ is a P-point and a Q-point.

\textit{Proof.} ($\Leftarrow$) This follows immediately from \textsc{Proposition} 10.7 (b) and the definitions of P-points and Q-points.

($\Rightarrow$) Let $\mathcal{U}$ be a P-point and a Q-point and let $\{u_n \subseteq \omega : n \in \omega\}$ be a partition of $\omega$. We have to show that either $u_n \in \mathcal{U}$ for a (unique) $n \in \omega$, or there exists an $x \in \mathcal{U}$ such that for each $n \in \omega$, $x \cap u_n$ has at most one element. If there is a $u_n \not\in \mathcal{U}$, then we are done. So, assume that for all $n \in \omega$, $u_n \notin \mathcal{U}$. Since $\mathcal{U}$ is a P-point, there exists a $y_0 \in \mathcal{U}$ such that for each $n \in \omega$, $y_0 \cap u_n$ is finite. For $n \in \omega$ let $I_{2n} := y_0 \cap u_n$. Further, let $\{a_i : i \in \omega\} = \omega \setminus \bigcup_{n \in \omega} \{I_{2n} : n \in \omega\}$ and for $n \in \omega$ let $I_{2n+1} := \{a_n\}$. Then $\{I_n : n \in \omega\}$ is a partition of $\omega$ into finite pieces. Since $\mathcal{U}$ is a Q-point, there exists a $y_1 \in \mathcal{U}$ such that for each $n \in \omega$, $y_1 \cap I_n$ has at most one element. Now, let $x = y_0 \cap y_1$. Then $x \in \mathcal{U}$ and for each $n \in \omega$, $x \cap u_n$ has at most one element.

Below, we give a few other characterisations of P-points and Q-points.

\textbf{Fact 10.11.} For every free ultrafilter $\mathcal{U}$, the following conditions are equivalent:

\begin{itemize}
    \item[(i)] $\mathcal{U}$ is a P-point.
    \item[(ii)] $\mathcal{U}$ is a Q-point.
    \item[(iii)] For each $n \in \omega$, $x \in \mathcal{U}$ has at most one element.
\end{itemize}
(a) $\mathcal{U}$ is a $P$-point.

(b) For every family $\{x_n : n \in \omega\} \subseteq \mathcal{U}$ there is an $x \in \mathcal{U}$ such that for all $n \in \omega$, $x \subseteq^* x_n$ (i.e., $x \setminus x_n$ is finite).

(c) For every family $\{x_n : n \in \omega\} \subseteq \mathcal{U}$ there is a function $f \in \omega^\omega$ and a set $x \in \mathcal{U}$ such that for all $n \in \omega$, $x \setminus f(n) \subseteq x_n$.

**Fact 10.12.** For every free ultrafilter $\mathcal{U}$, the following conditions are equivalent:

(a) $\mathcal{U}$ is a $Q$-point.

(b) For every family $\{x_n : n \in \omega\} \subseteq \mathcal{U}$ there is an $x \in \mathcal{U}$ such that for all $n \in \omega$, $x \cap (\omega \setminus x_n)$ is finite.

There are also characterisations of $P$-points which are not so obvious:

**Proposition 10.13.** For a free ultrafilter $\mathcal{U}$, the following conditions are equivalent:

(a) $\mathcal{U}$ is a $P$-point.

(b) For every family $\{x_n : n \in \omega\} \subseteq \mathcal{U}$ there is an $x \in \mathcal{U}$ such that for infinitely many $n \in \omega$, $x \setminus n \subseteq x_n$.

**Proof.** Since (b) $\Rightarrow$ (a) is obvious, we just prove (a) $\Rightarrow$ (b): Since $\mathcal{U}$ is a $P$-point, by Fact 10.11(c) there exists a function $f \in \omega^\omega$ and a set $y \in \mathcal{U}$ such that for all $n \in \omega$, $y \setminus f(n) \subseteq x_n$. Hence, there exists also a function $g \in \omega^\omega$ such that $g(0) = 0$ and for all $k \in \omega$ we have $y \setminus g(k+1) \subseteq x_{g(k)}$. Since $\mathcal{U}$ is an ultrafilter, either $y_0 = \bigcup_{k \in \omega} [g(2k+1), g(2k+2))$ or $y_1 = \bigcup_{k \in \omega} (g(2k), g(2k+1))$ belongs to $\mathcal{U}$. Let $x = y \cap y_\varepsilon$, where $\varepsilon \in \{0, 1\}$ is such that $y_\varepsilon \in \mathcal{U}$. Then for every $k \in \omega$ we have $x \setminus g(2k + \varepsilon) = x \setminus g(2k + \varepsilon + 1) \subseteq x_{2k+\varepsilon}$. \hfill $\dashv$

$P$-points and $Q$-points, and consequently Ramsey ultrafilters, can also be characterised in terms of functions, but before we have to introduce the notion of finite-to-one functions: A function $f \in \omega^\omega$ is **finite-to-one** if for every $k \in \omega$, the set $\{n \in \omega : f(n) = k\}$ is finite.

**Proposition 10.14.** Let $\mathcal{U}$ be a free ultrafilter.

(a) $\mathcal{U}$ is a $P$-point if and only if for every function $f \in \omega^\omega$ there exists an $x \in \mathcal{U}$ such that $f|_x$ is constant or finite-to-one.

(b) $\mathcal{U}$ is a $Q$-point if and only if for every finite-to-one function $f \in \omega^\omega$ there exists an $x \in \mathcal{U}$ such that $f|_x$ is one-to-one.

(c) $\mathcal{U}$ is a Ramsey ultrafilter if and only if for every function $f \in \omega^\omega$ there exists an $x \in \mathcal{U}$ such that $f|_x$ is constant or one-to-one.

**Proof.** Let $f \in \omega^\omega$ be an arbitrary but fixed function. For $k \in \omega$ define $u_k := \{n \in \omega : f(n) = k\}$. Then $\{u_k : k \in \omega\}$ is a partition of $\omega$. The proof now follows from Fact 10.10 and the following observations (the details are left to the reader):
• For any $x \in [\omega]^\omega$, $f|_x$ is constant iff there is a $k \in \omega$ such that $x \subseteq u_k$.
• For any $x \in [\omega]^\omega$, $f|_x$ is finite-to-one iff for all $k \in \omega$ we have $x \cap u_k$ is finite.
• The function $f$ is finite-to-one iff each $u_k$ is finite.
• For any $x \in [\omega]^\omega$, $f|_x$ is one-to-one iff for all $k \in \omega$, $x \cap u_k$ has at most one element.  

The next result shows that ultrafilters, and especially $\mathbb{Q}$-points, must contain quite “sparse” sets.

**Proposition 10.15.** For free families $\mathcal{U} \subseteq [\omega]^\omega$ we have:

(a) If $\mathcal{U}$ is a free ultrafilter, then the family $\{f_x \in {\omega}: x \in \mathcal{U}\}$ is unbounded.

(b) If $\mathcal{U}$ is a $\mathbb{Q}$-point, then the family $\{f_x \in {\omega}: x \in \mathcal{U}\}$ is dominating.

**Proof.** (a) Let $f \in {\omega}$ be arbitrary. Define $g(0) = \max \{f(0), 1\}$ and for $k \in \omega$ define $g(k + 1) := g(k) + f(g(k))$. Further, let $x_0 = [0, g(0)]$, and in general, for $n \in \omega$ let $x_n = [g(2n), g(2n + 1)]$ and $y_n = [g(2n + 1), g(2n + 2)]$. Finally, let $x = \bigcup_{n \in \omega} x_n$ and $y = \bigcup_{n \in \omega} y_n$. We leave it as an exercise to the reader to verify that $f_x \not\leq^* f$ and $f_y \not\leq^* f$. Hence, $f$ dominates neither $f_x$ nor $f_y$. Now, since $\mathcal{U}$ is an ultrafilter, either $x$ or $y$ belongs to $\mathcal{U}$. Hence, $f$ does not dominate the family $\mathcal{B} = \{f_x \in {\omega}: x \in \mathcal{U}\}$, and since $f$ was arbitrary, $\mathcal{B}$ is unbounded.

(b) Let $g \in {\omega}$ be arbitrary. Without loss of generality we may assume that $g$ is strictly increasing. For $n \in \omega$ let $I_n = [g(2n), g(2n + 2)]$. Then $\{I_n : n \in \omega\}$ is a partition of $\omega$ into finite pieces. Since $\mathcal{U}$ is a $\mathbb{Q}$-point, there exists an $x \in \mathcal{U}$ such that for each $n \in \omega$, $x \cap I_n$ has at most one element which implies that $g <^* f_x$. Hence, $f_x$ dominates $g$, and since $g$ was arbitrary, the family $\{f_x \in {\omega}: x \in \mathcal{U}\}$ is dominating.  

As we have seen above (Proposition 10.9), $\mathcal{U} = \mathcal{U}$ implies the existence of a Ramsey ultrafilter. On the other hand, one can show that $\mathcal{U} = \mathcal{U}$ is not sufficient to prove the existence of Ramsey ultrafilters (see Proposition 25.11).

However, as a consequence of the next result, we get that $\mathcal{U} = \mathcal{U}$ is sufficient to prove the existence of $P$-points—which shows that $P$-points are easier to get than Ramsey ultrafilters (cf. RELATED RESULTS 66&67).

**Theorem 10.16.** $\mathcal{U} = \mathcal{U}$ if and only if every free filter over a countable set which is generated by less than $\mathcal{U}$ sets can be extended to a $P$-point. In particular, $\mathcal{U} = \mathcal{U}$ implies the existence of $P$-points.

**Proof.** ($\Rightarrow$) Suppose that $\mathcal{E} \subseteq [\omega]$ is a family of cardinality less than $\mathcal{U}$. For $f \in \mathcal{E}$ and $n \in \omega$ define

$$x_f = \{ (m, k) \in \omega \times \omega : f(n) < k \} \text{ and } x_n = \{ (m, k) \in \omega \times \omega : n \leq m \},$$

and let
\[
\mathcal{C} = \{x_f : f \in \mathcal{E}\} \cup \{x_n : n \in \omega\} \cup \{z : z \subseteq \omega \times \omega : (\omega \times \omega) \setminus z \text{ is finite}\}.
\]

Notice that \(|\mathcal{C}| < \epsilon\) and that each set in \(\mathcal{C}\) is an infinite subset of the countable set \(\omega \times \omega\). Moreover, for any finitely many members \(g_0, \ldots, g_n \in \mathcal{C}\) we have \(g_0 \cap \cdots \cap g_n\) is infinite. Now, the family \(\mathcal{C}\) generates a free filter over \(\omega \times \omega\), which, by assumption, can be extended to a \(P\)-point \(\mathcal{W} \subseteq [\omega \times \omega]^\omega\).

Consider the partition \(\{u_n : n \in \omega\}\) of \(\omega \times \omega\), where for \(n \in \omega\), \(u_n := \{(n) \times \omega\}\). Notice that no \(u_n\) (for \(n \in \omega\)) belongs to \(\mathcal{W}\). Since \(\mathcal{W}\) is a \(P\)-point, there exists a \(y \in \mathcal{W}\) such that for all \(n \in \omega\), \(y \cap u_n\) is finite. Let us define the function \(g \in \omega^\omega\) by stipulating \(g(n) = \bigcup \{k \in \omega : (n, k) \in y \cap u_n\}\). Since \(y \in \mathcal{W}\), for all \(f \in \mathcal{E}\) we have \(y \cap x_f\) is infinite. Hence, for every \(f \in \mathcal{E}\) there are infinitely many \(n \in \omega\) such that \(g(n) > f(n)\). In other words, \(g\) is not dominated by any function \(f \in \mathcal{E}\), which shows that no family of cardinality less than \(\epsilon\) is dominating.

\((\Rightarrow)\) The proof is by induction using the following

**Claim.** Suppose that the free filter \(\mathcal{F} \subseteq [\omega]^\omega\) is generated by less than \(\mathfrak{d}\) sets and let \(\{x_n : n \in \omega\} \subseteq \mathcal{F}\). Then there exists \(x \in [\omega]^\omega\) such that for all \(n \in \omega\), \(x \subseteq^* x_n\), and for all \(y \in \mathcal{F}\), \(x \cap y\) is infinite.

**Proof of Claim.** Without loss of generality we may assume that for all \(n \in \omega\), \(x_{n+1} \subseteq x_n\). For \(y \in \mathcal{F}\) define \(g_y \in \omega^\omega\) by stipulating \(g_y(n) = \bigcap (y \cap x_n)\). Notice that the set \(y \cap x_n\) is non-empty, and that if \(y \subseteq y'\), then for all \(n \in \omega\), \(g_y(n) \leq g_{y'}(n)\). Now, since \(\mathcal{F}\) is generated by less than \(\mathfrak{d}\) sets, and since every free ultrafilter generated by less than \(\mathfrak{d}\) sets has a basis of less than \(\mathfrak{d}\) sets, there exists a function \(f \in \omega^\omega\) such that for all \(y \in \mathcal{F}\) we have \(f \neq^* g_y\). Finally let

\[
x = \bigcup_{n \in \omega} (x_n \cap f(n))
\]

We leave it to the reader to verify that \(x\) has the required properties. \(\text{Claim}\)

By the claim and the assumption that \(\mathfrak{d} = \epsilon\) we inductively construct a \(P\)-point as follows: Let \(\{X_\alpha : \alpha \in \epsilon\} \subseteq [\omega]^\omega\) be an enumeration of all countable subsets of \([\omega]^\omega\). Let \(\mathcal{F}_0\) be any free filter which is generated by less than \(\mathfrak{d}\) sets and assume that we have already constructed \(\mathcal{F}_\alpha\) for some \(\alpha \in \epsilon\). If \(X_\alpha \cup \mathcal{F}_\alpha\) has the finite intersection property, then we use the claim to obtain a set \(x_{\alpha+1}\) such that \(\{x_{\alpha+1}\} \cup \mathcal{F}_\alpha\) has the finite intersection property and \(x_{\alpha+1}\) is a pseudo-intersection of \(X_\alpha\); and let \(\mathcal{F}_{\alpha+1}\) be the filter generated by \(\mathcal{F}_\alpha\) and \(x_{\alpha+1}\). If \(X_\alpha \cup \mathcal{F}_\alpha\) does not have the finite intersection property, then let \(\mathcal{F}_{\alpha+1} = \mathcal{F}_\alpha\). Further, if \(\alpha \in \epsilon\) is a limit ordinal and for all \(\beta \in \alpha\) we have already constructed \(\mathcal{F}_\beta\), then let \(\mathcal{F}_\alpha = \bigcup_{\beta \in \alpha} \mathcal{F}_\beta\). Finally, let \(\mathcal{F} = \bigcup_{\alpha \in \epsilon} \mathcal{F}_\alpha\).

Then \(\mathcal{F}\) is a \(P\)-point: Firstly, by construction, \(\mathcal{F}\) is a filter, and since the free filter \(\mathcal{F}_0\) is contained in \(\mathcal{F}\), \(\mathcal{F}\) is even a free filter. Secondly, for any \(x \in [\omega]^\omega\) there exists a \(\beta \in \epsilon\) such that \(X_\beta = \{x\}\). Thus, either \(x \in \mathcal{F}_{\beta+1}\) or there is a \(y \in \mathcal{F}_\beta\) such that \(x \cap y\) is finite, which implies that \(x \notin \mathcal{F}_\beta\). Hence, \(\mathcal{F}\) is a free ultrafilter. Finally, for every set \(\{x_n : n \in \omega\} \subseteq \mathcal{F}\) there exists a
$\gamma \in \mathfrak{c}$ such that $X_{\gamma} = \{x_n : n \in \omega\}$. Since $X_{\gamma} \cup \mathcal{F}_{\gamma}$ has the finite intersection property, there is an $x_{\gamma+1} \in \mathcal{F}_{\gamma+1}$ such that for all $n \in \omega$, $x_{\gamma+1} \subseteq \star x_n$.

### Ramsey families and $P$-families

Below, we give characterisations of Ramsey ultrafilters and $P$-points in terms of games, which lead to so-called Ramsey families and $P$-families respectively.

The two games we shall consider are infinite and played between two players. Now, a run of an infinite two-player game consists of an infinite sequence $\langle x_0, y_0, x_1, y_1, \ldots \rangle$ which is constructed alternately by the two players. More precisely, the first player starts the game with $x_0$ and the second player responds with $y_0$. Then the first player plays $x_1$ and the second player responds with $y_1$, and so on. Of course, in order to get a proper game we have to introduce also some rules defining legal moves and telling which player wins a particular run of the game.

Before we introduce some further game-theoretical notions, let us illustrate the notion of rules by the following infinite two-player game, played between Death and the Maiden.

Let $\mathcal{E}$ be an arbitrary free family. Associated with $\mathcal{E}$ we define two quite similar games, denoted $\mathcal{G}_\mathcal{E}$ and $\mathcal{G}_\mathcal{E}^*$, between two players, say Death and the Maiden.

In the game $\mathcal{G}_\mathcal{E}$, the Maiden always plays members of $\mathcal{E}$ and then Death responds with an element of Maiden’s move. Thus, a run of $\mathcal{G}_\mathcal{E}$ can be illustrated as follows:

\[ \begin{array}{cccc}
\text{Maiden} & x_0 & \geq & x_1 & \geq & x_2 & \geq & \ldots \\
\text{Death} & a_0 & < & a_1 & < & a_2 & < & \ldots
\end{array} \]

More formally, the rules for $\mathcal{G}_\mathcal{E}$ are as follows: For each $i \in \omega$, $x_i \in \mathcal{E}$ and $a_i \in x_i$. Furthermore, we require that for each $i \in \omega$, $x_{i+1} \subseteq x_i$ and $a_i < a_{i+1}$. Finally, Death wins the game $\mathcal{G}_\mathcal{E}$ if and only if $\{a_i : i \in \omega\}$ belongs to the family $\mathcal{E}$.

In the game $\mathcal{G}_\mathcal{E}^*$, Death has slightly more freedom, since he can play now finite sequences instead of just singletons. A run of $\mathcal{G}_\mathcal{E}^*$ can be illustrated as follows:

\[ \begin{array}{cccc}
\text{Maiden} & x_0 & \geq & x_1 & \geq & x_2 & \geq & \ldots \\
\text{Death} & s_0 & < & s_1 & < & s_2 & < & \ldots
\end{array} \]

Again, the sets $x_i$ played by the Maiden must belong to the free family $\mathcal{E}$ and each finite set $s_i$ played by Death must be a subset of the corresponding
$x_i$. Furthermore, for each $i \in \omega$ we require that $x_{i+1} \subseteq (x_i \setminus \bigcup_{j \leq i} s_j)$. Notice that the finite sets $s_i$ may be empty. Finally, Death wins the game $G^n$ if and only if $\bigcup\{s_i : i \in \omega\}$ belongs to the family $\mathcal{E}$.

Now we define the notion of a strategy for the Maiden. Roughly speaking, a strategy for the Maiden is a “rule” that tells her how to play, for each $n \in \omega$, her $n^{th}$ move $x_n$, given Death’s previous moves $m_0, \ldots, m_n$. In fact, a strategy for the Maiden in the game $G^n$ is a certain mapping from $\text{seq}(\mathcal{E} \cup \omega)$ to $\mathcal{E}$. Intuitively, with respect to $G^n$, a strategy $\sigma$ for the Maiden works as follows: The Maiden starts playing $x_0 \in \mathcal{E}$, where $x_0 = \sigma(\emptyset)$ and then Death responds by playing an element $a_0 \in x_0$. Then the Maiden plays $x_1 = \sigma(x_0, a_0)$, which — by the rules of the game — is a set in $\mathcal{E}$ and a subset of $x_0$, and Death responds with an element $a_1 \in x_1$ where $a_1 > a_0$. In general, for positive integers $n$, $x_n = \sigma(x_0, a_0, \ldots, x_{n-1}, a_{n-1})$, where $x_n \in \mathcal{E}$, $x_n \subseteq x_{n-1}$, $a_0, \ldots, a_{n-1}$ are the moves of Death, and $x_0, \ldots, x_{n-1}$ are the previous moves of the Maiden.

A strategy $\sigma$ for the Maiden is a winning strategy if, whenever the Maiden follows the strategy $\sigma$, she wins the game — no matter how sophisticated Death plays. For example, $\sigma$ is a winning strategy for the Maiden in the game $G^n$, if whenever $\{a_n : n \in \omega\} \subseteq \omega$ is such that $a_0 \in \sigma(\emptyset)$ and for all $n \in \omega$, $a_n < a_{n+1}$ and $a_{n+1} \in \sigma(x_0, a_0, \ldots, x_{n+1})$, then $\{a_n : n \in \omega\} \notin \mathcal{E}$.

Now, a free family $\mathcal{E}$ is called a Ramsey family if $\mathcal{E}$ has no winning strategy in the game $G^n$. In other words, no matter how sophisticated her strategy is, if $\mathcal{E}$ is a Ramsey family, then Death can win the game. Ramsey families will play an important role in the investigation of Mathias forcing notions (see Chapter 24).

Furthermore, a free family $\mathcal{E}$ is called a $P$-family if $\mathcal{E}$ has no winning strategy in the game $G^n$. $P$-families will play an important role in the investigation of restricted Silver forcing. In fact, in Chapter 22 it will be shown that Silver forcing restricted to a $P$-family (called Silver-like forcing) has the same combinatorial properties as unrestricted Silver forcing and as Grigorieff forcing, which is Silver forcing restricted to a $P$-point.

Obviously, the family $[\omega]^\omega$ is a Ramsey family and every Ramsey family is also a $P$-family. Now, the reader might guess that $[\omega]^\omega$ is not the only example and that there must be some relation between Ramsey families and Ramsey ultrafilters, as well as between $P$-families and $P$-points; this is indeed the case.

**Theorem 10.17.** For free ultrafilters $\mathcal{U} \subseteq [\omega]^\omega$ we have:

1. $\mathcal{U}$ is a Ramsey ultrafilter if and only if $\mathcal{U}$ is a Ramsey family.
2. $\mathcal{U}$ is a $P$-point if and only if $\mathcal{U}$ is a $P$-family.

**Proof.** (a) We have to show that $\mathcal{U} \subseteq [\omega]^\omega$ is a Ramsey ultrafilter if and only if the Maiden plays the game $G_\mathcal{U}$ by following a strategy, Death can win.

   (⇒) Under the assumption that the free ultrafilter $\mathcal{U}$ is not Ramsey we construct a winning strategy for the Maiden in the game $G_\mathcal{U}$. If $\mathcal{U}$ is not a Ramsey ultrafilter, then, by Proposition 10.6, there exists a set $\{x_n : n \in$
ω\} \subseteq \mathcal{U} such that for each function \( f \in \omega \) with \( f(0) \in x \) and \( f(n+1) \in x_{f(n)} \) we have \( f[a] \notin \mathcal{U} \). Let \( \sigma(\emptyset) := x_0 \), and for \( n \in \omega \) let \( \sigma(x_0, a_0, \ldots, x_n, a_n) := x_{a_n} \). By the rules of \( G_{xy} \), \( a_{n+1} \in x_{a_n} \). Define \( f \in \omega \) by stipulating \( f(n) = a_n \). Then \( f(0) \in x_0 \) and for all \( n \in \omega \) we have \( f(n+1) \in x_{f(n)} \), and therefore \( \{ f(n) : n \in \omega \} \notin \mathcal{U} \). Thus, \( \{ a_n : n \in \omega \} \notin \mathcal{U} \), which shows that Death loses the game (i.e., \( \sigma \) is a winning strategy for the Maiden), and consequently, \( \mathcal{U} \) is not a Ramsey family.

\((\Rightarrow)\) Under the assumption that the free ultrafilter \( \mathcal{U} \) is Ramsey we show that no strategy for the Maiden is a winning strategy. Let \( \sigma \) be any strategy for the Maiden, let \( x_0 := \sigma(\emptyset) \), and for \( s = \{ c_0, \ldots, c_n \} \in \text{fin}(\omega) \) let

\[
x_s = \begin{cases} \sigma(x_0, c_0, \ldots, x_n, c_n) & \text{if } \forall k \leq n \in \mathcal{U} \\ \omega & \text{otherwise} \end{cases}
\]

Notice that in the first case, \( \sigma(x_0, c_0, \ldots, x_n, c_n) = x_{n+1} \). If \( \mathcal{U} \) is a Ramsey ultrafilter, then \( \mathcal{U} \) is happy. Thus, there exists an \( x \in \mathcal{U} \) such that \( x \subseteq x_0 \) and \( x \setminus s^+ \subseteq x \), whenever \( s \in x \). In particular, if \( x = \{ a_n : n \in \omega \} \) with \( a_n < a_{n+1} \) (for all \( n \in \omega \)), then \( a_0 \in x_0 \) and for all \( n \in \omega \), \( x \setminus \{ a_0, \ldots, a_n \} = \{ a_{n+1}, a_{n+2}, \ldots \} \subseteq x_{a_0, \ldots, a_n} = x_{a_n+1} \). Hence, for all \( n \in \omega \) we have \( a_n \in x_n \). In particular, whenever the Maiden follows the strategy \( \sigma \), Death wins the game by playing the sequence \( \{ a_n : n \in \omega \} \). So, \( \sigma \) is not a winning strategy for the Maiden, and since \( \sigma \) was arbitrary, the Maiden does not have a winning strategy.

\((\Leftarrow)\) The proof is similar to that of \((\alpha)\), i.e., we show that the Maiden has a winning strategy in the game \( G_{xy} \) iff the free ultrafilter \( \mathcal{U} \) is not a \( P \)-point.

\((\Rightarrow)\) Suppose that \( \mathcal{U} \) is not a \( P \)-point. Then, by Fact 10.11.(b), there exists a set \( \{ y_n : n \in \omega \} \subseteq \mathcal{U} \) such that whenever \( y \in [\omega]^{\omega} \) has the property that for all \( n \in \omega \), \( y \setminus y_n \) is finite, then \( y \notin \mathcal{U} \). Let \( \sigma(\emptyset) := y_0 \) (i.e., \( x_0 = y_0 \)), and for any \( k \in \omega \) and \( \{ s_0, \ldots, s_k \} \subseteq \text{fin}(\omega) \) let \( \sigma(x_0, s_0, \ldots, x_k, s_k) := \bigcap_{i \leq k} y_i \setminus \bigcup_{i \leq k} s_i \). If the Maiden follows that strategy \( \sigma \) and the sequence \( \{ s_k : k \in \omega \} \) represents the moves of Death, then for all \( n \in \omega \) we have \( \bigcup_{k \in \omega} s_k \notin \mathcal{U} \), which shows that Death loses the game, or in other words, \( \sigma \) is a winning strategy for the Maiden.

\((\Leftarrow)\) Under the assumption that \( \mathcal{U} \) is a \( P \)-point we show that no strategy for the Maiden is a winning strategy. Let \( \sigma \) be any strategy for the Maiden. We have to show that Death can win. Define \( X_n \) as the family of sets played by the Maiden in her first \( n+1 \) moves, assuming that she is following the strategy \( \sigma \) and Death plays in his first \( n \) moves only sets \( s_k \subseteq n \) (for \( k < n \)). More formally, \( x_0 = \sigma(\emptyset) \), and for positive integers \( k \leq n \), \( x_k \in X_n \) iff there are \( s_0, \ldots, s_{k-1} \subseteq n \) such that for all \( i < k \), \( s_i \subseteq x_i \cap n+1 \), where \( x_{i+1} = \sigma(x_0, s_0, \ldots, x_i, s_i) \). Clearly, for every \( n \in \omega \), \( X_n \) is finite, and since \( \mathcal{U} \) is an ultrafilter, \( y_n := \bigcap X_n \) belongs to \( \mathcal{U} \). Moreover, since \( \mathcal{U} \) is a \( P \)-point, by Fact 10.11.(c) there is a set \( y \in \mathcal{U} \) and a strictly increasing function \( f \in [\omega]^{\omega} \) such that for all \( n \in \omega \), \( y \setminus f(n) \subseteq y_n \). Let \( k_0 := f(0) \), and in general, for \( n \in \omega \) let \( k_{n+1} := f(k_n) \). Since \( \mathcal{U} \) is an ultrafilter, either
Ramsey families and $P$-families

$$y_0 = \bigcup_{n \in \omega} [k_{2n}, k_{2n+1}) \text{ or } y_1 = \omega \setminus y_0$$

belongs to $\mathcal{V}$. Without loss of generality we may assume that $y_1 \in \mathcal{V}$, in particular, $y_1 \cap y \in \mathcal{V}$. Consider the run

$$(x_0, s_0, x_1, x_1, \ldots)$$

of the game $\mathcal{G}_y^*$, where the Maiden plays according to the strategy $\sigma$ and Death plays

$$s_n = \begin{cases} [k_{2j+1}, k_{2j+2}) \cap y & \text{if } n = k_{2j} \text{ (for some } j \in \omega), \\ \emptyset & \text{otherwise.} \end{cases}$$

It is clear that the Maiden loses the game (i.e., $\bigcup_{n \in \omega} s_n \in \mathcal{V}$). It remains to check that the moves of Death are legal (i.e., satisfy the rules of the game $\mathcal{G}_y^*$). First notice that for all positive integers $j$, $s_{k_{2j-2}} \subseteq k_{2j}$. Thus, if $n = k_{2j}$, then for all $k < n$ we have $s_k \subseteq n$. Now, if $n = k_{2j}$ for some $j \in \omega$, then $s_n = s_{k_{2j}} = [k_{2j+1}, k_{2j+2}) \cap y$. Further, we have

$$y \setminus k_{2j+1} = y \setminus f(k_{2j}) \subseteq y_{k_{2j}} = \bigcap \{x_0, \ldots, x_{k_{2j}}\},$$

and in particular, for $n = k_{2j}$ we get $s_n = s_{k_{2j}} \subseteq x_{k_{2j}} = x_n$. Hence, for all $n \in \omega$, $s_n \subseteq x_n$. \hfill \dagger

Roughly speaking, Ramsey families are a kind of generalised Ramsey ultrafilters and $P$-families are a kind of generalised $P$-points.

Let us turn back to happy families and let us compare them with Ramsey families. At a first glance, happy families and Ramsey families look very similar. However, it turns out that the conditions for a Ramsey family are slightly stronger than for a happy family. This is because in the definition of happy families we require that they contain sets which diagonalise certain subfamilies having the finite intersection property. On the other hand, a strategy of the Maiden in the game $\mathcal{G}_x$ can be quite arbitrary: Even though the sets played by her in a run of $\mathcal{G}_x$ form a decreasing sequence, the family of possible moves of the Maiden does not necessarily have the finite intersection property. Of course, by restricting the set of strategies the Maiden can choose from, we could make sure that all happy families are Ramsey. In fact we just have to require that all the moves of the Maiden—no matter what Death is playing—belong to some family which has the finite intersection property. However, the definition of Ramsey families given above has the advantage that the Maiden is able—by a winning strategy—to defeat Death in the game $\mathcal{G}_x$ even in some cases when $\mathcal{E}$ is happy (see Proposition 10.19).

Below, we show first that every Ramsey family is happy, and then we show that there are happy families which are not even $P$-families. Thus, Ramsey families are smaller “class” (i.e., families who originate from the same family and have the same name) than happy families.
FACT 10.18. Every Ramsey family is happy.

Proof. Let \( \mathcal{F} \) be a free family which is not happy. Thus, there exists a set \( \mathcal{C} = \{ y_s : s \in \text{fin}(\omega) \} \subseteq \mathcal{F} \) such that \( \text{fil}(\mathcal{C}) \subseteq \mathcal{F} \) but no \( \mathcal{G} \in \mathcal{F} \) diagonalises \( \mathcal{C} \). Let \( \sigma(0) := x_0 \) and for \( n \in \omega \) and \( s = \{ a_0, \ldots, a_n \} \in \text{fin}(\omega) \) let \( \sigma(x_0, a_0, \ldots, x_n, a_n) := \bigcap_{i \subseteq s} y_i \). It is not hard to verify that in the game \( \mathcal{G}_{\mathcal{C}} \), \( \sigma \) is a winning strategy for the Maiden.

By Proposition 10.5 we know that every mad family induces a happy family. This type of happy families provides examples of happy families which are not Ramsey families, in fact, which are not even \( P \)-families.

PROPOSITION 10.19. Not every happy family is Ramsey; moreover, not every happy family is a \( P \)-family.

Proof. It is enough to construct a happy family which is not a \( P \)-family: Let \( \{ t_k : k \in \omega \} \) be an enumeration of \( \bigcup_{n \in \omega} t_n \omega \) such that for all \( i, j \in \omega, t_i \subseteq t_j \) implies \( i \leq j \), in particular, \( t_0 = \emptyset \). For functions \( f \in \omega^\omega \) define the set \( x_f \in [\omega]^\omega \) by stipulating

\[
x_f := \{ k : \exists n, i, j \in \omega (f|_n = t_i \land f|_{n+1} = t_j \land i \leq k < j \land t_i \subseteq t_k) \}.
\]

Obviously, for any distinct functions \( f, g \in \omega^\omega \), \( x_f \cap x_g \) is finite (compare with the sets constructed in the proof of Proposition 8.6). Now, let \( \mathcal{A}_0 := \{ x_f : f \in \omega^\omega \} \). Then \( \mathcal{A}_0 \subseteq [\omega]^\omega \) is a set of pairwise almost disjoint sets which can be extended to a mad family, say \( \mathcal{A} \). Recall that by Proposition 10.5, \( \mathcal{F}_\mathcal{A}^+ \) is a happy family.

We show that \( \mathcal{F}_\mathcal{A}^+ \) is not a \( P \)-family: Let \( k_0 := 0 \) and let \( x_0 := \omega \) be the first move of the Maiden, and let \( s_0 \) be Death’s response. In general, if \( s_n \) is the \( n^{th} \) move of Death, then the Maiden chooses \( k_{n+1} \) such that \( k_{n+1} \geq \max(s_n) \), \( |t_{k_{n+1}}| = n + 1 \), and \( t_{k_n} \subseteq t_{k_{n+1}} \), and then she plays

\[
x_{n+1} = \{ i \in \omega : t_{k_{n+1}} \subseteq s_i \}.
\]

Obviously, for every \( n \in \omega \) we have \( x_{n+1} \subseteq x_n \). Moreover, all moves of the Maiden are legal:

Claim. For every \( n \in \omega, x_n \in \mathcal{F}_\mathcal{A}^+ \).

Proof of Claim. Firstly, for every \( n \in \omega \), \( x_n \) has infinite intersection with infinitely many members of \( \mathcal{A}_0 \). Indeed, \( x_n \cap x_f \) is infinite whenever \( f|_n = t_{k_n} \). Secondly, for every \( z \in \mathcal{F}_\mathcal{A} \) there are finitely many \( y_0, \ldots, y_k \in \mathcal{A} \) such that \( (y_0 \cup \ldots \cup y_k)^\omega \subseteq^* z \). Now, for \( x_n \) let \( x_f \in \mathcal{A}_0 \setminus \{ y_0, \ldots, y_k \} \) such that \( x_f \cap x_n \) is infinite. Then, since \( x_f \cap (y_0 \cup \ldots \cup y_k) \) is finite, \( x_f \subseteq^* z \). Hence, \( x_n \cap z \) is infinite which shows that \( x_n \in \mathcal{F}_\mathcal{A}^+ \). By the Maiden’s strategy, \( \bigcup_{n \in \omega} t_{k_n} = f \) for some particular function \( f \in \omega^\omega \). Moreover, \( \bigcup_{n \in \omega} s_t \subseteq x_f \in \mathcal{A}_0 \), and since subsets of members of \( \mathcal{A}_0 \) do not belong to \( \mathcal{F}_\mathcal{A}^+ \), \( \bigcup_{n \in \omega} s_t \notin \mathcal{F}_\mathcal{A}^+ \). Hence, Death loses the game, no matter what he is playing, which shows that the Maiden has a winning strategy in the game \( \mathcal{G}_{\mathcal{F}_\mathcal{A}}^+ \). In other words, the happy family \( \mathcal{F}_\mathcal{A}^+ \) is not a \( P \)-family. \( \square \)
Related Results

Notes

Happy families and Ramsey ultrafilters. Happy families were introduced by Mathias [8] in order to investigate the Ramsey property as well as Ramsey ultrafilters. Furthermore, happy families are closely related to Mathias forcing — also introduced in [8] — which will be discussed in Chapter 24. Fact 10.3 and Proposition 10.5 are taken from Mathias [8, p. 61ff.]. Proposition 10.6 is due to Mathias [8, Proposition 0.8] and the characterisation of Ramsey ultrafilters (i.e., Proposition 10.7 and Fact 10.8) is taken from Bartoszyński and Judah [1, Theorem 4.5.2] and Booth [3, Theorem 4.9] (according to Booth [3, p. 19], most of [3, Theorem 4.9] is due to Kunen).

On P-points. A point x of a topological space X is called a P-point if every intersection of countably many open sets containing x, contains an open set containing x. Now, the ultrafilters we called P-points are in fact the P-points of the topological space βω \ ω (defined on page 222). The existence of P-points of the space βω \ ω cannot be shown in ZFC (see Related Result 68). However, by Theorem 10.16, which is due to Ketten [6] (see also Bartoszyński and Judah [1, Theorem 4.4.5]), it follows that P-points exist if we assume CH — which was first proved by Rudin [10].

Ramsey families and P-families. Ramsey families and P-families were first introduced and studied by Laflamme in [7], where the filters associated to a Ramsey family are called +-Ramsey filters, and the filters associated to a P-family are called P+-filters. However, Theorem 10.17 is due to Galvin and Shelah (see Bartoszyński and Judah [1, Theorems 4.5.3 & 4.4.4]), and Proposition 10.19 is a generalisation of Balbesien [4, Proposition 6.2].

Related Results

64. On the existence of Ramsey ultrafilters. Mathias showed that under CH, every happy family contains a Ramsey ultrafilter (see Mathias [8, Proposition 0.11]). In particular, this shows that Ramsey ultrafilters exist if we assume CH (according to Booth [3, p. 23], this was first shown by Galvin). However, by Proposition 10.9 we know that p = c is sufficient for the existence of Ramsey ultrafilters. With Martin’s Axiom in place of p = c, this result is due to Booth [3, Theorem 4.14]. Furthermore, Keisler showed that if we assume CH, then there are ν’ mutually non-isomorphic Ramsey ultrafilters (see Blass [2, p. 148]). Finally, by combining the proofs of Keisler and Booth, Blass [2, Theorem 2] showed that t = c (for t see Chapter 8 [Related Result 52] is enough to get ν’ mutually non-isomorphic Ramsey ultrafilters (see Proposition 13.9 for a slightly more general result). On the other hand, we shall see in Chapter 25 that the existence of Ramsey ultrafilters is independent of ZFC (see also Chapter 21 [Related Result 114]).

65. There may exist a unique Ramsey ultrafilter. We have seen above that we can have infinitely many Ramsey ultrafilters or none. So, it is natural to ask whether it is also consistent with ZFC that there exists, up to permutations of ω, a unique Ramsey ultrafilter. Now, Shelah [12, VI §5] proved that this is indeed the case.
Moreover, it is even consistent with ZFC that there are, up to permutations of \( \omega \), exactly two Ramsey ultrafilters (see Shelah [12, p. 335]).

66. There may be \( P \)-points which are not Ramsey. Booth [3, Theorem 4.12] showed that if we assume CH (or Martin’s Axiom), there are \( P \)-points which are not Ramsey (i.e., which are not \( Q \)-points). For examples of \( P \)-points which are not \( Q \)-points see Proposition 25.11.

67. On the existence of \( Q \)-points. Mathias [Proposition 10][9] showed that \( \mathfrak{d} = \omega_1 \) implies the existence of \( Q \)-points. Recall that by Proposition 10.9, \( \mathfrak{p} = \mathfrak{c} \) implies the existence of Ramsey ultrafilters; in particular the existence of \( P \)-points and \( Q \)-points. Thus, the existence of \( Q \)-points is consistent with \( \mathfrak{d} > \omega_1 \). However, if there are just \( P \)-points but no \( Q \)-points, then we must have \( \mathfrak{d} > \omega_1 \).

68. On the existence of \( P \)-points. \( P \)-points were studied by Rudin [10], who proved, assuming CH, that they exist and that any of them can be mapped to any other by a homeomorphism of \( \beta \omega \setminus \omega \) onto itself. In particular, CH implies the existence of \( P \)-points. Of course, this follows from the fact that CH implies the existence of Ramsey ultrafilters, and Ramsey ultrafilters are \( P \)-points. However, as we have seen above, the converse is not true (and there are models of ZFC in which there are \( P \)-points but no Ramsey ultrafilters). Now, it is natural to ask whether there are models of ZFC in which there are no \( P \)-points. Let us consider how models of ZFC are constructed in which there are no Ramsey ultrafilters. In order to construct a model of ZFC in which there are no Ramsey ultrafilters, one usually makes sure that the model does not contain any \( Q \)-points (see for example the proof of Proposition 25.11). To some extent, \( P \)-points are weaker than \( Q \)-points and therefore it is more difficult to construct a model in which there are no \( P \)-points. However, Shelah constructed such a model in [11] (see also Shelah [12, VI §4], Wimmers [14], or Bartoszyński and Judah [1, 4.4.7]). Moreover, like for Ramsey ultrafilters, it is consistent with ZFC that, up to up permutations of \( \omega \), there exists a single \( P \)-point (see Shelah [12, XVIII §4]).

69. Simple \( P_\kappa \)-points. For any regular uncountable cardinal \( \kappa \), a free ultrafilter \( \mathcal{U} \subseteq [\omega]^\omega \) is called a simple \( P_\kappa \)-point if \( \mathcal{U} \) is generated by an almost decreasing (i.e., modulo finite) \( \kappa \)-sequence of infinite subsets of \( \omega \). Clearly, every simple \( P_\kappa \)-point is a \( P \)-point. It is conjectured that the existence of both, a simple \( P_\omega \)-point and a \( P_\omega \)-point, is consistent with ZFC. (For weak \( P \)-points and other points in \( \beta \omega \setminus \omega \) see for example van Mill [13, Section 4].)

70. Rapid and unbounded filters. A free filter \( \mathcal{F} \subseteq [\omega]^\omega \) is called a rapid filter if for each \( f \in [\omega]^{\omega} \) there exists an \( \varepsilon \in \mathcal{F} \) such that for all \( n \in \omega \), \( |x \cap f(n)| \leq n \). By definition, if \( \mathcal{F} \) is a rapid filter, then \( \{ f_x : x \in \mathcal{F} \} \) is a dominating family. It is not hard to verify that all \( Q \)-points are rapid (see Fact 25.10), but the converse does not hold (see for example Bartoszyński and Judah [1, Lemma 4.6.3] and in particular the remark after the proof of that lemma). However, like for \( P \)-points or \( Q \)-points, the existence of rapid filter is independent of ZFC (see Proposition 25.11). A weaker notion than that of rapid filters is the notion of unbounded filters, where a free filter \( \mathcal{F} \subseteq [\omega]^\omega \) is called unbounded if the family \( \{ f_x : x \in \mathcal{F} \} \) is unbounded. Since every free ultrafilter induces an unbounded family (cf. Proposition 10.15(a)), unbounded filters always exist. Furthermore, one can show that every unbounded filter induces a set which does
not have the Ramsey property (for a slightly more general result see Judah [5, Fact 8]).

71. Another characterisation of Ramsey ultrafilters. Let \( \mathcal{U} \subseteq [\omega]^\omega \) be an ultrafilter. The game \( G'_\mathcal{U} \) is defined as follows.

\[
G'_\mathcal{U} : \quad \text{Maiden} \quad (a_0, x_0) \quad (a_1, x_1) \quad (a_2, x_2) \quad \ldots \\
\text{Death} \quad y_0 \quad y_1 \quad y_2 \quad \ldots
\]

The sets \( y_i \) and \( x_i \) played by Death and the Maiden respectively must belong to the ultrafilter \( \mathcal{U} \), and for each \( i \in \omega \), \( a_{i+1} \) must be a member of \( y_i \). Furthermore, for each \( i \in \omega \) we require that \( x_{i+1} \subseteq y_i \subseteq x_i \) and that \( a_i < \min(x_i) \). Finally, the Maiden wins the game \( G'_\mathcal{U} \) if and only if \( \{a_i : i \in \omega \} \) does not belong to the ultrafilter \( \mathcal{U} \).

In 2002, Claude Laflamme showed me that \( \mathcal{U} \) is a Ramsey ultrafilter if and only if the Maiden has no winning strategy in the game \( G'_\mathcal{U} \).

72. On strongly maximal almost disjoint families. A mad family \( \mathcal{F} \) is called strongly maximal almost disjoint if given countably many members of \( \mathcal{F}' \), then there is a member of \( \mathcal{F} \) that meets each of them in an infinite set.

For a free family \( \mathcal{E} \), consider the following game. The moves of the Maiden are members of \( \mathcal{E} \) and the Death responses like in the game \( G'_\mathcal{U} \). Furthermore, Death wins if and only if the set of integers played by Death belongs to \( \mathcal{F} \), but has infinite intersection with each set played by the Maiden.

If \( \mathcal{F} \) is a mad family, then obviously, in the game described above, the Maiden has a winning strategy if and only if \( \mathcal{F} \) is not strongly maximal almost disjoint, which motivates the following question: Is it the case that for a mad family \( \mathcal{F} \), \( \mathcal{F}' \) is Ramsey if and only if \( \mathcal{F} \) is strongly maximal almost disjoint?

References


Coda: A Dual Form of Ramsey’s Theorem

Musicians wanted compositions to end on a perfect consonance, because they correctly say that the perfection of anything depends upon and is judged by its end. Since they found that among consonances no greater perfection could be found than in the octave, they made it a fixed rule that each composition should terminate on the octave or unison and no other interval.

Giuseppe Zarlino
Le Istitutioni Harmoniche, 1558

In this chapter we shall present some results in dual Ramsey Theory, i.e., Ramsey type results dealing with partitions of ω. The word “dual” is motivated by the following fact: Each infinite subset of ω corresponds to the image of an injective function from ω into ω, whereas each infinite partition of ω corresponds to the set of pre-images of elements of ω of a surjective function from ω onto ω. Similarly, n-element subsets of ω correspond to images of injective functions from n into ω, whereas n-block partitions of ω correspond to pre-images of surjective functions from ω onto n. Thus, to some extent, subsets of ω and partitions of ω are dual to each other.

The Hales-Jewett Theorem

Since we introduced Ramsey’s Theorem in Chapter 2, we have used different forms of this powerful combinatorial tool in various applications. However, Ramsey’s Theorem is neither the only nor the earliest Ramsey-type result. In fact, the following theorem is one of the earliest results in Ramsey Theory.
Theorem 11.1 (van der Waerden). For any positive integers $r$ and $n$, there is a positive integer $N$ such that for every $r$-colouring of the set $\{0, 1, \ldots, N\}$ we find always a monochromatic (non-constant) arithmetic progression of length $n$.

Instead of a proof, let us consider van der Waerden’s Theorem from a more combinatorial point of view: Firstly, for some positive integer $l$, identify the integers $a \in [0, n^l]$ with the $l$-tuples $(a_0 \ldots a_{l-1})$ formed from the base-$n$ representation of $a$, i.e., $a = \sum_{i \leq l} a_i n^i$ and for all $i \leq l$, $0 \leq a_i < n$. Concerning arithmetic progressions, notice that for example the $l$-tuples

\begin{align*}
(a_0 \ldots a_{l-1} & \quad 0 \quad a_{i+1} \ldots a_{j-1} \quad 0 \quad a_{j+1} \ldots a_{l-1}) \\
(a_0 \ldots a_{l-1} & \quad 1 \quad a_{i+1} \ldots a_{j-1} \quad 1 \quad a_{j+1} \ldots a_{l-1}) \\
& \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
(a_0 \ldots a_{l-1} & \quad n - 2 \quad a_{i+1} \ldots a_{j-1} \quad n - 2 \quad a_{j+1} \ldots a_{l-1}) \\
(a_0 \ldots a_{l-1} & \quad n - 1 \quad a_{i+1} \ldots a_{j-1} \quad n - 1 \quad a_{j+1} \ldots a_{l-1})
\end{align*}

correspond to an arithmetic progression of length $n$ with common difference $n^i + n^j$. Let us call for the moment arithmetic progressions of length $n$ of that type special arithmetic progressions. Notice that not every arithmetic progression of length $n$ is special. However, if we could show that for all positive integers $n$ and $r$ there exists a positive integer $l$ such that for every $r$-colouring of $[0, n^l]$ we find a monochromatic special arithmetic progression, then this would obviously prove van der Waerden’s Theorem.

Now, identify the set of $l$-tuples $(a_0 \ldots a_{l-1})$ with the set of functions $f$ from $l$ to $n$, denoted $^lf$, by stipulating $f(k) = a_k$ (for all $k \in l$). Consequently, we can identify every $r$-colouring of $[0, n^l]$ with an $r$-colouring of $^lf$. Notice that for a non-empty set $s \subseteq l$ and a function $g : l \setminus s \rightarrow r$, the set $\{ f \in ^lf : f|_{l \setminus s} = g \land f|_s \text{ is constant} \}$ corresponds to a special arithmetic progression. In the example of a special arithmetic progression given above we have $s = \{i, j\}$ and $g(m) = a_m$ (for all $m \in l \setminus s$). Hence, in terms of functions from $l$ to $n$, van der Waerden’s Theorem is just a corollary of the following Ramsey-type theorem.

Theorem 11.2 (Hales-Jewett Theorem). For all positive integers $n, r \in \omega$ there exists a positive integer $l \in \omega$ such that for any $r$-colouring of $^lf$ there is always a non-empty set $s \subseteq l$ and a function $g : l \setminus s \rightarrow n$ such that $\{ f \in ^lf : f|_{l \setminus s} = g \land f|_s \text{ is constant} \}$ is monochromatic.

For given positive integers $n, r \in \omega$, the Hales-Jewett function $HJ(n, r)$ denotes the smallest such integer $l$. In particular, for all positive integers $r$, $HJ(1, r) = 1$.

Hales and Jewett proved their theorem almost 40 years after van der Waerden proved his. In the original proof, they used — like van der Waerden — a
double induction which led to an extremely fast growing upper bound for the Hales-Jewett function \( HJ(n, r) \). The proof of the HALE-JEWETT THEOREM given here—which is due to Shelah and modified by Matet involving the Finite Ramsey Theorem—uses just simple induction on \( n \) and provides a much better bound for the associated function \( HJ(n, r) \).

Before we can give a proof of the HALE-JEWETT THEOREM, including the bounds for \( HJ(n, r) \), we have to introduce a kind of Ramsey number (cf. Chapter 2 | Related Result 1): By the Finite Ramsey Theorem 2.3 we know that for any positive integers \( r, p, \) and \( q \), where \( q \leq p \), there exists a positive integer \( m \) such that for every \( r \)-colouring \( \pi : [m]^q \to r \) we find a \( p \)-element set \( t \in [m]^p \) such that \( \pi|_{[t]} \) is constant; let \( R^2(p) \) denote the least such \( m \).

**Theorem 11.3.** For positive integers \( n \) and \( r \) let \( l = HJ(n, r) \), \( a = (n+1)^l - n^l \), \( k = r^a \), and \( m = R^{2l-1}(2l) \). Then \( HJ(n+1, r) < m \).

**Proof.** Let \( \mathcal{F} \) be the set of all non-decreasing functions \( f \in \mathbb{N}^m \) (i.e. \( f(0) \leq f(1) \leq \ldots \leq f(2l-1) \)) such that \( 2l - 1 \leq |f| \) (i.e. \( f(i) = f(i+1) \)) for at most one \( i \leq 2l - 2 \). Let \( \mathcal{F}_0 = \{ f \in \mathcal{F} : |f| = 2l \} \) and let \( \mathcal{F}_1 = \mathcal{F} \setminus \mathcal{F}_0 \).

Notice that for each \( f \in \mathcal{F}_1 \) there exists a unique \( i \leq 2l - 2 \) such that \( f(i) = f(i+1) \). So, for every \( i \leq 2l - 2 \) let \( F_i = \{ f \in \mathcal{F}_1 : f(i) = f(i+1) \} \).

Then \( \mathcal{F}_1 = \bigcup_{0 \leq i \leq 2l-2} F_i \).

For \( f \in \mathcal{F} \) and \( i \in [1, 2l-1] \) let \( I^f_i = [f(i-1), f(i)] \), and let \( I^f_0 = [0, f(0)] \) and \( I^f_{2l} = [f(2l-1), m] \). Notice, if \( f(0) = 0 \) then \( I^f_0 = \emptyset \), if \( f(2l-1) = m-1 \) then \( I^f_{2l} = \{ m \} \), and if \( f \in F_i \), for some \( i \leq 2l - 2 \), then \( I^f_{i+1} = \emptyset \). Define \( g : I((n+1) \times \mathcal{F} \to m^{-1}n + 1 \) such that for each \( j \leq 2l \), \( g(h, f)|_{I^f_j} \) is constant, where

\[
g(h, f)|_{I^f_j} \text{ is constantly } \begin{cases} n - 1 & \text{if } j \equiv 0 \mod 4, \\ n & \text{if } j \equiv 2 \mod 4, \\ h((j-1)/2) & \text{otherwise.} \end{cases}
\]

For \( h \in I(n+1) \) and \( f \in \mathcal{F} \), \( g(h, f) \) is visualised by the following figure:

\[
0 \begin{array}{ccc} f(0) & f(1) & f(2) & f(3) & f(2l-2) & f(2l-1) & m-1 \\
I^f_0 & I^f_1 & I^f_2 & I^f_3 & I^f_{2l-2} & I^f_{2l-1} & I^f_{2l} \\
in-1 & n & h(0) & h(1) & h(t-1) & h(t) & n \text{ or } n-1 \end{array}
\]

Notice that for \( f \in F_{2i} \) and \( h \in \mathcal{H} \) we have the following situation.

\[
g(h, t) : I^f_{2i} \begin{array}{c} f(2i) = f(2i+1) \end{array} I^f_{2i+2} \]

For \( i \in I \), let \( H_i \subseteq I(n+1) \) be the set of all functions \( h : l \to (n+1) \) such that \( h(i) = n \) and for all \( j < n \), \( h(j) < n \). Let \( \mathcal{H} = \bigcup_{i \in I} H_i \). Notice that \( \mathcal{H} \)
is the set of all functions $h \in \iota(n + 1)$ such that $h(i) = n$ for some $i \in l$. For each $i \in l$ define a function $g_i : H_i \times [m]^{2l-1} \to m^{-1}(n + 1)$ by stipulating

$$g_i(h, s) = g(h, f_{s,i}),$$

where $f_{s,i} \in F_{2l}$ is such that $f_{s,i}[2l] = s$.

Fix a colouring $\pi : (m^{-1}(n + 1)) \to r$. Notice that we can apply $\pi$ to $g(h, f)$ (where $h \in \iota(n + 1)$ and $f \in F$) as well as to $g_i(h, s)$ (where $h \in H_i$ and $s \in [m]^{2l-1}$). Recall that we want to show $HJ(n + 1, r) \leq m - 1$, where $m = R_{2l-1}^{2l-1}(2l)$. By definition of $m$, for every colouring $\tau : [m]^{2l-1} \to k$ we find a 2l-element set $t \in [m]^{2l}$ such that $\tau|_{[t]^{2l-1}}$ is constant. In order to apply the properties of $m$, we have to find a suitable $k$-colouring of $[m]^{2l-1}$. Firstly, recall that $k = r^a$, where $a = (n + 1)l - n^l$. Now, $|\iota(n + 1) \setminus H| = n^l$, and since $|\iota(n + 1)| = (n + 1)^l$ we get $|H| = (n + 1)^l - n^l$. Thus, $a = |H|$, and therefore $k = |H_r|$. Now, define the colouring $\tau : [m]^{2l-1} \to H_r$ by stipulating

$$\tau(s)(h) = \pi(g_i(h, s)) \text{ whenever } h \in H_i \text{ for some } i \in l.$$  

By definition of $m$, there exists a 2l-element set $t \in [m]^{2l}$ such that $\tau|_{[t]^{2l-1}}$ is constant. In particular, for any $s_0, s_1 \in [t]^{2l-1}$ and any $h \in H_i$ we have

$$\pi(g_i(h, s_0)) = \pi(g_i(h, s_1)). \quad (*)$$

Let $f_t \in F_0$ be such that $f_t[2l] = t$ and define the colouring $\pi' : \iota_n \to r$ by stipulating $\pi'(h) := \pi(g(h, f_t))$. Since $l = HJ(n, r)$, there exists a non-empty set $u_0 \subseteq l$ and a function $\tilde{h} : l \setminus u_0 \to n$ such that

$$\tilde{H} = \{h \in \iota_n : h|_{\setminus u_0} = \tilde{h} \wedge h|_{u_0} \text{ is constant}\}$$

is monochromatic. Notice that $\tilde{H} \subseteq \iota_n \subseteq \iota(n + 1)$ and that $\pi'|_{\{g(h, f_t) : h \in \tilde{H}\}}$ is constant. Let $h_0 \in \iota(n + 1)$ be such that $h_0|_{\setminus u_0} = \tilde{h}$ and $h_0|_{u_0}$ is constant $n$. If we can show that $\{g(h, f_t) : h \in \tilde{H} \cup h = h_0\}$ is monochromatic, then we are done. In fact, it is enough to show that $\pi(g(h_0, f_t)) = \pi(g(h_0, f_t))$, where $h_0 \in \tilde{H}$ is such that for all $i \in l$, $h_0(i) := \min \{h_0(i), n - 1\}$. This is done by induction on the size of $u_0$, but first we have to do some preliminary work: For $i \in l$ and $h \in H_i$ define $h' \in \iota(n + 1)$ by stipulating

$$h'(j) = \begin{cases} n - 1 & \text{if } j = i, \\ h(j) & \text{otherwise.} \end{cases}$$

Notice that either $h' \in H_i$ for some $i' > i$, or $h' \in \iota_n$. We show now that for every $h \in H_i$, $\pi'(g(h, f_t)) = \pi(g(h', f_t))$. We consider the cases $i$ odd and $i$ even separately.

For $i$ odd and $h \in H_i$ we have the following situation:
\[ g(h, f_i) : \begin{array}{ccc}
 f_i(2i) & f_i(2i + 1) \\
 n & h(i) = n & n - 1
\end{array} \]

Similarly, for \( i \) even and \( h \in H_i \) we get:

\[ g(h, f_i) : \begin{array}{ccc}
 f_i(2i) & f_i(2i + 1) \\
 n - 1 & h(i) = n & n
\end{array} \]

By (\pm) we have \( \pi(g(h, t \setminus \{f_i(2i)\})) = \pi(g(h, t \setminus \{f_i(2i + 1)\})) \), and since we obviously have

\[
\begin{align*}
  g(h, f_i) &= g_i(h, t \setminus \{f_i(2i)\}) & \text{if } i \text{ is odd,} \\
  g(h', f_i) &= g_i(h, t \setminus \{f_i(2i + 1)\}) & \text{if } i \text{ is even,}
\end{align*}
\]

we get

\[ \pi(g(h, f_i)) = \pi(g(h', f_i)). \]

Now we are ready to show that \( \pi(g(h_0, f_i)) = \pi(g(h_0, f_i)) \): For \( j < |u_0| \) let \( h_{j+1} := h'_j \). Then, by the preceding fact we have

\[ \pi(g(h_0, f_i)) = \pi(g(h_1, f_i)) = \ldots = \pi(g(h_{|u_0|}, f_i)) \],

and since \( h_{|u_0|} = h_0 \), we finally get \( \pi(g(h_0, f_i)) = \pi(g(h_0, f_i)) \), which completes the proof of Theorem 11.3 as well as of the Hales-Jewett Theorem.
The Hales-Jewett Theorem will be used to start the induction in the proof of Carlson’s Lemma (see Claim 2), where Carlson’s Lemma is the crucial part in the proof of a generalisation of Ramsey’s Theorem in terms of partitions—the main result of this chapter which will be called Partition Ramsey Theorem.

The Partition Ramsey Theorem is a very strong combinatorial result which implies the Hales-Jewett Theorem as well as some other Ramsey-type results like the Weak Halpern-Läuchli Theorem 11.6. However, before we can formulate and prove the Partition Ramsey Theorem, we have to introduce first the corresponding terminology.

Families of Partitions

Even though partitions have already been used in Chapter 10, let us introduce the notion of partition in a more formal way.

A set $P \subseteq \mathcal{P}(S)$ is a partition of the set $S$, if $\emptyset \notin P$, $\bigcup P = S$, and for all distinct $p_1, p_2 \in P$ we have $p_1 \cap p_2 = \emptyset$. A member of a partition $P$ is called a block of $P$ and $\text{Dom}(P) := \bigcup P$ is called the domain of $P$. A partition $P$ is called infinite, if $|P|$ is infinite (where $|P|$ denotes the cardinality of the set $P$); otherwise, the partition $P$ is called finite.

If $P$ and $Q$ are two partitions with the same domain, then $P$ is coarser than $Q$, or equivalently $Q$ is finer than $P$, if each block of $P$ is the union of blocks of $Q$. Notice that the relation “coarser” is a partial ordering on the set of partitions with a given domain, and that there are unique finest and coarsest partitions. For example with respect to partitions of $\omega$, the finest partition is $\{\{n\} : n \in \omega\}$ and the coarsest partition is $\{\}$.

Below, we are mainly interested in infinite partitions of $\omega$, denote by capital letters like $X, Y, Z, \ldots$, as well as in (finite) partitions of natural numbers, usually denoted by capital letters like $S, T, U, \ldots$. So, let $(\omega)^n$ denote the set of all infinite partitions of $\omega$ and let $(\mathbb{N})$ denote the set of all (finite) partitions $S$ with $\text{Dom}(S) \in \omega$. Notice that $S \in (\mathbb{N})$ iff $S$ is a partition of some natural number $n \in \omega$.

The following notation allows us to compare partitions with different domains: For partitions $P$ and $Q$ (e.g., $P \in (\mathbb{N})$ and $Q \in (\omega)^n$) we write $P \subseteq Q$ if for all blocks $p \in P$ the set $p \cap \text{Dom}(Q)$ is the union of some sets $q_i \cap \text{Dom}(P)$, where each $q_i$ is a block of $Q$. Notice that in general, $P \subseteq Q \subseteq P$ does not imply $P = Q$, except when $\text{Dom}(P) = \text{Dom}(Q)$. Furthermore, let $P \cap Q$ ($P \cup Q$) denote the finest (coarsest) partition $R$ such that $\text{Dom}(R) = \text{Dom}(P) \cup \text{Dom}(Q)$ and $R$ is coarser (finer) than $P$ and $Q$. In particular, if $\text{Dom}(P) \subseteq \text{Dom}(Q)$ then $P \cap Q \subseteq Q \subseteq P \cup Q$.

Let $S \in (\mathbb{N})$ and $X \in (\omega)^n$. If for each $s \in S$ there exists an $x \in X$ such that $x \cap \text{Dom}(S) = s$, we write $S \preceq X$. Similarly, for $S, T \in (\mathbb{N})$, where $\text{Dom}(S) \subseteq \text{Dom}(T)$, we write $S \preceq T$ if for each $s \in S$ there exists a $t \in T$ such that $t \cap \text{Dom}(S) = s$. Roughly speaking, $P \preceq Q$ is the same as saying “$Q$
Families of partitions

restricted to \(\text{Dom}(P)\) is equal to \(P''\). Notice that for \(S \subseteq X\), where \(S \in (\mathbb{N})\) and \(X \in (\omega)^\omega\), we have \(S \preceq (S \cap X) \subseteq X\).

At a first glance, the set of partitions of \(\omega\) with the partitions \(\{\omega\}\) and \(\{\{n\} : n \in \omega\}\) and the operations \("\cup"\) and \("\cap"\), looks similar to the Boolean algebra \((\mathcal{P}(\omega), \cup, \cap, \_\_, \emptyset, \omega)\). However, partitions of \(\omega\) behave differently than subsets of \(\omega\). The main difference between partitions and subsets is that partitions do not have proper complements. For example if \(x, y, z \in [\omega]^\omega\) are such that \(x \cup y = x \cup z = \omega\) and \(x \cap y = x \cap z = \emptyset\), then \(y = z\). This is not the case for partitions. It is not hard to find partitions \(X, Y, Z \in (\omega)^\omega\) such that \(X \cup Y = X \cup Z = Y \cup Z = \{\omega\}\) and \(X \cap Y = X \cap Z = Y \cap Z = \emptyset\), e.g., let \(X = \{\{3i, 3i+1\} : i \in \omega\}\cup\{\{3i+2\} : i \in \omega\}, Y = \{\{3i+1, 3i+2\} : i \in \omega\}\cup\{\{3i\} : i \in \omega\}\), and \(Z = \{\{3i, 3i+2\} : i \in \omega\}\cup\{\{3i+1\} : i \in \omega\}\). We leave it as an exercise to the reader to construct infinite partitions \(X, Y, Z \in (\omega)^\omega\) with the same property but such that all blocks of \(X, Y, Z\) are infinite.

Now, let us define a topology on \((\omega)^\omega\) which is similar to the Ellentuck topology on \([\omega]^\omega\) (defined on page 216). For \(S \in (\mathbb{N})\) and \(X \in (\omega)^\omega\) with \(S \subseteq X\), let

\[(S, X)^\omega = \{Y \in (\omega)^\omega : S \preceq Y \subseteq X\}\.

A set \((S, X)^\omega\) where \(S\) and \(X\) are as above, is usually called a dual Ellentuck neighbourhood. We leave it as an exercise to the reader to show that the intersection of finitely many dual Ellentuck neighbourhoods is either empty or a dual Ellentuck neighbourhood. The topology on \((\omega)^\omega\) generated by the dual Ellentuck neighbourhoods is called dual Ellentuck topology.

The usual trick to get subsets of \(\omega\) from partitions is as follows: For a partition \(P\) of a subset of \(\omega\), e.g., \(P \in (\omega)^\omega\) or \(P \in (\mathbb{N})\), let

\[
\text{Min}(P) = \{\min(p) : p \in P\}.
\]

Obviously, if \(X \in (\omega)^\omega\) then \(\text{Min}(X) \in [\omega]^\omega\) and if \(S \in (\mathbb{N})\) then \(\text{Min}(S) \in \text{fin}(\omega)\). Further we have that for any \(X, Y \in (\omega)^\omega\), \(X \subseteq Y\) implies \(\text{Min}(X) \subseteq \text{Min}(Y)\).

A non-empty family \(\mathcal{C} \subseteq (\omega)^\omega\) is called free, if for every \(X \in \mathcal{C}\) there is a \(Y \in \mathcal{C}\) such that \(Y \subseteq X\), but for all \(S \in (\mathbb{N})\), \((S \cap X) \not\subseteq Y\).

A family \(\mathcal{C} \subseteq (\omega)^\omega\) is closed under refinement if \(X \subseteq Y\) and \(X \in \mathcal{C}\) implies \(Y \in \mathcal{C}\), and it is closed under finite coarsening if \(S \in (\mathbb{N})\) and \(X \in \mathcal{C}\) implies \((S \cap X) \in \mathcal{C}\). Notice that a family \(\mathcal{C} \subseteq (\omega)^\omega\) is closed under refinement and finite coarsening iff for all \(S \in (\mathbb{N})\) and \(Y \in (\omega)^\omega\), \(X \subseteq (S \cap Y)\) and \(X \in \mathcal{C}\) implies \(Y \in \mathcal{C}\).

A family \(\mathcal{C} \subseteq (\omega)^\omega\) is called complete, if \(\mathcal{C}\) is free and closed under refinement and finite coarsening.

In order to define the game which plays a key role in the proof of the Partition Ramsey Theorem, we have to introduce the following notation. For \(S \in (\mathbb{N})\), let \(S^*\) denote the partition \(S \cup \{\text{Dom}(S)\}\). Notice that \(|S^*| = |S| + 1\). Further, notice that whenever \((S^*, X)^\omega\) is a dual El-
lentuck neighbourhood, then every $Y \in (S^*, X)^\omega$ has a block $y$ such that $y \cap \text{Dom}(S) = \emptyset$ and $y \cap \text{Dom}(S^*) = \{ \text{Dom}(S) \}$.

With respect to a complete family $\mathcal{E} \subseteq (\omega)^\omega$ we define the infinite two-player game $\mathcal{G_\mathcal{E}}$ as follows.

$$
\text{MAIDEN} \quad (S_0, X_0) \quad (S_1, X_1) \quad (S_2, X_2) \quad \ldots
$$

$\text{DEATH} \quad Y_0 \quad Y_1 \quad Y_2 \quad \ldots$

We require that the first move $(S_0, X_0)$ of the MAIDEN is such that $X_0 \in \mathcal{E}$ and that $(S_0^*, X_0)^\omega$ is a dual Ellentuck neighbourhood. Further, we require that for each $n \in \omega$, the $n^{\text{th}}$ move of DEATH $Y_n$ is such that $Y_n \in (S_n^*, X_n)^\omega$ and $Y_n \in \mathcal{E}$, and that the MAIDEN plays $(S_{n+1}, X_{n+1})$ such that

- $S_n^* \subseteq S_{n+1}$, $|S_{n+1}| = |S_n| + 1$, $S_{n+1}^* \subseteq Y_n$, and
- $X_{n+1} \in (S_{n+1}^*, Y_n)^\omega \cap \mathcal{E}$.

Finally, the MAIDEN wins the game $\mathcal{G_\mathcal{E}}$ if and only if $\bigcap_{n \in \omega} (S_n, X_n)^\omega \cap \mathcal{E} = \emptyset$, i.e., the (unique) infinite partition $X \in (\omega)^\omega$ such that $S_n \prec X$ (for all $n \in \omega$) does not belong to the family $\mathcal{E}$.

Now, a complete family $\mathcal{E} \subseteq (\omega)^\omega$ is called a **Ramsey partition-family** if the MAIDEN has no winning strategy in the game $\mathcal{G_\mathcal{E}}$ (compare with the game introduced in Chapter 10 | Related Result 71).

Obviously, the set $(\omega)^\omega$ is an example for a Ramsey partition-family and it is not hard to construct Ramsey partition-families which are proper subsets of $(\omega)^\omega$, e.g., for any partition $X \in (\omega)^\omega$, $(X)^\omega$ is a Ramsey partition-family.

For a non-trivial example of a Ramsey partition-family take a Ramsey ultrafilter $\mathcal{F} \subseteq [\omega]^\omega$ and let $\mathcal{E} = \{ X \in (\omega)^\omega : \text{Min}(X) \in \mathcal{F} \}$. Then, by Chapter 10 | Related Result 71, we get that $\mathcal{E}$ is a Ramsey partition-family (for other non-trivial examples of Ramsey partition-families see Chapter 26).

It turns out that Ramsey partition-families have very strong combinatorial properties, and to some extent, they are proper generalisations of Ramsey families (see also Chapter 26). The combinatorial strength of Ramsey partition-families is used for example in the proof of Carlson’s Lemma, which is — as mentioned above — the crucial part in the proof of the Partition Ramsey Theorem.

**Carlson’s Lemma and the Partition Ramsey Theorem**

Before we formulate and prove the **Partition Ramsey Theorem**, let us first consider a few possible generalisations of Ramsey’s Theorem in terms of partitions: **Ramsey’s Theorem** states that whenever we colour $[\omega]^n$ (i.e., the $n$-element subsets of $\omega$) with finitely many colours, then we find an $x \in [\omega]^n$ (i.e., an infinite subsets of $\omega$) such that $[x]^n$ is monochromatic (i.e., all whose
n-element subsets have the same colour). If we try to formulate Ramsey’s Theorem in terms of partitions, we first have to decide which partitions correspond to the “n-element subsets of ω” and “infinite subsets of ω” respectively. It seems natural that infinite subsets of ω correspond to infinite partitions of ω, i.e., x ∈ [ω]ω is replaced by X ∈ (ω)ω. Similarly, we could say that n-element subsets of ω correspond to n-block partitions of ω, and therefore we would replace [ω]n by (ω)n := {X ∈ (ω)ω : |X| = n}. This leads to the following first attempt of a generalisation of Ramsey’s Theorem in terms of partitions:

Generalisation 1. For every colouring of (ω)n with finitely many colours, there exists an infinite partition X ∈ (ω)ω such that (X)n is monochromatic, where (X)n := {Y ∈ (ω)n : Y ⊆ X ∧ |Y| = n}.

Unfortunately, this generalisation of Ramsey’s Theorem fails. In fact, by transfinite induction we can construct a counterexample even for the case when n = 2: Firstly notice that for each X ∈ (ω)ω, |(X)2| = |(ω)ω| = ω. Let {Xα : α ∈ ω} be an enumeration of (ω)ω. For each α ∈ ω choose two distinct partitions
\[ Y_α^0, Y_α^1 \in \left( X_α \right)^2 \setminus \{ Y_β^0, Y_β^1 : \beta < \alpha \} \].

Finally, define \( \pi : (ω)^2 \to \{0, 1\} \) by stipulating \( \pi(Y) = 0 \) iff there is an \( \alpha \in \omega \) such that \( Y = Y_α^0 \). By construction, for every \( X \in (ω)^ω \) we find \( Y_0 \) and \( Y_1 \) in \( (X)^2 \) such that \( \pi(Y_0) = 0 \) and \( \pi(Y_1) = 1 \). Thus, for every \( X \in (ω)^ω \), \( (X)^n \) is dichromatic.

One might ask why is it not possible to construct a similar counterexample for Ramsey’s Theorem? The reason is simple: For any partition \( X \in (ω)^ω \), \( (X)^2 \) is of cardinality \( ω \), whereas for any \( x \in [ω]^ω \) and \( n \in ω \), the set \([x]^n\) is countable.

Now, one might ask why are n-element subsets of ω so different from n-block partitions? A reason is that n-element subsets of ω are proper finitary objects, whereas an n-block partition \( Y \in (ω)^n \) necessarily contains infinite sets. Furthermore, every n-element subset of ω is a subset of some \( k \in ω \), which is not the case for partitions \( Y \in (ω)^n \). However, it is true for partitions \( S \in (N) \). So, let us replace now \([ω]^n \) and \([x]^n \) by \( (ω)^{(n)} \) and \( (X)^{(n)} \) respectively, where
\[ (ω)^{(n)} = \{ S \in (N) : |S| = n \} \],
and for \( X \in (ω)^ω \),
\[ (X)^{(n)} = \{ S \in (ω)^{(n)} : S \subseteq X \} \].

Generalisation 2. For every colouring of \( (ω)^{(n)} \) with finitely many colours, there exists an infinite partition \( X \in (ω)^ω \) such that \( (X)^{(n)} \) is monochromatic.

Unfortunately, this generalisation fails as well. Again, we can construct a counterexample even for the case when \( n = 2 \): For this, consider the colouring \( \pi : (ω)^{(2)} \to \{0, 1\} \) defined by stipulating...
\[ \pi(\{s_0, s_1\}) = 0 \iff 0 \in s_0 \land \max(s_0) < \max(s_1). \]

We leave it as an exercise to the reader to show that for every \( X \in (\omega)^\omega \), \((X)^{(n)}\) is dichromatic.

After these two failures, let us now formulate **Ramsey's Theorem** directly in terms of partitions of subsets of \( \omega \): A partition \( P \) of a subset of \( \omega \) is **segmented** if for any distinct \( p_0, p_1 \in P \), either \( \max(p_0) < \min(p_1) \) or \( \max(p_1) < \min(p_0) \). Let \((\omega)^\omega\) denote the set of all segmented partitions of \( \omega \). Notice that if \( P \in (\omega)^\omega \), then all blocks \( P \) are finite. For the moment let \( \omega := \omega \setminus \{0\} \). For an infinite set of positive integers \( x = \{k_i : i \in \omega\} \in [\omega]^{\omega} \), where \( k_i < k_{i+1} \) for all \( i \in \omega \), we define \( P_x \in (\omega)^\omega \) by stipulating

\[ P_x = \{ [k_i, k_{i+1}) : i \in \omega \}, \]

where \( k_0 := 0 \). Notice that \((\omega)^\omega = \{ P_x : x \in [\omega]^{\omega} \} \). Similarly, for an \( n \)-element set \( s = \{k_1, \ldots, k_n\} \in [\omega]^n \), where \( k_i < k_{i+1} \) for \( 1 \leq i \leq n \), we define

\[ Q_s = \{ [k_i, k_{i+1}) : i \in n \}, \]

where again \( k_0 = 0 \). Notice that for all \( s \in \text{fin}(\omega) \), \( Q_s \) is a segmented partition with \( \text{Dom}(Q_s) = \max(s) \). Now, let \((\omega)^{(n)} = \{ Q_s : s \in [\omega]^n \} \) and for \( P \in (\omega)^\omega \) let

\[ (P)^{(n)} = \{ Q \in (\omega)^{(n)} : Q \subseteq P \}. \]

Recall that for \( s \in \text{fin}(\omega) \), \( Q_s^* = Q_s \cup \{ \text{Dom}(Q_s) \} = Q_s \cup \{ \max(s) \} \), and notice that for all \( x \in [\omega]^{\omega} \), \((P_x)^{(n)} = \{ Q_s^* : s \in [x]^n \} \). We are now ready to formulate **Ramsey’s Theorem** in terms of segmented partitions — we leave it as an exercise to the reader to show that **Ramsey’s Theorem** is indeed equivalent to the following statement.

**Ramsey’s Theorem.** For every colouring of \((\omega)^{(n)}\) with finitely many colours, there exists an infinite segmented partition \( P \in (\omega)^\omega \) such that \((P)^{(n)}\) is monochromatic.

So, we finally found a formulation of **Ramsey’s Theorem** in terms of segmented partitions. The next step is to find a general formulation which works for all, and not just for segmented partitions. For this, we only have to replace the angle brackets by round brackets and define the meaning of \((X)^{(n)}\): For \( n \in \omega \) and \( X \in (\omega)^\omega \) let

\[ (X)^{(n)} = \{ S \in (\omega)^{(n)} : S^* \subseteq X \}. \]

Similarly, for a dual Ellenbuck neighbourhood \((S, X)^\omega\), where \( |S| \leq n \), let

\[ (S, X)^{(n)} = \{ U \in (\omega)^{(n)} : S \preceq U \land U^* \subseteq X \}. \]

Now we are ready to state the sought partition form of **Ramsey’s Theorem**:
Theorem 11.4 (Partition Ramsey Theorem). For any Ramsey partition-family $\mathcal{C} \subseteq (\omega)^n$ and for any colouring of $(\omega)^{(n)}$ with $r$ colours, where $r$ and $n$ are positive integers, there is an $X \in \mathcal{C}$ such that $(X)^{(n)}$ is monochromatic.

The Partition Ramsey Theorem will follow from Carlson’s Lemma. With respect to Ramsey partition-families, Carlson’s Lemma states as follows:

Lemma 11.5 (Carlson’s Lemma). Let $\mathcal{C} \subseteq (\omega)^n$ be an arbitrary but fixed Ramsey partition-family. For any colouring $\pi : (\omega)^{(n)} \to r$, where $r$ and $n$ are positive integers, and for any dual Ellentuck neighbourhood $(S_0, X_0)^n$, where $|S_0| = n$ and $X \in \mathcal{C}$, there is a $X \in (S_0, X_0)^n$ which belongs to $\mathcal{C}$ such that $(S, X)^{(n)}$ is monochromatic.

Proof. Before we begin with the proof, let us first introduce the following notion: For a dual Ellentuck neighbourhood $(S, X)^n$ and for a positive integer $m \in \omega$, a set $D \subseteq (\omega)^{(m)}$ is called $\mathcal{C}$-dense in $(S, X)^{(m)}$ if for all $Y \in (S, X)^{(m)} \cap \mathcal{C}$, $(S, Y)^{(m)} \cap D \neq \emptyset$. Notice that for every colouring $\pi : (\omega)^{(n)} \to r$, there exists a colour $c \in r$ and a partition $X_0' \in (S_0, X_0)^n$ such that the set $D_c := \{S \in (\omega)^{(n)} : \pi(S) = c\}$ is $\mathcal{C}$-dense in $(S_0, X_0')^{(n)}$. Indeed, if $D_0$ is $\mathcal{C}$-dense in $(S_0, X_0)^{(n)}$, then we are done. Otherwise, there exists an $X_1 \in (S_0, X_0)^n \cap \mathcal{C}$ such that $(S_0, X_1)^{(n)} \cap D_0 = \emptyset$. Now, either $D_1$ is $\mathcal{C}$-dense in $(S_0, X_1)^{(n)}$, or there exists an $X_2 \in (S_0, X_1)^n \cap \mathcal{C}$ such that $(S_0, X_2)^{(n)} \cap D_1 = \emptyset$. Proceeding this way, we finally find a $c \in r$ such that for all $Y \in (S_0, X_c)^n \cap \mathcal{C}$, $(S_0, Y)^{(n)} \cap D_c \neq \emptyset$; let $X_0' = X_c$.

After this preliminary remark, we can now begin with the proof: Without loss of generality we may assume that the dual Ellentuck neighbourhood $(S_0, X_0)^n$ is such that $D_0$ is $\mathcal{C}$-dense in $(S_0, X_0)^{(n)}$.

The proof is now given in several steps. Firstly we show that there exists an $S \in (\mathbb{N})$ with $S_0 \subseteq S \subseteq X_0$ such that for all $T \in (\mathbb{N})$ with $S \subseteq T \subseteq X_0$, there is a $T' \subseteq T$ such that $\text{Dom}(T') = \text{Dom}(T)$, $|T'| = n$, $S_0 \preceq T'$, and $T' \in D_0$. To state this in a more formal way, we introduce the following two notations: For $S, T \in (\mathbb{N})$, where $S \preceq T$ and $|S| \leq m$, let

$$(S, T)^m = \{ U \in (\mathbb{N}) : \text{Dom}(U) = \text{Dom}(T) \wedge S \preceq U \subseteq T \wedge |U| = m \},$$

and for a dual Ellentuck neighbourhood $(U, Z)^n$, let

$$(U, Z)^{(n)} = \bigcup_{k \in \omega} (U, Z)^{(k)}.$$ 

In other words, $(U, Z)^{(\omega)} = \{ S \in (\mathbb{N}) : U \preceq S \preceq Z \}$ and $(S, T)^m$ is the set of all $m$-block partitions of $\text{Dom}(T)$ which contain $S$ as a “sub-partition” and are coarser than $T$.

Claim 1. There is a $Z_0 \in (S_0, X_0)^n \cap \mathcal{C}$ and an $S_0 \preceq (S, Z_0)^{(n)}$ such that for all $S \in (\tilde{S}, Z_0)^{(n)}$, $(S, S_0)^n \cap D_0 \neq \emptyset.$
Proof of Claim 1. If the claim fails, then for every $Y \in (S_0, X_0)^\omega \cap \mathcal{C}$ and each $T \in (S_0, Y)^{(\omega^0)}$ there is an $S \in (T, Y)^{(\omega^0)}$ such that $(S_0, S)^n \cap D_0 = \emptyset$; in particular, there is an $S' \in (T, Y)^{(\omega^1)}$ such that $(S_0, S')^n \cap D_0 = \emptyset$. We define a strategy for the Maiden in the game $\mathcal{G}_\mathcal{C}$. The Maiden starts the game with $(S_0, X_0)$ and replies the $i$th move $Y_i$ of Death with $(S_{i+1}, X_{i+1})$, where $X_{i+1} = X_i$ and $S_{i+1}$ is constructed as follows: Take any $T_{i+1} \in (S_i, Y_i)^{(n+i+1)}$ and let $S_{i+1} \in (T_{i+1}, Y_i)^{(n+i+1)}$ be such that $(S_0, S_{i+1})^n \cap D_0 = \emptyset$. As $\mathcal{C}$ is a Ramsey-partition family, fix a play where the Maiden follows this strategy but Death wins. Let $Z \in (\omega)^\omega$ be the unique infinite partition such that for all $i \in \omega$ we have $S_i \prec Z$. Since $\mathcal{C}$ is a Ramsey-partition family, the partition $Z$ belongs to $\mathcal{C}$. By construction, $S_0 \prec Z$ and $(S_0, Z)^{(n^2)} \cap D_0 = \emptyset$. Thus, $D_0$ is not $\mathcal{C}$-dense in $(S_0, X_0)^{(n^2)}$, a contradiction. \hfill ($\text{Claim 1}$)

The next step is where the Hales-Jewett Theorem comes in:

**Claim 2.** Let $Z_0 \in (S_0, X_0)^\omega \cap \mathcal{C}$ be as in Claim 1. Then there is a $U \in (S_0, Z_0)^{(n+1)}$ such that $(S_0, U)^n \subseteq D_0$.

**Proof of Claim 2.** Let $\bar{S} \in (S_0, Z_0)^{(\omega^0)}$ be as in Claim 1, i.e., for all $W \in (\bar{S}, Z_0)^{(\omega^0)}$ there is a $W' \in (S_0, W)^n$ such that $W' \in D_0$. Let $m = |\bar{S}|$, $r_0 = |(S_0, \bar{S})^n|$, and let $\{U_k : k \in r_0\}$ be an enumeration of $(S_0, \bar{S})^n$. By the Hales-Jewett Theorem 11.2, or more precisely by a partition form of it, there is a positive integer $T = HJ(m, r_0)$ such that for any $T \in (\bar{S}, Z_0)^{(m+1)}$ and any $r_0$-colouring of $(\bar{S}, T)^m$ there is a $W_0 \in (\bar{S}, T)^{m+1}$ such that $(\bar{S}, W_0)^m$ is monochromatic (the details are left to the reader). Fix an arbitrary $\bar{T} \in (\bar{S}, Z_0)^{(m+1)}$. Then, by the choice of $\bar{S}$, for all $W \in (\bar{S}, T)^{m+1}$ there is a $U_0 \in (S_0, W)^m$ such that $U_0 \in D_0$. Moreover, there is a $k \in r_0$ such that $U_k \subset U$, and since $|U_k| = |U| = n$ we have $U = U_k \cap W$. Hence, for every $W \in (\bar{S}, T)^{m+1}$ there is a $k \in r_0$ such that $U_k \cap W \in D_0$. Now, for each $W \in (\bar{S}, T)^{m+1}$ let

$$\tau(W) = \min\{k \in r_0 : U_k \cap W \in D_0\}$$

Then $\tau$ is an $r_0$-colouring of $(\bar{S}, T)^m$. Since $\bar{T} \in (\bar{S}, Z_0)^{(m+1)}$, there is a $W_0 \in (\bar{S}, W_0)^m$ such that $(\bar{S}, W_0)^m$ is monochromatic, say of colour $k_0$. Thus, for all $W \in (\bar{S}, W_0)^m$, $U_{k_0} \cap W \in D_0$. Finally, let $U = U_{k_0} \cap W_0$. Then $U \in (S_0, W_0)^{(n+1)}$, hence $U \in (S_0, Z_0)^{(n+1)^2}$, and $(S_0, U)^n \subseteq D_0$ as required. \hfill ($\text{Claim 2}$)

As an obvious generalisation of Claim 2 we get

**Claim 2*. For each $X \in (S_0, X_0)^\omega \cap \mathcal{C}$ there is a $U \in (S_0, X)^{(n+1)}$ such that $(S_0, U)^n \subseteq D_0$.

The next step is crucial in the construction of $\bar{X}$:

**Claim 3.** Let $Z_0 \in (S_0, X_0)^\omega \cap \mathcal{C}$ be as in Claim 1. Then there are $S \in (S_0, Z_0)^{(n+1)}$ and $X \in (S, Z_0)^\omega \cap \mathcal{C}$ such that the set

$$\{T \in (S, X)^{(n+1)} : (S_0, T)^n \subseteq D_0\}$$

is $\mathcal{C}$-dense in $(S, X)^{(n+1)}$. 

Proof of Claim 3. Assume towards a contradiction that the claim fails. Then, for any \( S \in (S_0, Z_0)^{(n+1)^*} \) and each \( Y \in (S, X_0)^{(\omega)} \cap \mathcal{C} \) there exists a \( Z \in (S, Y)^{(\omega)} \cap \mathcal{C} \) such that for all \( T \in (S, Z)^{(n+1)^*} \) we have \((S_0, T)^n \notin D_0\). We define a strategy for the Maiden in the game \( \mathcal{G}_\mathcal{C} \). The Maiden starts the game with \((S_0, Z_0)\) and replies the \( i^{th} \) move \( Y_i \) of Death with \((S_{i+1}, Z_{i+1})\), where \( Z_{i+1} \in (S_i^*, Y_i)^{(\omega)} \cap \mathcal{C} \) and \( S_{i+1} \in (S_i^*, Z_{i+1})^{(n+1)^*} \) are such that for all \( S \in (S_0, S_{i+1})^{n+1} \) and all \( T \in (S, Z_{i+1})^{(n+1)^*} \) we have \((S_0, T)^n \notin D_0\). For \( i = 0 \), let \( S_1 \in (S_0, Y_0)^{(n+1)^*} \) be arbitrary and let \( Z_1 \in (S_1^*, Z_0)^{(\omega)} \cap \mathcal{C} \) be such that for all \( T \in (S_1, Z_1)^{(n+1)^*} \) we have \((S_0, T)^n \notin D_0\). For \( i > 0 \), we construct \( S_{i+1} \) and \( Z_{i+1} \) as follows. Firstly, let \( \{T_{i,k} : k \in h_i\} \) be an enumeration of \((S_0, S_i)^{n+1}\). Secondly, let \( Z_{i,0} = Y_i \) and for \( k \in h_i \) let \( Z_{i,k+1} \in (S_i, Z_{i,k})^{(\omega)} \cap \mathcal{C} \) be such that for all \( T \in (T_{i,k}, Z_{i,k+1})^{(<\omega)^*} \) we have \((S_0, T)^n \notin D_0\). Finally, let \( Z_{i+1} = Z_{i,h_i} \) and let \( S_{i+1} \in (S_i^*, Z_{i+1})^{(n+1)^*} \). Fix a play where the Maiden follows this strategy but Death wins. Since \( \mathcal{C} \) is a Ramsey partition-family, the unique infinite partition \( Z \in (\omega)^{\omega} \) such that for all \( i \in \omega \) we have \( S_i < Z \) belongs to \( \mathcal{C} \). Now, by construction, for any \( U \in (S_0, Z)^{(n+1)^*} \) we find a positive integer \( i \in \omega \) as well as a \( k \in h_i \) such that \( U \in (T_{i,k}, Z_{i,k+1})^{(n+1)^*} \). Thus, for all \( U \in (S_0, Z)^{(n+1)^*} \) we have \((S_0, U)^n \notin D_0\), but since \((S_0, Z)^{(\omega)} \subseteq (S_0, Z_0)^{(\omega)} \), this contradicts Claim 2*.

The following claim is just a generalisation of Claim 3:

Claim 3*. Let \((T_0, Y_0)^{(\omega)} \subseteq (S_0, X_0)^{(\omega)} \) be a dual Ellentuck neighbourhood, where \( Y_0 \in \mathcal{C} \) and \( |T_0| = m \). If \( E \subseteq (\omega)^{(m)} \) is \( \mathcal{C} \)-dense in \((T_0, Y_0)^{(m)} \), then there exist \( S \in (T_0, Y_0)^{(m+1)^*} \) and \( X \in (S, Y_0)^{(\omega)} \cap \mathcal{C} \) such that the set \( \{T \in (S, Y)^{(n+1)^*} : (T_0, T)^n \subseteq E\} \) is \( \mathcal{C} \)-dense in \((S, X)^{(m+1)^*} \).

Proof of Claim 3*. In the proofs of the preceding claims, just replace \( S_0 \) by \( T_0 \), \( X_0 \) by \( Y_0 \), and \( D_0 \) by \( E \).

Now we construct the first piece of the sought partition \( \tilde{X} \):

Claim 4. There is a \( U_0 \in (S_0, X_0)^{(m)} \) such that \( \pi(U_0) = 0 \), i.e., \( U_0 \in D_0 \), and in addition there is an \( X \in (U_0^*, X_0)^{(\omega)} \cap \mathcal{C} \) such that the set \( \{T \in (U_0, X)^{(n+1)^*} : (S_0, T)^n \subseteq D_0\} \) is \( \mathcal{C} \)-dense in \((U_0, X)^{(n+1)^*} \).

Proof of Claim 4. We define a strategy for the Maiden in the game \( \mathcal{G}_\mathcal{C} \). The Maiden starts the game with \((S_0, X_0)\) and replies the \( i^{th} \) move \( Y_i \) of Death with \((S_{i+1}, X_{i+1})\), where \( S_{i+1} \) and \( X_{i+1} \) are constructed as follows: For \( i = 0 \), let \( S_1 \in (S_0, Y_0)^{(n+1)^*} \) and \( X_1 \in (S_1, Y_0)^{(\omega)} \cap \mathcal{C} \) be such that the set \( E_1 = \{T \in (S_1, X_1)^{(n+1)^*} : (S_0, T)^n \subseteq D_0\} \) is \( \mathcal{C} \)-dense in \((S_1, X_1)^{(n+1)^*} \). Notice that by Claim 3*, \( S_1 \) and \( X_1 \) exist. Similarly, for \( i > 0 \) let \( S_{i+1} \in (S_i, Y_i)^{(n+1)^*} \) and \( X_{i+1} \in (S_i, Y_i)^{(\omega)} \cap \mathcal{C} \) be such that the set \( E_{i+1} = \{T \in (S_{i+1}, X_{i+1})^{(n+1)^*} : (S_i, T)^{n+i} \subseteq E_i\} \) is \( \mathcal{C} \)-dense in \((S_{i+1}, X_{i+1})^{(n+1+i)^*} \).
is \( \mathcal{G} \)-dense in \((S_{i+1}, X_{i+1})^{(n+1)} \)^*\(^*\). By induction on \(i\) one verifies that for all \(i \in \omega\) we have

\[
E_{i+1} \subseteq \{ T \in (S_{i+1}, X_{i+1})^{(n+1)} : (S_0, T)^n \subseteq D_0 \},
\]

where \(E_0 := D_0\) (the details are left to the reader). Finally, fix a play where the MAIDEN follows this strategy but DEATH wins, and let \(X \in (\omega)^\omega\) be the unique infinite partition such that for all \(i \in \omega\) we have \(S_i \prec X\). Since \(\mathcal{G}\) is a Ramsey partition-family, \(X\) belongs to \(\mathcal{G}\). Now, since \(D_0\) is \(\mathcal{G}\)-dense in \((S_0, X_0)^\omega\) and \(X \in (S_0, X_0)^\omega \cap \mathcal{G}\), there is a \(U_0 \in (S_0, X)^{\omega^n}\) such that \(U_0 \subseteq D_0\). Choose \(i_0 \in \omega\) large enough such that there is an \(S \in (S_0, S_{i_0})^{(n+1)}\) for which we have \(U_0^* \subseteq S\). Since \((S_0, S)^n \subseteq (S_0, S_{i_0})^{\omega^n}\) we get that \(\{ T \in (S, X)^{(n+1)} : (S_0, T)^n \subseteq D_0 \}\) is \(\mathcal{G}\)-dense in \((S, X)^{(n+1)}\). In particular, the set \(\{ T \in (S, X)^{(n+1)} : (S_0, T)^n \subseteq D_0 \}\) is \(\mathcal{G}\)-dense in \((S, X)^{(n+1)}\), and since \(\pi(U_0) = 0\) and \(U_0^* \ll S\), \(U_0\) has the required properties. \(\dashv\)

We leave it as an exercise to the reader to prove the following generalisation of CLAIM 4:

CLAIM 4 *. If \(U_i \in (S_0, X_0)^{(n+i)}\) is such that \((S_0, U_i)^n \subseteq D_0\) and \(Y \in (U_i^*, X_0)^\omega \cap \mathcal{G}\) is such that \(\{ T \in (U_i, Y)^{(n+i+1)} : (S_0, T)^n \subseteq D_0 \}\) is \(\mathcal{G}\)-dense in \((U_i, Y)^{(n+i+1)}\), then there are \(U_{i+1} \in (U_i^*, Y)^{(n+i+1)}\) and \(X \in (U_{i+1}^*, Y)^\omega \cap \mathcal{G}\) such that

\[
\{ T \in (U_{i+1}, X)^{(n+i+2)} : (S_0, T)^n \subseteq D_0 \}\]

is \(\mathcal{G}\)-dense in \((U_{i+1}, X)^{(n+i+1)}\) and \((S_0, U_{i+1})^n \subseteq D_0\).

Now we are ready to construct an infinite partition \(\bar{X} \in (S_0, X_0)^\omega \cap \mathcal{G}\) such that for every \(U \in (S_0, X)^\omega\) we have \(\pi(U) = 0\), i.e., \((S_0, X)^\omega \subseteq D_0\). Indeed, by defining a suitable strategy for the MAIDEN in the game \(G\) (applying CLAIM 4 *), we can construct partitions \(U_i \in (S_0, X_0)^{\omega^n}\) such that for all \(i \in \omega\) we have

\[
|U_i| = n + i, \quad U_i^* \ll U_{i+1}, \quad (S_0, U_i)^n \subseteq D_0, \quad \forall i \in \omega
\]

and the unique partition \(\bar{X} \in (\omega)^\omega\) such that \(U_i \prec \bar{X}\) (for all \(i \in \omega\)) belongs to the Ramsey partition-family \(\mathcal{G}\). By (\(\ddagger\)), for all \(U \in (S_0, X)^{\omega^n}\) we have \((S_0, U)^n \subseteq D_0\), i.e., \((S_0, X)^{\omega^n}\) is monochromatic, which completes the proof of CARLSON’S LEMMA.

Having CARLSON’S LEMMA at hand, we are now able to prove the main result of this chapter:

**Proof of the Partition Ramsey Theorem.** The proof is by induction on \(n\).

For \(n = 1\), the **PARTITION RAMSEY THEOREM** follows immediately by the Pigeon-Hole Principle. So, let \(n, r \in \omega\) be given, where \(r\) is positive and \(n > 1\), and assume that the **PARTITION RAMSEY THEOREM** is already proved for all positive integers \(n' < n\).
A weak form of the Halpern-Läuchli Theorem

Fix an arbitrary colouring \( \pi : (\omega)^n \to r \). Take an arbitrary partition \( X_0 \in \mathcal{C} \) and let \( S_0 \in (\mathbb{N}) \) be such that \( |S_0| = n - 1 \) and \( S_0 \not\prec X_0 \).

We define a strategy for the Maiden in the game \( \mathcal{G}_\mathcal{C} \) and as by-product we get a partial mapping \( \tau \) from \( (\omega)^{n-1} \) to \( r \). The Maiden starts the game with \( (S_0, X_0) \) and replies the 2th move \( Y_0 \) of Death with \( (S_{i+1}, X_{i+1}) \), where \( S_{i+1} \) and \( X_{i+1} \) are constructed as follows: Let \( \{ T_k \in (\mathbb{N}) : k \in h_i \} \) be an enumeration of all \( T \subseteq S_i \) with \( \text{Dom}(T) = \text{Dom}(S_i) \) and \( |T| = n - 1 \). Let \( Z_0 := Y_0 \) and for each \( k \in h_i \), let \( Z_{k+1} \in (S^*_i, Z_k)^\omega \cap \mathcal{C} \) be such that \( \pi|_{(T^*_k, Z_{k+1})^\omega} \) is constant and define

\[
\tau(T_k) = \pi(U) \quad \text{for some} \quad U \in (T^*_k, Z_{k+1})^\omega.
\]

Now, the partition \( Z_{k+1} \in \mathcal{C} \) we construct by applying first Carlson’s Lemma 11.5 with respect to the dual Ellentuck neighbourhood \( (T^*_k, Z_k)^\omega \) and then by refining the resulting partition such that it belongs to the dual Ellentuck neighbourhood \( (S^*_i, Z_k)^\omega \). Let \( X_{i+1} := Z_{h_i} \) and let \( S_{i+1} \in (\mathbb{N}) \) be such that \( S^*_i \not\prec X_{i+1} \) and \( |S_{i+1}| = (n - 1) + (i + 1) \). Finally, fix a play where the Maiden follows this strategy but Death wins, and let \( Z \in (\omega)^\omega \) be the unique infinite partition such that for all \( i \in \omega \) we have \( S_i \not\prec Z \). Since \( \mathcal{C} \) is a Ramsey partition-family, the partition \( Z \) belongs to \( \mathcal{C} \). For each \( T \in (Z)^{(n-1)} \) there exist unique numbers \( i, k \in \omega \) such that \( k \in h_i \) and \( T = T_k \). Thus, \( \tau \) is an \( r \)-colouring of \( (Z)^{(n-1)} \). By the induction hypothesis we find an \( X \in (Z)^\omega \cap \mathcal{C} \) such that \( \pi|_{(X)^{(n-1)}} \) is constant, say \( \tau(X) = j \) for all \( T \in (X)^{(n-1)} \). Now, take any \( S \in (X)^\omega \) and let \( \tilde{S} \prec S \) be such that \( |\tilde{S}| = n - 1 \). Notice that the domain of \( \tilde{S} \) is equal to \( \text{Dom}(S_i) \) for some \( i \in \omega \). Consider the partition \( X_{i+1} \). By the construction of \( X_{i+1} \) we know that \( (T^*_k, X_{i+1})^\omega \) is monochromatic whenever \( T \subseteq S_i \) with \( |T| = n - 1 \) and \( \text{Dom}(T) = \text{Dom}(S_i) \), and by the construction of the partition \( X, \pi|_{(T^*_k, X_{i+1})^\omega} \) is constantly \( j \). In particular, \( \pi(U) = j \) whenever \( U \in (\tilde{S}^*, X_{i+1})^\omega \), and since \( S \in (\tilde{S}^*, X_{i+1})^\omega \), we get \( \pi(S) = j \), which completes the proof.

A Weak Form of the Halpern-Läuchli Theorem

One can show that for example the Hales-Jewett Theorem, a weakened form of the Halpern-Läuchli Theorem, Ramsey’s Theorem, as well as the Finite Ramsey Theorem and a partition form of it, are all derivable from the Partition Ramsey Theorem. Below, we just give the proof of the Weak Halpern-Läuchli Theorem (for the other results see related Result 75).

To state this weakened form of the Halpern-Läuchli Theorem, we have to give first some notations: A set \( T \subseteq \text{seq}(2) \), where \( \text{seq}(2) = \bigcup_{n \in \omega} \omega_2, \) is a tree if for every \( s \in T \) and \( k \in \text{dom}(s) \) we have \( s|_k \in T \). In particular, \( \text{seq}(2) \) is a tree. For a tree \( T \subseteq \text{seq}(2) \) and \( l \in \omega \) let

\[
T(l) = \{ s \in T : \text{dom}(s) = l \}.
\]
For a finite product of trees $\mathcal{T} = T_0 \times \ldots \times T_{d-1} \subseteq (\text{seq}(2))^d$ (i.e., for all $k \in d$, where $d \in \omega$, $T_k \subseteq \text{seq}(2)$ is a tree), and for $l \in \omega$, let

$$\mathcal{T}(l) = \{ s \in \mathcal{T} : s \in T_0(l) \times \ldots \times T_{d-1}(l) \}.$$ 

A tree $T \subseteq \text{seq}(2)$ is perfect if for each $s \in T$ there is an $n > \text{dom}(s)$ and two distinct functions $t_0, t_1 \in n \cap T$ such that $t_0|_{\text{dom}(s)} = t_1|_{\text{dom}(s)} = s$. In other words, for each $s \in T$ there are $t_0, t_1 \in T$ and $k \in \text{dom}(t_0) \cap \text{dom}(t_1)$ such that $t_0|_{\text{dom}(s)} = t_1|_{\text{dom}(s)} = s$ and $t_0(k) = 1 - t_1(k)$.

Now we are ready to state and proof the following result.

**Theorem 11.6 (Weak Halpern-Läuchli Theorem).** For every positive $d \in \omega$ and for every colouring of $\bigcup_{l \in \omega} (\text{seq}(2))^d$ with finitely many colours, there exists a product of perfect trees $\mathcal{T} = T_0 \times \ldots \times T_{d-1}$ and an infinite set $H \subseteq \omega$ such that $\bigcup_{l \in H} \mathcal{T}(l)$ is monochromatic.

**Proof.** Let $d$ be a fixed positive integer and let $n := 2^d$. Because $|d^2| = 2^d$, there exists a one-to-one correspondence $\zeta$ between $n$ and $2^d$. For any $l \in \omega$, an element $\langle s_0, \ldots, s_{d-1} \rangle \in (\text{seq}(2))^d$ is a sequence of length $d$ of functions $s_i : l \to 2$. For any $l \in \omega$, define the function $\xi : (\text{seq}(2))^d \to (\text{seq}(2))^d$ by stipulating

$$\xi((s_0, \ldots, s_{d-1})) = \langle t_0, \ldots, t_{l-1} \rangle \text{ where } t_j(i) := s_i(j),$$

in other words, for any function $s : d \to 2$, $\xi(s)(j)(i) = s(i)(j)$. Notice that for each $l \in \omega$, the function $\xi$ is a one-to-one function from $\text{seq}(2)^d$ onto $\text{seq}(2)^d$. Let $S = \{ u_k : k \in n \}$ be such that $\min(u_0) < \min(u_1) < \ldots \min(u_{n-1})$.

For $j \in n$ let $t_j^S(i) := \xi(k)(i)$. Now, define the function $\eta : (\omega)^n \to (\text{seq}(2))^d$ by stipulating

$$\eta(S) = \xi^{-1}(\langle t_0^S, \ldots, t_{\text{dom}(S)-1}^S \rangle).$$

Notice that for $S \in \omega^n$ with $\text{Dom}(S) = l$, $\eta(S) \in (\text{seq}(2))^d$. Finally, for any colouring $\tau : \bigcup_{l \in \omega} (\text{seq}(2))^d \to r$, where $r$ is a positive integer, we define the colouring $\tau : (\omega)^n \to r$ by stipulating $\tau(S) := \tau(\eta(S))$. Let $X \in (\omega)^n$ be as in the conclusion of the Partition Ramsey Theorem 11.4 (with respect to the colouring $\tau$). Let $S_0^* < X$ be such that $|S_0| = n$ and let $H := \text{Min}(X) \setminus \text{Min}(S_0)$. Further, let

$$\mathcal{F} = \{ S \in (\omega)^n : S \not< S_0 \lor S_0 \not< S \subseteq X \}$$

and define

$$\mathcal{F} = \{ s \in (\text{seq}(2))^d : \exists S \in \mathcal{F} (s = \eta(S)) \}.$$ 

We leave it as an exercise to the reader to check that $\mathcal{F}$ and $H$ are as desired and that they have the desired properties.

For the full version of the Halpern-Läuchli Theorem see Related Result 77. However, in many applications the Weak Halpern-Läuchli
Theorem is strong enough. For example the Weak Halpern-Läuchli Theorem is sufficient to prove that a finite product of Sacks forcing does not add splitting reals (see Chapter 22 | Related Result 121).

Notes

Van der Waerden's Theorem. The theorem of van der Waerden can be considered as the beginning of Ramsey Theory and it was first proved by van der Waerden in [34]. For a short but easy proof of van der Waerden's Theorem see Graham and Rothschild [8], and for a combinatorial proof of a slightly more general result see Pin [22, Chapter 3]. For a description of how van der Waerden found his proof we refer the reader to [35].

The Hales-Jewett Theorem. In Graham, Rothschild, and Spencer [9, p. 35ff.] we can read the following remark: Van der Waerden's Theorem should be regarded, not as a result dealing with integers, but rather as a theorem about finite sequences formed from finite sets. The Hales-Jewett Theorem strips van der Waerden's Theorem of its unessential elements and reveals the heart of Ramsey theory. As mentioned above, the original proof of Hales and Jewett [13] (cf. Prömel and Voigt [28, p. 117ff.]) uses a double induction which leads to an extremely fast growing upper bound for the Hales-Jewett function $H_J(n, r)$. In 1987, Shelah [30] found a fundamentally new proof of the Hales-Jewett Theorem which just uses simple induction on $n$ and provides a much better bound for $H_J(n, r)$. The proof of the Hales-Jewett Theorem (i.e., of Theorem 11.3) presented here is Shelah's proof modified by Matet [23], who replaced what is sometimes called "Shelah's pigeonhole lemma" by the Finite Ramsey Theorem. For the Hales-Jewett Theorem, and in particular for Shelah's proof, see also Graham, Rothschild, and Spencer [9, Chapter 2]. Nill [25], Prömel and Voigt [28, p. 119ff.], and Jukna [19, Chapter 29].

Carlson's Lemma and the Partition Ramsey Theorem. According to Carlson and Simpson [4, p. 268], Carlson proved Lemma 2.4 of [4] in 1982. In fact, he proved a stronger result involving so-called "special partitions", which are essentially segmented partitions where finitely many blocks may be infinite; and in the proof of Lemma 11.5 we essentially followed Carlson's proof of that stronger result, which is Theorem 6.3 of [4]. Carlson's Lemma, or more precisely Lemma 2.4 of [4], plays a key role in the proof of the Dual Ramsey Theorem, which is the main result of Carlson and Simpson [4]. The Dual Ramsey Theorem corresponds to our Generalisation 1 — where the set $(\omega)^n$ is coloured with finitely many colours — except that the set of admissible colours of $(\omega)^n$ is restricted to Borel colourings. Thus, the Dual Ramsey Theorem is in a certain sense the dual of Ramsey's Theorem. However, it was natural to seek a partition form (i.e., dual form) of Ramsey's Theorem which works for arbitrary colourings. Such a result we found in the Partition Ramsey Theorem (see also Related Result 75). The proof of the Partition Ramsey Theorem 11.4 is taken from Halbeisen [10, Chapter IV.2] (for the relation between the Partition Ramsey Theorem and other Ramsey-type results we refer the reader to Halbeisen [10, Chapter IV.4]).

The Halpern-Läuchli Theorem. What we stated as Weak Halpern-Läuchli Theorem 11.6 is just a consequence of the Halpern-Läuchli Theorem (see Related Result 77), which was first proved by Halpern and Läuchli in [15] and later
by Halpern in [14] (see also Argyros, Felouzis and Kanellopoulos [1]), Todorcević [32, Chapter 3], or Todorcević and Farah [33]). According to Pincus and Halpern [26, p. 549] (cf. [16, p. 97]) the original purpose of the Halpern-Läuchli Theorem was to show that in ZF, the Prime Ideal Theorem does not imply the Axiom of Choice, which was proved by Halpern and Lévy in [16] (cf. Theorem 7.16, where it is shown that in ZFA, PIT does not imply AC). As mentioned above, in many applications, a weak form or a particular case of the Halpern-Läuchli Theorem is sufficient (e.g., Halpern and Lévy [16, p. 97]). The version of the Halpern-Läuchli Theorem given above—as well as the idea of proof—is taken from Carlson and Simpson [4, p. 272]. For some applications and other weak forms of the Halpern-Läuchli Theorem see Related Result 77.

Related Results

73. Van der Waerden numbers. For positive integers \( r \) and \( l_1, l_2, \ldots, l_r \), the van der Waerden number \( w(l_1, l_2, \ldots, l_r; r) \) is the least positive integer \( N \) such that for every \( r \)-colouring of set \{1, 2, \ldots, N\}, there is a monochromatic arithmetic progression of length \( l_i \) of colour \( i \) for some \( i \). In [3], Brown, Landman, and Robertson gave asymptotic lower bounds for \( w(l, m; 2) \) for fixed \( m \), as well as for \( w(4, 4, \ldots, 4; r) \).

74. Non-repetitive sequences and van der Waerden’s Theorem. A finite set of one or more consecutive terms in a sequence is called a segment of the sequence. A sequence on a finite set of symbols is called non-repetitive if no two adjacent segments are identical, where adjacent means abutting but not overlapping. It is known that there are infinite non-repetitive sequences on three symbols (see Pleasants [27]), and on the other hand, it is obvious that a non-repetitive sequence is at most of length 3. Erdős has raised in [6] the question of the maximum length of a sequence on \( k \) symbols, such that no two adjacent segments are permutations of each other. Such a sequence is called strongly non-repetitive. Keränen [30] has shown that four symbols are enough to construct an infinite strongly non-repetitive sequence. Now, replacing the finite set of symbols of an infinite strongly non-repetitive sequence by different prime numbers, one gets an infinite sequence on a finite set of integers such that no two adjacent segments have the same product. It is natural to ask whether one can replace in the statement above “product” by “sum”, which leads to the following question: Is it possible to construct an infinite sequence on a finite set of integers such that no two adjacent segments have the same sum? By an application of van der Waerden’s Theorem, it is not hard to show that the answer to this question is negative. Moreover, in any infinite sequence on a finite set of integers we always find arbitrary large finite sets of adjacent segments such that all these segments have the same sum (see Hungerbühler and Halbeisen [13]). However, it is still open whether there exists an infinite sequence on a finite set of integers such that no two adjacent segments of the same length have the same sum. It seems that van der Waerden’s Theorem alone is not strong enough to solve this problem.

75. Corollaries of the Partition Ramsey Theorem. Below, we present a few corollaries of the Partition Ramsey Theorem. We would like to mention that these
corollaries — like for example the Weak Halpern-Läuchli Theorem — also follow from the so-called Dual Ramsey Theorem, which is due to Carlson and Simpson [4].

Firstly we derive Ramsey’s Theorem from the Partition Ramsey Theorem: To every $r$-colouring $\pi : [\omega]^n \to r$ of the $n$-element subsets of $\omega$ we can assign an $r$-colouring $\tau : [\omega]^n \to r$ by stipulating $\tau(S) := \pi(\text{Min}(S^n) \setminus \{0\})$. Now, if $(X)^n$ is monochromatic for $\pi$ for some $X \in [\omega]^n$, then $\text{Min}(X) \setminus \{0\}$ is monochromatic for $\pi$, and since $\text{Min}(X) \in [\omega]^n$, this shows that Ramsey’s Theorem 2.1 is just a special case of the Partition Ramsey Theorem. Similarly, the finite Ramsey Theorem 2.3 as well as the Hales-Jewett Theorem 11.2 follows from the following finite version of the Partition Ramsey Theorem which is originally due to Graham and Rothschild [7, Corollary 10].

Graham-Rothschild Result: For all $m, n, r \in \omega$, where $r \geq 1$ and $n \leq m$, there exists an $N \in \omega$, where $N \geq m$, such that for every $r$-colouring of $(N)^n$ there exists a partition $H \in (N)^m$, all of whose $r$-block coarsenings have the same colour.

The relation between these results is illustrated by the following figure.

 related results

As a matter of fact we would like to remind the reader that we used the Finite Ramsey Theorem to prove the Hales-Jewett Theorem, that we used the Hales-Jewett Theorem to start the induction in the proof of Carlson’s Lemma 11.5, and that Carlson’s Lemma was crucial in the proof of the Partition Ramsey Theorem.

67. A generalisation of the Partition Ramsey Theorem. By combining Carlson’s Lemma with the Graham-Rothschild Result, Halbeisen and Matet [12] proved a result which is even stronger than the Partition Ramsey Theorem.

77. The Halpern-Läuchli Theorem. Before we can state the full Halpern-Läuchli Theorem of Halpern and Läuchli [10], we have to introduce some terminology. A set $T \subseteq \omega^\omega$, where $\omega^\omega = \bigcup_n \omega^n$, is a finitely branching tree if $T$ is a tree (i.e., for every $s \in T$ and $k \in \text{dom}(s)$, $s(k) \in T$) such that for all $s \in T$, the set \( \{ t \in T : s \subseteq t \land |t| = |s| + 1 \} \) is finite. An element $s \in T$ of a tree $T \subseteq \omega^\omega$ is a leaf if $\{ t \in T : s \nsubseteq t \} = \emptyset$. If $A$ and $B$ are subsets of a tree $T \subseteq \omega^\omega$, then we say that $A$ supports (dominates) $B$ if for all $t \in B$ there exists an $s \in A$ such that $s \subseteq t$ ($t \subseteq s$). A subset $D$ of a tree $T \subseteq \omega^\omega$ is said to be $(h, k)$-dense if there is an $s \in T$ with $|s| = h$ such that $\{ t \in T : s \subseteq t \land |t| = h + k \}$ is dominated by $D$. Let $\prod_{i \in d} T_i = T_0 \times \ldots \times T_{d-1}$ be a product of trees $T_i \subseteq \omega^\omega$. 
A product $\prod_{i \in I} A_i \subseteq \prod_{i \in I} T_i$, where each $A_i$ is $(h,k)$-dense in $T_i$, is called a $(h,k)$-matrix. Now we can state Theorem 1 of Halpern and Läuchli [15].

**Halpern-Läuchli Theorem:** Let $\prod_{i \in I} T_i$ be a finite product of finitely branching trees $T_i \subseteq \omega^\omega$ without leaves, and let $Q \subseteq \prod_{i \in I} T_i$. Then either

(a) for each $k$, $Q$ contains a $(0,k)$-matrix, or

(b) there exists $h$ such that for each $k$, $(\prod_{i \in I} T_i) \setminus Q$ contains an $(h,k)$-matrix.

There exist many reformulations, weakenings, and generalised forms of the Halpern-Läuchli Theorem. For example, Hans Läuchli proved in a student seminar at the ETH Zürich a weak form of the Halpern-Läuchli Theorem in which the trees $T_i \subseteq \omega^\omega$ were replaced by $\bigcup_{n \in \omega} \{ (\frac{1}{n}, \frac{1}{n}) : k \in \mathbb{N} \}$, and in which the set $\{0,1\}^2$ was coloured with two colours. The Halpern-Läuchli Theorem is a very strong combinatorial statement and even weak forms of it have interesting applications (see for example Chapter 22 Related Result 121, Blass [2], Polarized Theorem, or Milliken [24]). However, there are also some generalisations of the Halpern-Läuchli Theorem: For example, Laver [21] generalised the perfect tree version of the Halpern-Läuchli Theorem to infinite products (see also Ramović [29]), and Shelah [31] replaced the trees $T \subseteq \omega^\omega$ of height $\omega$ by trees of uncountable height (see also Díaz-Jarca, Larson, and Mitchell [9]).

78. **Partition regularity.** A finite or infinite matrix $A$ with rational entries in which there are only a finite number of non-zero entries in each row is called partition regular if, whenever the natural numbers are finitely coloured, there is a monochromatic vector $x$ (i.e., all entries of $x$ have the same colour) with $Ax = 0$.

Many of the classical theorems of Ramsey Theory may naturally be interpreted as assertions that particular matrices are partition regular. For example, Schur’s Theorem (i.e., Corollary 2.5) is the assertion that the $1 \times 3$-matrix $\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ is partition regular; or van der Waerden’s Theorem is (with the strengthening that we may also choose the common difference of the arithmetic progression to have the same colour) exactly the statement that a certain $(m-1) \times (m+1)$-matrix is partition regular (see Hindman, Leader, and Strauss [18]). While in the finite case partition regularity is well understood, very little is known in the infinite case. For a survey of results on partition regularity of matrices see Hindman [17].

**References**

References

10. Lorenz Halbeisen, Combinatorial properties of sets of partitions, Habilitationsschrift (2003/2009), University of Bern/Zürich (Switzerland).


Part II

From Martin’s Axiom to Cohen’s Forcing
...changes of genus are brought about not by the introduction of major or minor thirds, divided or undivided, but by a melodic progression through intervals proper to certain genera. It remains to be noted that the change from one genus to another is also accompanied by a change in melodic style.

...a difference of genus may be assumed when a notable divergence in melodic style is heard, with rhythm and words suitably accommodated to it.

GIOSEFFO ZARLINO

Le Istitutioni Harmoniche, 1558
The Idea of Forcing

Forcing is a technique— invented by Cohen in the early 1960s—for proving the independence, or at least the consistency, of certain statements relative to ZFC. In fact, starting from a model of ZFC, Cohen constructed in 1962 models of ZF in which the Axiom of Choice fails as well as models of ZFC in which the Continuum Hypothesis fails. On the other hand, starting from a model of ZF, Gödel constructed a model of ZFC in which the Continuum Hypothesis holds (cf. Chapter 5). By combining these results we get that the Axiom of Choice is independent of ZF and that the Continuum Hypothesis is independent of ZFC.

Before we discuss Cohen's forcing technique, let us briefly recall what it means for a sentence \( \varphi \) to be independent of ZFC. From a syntactical point of view it means that neither \( \varphi \) nor its negation is provable from ZFC. From a semantical point of view it means that there are models of ZFC in which \( \varphi \) holds and some in which \( \varphi \) fails. Equivalently we can say that \( \varphi \) is independent of ZFC iff \( \varphi \) as well as its negation is consistent with ZFC (i.e., ZFC + \( \varphi \) as well as ZFC + \( \neg \varphi \) has a model).

Now, in order to prove that a given sentence \( \varphi \) is consistent with ZFC, we have to show that ZFC + \( \varphi \) is consistent—tacitly assuming the consistency of ZFC. This can be done in different ways: For example one could apply the Compactness Theorem 3.7 and show that whenever ZFC* \( \subseteq \) ZFC is a finite set of axioms, then ZFC* + \( \varphi \) has a model (i.e., ZFC + \( \varphi \) is consistent); or, starting from a model of ZFC, one could construct directly a model of ZFC + \( \varphi \).

These two approaches correspond to two different ways to look at forcing: In the latter point of view we consider forcing as a technique for extending models of ZFC in such a way that \( \varphi \) holds in the extended model. Except for Chapter 16, we will mainly take this approach which will be demonstrated in Chapter 14. Before we discuss the former approach, let us give two examples how a model of a given theory can be extended.

An example from Group Theory: Consider the group \( G = (Q^+, \cdot) \) (i.e., \( G \models \text{GT} \), the domain of \( G \) is the set of all positive rational numbers with
multiplication as operation), and let \( \varphi \) be the statement \( \exists x (x \cdot x = 2) \). Obviously we have \( G \not\models \varphi \).

Now, extend the domain of \( G \) by elements of the form \( qX \), where \( q \in Q^+ \), and for all \( p, q \in Q^+ \) define:

- \( p \cdot q := p \cdot q \)
- \( p \cdot qX := (p \cdot q)X \)
- \( pX \cdot q := (p \cdot q)X \)
- \( pX \cdot qX := 2 \cdot p \cdot q \)
- \( (pX)^{-1} := \left( \frac{1}{2} \cdot p^{-1} \right)X \)

Let \( Q^+[X] = Q^+ \cup \{ pX : p \in Q^+ \} \) and \( G[X] = (Q^+[X], *) \). We leave it as an exercise to the reader to show that \( G[X] \not\models GT \). Now, \( G[X] \models 1X \cdot 1X = 2 \), and therefore, \( G[X] \not\models \varphi \). Thus, the extended model \( G[X] \) is a model of \( GT \) and the statement \( \varphi \), which failed in \( G \), holds in \( G[X] \). So, by extending an existing model we were able to “force” that a given statement became true.

An example from Peano Arithmetic: Assume that \( PA \) is consistent and let \( N = (\mathbb{N}, 0, s, +, \cdot) \) — where for \( n \in \mathbb{N} \), \( s(n) := n + 1 \) — be a model of \( PA \). Let \( \psi \) be the statement \( \exists x (x + x = 1) \), where \( 1 := s(0) \). Obviously we have \( N \not\models \psi \). Now, let us try the same trick as above: So, extend the domain of \( N \) by elements of the form \( n + X \), where \( n \in \mathbb{N} \), and extend the operation “+” by stipulating \( X + X := 1 \). Now, the corresponding model \( N[X] \) is surely a model of \( \psi \), but do we also have \( N[X] \models PA? \)

By setting \( \varphi(x) \equiv (x = 0) \lor \exists y (x = s(y)) \) in \( PA \), we get that each number is either equal to 0 or a successor. Now, since \( X \neq 0 \), it must be a successor. Thus, there is a \( y \) such that \( X = y + 1 \), and since \( X \neq 1 \), by \( PA \) we get \( y \neq 0 \). Similarly we can show that there is a \( z \) such that \( y = z + 1 \), and consequently \( X = (z + 1) + 1 \). Now, \( 1 = X + X = X + ((z + 1) + 1) \) and by \( PA \) we get \( X + ((z + 1) + 1) = (X + (z + 1)) + 1 \), which implies (by \( PA \)) that \( X + (z + 1) = 0 \). Applying again \( PA \) we finally get \( X + z + 1 = 0 \), which contradicts \( PA \). Thus, \( N[X] \) is not a model of \( PA \).

This example shows that just extending an existing model of a theory \( T \) in order to “force” that a given statement becomes true may result in a model which is no longer a model of \( T \).

Let us now discuss the other approach to forcing (demonstrated in Chapter 16), where one shows that whenever \( ZFC^* \) is a finite set of axioms of \( ZFC \), then \( ZFC^* + \varphi \) is consistent (as always, we tacitly assume the consistency of \( ZFC \)). Let \( ZFC^* \) be an arbitrary finite set of axioms of \( ZFC \) and let \( V \) be a model of \( ZFC \) (e.g., \( V = L \)). The so-called Reflection Principle (discussed in Chapter 15) tells us that for every finite fragment \( ZFC^* \) of \( ZFC \) (i.e., for every finite set of axioms of \( ZFC \)) there is a set model \( M \) such that \( M \models ZFC^* \) where the domain of \( M \) is a set \( M \) in the model \( V \). The goal is now to show that for any finite set \( \Phi \) of axioms of \( ZFC \), there is a finite fragment \( ZFC^* \) of \( ZFC \) such that it is possible to extend any set model \( M \) of \( ZFC^* \) to a set model \( M[X] \) of \( \Phi + \varphi \) (i.e., we “force” that \( \varphi \) as well as the formulae in \( \Phi \)
become true in $M[X]$). Then, since $\Phi$ was arbitrary, by the Compactness Theorem 3.7 we get the consistency of $\text{ZFC} + \varphi$.

The advantage of this approach is that the entire forcing construction can be carried out in the model $V$. Because $M$, the domain of $M$, is a set in the model $V$ (but not in the model $M$), we can extend the model $M$ within $V$ to the desired model $M[X]$, such that the domain of $M[X]$ is still a set in $V$. So, all takes place within the model $V$.

To illustrate this approach let us consider again the group-theoretic example from above: Let us work with the group $\mathcal{G} = (\mathbb{R}^+, \cdot)$, where $\mathbb{R}^+$ is the set of positive real numbers. Now, the group $G = (\mathbb{Q}^+, \cdot)$ is just a subgroup of $\mathcal{G}$, and in $\mathcal{G}$ we can extend $G$ to the group $G[\sqrt{2}]$ with domain $\mathbb{Q}^+ \cup \{p \cdot \sqrt{2} : p \in \mathbb{Q}^+\}$, which is still a subgroup of $\mathcal{G}$.

A difference to the other approach is that we look now at the model $G$ from the larger model $\mathcal{G}$ (i.e., from "outside"), and extend $G$ within this model. Another difference is that in the former example, the symbol $X$ — at least for people living in $G$ — is just a symbol with some specified properties, whereas in the latter example, $\sqrt{2}$ — at least for people living in $\mathcal{G}$ — is a proper real number. Of course, for people living in $G$, $\sqrt{2}$ is also just a symbol and is not more real than any other symbol. On the other hand, in the latter example the people living in $\mathcal{G}$ know already that $\sqrt{2}$ exists, whereas in the former example there are no such people, since our universe is just $G$.

Before the notion of forcing is introduced in Chapter 14, we present in the next chapter the so-called Martin's Axiom. We do so because on the one hand, Martin's Axiom is a statement closely related to forcing, involving also partially ordered sets and certain generic filters, but on the other hand, unlike forcing, it does not involve any model-theoretic or even metamathematical arguments. Furthermore, Martin's Axiom is a proper set-theoretical axiom which is widely used in other branches of Mathematics, especially in Topology.
Martin’s Axiom

In this chapter, we shall introduce a set-theoretic axiom, known as Martin’s Axiom, which is independent of ZFC. In the presence of the Continuum Hypothesis, Martin’s Axiom becomes trivial, but if the Continuum Hypothesis fails, then Martin’s Axiom becomes an interesting combinatorial statement as well as an important tool in Combinatorics. Furthermore, Martin’s Axiom provides a good introduction to the forcing technique which will be introduced in the next chapter.

Filters on Partially Ordered Sets

Below, we introduce (and recall respectively) some properties of partially ordered sets, which will play an important role in the development and investigation of forcing constructions.

Let \( P = (P, \leq) \) be a partially ordered set. The elements of \( P \) are usually called **conditions**, since in the context of forcing, elements of partially ordered sets are conditions for sentences to be true in generic extensions. Two conditions \( p_1 \) and \( p_2 \) of \( P \) are called **compatible**, denoted \( p_1 \parallel p_2 \), if there exists a \( q \in P \) such that \( p_1 \leq q \geq p_2 \); otherwise they are called **incompatible**, denoted \( p_1 \perp p_2 \).

A typical example of a partially ordered set is the set of finite partial functions with inclusion as partial ordering: Let \( I \) and \( J \) be arbitrary sets. Then \( \text{Fn}(I, J) \) is the set of all functions \( p \) such that

- \( \text{dom}(p) \in \text{fin}(I) \), i.e., \( \text{dom}(p) \) is a finite subset of \( I \), and
- \( \text{ran}(p) \subseteq J \).

For \( p, q \in \text{Fn}(I, J) \) define:

\[
p \leq q \iff \text{dom}(p) \subseteq \text{dom}(q) \land q|_{\text{dom}(p)} \equiv p
\]

If we consider functions as sets of ordered pairs, as we usually do, then \( p \leq q \) is just \( p \subseteq q \). We leave it as an exercise to the reader to verify that \( (\text{Fn}(I, J), \subseteq) \) is indeed a partially ordered set.
Let \( \mathbb{P} = (P, \leq) \) be a partially ordered set, and for the moment let \( C \subseteq P \). Then \( C \) is called **directed** if for any \( p_1, p_2 \in C \) there is a \( q \in C \) such that \( p_1 \leq q \leq p_2 \). \( C \) is called **open** if \( p \in C \) and \( q \geq p \) implies \( q \in C \), and \( C \) is called **downwards closed** if \( p \in C \) and \( q \leq p \) implies \( q \in C \). Furthermore, \( C \) is called **dense** if for every condition \( p \in P \) there is a \( q \in C \) such that \( q \geq p \). For example with respect to \((\text{Fn}(I, J), \subseteq)\), for every \( x \in I \) the set \( \{ p \in \text{Fn}(I, J) : x \in \text{dom}(p) \} \) is open and dense. Finally, a non-empty set \( F \subseteq P \) is a **filter** (on \( P \)) if it is directed and downwards closed. Notice that this definition of “filter” reverses the ordering from the definition given in Chapter 5. Let \( \mathcal{D} \subseteq \mathcal{P}(P) \) be a set of open dense subsets of \( P \). A filter \( G \subseteq P \) is a **\( \mathcal{D} \)-generic filter** on \( P \) if \( G \cap D \neq \emptyset \) for every open dense set \( D \in \mathcal{D} \). As an example consider again \((\text{Fn}(I, J), \subseteq)\): If \( \mathcal{D} \) is a filter on \( \text{Fn}(I, J) \), then \( \bigcup \mathcal{D} : X \to J \) is a function, where \( X \) is some (possibly infinite) subset of \( I \).

**Proposition 13.1.** If \((P, \leq)\) is a partially ordered set and \( \mathcal{D} \) is a countable set of open dense subsets of \( P \), then there exists a \( \mathcal{D} \)-generic filter on \( P \). Moreover, for every \( p \in P \) there exists a \( \mathcal{D} \)-generic filter \( G \) on \( P \) which contains \( p \).

**Proof.** For \( \mathcal{D} = \{ D_n : n \in \omega \} \) and \( p_{-1} := p \), choose for each \( n \in \omega \) a \( p_n \in D_n \) such that \( p_n \geq p_{n-1} \), which is possible since \( D_n \) is dense. Then the set

\[
G = \{ q \in P : \exists n \in \omega (q \leq p_n) \}
\]

is a \( \mathcal{D} \)-generic filter on \( P \) and \( p \in G \).

A set \( A \subseteq P \) is an **anti-chain** in \( P \) if any two distinct elements of \( A \) are incompatible. As mentioned in Chapter 5, this definition of “anti-chain” is different from the one used in Order Theory. A partially ordered set \( \mathbb{P} = (P, \leq) \) satisfies the **countable chain condition**, denoted \( ccc \), if every anti-chain in \( P \) is at most countable (i.e., finite or countably infinite).

As a consequence of the following lemma we get that \( \text{Fn}(I, J) \) satisfies \( ccc \) whenever \( J \) is countable.

**Lemma 13.2 (\( \Delta \)-System Lemma).** Let \( \mathcal{E} \) be an uncountable family of finite sets. Then there exist an uncountable family \( \mathcal{C} \subseteq \mathcal{E} \) and a finite set \( \Delta \) such that for any distinct elements \( x, y \in \mathcal{C} : x \cap y = \Delta \).

**Proof.** We shall consider two cases.

**Case 1:** There exists an uncountable \( \mathcal{E}' \subseteq \mathcal{E} \) such that for every \( \alpha \in \bigcup \mathcal{E}' \), \( \{ x \in \mathcal{E}' : x \cap \alpha = \emptyset \} \) is countable. Firstly notice that for such a set \( \mathcal{E}' \), \( \bigcup \mathcal{E}' \) is uncountable, and that for any countable set \( C \subseteq \bigcup \mathcal{E}' \), also the set \( \{ x \in \mathcal{E}' : x \cap C = \emptyset \} \) must be uncountable. By transfinite induction we construct an uncountable family \( \{ x_\alpha : \alpha \in \omega_1 \} \subseteq \mathcal{E}' \) of pairwise disjoint sets as follows: Let \( x_0 \) be any member of \( \mathcal{E}' \). If we have already constructed a set \( C_\alpha = \{ x_\xi : \xi < \alpha \in \omega_1 \} \subseteq \mathcal{E}' \) of pairwise disjoint sets, let \( x_\alpha \in \mathcal{E}' \) be such that \( x_\alpha \cap \bigcup C_\alpha = \emptyset \). Then \( \mathcal{C} = \{ x_\alpha : \alpha \in \omega_1 \} \) and \( \Delta = \emptyset \) are as required.
Case 2: For every uncountable \( \mathcal{E}' \subseteq \mathcal{E} \) there exists an \( a \in \bigcup \mathcal{E}' \) such that 
\( \{ x \in \mathcal{E}' : a \in x \} \) is uncountable. In this case, consider the function \( \nu : \mathcal{E} \to \omega \), where for all \( x \in \mathcal{E} \), \( \nu(x) := |x| \). Since \( \mathcal{E} \) is uncountable, there is an \( n \in \omega \) and an uncountable set \( \mathcal{E}' \subseteq \mathcal{E} \) such that \( \nu|_{\mathcal{E}'} \equiv n \), i.e., for all \( x \in \mathcal{E}' \) we have \( \nu(x) = n \).

The proof is now by induction on \( n \): If \( n = 1 \), then for any two distinct elements \( x, y \in \mathcal{E}' \) we have \( x \cap y = \emptyset \), thus, \( \Delta = \emptyset \) and in this case \( \mathcal{C} = \mathcal{E}' \).

Now, let us assume that \( \nu|_{\mathcal{E}'} \equiv n + 1 \) for some \( n \geq 1 \) and that the lemma holds for \( n \). Since we are in Case 2, there is an \( a \in \bigcup \mathcal{E}' \) such that 
\( \{ x \in \mathcal{E}' : a \in x \} \) is uncountable. Thus, we can apply the induction hypothesis to the family \( \mathcal{E}_a' := \{ x \setminus \{ a \} : x \in \mathcal{E}' \land a \in x \} \) and obtain an uncountable family \( \mathcal{E}_a \subseteq \mathcal{E}_a' \) and a finite set \( \Delta_a \) such that for any distinct elements \( x, y \in \mathcal{E}_a \) we have \( x \cap y = \Delta_a \). Then \( \mathcal{C} := \{ x \cup \{ a \} : x \in \mathcal{E}_a \} \) and \( \Delta := \Delta_a \cup \{ a \} \) are as required.

**Corollary 13.3.** If \( I \) is arbitrary and \( J \) is countable, then \( \text{Fn}(I, J) \) satisfies the countable chain condition.

**Proof.** Let \( \mathcal{F} \subseteq \text{Fn}(I, J) \) be an uncountable family of partial functions. We have to show that \( \mathcal{F} \) is not an anti-chain, i.e., we have to find at least two distinct conditions in \( \mathcal{F} \) which are compatible. Let \( \mathcal{E} := \{ \text{dom}(p) : p \in \mathcal{F} \} \). Then \( \mathcal{E} \) is obviously a family of finite sets. Further, since \( J \) is assumed to be countable, for every finite set \( K \in \text{fin}(I) \) the set \( \{ p \in \mathcal{E} : \text{dom}(p) = K \} \) is countable, and therefore, since \( \mathcal{F} \) is uncountable, \( \mathcal{E} \) is uncountable as well.

Applying the \( \Delta \)-System Lemma 13.2 to the family \( \mathcal{E} \) yields an uncountable family \( \mathcal{C} \subseteq \mathcal{F} \) and a finite set \( \Delta \subseteq I \), such that for all distinct \( p, q \in \mathcal{C} \), \( \text{dom}(p) \cap \text{dom}(q) = \Delta \).

Since \( J \) is countable and \( \Delta \) is finite, uncountably many conditions of \( \mathcal{C} \) must agree on \( \Delta \), i.e., for some \( p_0 \in \text{Fn}(I, J) \) with \( \text{dom}(p_0) = \Delta \), the set \( \mathcal{C}' = \{ q \in \mathcal{C} : q|_{\Delta} = p_0 \} \) is uncountable. So, \( \mathcal{C}' \) is an uncountable subset of \( \mathcal{F} \) consisting of pairwise compatible conditions, hence, \( \mathcal{F} \) is not an anti-chain.

The following hypothesis can be regarded as a generalisation of Proposition 13.1 — for the reason why \( \mathcal{P} \) must satisfy ccc see Proposition 13.4.

**MA(\( \kappa \)):** If \( \mathcal{P} = (P, \leq) \) is a partially ordered set which satisfies ccc, and \( \mathcal{D} \) is a set of at most \( \kappa \) open dense subsets of \( P \), then there exists a \( \mathcal{D} \)-generic filter on \( P \).

On the one hand, MA(\( \omega \)) is just Proposition 13.1, and therefore, MA(\( \omega \)) is provable in ZFC. On the other hand, MA(\( c \)) is just false as we will see in Proposition 13.5. However, the following statement can neither be proved nor disproved in ZFC and can therefore be considered as a proper axiom of Set Theory (especially when CH fails):
Martin’s Axiom (MA): If \( P = (P, \leq) \) is a partially ordered set which satisfies ccc, and \( \mathcal{D} \) is a set of less than \( \kappa \) open dense subsets of \( P \), then there exists a \( \mathcal{D} \)-generic filter on \( P \). In other words, MA(\( \kappa \)) holds for each cardinal \( \kappa < \kappa \).

If we assume CH, then \( \kappa < \kappa \) is the same as saying \( \kappa \leq \omega \), thus, by Proposition 13.1, CH implies MA. On the other hand, MA can replace the Continuum Hypothesis in many proofs that use CH, which is important since MA is consistent with ZFC + ~CH (see Chapter 19).

It might be tempting to generalise Martin’s Axiom by weakening its premise: Firstly, one might try to omit ccc, and secondly, one might try to allow larger families of open dense subsets of \( P \). However, both attempts to generalise MA fail.

**Proposition 13.4.** There exist a (non ccc) partially ordered set \( P = (P, \leq) \) and a set \( \mathcal{D} \) of cardinality \( \omega_1 \) of open dense subsets of \( P \) such that no filter on \( P \) is \( \mathcal{D} \)-generic.

**Proof.** Consider the partially ordered set \( (\text{Fn}(\omega, \omega_1), \subseteq) \). For each \( \alpha \in \omega_1 \), the set

\[ D_\alpha = \{ p \in \text{Fn}(\omega, \omega_1) : \alpha \in \text{ran}(p) \} \]

is an open dense subset of \( \text{Fn}(\omega, \omega_1) \): Obviously, \( D_\alpha \) is open. To see that \( D_\alpha \) is also dense, take any \( p \in \text{Fn}(\omega, \omega_1) \). If \( \alpha \in \text{ran}(p) \), then \( p \in D_\alpha \) and we are done. Otherwise, let \( n \in \omega \) be such that \( n \notin \text{dom}(p) \) (notice that such an \( n \) exists since \( \text{dom}(p) \) is finite). Now, let \( q := p \cup \{(n, \alpha)\} \); then \( q \in D_\alpha \) and \( q \geq p \). Similarly, for each \( n \in \omega \), the set \( E_n = \{ p \in \text{Fn}(\omega, \omega_1) : n \in \text{dom}(p) \} \) is open dense.

Let \( \mathcal{D} = \{ D_\alpha : \alpha \in \omega_1 \} \cup \{ E_n : n \in \omega \} \); then \( |\mathcal{D}| = \omega_1 \). Assume that \( G \subseteq \text{Fn}(\omega, \omega_1) \) is a \( \mathcal{D} \)-generic filter on \( \text{Fn}(\omega, \omega_1) \). Since for each \( n \in \omega \), \( G \cap E_n \neq \emptyset \), \( f_G = \bigcup G \) is a function from \( \omega \) to \( \omega_1 \). Further, since for each \( \alpha \in \omega_1 \), \( G \cap D_\alpha \neq \emptyset \), the function \( f_G : \omega \to \omega_1 \) is even surjective, which contradicts the definition of \( \omega_1 \).

**Proposition 13.5.** MA(\( \kappa \)) is false.

**Proof.** Consider the partially ordered set \( (\text{Fn}(\omega, 2), \subseteq) \). Then \( \text{Fn}(\omega, 2) \) is countable and consequently satisfies ccc. For each \( g \in \omega_2 \), the set

\[ D_g = \{ p \in \text{Fn}(\omega, 2) : \exists n \in \omega (p(n) = 1 - g(n)) \} \]

is an open dense subset of \( \text{Fn}(\omega, 2) \): Obviously, \( D_g \) is open, and for \( p \notin D_g \) let \( q := p \cup \{(n, 1 - g(n))\} \) where \( n \notin \text{dom}(p) \). Then \( q \in D_g \) and \( q \geq p \). Similarly, for each \( n \in \omega \), the set \( D_n = \{ p \in \text{Fn}(\omega, 2) : n \in \text{dom}(p) \} \) is open dense.

Let \( \mathcal{D} = \{ D_g : g \in \omega_2 \} \cup \{ D_n : n \in \omega \} \). Then \( |\mathcal{D}| = \omega_2 = \kappa \). Assume that \( G \subseteq \text{Fn}(\omega, 2) \) is a \( \mathcal{D} \)-generic filter on \( \text{Fn}(\omega, 2) \). Since for each \( n \in \omega \), \( G \cap D_n \neq \emptyset \), \( f_G = \bigcup G \) is a function from \( \omega \) to \( 2 \). Further, since for each \( g \in \omega_2 \), \( G \cap D_g \neq \emptyset \), \( f_G \) would be a function from \( \omega \) to \( 2 \) which differs from every function \( g \in \omega_2 \), which is impossible. 

\[ \square \]
Weaker Forms of MA

Below, we introduce a few forms of Martin's Axiom which are in fact proper weakenings of MA (cf. Related Result 81).

Let $P = (P, \leq)$ be a partially ordered set. $P$ is said to be countable if the set $P$ is countable; and $P$ is said to be $\sigma$-centred if $P$ is the union of at most countably many centred sets, where a set $Q \subseteq P$ is called centred, if any finite set $q_1, \ldots, q_n \in Q$ has an upper bound in $Q$.

Let $P$ be any property of partially ordered sets, e.g., $P = \sigma$-centred, $P = \text{ccc}$, or $P = \text{countable}$. Then $\text{MA}(P)$ is the following statement.

$$\text{MA}(P): \text{If } P = (P, \leq) \text{ is a partially ordered set having the property } P, \text{ and } \mathcal{D} \text{ is a set of less than } \kappa \text{ open dense subsets of } P, \text{ then there exists a } \mathcal{D}\text{-generic filter on } P.$$ 

Since every countable partially ordered set is $\sigma$-centred, and every $\sigma$-centred partially ordered set satisfies $\text{ccc}$, we obviously get:

$$\text{MA} \Rightarrow \text{MA}(\sigma\text{-centred}) \Rightarrow \text{MA}(\text{countable})$$

Below, we present some consequences of Martin's Axiom for countable and $\sigma$-centred partially ordered sets.

Some consequences of $\text{MA}(\sigma\text{-centred})$

Theorem 13.6. $\text{MA}(\sigma\text{-centred})$ implies $\mathfrak{p} = \kappa$.

Proof. Let $\kappa < \kappa$ be an infinite cardinal and let $\mathcal{F} = \{x_\alpha : \alpha \in \kappa\} \subseteq [\omega]^{<\omega}$ be a family with the strong finite intersection property (i.e., intersections of finitely many members of $\mathcal{F}$ are infinite) of cardinality $\kappa$. Under the assumption of $\text{MA}(\sigma\text{-centred})$ we construct an infinite pseudo-intersection of $\mathcal{F}$.

Let $P$ be the set of all ordered pairs $\langle s, E \rangle$ such that $s \in [\omega]^{<\omega}$ and $E \in \text{fin}(\kappa)$; and for $\langle s, E \rangle, \langle t, F \rangle \in P$ define

$$\langle s, E \rangle \leq \langle t, F \rangle \iff s \subseteq t \land E \subseteq F \land (t \setminus s) \subseteq \bigcap \{x_\alpha : \alpha \in E\}.$$ 

For $s \in [\omega]^{<\omega}$ let $P_s := \{\langle s, E \rangle \in P : E \in \text{fin}(\kappa)\}$. Then any finite set $\langle s, E_1 \rangle, \ldots, \langle s, E_n \rangle \in P_s$ has an upper bound, namely $\langle s, \bigcup_{i=1}^n E_i \rangle$, and since $[\omega]^{<\omega}$ is countable and $P = \bigcup \{P_s : s \in [\omega]^{<\omega}\}$, the partially ordered set $(P, \leq)$ is $\sigma$-centred. For each $\alpha \in \kappa$ and $n \in \omega$, the set

$$D_{\alpha, n} = \{\langle s, E \rangle \in P : \alpha \in E \land |s| > n\}$$

is an open dense subset of $P$. Let $\mathcal{G} = \{D_{\alpha, n} : \alpha \in \kappa \land n \in \omega\}$. Then $|\mathcal{G}| = \kappa$, in particular, $|\mathcal{G}| < \kappa$. So, by $\text{MA}(\sigma\text{-centred})$ there exists a $\mathcal{G}$-generic filter
G on P. Let \( x_G := \bigcup \{ s \in [\omega]^{<\omega} : \exists E \in \text{fin}(\kappa)(\langle s, E \rangle \in G) \} \). Then, by construction, \( x_G \) is infinite. Moreover, for every \( \alpha \in \kappa \) there is a condition \( \langle s, E \rangle \in G \) such that \( \alpha \in E \), which implies that \( x_G \setminus s \subseteq x_\alpha \). Hence, for each \( \alpha \in \kappa \) we have \( x_G \subseteq x_\alpha \), and therefore, \( x_G \) is an infinite pseudo-intersection of \( \mathcal{F} \).

The key idea in the proof that MA(\( \sigma \)-centred) \( \implies \) \( 2^\kappa = \kappa^+ \) for all infinite cardinals \( \kappa < \kappa^+ \) is to encode subsets of an almost disjoint family of cardinality \( \kappa < \kappa^+ \) by subsets of \( \omega \). For the premise of the following lemma — in which the “codes” are constructed — recall that there is always an almost disjoint family of cardinality \( \kappa \), and therefore of any cardinality \( \kappa < \kappa^+ \) (cf. Proposition 8.6).

**Lemma 13.7.** Let \( \kappa < \kappa^+ \) be an infinite cardinal and let \( \mathcal{A} = \{ x_\alpha : \alpha \in \kappa \} \subseteq [\omega]^{<\omega} \) be an almost disjoint family of cardinality \( \kappa < \kappa^+ \). Furthermore, let \( \mathcal{B} \subseteq \mathcal{A} \) be any subfamily of \( \mathcal{A} \) and let \( \mathcal{C} = \mathcal{A} \setminus \mathcal{B} \). If we assume MA(\( \sigma \)-centred), then there exists a set \( c \subseteq \omega \) such that for all \( x \in \mathcal{A} \):

\[
|c \cap x| = \omega \iff x \in \mathcal{B}
\]

**Proof.** Similar as in the proof of Theorem 13.6, let \( P \) be the set of all ordered pairs \( (s, E) \) such that \( s \in [\omega]^{<\omega} \) and \( E \in \text{fin}(\mathcal{C}) \); and for \( (s, E), (t, F) \in P \) define

\[
(s, E) \leq (t, F) \iff s \subseteq t \land E \subseteq F \land (t \setminus s) \cap \bigcup E = \emptyset.
\]

Similar as above, one shows that the partially ordered set \( (P, \leq) \) is \( \sigma \)-centred.

Now, for each \( x_\gamma \in \mathcal{C} \), the set

\[
D_{x_\gamma} = \{ (s, E) \in P : x_\gamma \in E \}
\]

is an open dense subset of \( P \); and for each \( x_\beta \in \mathcal{B} \) and each \( k \in \omega \), the set

\[
D_{x_\beta,k} = \{ (s, E) \in P : |s \cap x_\beta| \geq k \}
\]

is also an open dense subset of \( P \). Notice that we do not require that \( \mathcal{C} \) or \( \mathcal{B} \) is non-empty. Finally, let \( \mathcal{D} = \{ D_{x_\gamma} : x_\gamma \in \mathcal{C} \} \cup \{ D_{x_\beta,k} : x_\beta \in \mathcal{B} \land k \in \omega \} \). Then \( |\mathcal{D}| = \kappa \), and since \( \kappa < \kappa^+ \) we get \( |\mathcal{D}| < \kappa^+ \). So, by MA(\( \sigma \)-centred) there exists a \( \mathcal{D} \)-generic filter \( G \) on \( P \). Let \( c = \bigcup \{ s \in [\omega]^{<\omega} : \exists E \in \text{fin}(\mathcal{C})(\langle s, E \rangle \in G) \} \).

Then for any \( x_\beta \in \mathcal{B} \), \( |c \cap x_\beta| = \omega \); and, like in the proof of Theorem 13.6, for any \( x_\gamma \in \mathcal{C} \), \( |c \cap x_\gamma| < \omega \). Thus, the set \( c \subseteq \omega \) has the required properties. \( \dagger \)

Now we are ready to prove the following consequences of MA(\( \sigma \)-centred):

**Theorem 13.8.** If we assume MA(\( \sigma \)-centred), then for all infinite cardinals \( \kappa < \kappa^+ \) we have \( 2^\kappa = \kappa^+ \), and as a consequence we get that \( \kappa \) is regular.
MA(countable) implies the existence of Ramsey ultrafilters

Proof. Let $\kappa < \infty$ be an infinite cardinal. We have to show that $2^\kappa = \infty$. For this, fix an almost disjoint family $\mathcal{A} = \{x_\alpha : \alpha \in \kappa\} \subseteq [\omega]^\kappa$ of cardinality $\kappa$, and for each $u \in \mathcal{P}(\kappa)$ let $\mathcal{B}_u := \{x_\alpha \in \mathcal{A} : \alpha \in u\}$. Then, by Lemma 13.7, there is a set $c_\kappa \subseteq \omega$ such that for each $x \in \mathcal{A}$ we have $|c_\kappa \cap x| = \omega \iff x \in \mathcal{B}_u$. Notice that for any distinct $u, v \in \mathcal{P}(\kappa)$ we have $c_u \neq c_v$. Indeed, if $u, v \in \mathcal{P}(\kappa)$ are distinct, then without loss of generality we may assume that there exists an $\alpha \in \kappa$ such that $\alpha \in u \setminus v$. So, $c_u \cap x_\alpha$ is infinite, whereas $c_v \cap x_\alpha$ is finite, and hence, $c_u \neq c_v$. Thus, the mapping

$$\mathcal{P}(\kappa) \to \mathcal{P}(\omega)
\quad u \mapsto c_u$$

is one-to-one, which implies that $2^\kappa \leq \infty$. Now, since $\omega \leq \kappa$, and consequently $\infty \leq 2^\kappa$, we finally get $2^\kappa = \infty$.

To see that $\infty$ is regular assume towards a contradiction that $\kappa = cf(\infty) < \infty$. Then, by Corollary 5.12, $\infty < \infty$, but since $\infty = 2^\kappa$ we get that $\infty = (2^\kappa)^{\infty} = 2^\kappa = \infty$, a contradiction.

MA(countable) implies the existence of Ramsey ultrafilters

As a consequence of MA(countable) we get that there are $2^\infty$ mutually non-isomorphic Ramsey ultrafilters. By Chapter 10 Related Result 64, it would be enough to show that MA(countable) implies $p = \infty$; however, this is not the case (cf. Related Results 79–81 and Corollary 21.11).

Proposition 13.9. MA(countable) implies that there exist $2^\infty$ mutually non-isomorphic Ramsey ultrafilters.

Proof. Since there are just $\infty$ permutations of $\omega$, in order to get $2^\infty$ mutually non-isomorphic Ramsey ultrafilters it is enough to find $2^\infty$ distinct Ramsey ultrafilters. The $2^\infty$ mutually distinct Ramsey ultrafilters are constructed by transfinite induction: For every $\gamma : \infty \to 2$ and every $\alpha \in \infty$ we construct a set $\mathcal{F}_\gamma|\alpha = \{x_\beta(\gamma(\beta)) : \beta \in \alpha\} \subseteq [\omega]^\omega$ with the finite intersection property such that the filter generated by $\bigcup_{\alpha \in \infty} \mathcal{F}_\gamma|\alpha$ is a Ramsey ultrafilter. In addition we make sure that for any two distinct $\gamma, \gamma' \in 2^\infty$, the filters generated by $\bigcup_{\alpha \in \infty} \mathcal{F}_\gamma|\alpha$ and $\bigcup_{\alpha \in \infty} \mathcal{F}_{\gamma'}|\alpha$ are distinct. In order to get Ramsey ultrafilters at the end, by Proposition 13.7(b) it is enough to make sure that for every partition $\{Y_n : n \in \omega\}$ of $\omega$, either there is an $n_0 \in \omega$ such that $Y_{n_0} \subseteq \bigcup_{\alpha \in \infty} \mathcal{F}_\gamma|\alpha$, or there exists an $x \in \bigcup_{\alpha \in \infty} \mathcal{F}_\gamma|\alpha$ such that for all $n \in \omega$, $|x \cap Y_n| \leq 1$.

Let $\{\mathcal{P}_\alpha : \alpha \in \infty\}$ be the set of all infinite partitions of $\omega$. Thus, for each $\alpha \in \infty$, $\mathcal{P}_\alpha = \{Y_n^\alpha : n \in \omega\}$ is a set of pairwise disjoint subsets of $\omega$ such that $\bigcup \mathcal{P}_\alpha = \omega$. Further, let $x_{0,0} := \{2n : n \in \omega\}$, $x_{0,1} := \{2n+1 : n \in \omega\}$, and for $\delta \in \{0, 1\}$ let $\mathcal{F}_{\delta}(\omega) := \{x_{\delta, \delta}\} \cup \{x \subseteq \omega : |\omega \setminus x| < \omega\}$. Obviously, both sets $\mathcal{F}_{\delta}(\omega)$ and $\mathcal{F}_{\delta}(\omega)$ have the finite intersection property. Let $\alpha \in \infty$ and assume that for each $\eta \in \alpha'e$ and each $\beta \in \alpha$ we already have constructed a set $\mathcal{F}_\beta|\gamma \subseteq [\omega]^\omega$ with the finite intersection property, and such that for any
\( \beta_0 \in \beta_1 \in \alpha \) we have \( \mathcal{F}_{\eta|\beta_0} \subseteq \mathcal{F}_{\eta|\beta_1} \). In order to construct \( \mathcal{F}_\eta \) we have to consider two cases:

**\( \alpha \) limit ordinal:** If \( \alpha \) is a limit ordinal, then let

\[
\mathcal{F}_\eta = \bigcup_{\beta \in \alpha} \mathcal{F}_{\eta|\beta}.
\]

Since the sets \( \mathcal{F}_{\eta|\beta} \) are increasing and each of these sets has the finite intersection property, \( \mathcal{F}_\eta \) has the finite intersection property as well.

**\( \alpha \) successor ordinal:** If \( \alpha \) is a successor ordinal, say \( \alpha = \beta_0 + 1 \), then we proceed as follows: Consider the partition \( \mathcal{P}_{\beta_0} = \{ Y_n : n \in \omega \} \) and notice that either there is an \( n_0 \in \omega \) such that \( \mathcal{F}_{\eta|\beta_0} \cup \{ Y_{n_0} \} \) has the finite intersection property, or for every \( n \in \omega \), \( Y_n \) belongs to the dual ideal of \( \mathcal{F}_{\eta|\beta_0} \), i.e., it is a subset of the complement of a finite intersection of members of \( \mathcal{F}_{\eta|\beta_0} \). We consider the two cases separately:

**Case 1:** Let \( n_0 \in \omega \) be such that \( \mathcal{F}_{\eta|\beta_0} \cup \{ Y_{n_0} \} \) has the finite intersection property. Let \( P_1 = \text{Fin}(Y_{n_0}, 2) \) and for \( p, q \in P_1 \) let \( p \leq q \iff p \subseteq q \). Then \( (P_1, \leq) \) is countable and for every finite set \( E \in \text{fin}(\beta_0) \), every \( n \in \omega \) and each \( \delta \in \{0, 1\} \), the set

\[
D_{E,n,\delta} = \left\{ p \in P_1 : \left| p^{-1}(\delta) \cap \bigcap_{i \in E} x_{\alpha, \eta(i)} \right| \geq n \right\}
\]

is an open dense subset of \( P_1 \). Now let \( \mathcal{Q} = \{ D_{E,n,\delta} : E \in \text{fin}(\beta_0) \land n \in \omega \land \delta \in \{0, 1\} \} \). Then \( |\mathcal{Q}| \leq \max\{ |\alpha|, \omega \} < \kappa \) and by MA(countable) there exists a \( \mathcal{Q} \)-generic filter \( G \) on \( P_1 \). For \( \delta \in \{0, 1\} \), let

\[
x_{\eta|\beta_0,\delta} := \bigcup \left\{ p^{-1}(\delta) : p \in G \right\}.
\]

For \( \delta \in \{0, 1\} \) we get that \( x_{\eta|\beta_0,\delta} \in [Y_{n_0}]^\omega \) and that \( \mathcal{F}_\eta := \mathcal{F}_{\eta|\beta_0} \cup \{ x_{\eta|\beta_0,\delta} \} \) has the finite intersection property. Finally, let \( \eta, \eta' \in {}^\omega \mathbb{2} \) be such that \( \eta(\beta_0) = 1 - \eta'(\beta_0) \). Since \( x_{\eta|\beta_0} \cap x_{\eta'|\beta_0} = \emptyset \) we obviously have \( \mathcal{F}_\eta \neq \mathcal{F}_{\eta'} \). Moreover, by construction we get that \( \mathcal{F}_\eta \cup \mathcal{F}_{\eta'} \) lacks the finite intersection property, and therefore no ultrafilter can extend both \( \mathcal{F}_\eta \) and \( \mathcal{F}_{\eta'} \).

**Case 2:** If for each \( n \in \omega \), \( Y_n \) belongs to the dual ideal of \( \mathcal{F}_{\eta|\beta_0} \), then each finite intersection of members of \( \mathcal{F}_{\eta|\beta_0} \) meets infinitely many sets of \( \mathcal{F}_{\eta|\beta_0} \). Let \( P_2 \subseteq \text{Fin}(\omega, 2) \) be such that \( p \in P_2 \) iff for every \( Y \in \mathcal{F}_{\eta|\beta_0} \) we have

\[
\max \left\{ |p^{-1}(0) \cap Y|, |p^{-1}(1) \cap Y| \right\} \leq 1,
\]

and for \( p, q \in P_2 \) let \( p \leq q \iff p \subseteq q \). Like before, \( (P_2, \leq) \) is countable and for every finite set \( E \in \text{fin}(\beta_0) \), every \( n \in \omega \) and each \( \delta \in \{0, 1\} \), the set

\[
D_{E,n,\delta} = \left\{ p \in P_2 : \left| p^{-1}(\delta) \cap \bigcap_{i \in E} x_{\alpha, \eta(i)} \right| \geq n \right\}
\]
is an open dense subset of $P_2$. Let $\mathcal{D} = \{ D_{E,n,\delta} : E \in \text{fin}(\beta_0) \land n \in \omega \land \delta \in \{ 0, 1 \} \}$ and let $G$ be a $\mathcal{D}$-generic filter on $P_2$. Finally, for $\delta \in \{ 0, 1 \}$ let $x_{\beta_0,\delta} := \bigcup \{ p^{-1}(\delta) : p \in G \}$. Then $\mathcal{F}_n := \mathcal{F}_{n|\beta_0} \cup \{ x_{\beta_0,n(\beta_0)} \}$ has the finite intersection property, and in addition there exists a set $x \in \mathcal{F}_n$ such that for all $n \in \omega$, $|x \cap Y_n| \leq 1$. Further, for $\eta, \eta' \in {}^\omega 2$ with $\eta(\beta_0) = 1 - \eta'(\beta_0)$, no ultrafilter can extend both $\mathcal{F}_n$ and $\mathcal{F}_{n'}$.

Finally, for each $\gamma \in {}^\omega 2$, let $\mathcal{F}_\gamma$ be the filter generated by the set $\bigcup_{\alpha \in \gamma} \mathcal{F}_{n|\alpha}$. By construction, for any two distinct $\gamma, \gamma' \in {}^\omega 2$, $\mathcal{F}_\gamma$ and $\mathcal{F}_{\gamma'}$ are two distinct Ramsey ultrafilters, and consequently there exist $2^\omega$ mutually non-isomorphic Ramsey ultrafilters.

\begin{notes}

\textbf{Martin’s Axiom.} MA was first discovered by Martin and Solovay [8]. The paper contains various equivalent formulations of MA and numerous applications (including Theorem 13.8). They also stress the usefulness of MA as a viable alternative to CH and point out that many of the traditional problems solved using CH can be solved using MA. Roughly speaking, this is because under MA, sets of cardinality less than $\omega$ usually behave like countable sets (but of course, there are exceptions).

For equivalents of MA, consequences, weaker forms, history, \textit{et cetera}, refer the reader to Kuenen [7, Chapter II, §2-§5], Fremlin [4], Weiss [12], Rudin [10], Blass [2, Section 7], and Jech [6, Chapter 10].

\textbf{MA(countable) and Ramsey ultrafilters.} Proposition 13.9 is due to Canjar [3] (who actually proved even more), but the proof given above was communicated to me by Michael Hrušík (compare Proposition 13.9 with Chapter 10 | Related Result 64).

\textbf{The $\Delta$-System Lemma.} This useful combinatorial result was first proved by Shanin [11] (see Kuenen [7, Chapter II, §1] for a slightly more general result).

\textbf{Related Results}

79. $\text{MA}(\sigma\text{-centred}) \iff p = \omega$. As we have seen above in Theorem 13.6, $\text{MA}(\sigma\text{-centred})$ implies $p = \omega$. On the other hand, also the converse is true, i.e., $p = \omega$ implies $\text{MA}(\sigma\text{-centred})$. This somewhat surprising result was first proved by Bell [1] (see also Fremlin [4, 14C] or the proof of Theorem 19.4).

80. $\text{MA}(\text{countable}) \iff \text{cov}(\mathcal{M}) = \omega$. Fremlin and Shelah showed in [5] that $\text{MA}(\text{countable})$ is equivalent to $\text{cov}(\mathcal{M}) = \omega$, where $\text{cov}(\mathcal{M})$ denotes the covering number of the meager ideal (defined in Chapter 21). See also Martin and Solovay [8, §4], Blass [2, Theorem 7.13], and Miller [9] for some further results concerning $\text{cov}(\mathcal{M})$.

81. $\text{MA}(\sigma\text{-linked})$. A partially ordered set $(P, \leq)$ is said to be $\sigma\text{-linked}$ if we can write $P = \bigcup_{n \in \omega} P_n$, where each set $P_n$ consists of pairwise compatible elements. On the one hand, it is easily verified that

$$\text{MA} \implies \text{MA}(\sigma\text{-linked}) \implies \text{MA}(\sigma\text{-centred}) \implies \text{MA}(\text{countable}),$$

\end{notes}
but on the other hand, to show that none of the converse implications hold requires quite-sophisticated techniques. For the corresponding references we refer the reader to Fremlin [4, Appendix B3].

References

The Notion of Forcing

In this chapter we present a general technique, called forcing, for extending models of ZFC. The main ingredients to construct such an extension are a model $\mathbf{V}$ of ZFC (e.g., $\mathbf{V} = L$), a partially ordered set $\mathbb{P} = (P, \leq)$ contained in $\mathbf{V}$, as well as a special subset $G$ of $P$ which will not belong to $\mathbf{V}$. The extended model $\mathbf{V}[G]$ will then consist of all sets which can be “described” or “named” in $\mathbf{V}$, where the “naming” depends on the set $G$. The main task will be to prove that $\mathbf{V}[G]$ is a model of ZFC as well as to decide (within $\mathbf{V}$) whether a given statement is true or false in a certain extension $\mathbf{V}[G]$.

To get an idea how this is done, think for a moment that there are people living in $\mathbf{V}$. For these people, $\mathbf{V}$ is the unique set-theoretic universe which contains all sets. Now, the key point is that for any statement, these people are able to compute whether the statement is true or false in a particular extension $\mathbf{V}[G]$, even though they have almost no information about the set $G$ (in fact, they would actually deny the existence of such a set).

The Language of Forcing

The notion of forcing notion. In fact, a forcing notion is just a partially ordered set $\mathbb{P} = (P, \leq)$ with a smallest element, i.e.,

$$\exists p \in P \forall q \in P \, (p \leq q).$$

Notice that this condition implies that $P$ is non-empty. Further notice that we do not require that $\mathbb{P}$ is anti-symmetric (i.e., $p \leq q$ and $q \leq p$ does not necessarily imply $p = q$), even though most of the forcing notions considered in this book are actually anti-symmetric. In fact, for every forcing notion $\mathbb{P}$ there exists an equivalent forcing notion $\hat{\mathbb{P}}$ which is anti-symmetric (see Fact 14.5 below).

In order to make sure that forcing with a forcing notion $\mathbb{P}$ yields a non-trivial extension, we require that a forcing notion $\mathbb{P} = (P, \leq)$ has the property that there are incompatible elements above each $p \in P$, i.e.,
∀p ∈ P 3q_1 ∈ P 3q_2 ∈ P (p ≤ q_1 ∧ p ≤ q_2 ∧ q_1 ⊥ q_2).

Notice that this property implies that there is no maximal element in P, i.e., ∀p ∈ P 3q ∈ P (p < q). Later on, when we shall be somewhat familiar with forcing, the second condition will be tacitly cancelled in order to allow also trivial forcing notions like for example \( P = (\{\emptyset\}, \subseteq) \).

Usually, forcing notions are named after the person who investigated first the corresponding partially ordered set in the context of forcing (e.g., the forcing notion defined below is called Cohen forcing). As in the previous chapter, the elements of P are called “conditions”. Furthermore, if p and q and two conditions and p ≤ q, then we say that p is weaker than q, or equivalently, that q is stronger than p.

Below, we give two quite different examples of forcing notions. The first one is the forcing notion which is used to prove that \( \neg \text{CH} \) is consistent with ZFC, and the second one is a forcing notion which will accompany us — in different forms — throughout this book.

1. Recall that \( \text{Fn}(I, J) \) is the set of all finite partial functions from I to J (defined in the previous chapter). Now, for cardinal numbers \( \kappa > 0 \) define the partially ordered set

\[ C_\kappa = (\text{Fn}(\kappa \times \omega, 2), \subseteq) , \]

i.e., for \( p, q ∈ \text{Fn}(\kappa \times \omega, 2) \), p is stronger than q iff the function p extends q. Obviously, the smallest (i.e., weakest) element of \( \text{Fn}(\kappa \times \omega, 2) \) is \( \emptyset \) (i.e., the empty function), thus, \( C_\kappa \) has a smallest element. Furthermore, for each condition (i.e., function) \( p ∈ \text{Fn}(\kappa \times \omega, 2) \) there is an ordered pair \( (α, n) ∈ \kappa \times \omega \) which does not belong to \( \text{dom}(p) \). Now, let \( q_1 := p ∪ \{⟨(α, n), 1⟩\} \) and \( q_2 := p ∪ \{⟨(α, n), 0⟩\} \). Obviously, \( q_1, q_2 ∈ \text{Fn}(\kappa \times \omega, 2) \), \( q_1 ⊥ q_2 \), and \( q_1 ≥ p ≥ q_2 \). This shows that there are incompatible elements above each \( p ∈ \text{Fn}(\kappa \times \omega, 2) \). Hence, \( C_\kappa \) is a forcing notion. The forcing notion \( C_1 \), denoted \( C \), is called Cohen forcing, and \( C_\kappa \) is in fact just a kind of product of \( \kappa \) copies of Cohen forcing (cf. Chapter 21).

2. A natural example of a partially ordered set is the set of infinite subsets of \( \omega \) together with the superset relation. However, let us consider a slightly different partially ordered set: Define an equivalence relation on \( [\omega]^\omega \) by stipulating

\[ x ∼ y ⇐⇒ x△y \text{ finite} \]

and let \( [\omega]^\omega/\text{fin} := \{[x]^- : x ∈ [\omega]^\omega\} \). On \( [\omega]^\omega/\text{fin} \) we define a partial ordering “\( ≤^* \)” by stipulating

\[ [x]^− ≤ [y]^− ⇐⇒ y ≤^* x , \]

i.e., \( [x]^− ≤ [y]^− \) iff \( y \setminus x \) is finite, and let

\[ U = ([\omega]^\omega/\text{fin}, ≤^* \) . \]
Then \( U \) is a partially ordered. Moreover, \( U \) is a forcing notion: Obviously, the weakest element of \( U \) is \([\omega]^\omega\) (the set of all \( \omega \)-finite subsets of \( \omega \)); thus, \( U \) has a smallest element. Furthermore, for each \( x \in [\omega]^\omega \) one easily finds disjoint sets \( y_1 \) and \( y_2 \) in \([x]^\omega\). This shows that there are incompatible elements above any condition \([x]_U\). This forcing notion — which does not have an established name — we shall call **ultrafilter forcing** (the name is motivated by Proposition 14.18).

**Making names for sets.** Let \( V \) be a model of \( \text{ZFC} \) and let \( P = (P, \leq) \) be a forcing notion which belongs to \( V \), i.e., the set \( P \) as well as the relation “\( \leq \)” (which is a subset of \( P \times P \)) belongs to the model \( V \). The goal is to extend the so-called ground model \( V \), by adding a certain subset \( G \subseteq P \) to \( V \), and then construct a model \( V[G] \) of \( \text{ZFC} \) which contains \( V \). In order to get a proper extension of \( V \), the set \( G \) — even though it is a subset of \( P \) — must not belong to \( V \). However, this seemingly paradoxical property of \( G \) does not affect the construction of the model \( V[G] \).

Roughly speaking, \( V[G] \) consists of all sets which can be constructed from \( G \) by applying set-theoretic processes definable in \( V \). In fact each set in the extension will have a name in \( V \), which tells how it has been constructed from \( G \). We use symbols like \( x \), \( y \), \( f \), \( X \), et cetera for ordinary names, but also \( x \), \( y \), \( G \), et cetera for some special names (e.g., names for sets in \( V \)).

Informally, a name, or more precisely a \( \mathbb{P} \)-name, is a possibly empty set of ordered pairs of the form \((x, p)\), where \( x \) is a \( \mathbb{P} \)-name and \( p \in P \). The class of all \( \mathbb{P} \)-names is denoted by \( V^P \).

Formally, \( V^P \) is defined by transfinite induction (similar to the cumulative hierarchy of sets defined in Chapter 3):

\[
\begin{align*}
V_0^P &= \emptyset \\
V_\alpha^P &= \bigcup_{\beta \in \alpha} V_\beta^P & \text{if } \alpha \text{ is a limit ordinal} \\
V_{\alpha+1}^P &= \mathcal{P}(V_\alpha^P \times P)
\end{align*}
\]

and let

\[
V^P = \bigcup_{\alpha \in \Omega} V_\alpha^P.
\]

Notice that \( V^P \) is a proper subclass of \( V \). The formal definition of \( V^P \) allows to define a rank-function on the class of names: For \( \mathbb{P} \)-names \( x \in V^P \) let

\[
\text{rk}(x) := \bigcup \{ \text{rk}(y) + 1 : \exists p \in P \ (\langle y, p \rangle \in x) \}.
\]

Consider for example the three \( \mathbb{U} \)-conditions \( u_1 = [\omega]^\omega \), \( u_2 = [\{2n : n \in \omega\}]^\omega \), and \( u_3 = [\{3n : n \in \omega\}]^\omega \), as well as the three \( \mathbb{U} \)-names \( x = \{\emptyset, u_2, \langle \emptyset, u_3 \rangle\} \), \( y = \{x, u_2, \langle \emptyset, u_1 \rangle\} \), and \( z = \{\langle y, u_1 \rangle, \langle x, u_2 \rangle, \langle \emptyset, u_2 \rangle, \langle \emptyset, u_3 \rangle, \langle y, u_3 \rangle\} \). Then

\[
\text{rk}(x) = 1, \text{rk}(y) = 2, \text{and } \text{rk}(z) = 3.
\]
Making sets from names. Names are objects in $\mathbf{V}$ intended to designate sets in the extension $\mathbf{V}[G]$ (where $G$ is a certain subset of $P$). In other words, names are special sets in $\mathbf{V}$ which stand for sets in the extension. So, the next step in the construction of $\mathbf{V}[G]$ is to transform the names to the sets they stand for: Let $G$ be a subset of $P$ (later, $G$ will always be a generic filter). Then by transfinite recursion on $\mathbb{P}$-names $x$ we define

$$x[G] = \{ y[G] : \exists q \in G \ (y, q) \in x \},$$

and in general let

$$\mathbf{V}[G] = \{ x[G] : x \in \mathbf{V}^G \}.$$

Notice that if $G = \emptyset$, then $\mathbf{V}[G] = \emptyset$. For example let us consider again the three $\mathbb{U}$-names $x, y, z$, and the three $\mathbb{U}$-conditions $u_1, u_2, u_3$, from above and let $G_1 = \{ u_1 \}, G_{1, 2} = \{ u_1, u_2 \},$ and $G_3 = \{ u_3 \}$. Then $x[G_1] = 0, x[G_{1, 2}] = 1,$ $x[G_3] = 1, y[G_1] = 1, y[G_{1, 2}] = 2, y[G_3] = 0, z[G_1] = \{ 1 \}, z[G_{1, 2}] = 3, z[G_3] = 2$ (recall that $0 = 0, 1 = \{ 0 \}, 2 = \{ 0, 1 \}$, et cetera).

A saucierful of names. Since $\mathbf{V}[G]$ is supposed to be an extension of $\mathbf{V}$, we have to show that $\mathbf{V}$ is in general a subclass of $\mathbf{V}[G]$. Furthermore, $G$ should belong to $\mathbf{V}[G]$, no matter whether $G$ belongs to $\mathbf{V}$ or not.

Firstly, let us show that $\mathbf{V}$ is a subclass of $\mathbf{V}[G]$ whenever $G \subseteq P$ is non-empty. Below, we always assume that $G$ contains $0$ where $0$ denotes the smallest element of $P$. For every set $x \in \mathbf{V}$ there is a canonical name $x \in \mathbf{V}[G]$ such that $x[G] = x$: By transfinite recursion define

$$x = \{ \langle y, 0 \rangle : y \in x \}.$$

For example $\emptyset = \emptyset, 1 = \{ \langle \emptyset, 0 \rangle \}, 2 = \{ \langle \emptyset, 0 \rangle, \{ 1, 0 \} \},$ et cetera. Notice that since $0 \in G$, for all $x \in \mathbf{V}$ we have $x[G] = \{ y[G] : y \in x \}$. It remains to show that for each $x \in \mathbf{V}$ we have $x[G] = x$.

**FACT 14.1.** *If $G \subseteq P$ with $0 \in G$, then for every $x \in \mathbf{V}$ we have $x[G] = x$.*

**Proof.** The proof is by transfinite induction on $\rk(x)$. If $\rk(x) = 0$, then $x = \emptyset \in \emptyset$, and

$$\emptyset[G] = \{ y[G] : y \in \emptyset \} = \emptyset.$$

Now let $\rk(x) = \alpha$ and assume that $y[G] = y$ for all $\mathbb{P}$-names $y$ with $\rk(y) \in \alpha$. Then

$$x[G] = \{ y[G] : y \in x \} = \{ y : y \in x \} = x$$

which completes the proof.

In order to make sure that $G$ belongs to $\mathbf{V}[G]$, we need a $\mathbb{P}$-name $G$ for $G$ such that $G[G] = G$. For example define

$$G = \{ \langle p, p \rangle : p \in P \}.$$

As an immediate consequence of Fact 14.1 we get the following
The language of forcing

Fact 14.2. For every $G \subseteq P$ which contains $0$ we have $G[G] = G$.

Proof. We just have to evaluate the $P$-name $G$:

$G[G] = \{p[G] : \exists q \in G \ (q, p) \in G\} = \{p[G] : p \in G\} = \{p : p \in G\} = G$

Hence, for any subset $G \subseteq P$ we have $G = G[G]$. Thus, the name $G$, usually denoted $G$, is the canonical name for $G$. Furthermore, we see that $G \in V[G]$, no matter whether $G$ — belonging to some set-theoretic universe — belongs to $V$.

We can also define names for unordered and ordered pairs of sets: For $P$-names $x$ and $y$ define

$$\text{up}(x, y) = \{\langle x, 0 \rangle, \langle y, 0 \rangle\}$$

and

$$\text{op}(x, y) = \{\{\langle x, 0 \rangle, 0\}, \{\langle x, 0 \rangle, \langle y, 0 \rangle, 0\}\}.$$ We leave it as an exercise to the reader to verify that for every $G \subseteq P$ with $0 \in G$ we have $\text{up}(x, y)[G] = \{x[G], y[G]\}$ and $\text{op}(x, y)[G] = \langle x[G], y[G]\rangle$.

The forcing language. We are now ready to introduce a kind of logical language, the so-called forcing language. A sentences $\psi$ of the forcing language is like a first-order sentence, except that the parameters appearing in $\psi$ are some names in $V^P$, i.e., specific sets in $V$. Sentences of the forcing language use the names in $V^P$ to assert something about $V[G]$ (for certain $G \subseteq P$). The people living in the ground model $V$ may not know whether a given sentence $\psi$ is true in $V[G]$. The truth or falsity of $\psi$ in $V[G]$ will in general depend on the set $G \subseteq P$. For example consider the $\mathbb{U}$-name $x = \{\langle 0, p_0 \rangle\}$ with $p_0 = \{\{2n : n \in \omega\}\}$, and the sentence $\psi \equiv \exists y (y \in x)$ of the forcing language which asserts that $x$ is non-empty. Now, $\psi$ is true in $V[G]$ if and only if $V[G] \models \exists y (y \in x[G])$, which is the case if and only if $p_0 \in G$. Hence, depending on $G \subseteq [\omega]^\omega$, $\psi$ becomes true or false in $V[G]$.

However, even though people living in $V$ do not know whether $V[G] \models \psi$, they know that $V[G] \models \psi$ iff $p_0 \in G$. Thus, in order to decide whether $V[G] \models \psi$ they just need to know whether $G$ contains the condition $p_0$.

This leads to one of the key features of forcing: By knowing whether a certain condition $p$ belongs to $G \subseteq P$, people living in $V$ can figure out whether a given sentence of the forcing language is true or false in $V[G]$. Moreover, it will turn out that people living in $V$ are able to verify that in certain models $V[G]$ all axioms of ZFC remain true. In the following section we shall see how this is done.
Generic Extensions

Let again \( P = (P, \leq) \) be an arbitrary forcing notion which belongs to a model \( V \) of ZFC. Below, we define first the notion of a generic filter (which is a special subset \( G \subseteq P \)) and the corresponding generic model \( V[G] \); then we introduce the forcing relation and show how people in \( V \) can decide whether a given sentence is true or false in a particular generic model. Finally we construct a generic model in which the Continuum Hypothesis fails and discuss the existence of generic filters.

**Generic filters and generic models.** Let us briefly recall some definitions from the previous chapter: A set \( D \subseteq P \) is open dense if \( p \in D \) and \( q \geq p \) implies \( q \in D \) (open), and if for every \( p \in P \) there is a \( q \in D \) such that \( q \geq p \) (dense). A set \( A \subseteq P \) is an anti-chain in \( P \) if any two distinct elements of \( A \) are incompatible, and it is maximal if it is not properly contained in any anti-chain in \( P \). A non-empty set \( G \subseteq P \) is a filter (on \( P \)) if \( p \in G \) and \( q \leq p \) implies \( q \in G \) (downwards closed), and if for any \( p_1, p_2 \in G \) there is a \( q \in G \) such that \( p_1 \leq q \geq p_2 \) (directed).

Now, a filter \( G \subseteq P \) is said to be \( P \)-generic over \( V \) if \( G \cap D \neq \emptyset \) for every open dense set \( D \subseteq P \) which belongs to \( V \) (compare with the notion of a \( \mathcal{P} \)-generic filter, which was introduced in the previous chapter). In other words, a filter \( G \) on \( P \) is \( P \)-generic over \( V \) if it meets every open dense subset of \( P \) which belongs to \( V \). The restriction that the open dense subsets have to belong to \( V \) — which at a first glance seems to be superficial — is in fact crucial.

**Equivalent forcing notions.** It may happen that two different forcing notions \( P = (P, \leq_P) \) and \( Q = (Q, \leq_Q) \) yield the same generic models, in which case we say that \( P \) and \( Q \) are equivalent, denoted \( P \cong Q \). More precisely, \( P \cong Q \) if for every \( G \subseteq P \) which is \( P \)-generic over \( V \), there exists an \( H \subseteq Q \) which is \( Q \)-generic over \( V \) such that \( V[G] = V[H] \), and vice versa, for every \( Q \)-generic \( H \) there is a \( P \)-generic \( G \) such that \( V[H] = V[G] \). Notice that “\( \cong \)” is indeed an equivalence relation on the class of forcing notions.

In order to prove that two forcing notions \( P = (P, \leq_P) \) and \( Q = (Q, \leq_Q) \) are equivalent, it is sufficient to show the existence of a so-called dense embedding from \( P \) to \( Q \) (or vice versa), where a function \( h : P \to Q \) is called a dense embedding if it satisfies the following conditions:

- \( \forall p_0, p_1 \in P \ (p_0 \leq_P p_1 \leftrightarrow h(p_0) \leq_Q h(p_1)) \)
- \( \forall q \in Q \exists p \in P \ (q \leq_Q h(p)) \)

Notice that the function \( h \) is not necessarily surjective, in particular, \( h \) is in general not an isomorphism. However, it is not hard to verify that the forcing notions \( P \) and \( Q \) are equivalent whenever there exists a dense embedding \( h : P \to Q \). The proof of the following fact is left to the reader.
Generic extensions

**Fact 14.3.** Let $\mathbb{P} = (P, \leq)$ and $\mathbb{Q} = (Q, \leq)$ be any forcing notions. If there exists a dense embedding $h : P \to Q$, then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. In fact, if $G \subseteq P$ is $\mathbb{P}$-generic over $V$, then the set

$$H = \{q \in Q : \exists p \in G \ (q \leq h(p))\}$$

is $\mathbb{Q}$-generic over $V$ and $V[H] = V[G]$. Conversely, if a set $H \subseteq Q$ is $\mathbb{Q}$-generic over $V$, then the set

$$G = \{p \in P : h(p) \in H\}$$

is $\mathbb{P}$-generic over $V$ and $V[H] = V[G]$.

The preceding fact implies that it is enough to consider forcing notions of the form $(\kappa, \leq, \emptyset)$, where $\kappa$ is a cardinal number, “$\leq$” is a partial ordering on $\kappa$, and $\emptyset$ is the smallest element (with respect to $\leq$) in $\kappa$. More precisely, we get the following

**Fact 14.4.** Every forcing notion $\mathbb{P} = (P, \leq, \emptyset)$, where $\emptyset$ is a smallest element in $P$, is equivalent to some forcing notion $(\kappa, \leq, \emptyset)$, where $\kappa = |P|$. In particular, we may always identify the smallest element of a forcing notion with the empty set.

**Proof.** Let $h : P \to \kappa$ be a bijection, where $h(\emptyset) = \emptyset$, and let

$$h(p) \leq h(q) \iff p \leq q.$$

Then $h$ is obviously a dense embedding. \(-\)

As another consequence of Fact 14.3 we get that every forcing notion is equivalent to some anti-symmetric forcing notion.

**Fact 14.5.** Let $\mathbb{P} = (P, \leq)$ be any forcing notion and let $\hat{\mathbb{P}} := (\hat{P}, \leq^\ast)$, where $p \sim q \iff p \leq q \land q \leq p$. $\hat{P} = \{\hat{p}^- : p \in P\}$, and $[\hat{p}]^\ast \leq^\ast [\hat{q}]^\ast \iff p \leq q$. Then $\hat{\mathbb{P}}$ is anti-symmetric and equivalent to $\mathbb{P}$.

**Proof.** Firstly notice that $\hat{\mathbb{P}}$ is a forcing notion. Now define $h : P \to \hat{P}$ by stipulating $h(p) := [\hat{p}]^\ast$. Then $h$ is obviously a dense embedding and therefore $\hat{\mathbb{P}} \simeq \mathbb{P}$. Finally, if we have $[\hat{p}]^\ast \leq^\ast [\hat{q}]^\ast$ and $[\hat{q}]^\ast \leq^\ast [\hat{p}]^\ast$, then $[\hat{p}]^\ast = [\hat{q}]^\ast$, which shows that $\hat{\mathbb{P}}$ is anti-symmetric. \(-\)

**Alternative definitions of generic filters.** It is sometimes useful to have a few alternative definitions of $\mathbb{P}$-generic filters at hand which are sometimes easier to apply.

**Fact 14.6.** Let $\mathbb{P} = (P, \leq)$ be a forcing notion which belongs to a model $V$ of ZFC. Then, for a filter $G$ on $P$, the following statements are equivalent:

(a) $G$ is $\mathbb{P}$-generic over $V$.
(b) $G$ meets every maximal anti-chain in $P$ which belongs to $V$.
(c) $G$ meets every dense subset of $P$ which belongs to $V$. 


**Proof.** (a)⇒(b) Let \( A \subseteq P \) be a maximal anti-chain in \( P \) which belongs to \( V \). Then \( D_A := \{ p \in P : \exists q \in A (p \geq q) \} \) is open dense in \( P \): \( D_A \) is obviously open, and since \( A \) is a maximal anti-chain in \( P \), for every \( p_0 \in P \) there is a condition \( q_0 \in A \) such that \( p_0 \) and \( q_0 \) are compatible, i.e., there is a \( p \in D_A \) such that \( q_0 \leq p \geq p_0 \), which implies that \( D_A \) is dense. Now, if \( G \) is \( \mathbb{P} \)-generic over \( V \), then \( G \) meets \( D_A \), and since \( G \) is downwards closed, it meets the maximal anti-chain \( A \).

(b)⇒(c) Let \( D \subseteq P \) be a dense subset of \( P \) which belongs to \( V \). Then by Kurepa’s Principle (introduced in Chapter 5) there is a maximal anti-chain \( A \) in \( D \). Since \( D \) is dense in \( P \), \( A \) is also a maximal anti-chain in \( P \) (otherwise, there would be a condition \( p \in P \) which is incompatible with all conditions of \( D \), contradicting the fact that \( D \) is dense in \( P \)). Now, if \( G \) meets every maximal anti-chain in \( P \) (which belongs to \( V \)), then \( G \) meets \( A \), and since \( A \) is a subset of \( D \), it meets the dense set \( D \).

(c)⇒(a) If \( G \) meets every dense subset of \( P \) which belongs to \( V \), then it obviously meets also every open dense subset of \( P \) which belongs to \( V \). ⊣

Let \( p \in P \); then a set \( D \subseteq P \) is **dense above** \( p \) if for any \( p' \geq p \) there is a \( q \in D \) such that \( q \geq p' \). Notice that if \( D \subseteq P \) is dense above \( p \) (for some \( p \in P \)) and \( q \geq p \), then \( D \) is also dense above \( q \).

The proof of the following characterisation of \( \mathbb{P} \)-generic filters is left to the reader.

**Fact 14.7.** Let \( \mathbb{P} = (P, \leq) \) be a forcing notion which belongs to a model \( V \) of ZFC, and let \( G \subseteq P \) be a filter on \( P \) which contains the condition \( p \). Then \( G \) is \( \mathbb{P} \)-generic over \( V \) if and only if \( G \) meets every set \( D \subseteq P \) which is dense above \( p \).

If the filter \( G \subseteq P \) is \( \mathbb{P} \)-generic over \( V \), then the class \( V[G] \) is called a **generic extension** of \( V \), or just a **generic model**.

---

**ZFC in Generic Models**

In order to prove that a generic model \( V[G] \) is indeed a model of ZFC, we first have to develop a technique which allows us to verify within \( V \) that all axioms of ZFC remain true in \( V[G] \).

**The forcing relationship.** In this section, we shall define a relationship, denoted \( \forces_p \), between conditions \( p \in P \) and sentences \( \psi \) of the forcing language. Even though the relationship \( \forces_p \) involves formulae and is therefore not expressible in the language of First-Order Logic, we write \( p \forces \psi \) ("\( p \) forces \( \psi \)") to mean that if \( G \) is \( \mathbb{P} \)-generic over \( V \) and contains \( p \), then \( \psi \) is true in \( V[G] \), where we tacitly assume that for every \( p \in P \) there is a \( \mathbb{P} \)-generic filter over \( V \) which contains \( p \). Surprisingly, the definition of the relationship \( \forces_p \) takes place in the model \( V \) without actually knowing any \( \mathbb{P} \)-generic filter.
**Definition 14.8.** Let \( p_0 \in P \) be a condition, let \( \psi(x_1, \ldots, x_n) \) be a first-order formula with all free variables shown, and let \( x_1, \ldots, x_n \in V^P \) be any \( P \)-names. The relationship \( p_0 \mid_\psi \psi(x_1, \ldots, x_n) \) is essentially defined by induction on the complexity of \( \psi \). However, for atomic formulae \( \psi \) we have to use a double induction on the ranks of the names that are substituted for the variables in \( \psi \):

(a) \( p_0 \mid_\psi \exists x \in \emptyset \) if and only if

\[
\{ q \geq p_0 : q \geq s_1 \rightarrow \exists (y_2, s_2) \in \emptyset (q \geq s_2 \land q \mid_\psi y_2 = y_2) \}
\]

is dense above \( p_0 \), and

(\( \beta \)) for all \( (y_2, s_2) \in \emptyset \), the set

\[
\{ q \geq p_0 : q \geq s_2 \rightarrow \exists (y_1, s_1) \in \emptyset (q \geq s_1 \land q \mid_\psi y_1 = y_2) \}
\]

is dense above \( p_0 \).

(b) \( p_0 \mid_\psi \emptyset \in \emptyset \) if and only if the set

\[
\{ q \geq p_0 : \exists (y, s') \in \emptyset (q \geq s \land q \mid_\psi y = \emptyset) \}
\]

is dense above \( p_0 \).

(c) \( p_0 \mid_\psi \neg \varphi(x_1, \ldots, x_n) \) if and only if for all \( q \geq p_0 \) we have

\[
q \mid_\psi \varphi(x_1, \ldots, x_n),
\]

i.e., for no \( q \geq p_0 \) we have \( q \mid_\psi \varphi(x_1, \ldots, x_n) \).

(d) \( p_0 \mid_\psi \varphi_1(x_1, \ldots, x_n) \land \varphi_2(x_1, \ldots, x_n) \) if and only if

\[
p_0 \mid_\psi \varphi_1(x_1, \ldots, x_n) \quad \text{and} \quad p_0 \mid_\psi \varphi_2(x_1, \ldots, x_n).
\]

(e) \( p_0 \mid_\psi \exists z \varphi(z, x_1, \ldots, x_n) \) if and only if the set

\[
\{ q \geq p_0 : \exists z \in V^P (q \mid_\psi \varphi(z, x_1, \ldots, x_n)) \}
\]

is dense above \( p_0 \).

As an immediate consequence of **Definition 14.8** we get the following

**Fact 14.9.** For any sentence \( \psi \) of the forcing language we have:

(a) If \( p \mid_\psi \psi \) and \( q \geq p \), then \( q \mid_\psi \psi \).

(b) The set \( \Delta_\psi := \{ p \in P : (p \mid_\psi \psi) \lor (p \mid_\psi \neg \psi) \} \) is open dense in \( P \).
Proof. Part (a) is obvious. For (b) notice that for every \( p \in P \), either there is a \( q \geq p \) such that \( q \Vdash_p \psi \), or for all \( q \geq p \) we have \( q \Vdash_p \neg \psi \). In the former case, \( q \in \Delta \psi \), and in the latter case we get \( p \Vdash_p \neg \psi \) and consequently \( p \in \Delta \psi \). \( \dashv \)

Until now, we did not prove that the forcing relationship is doing what we want, e.g., \( p \Vdash_p \psi \) should imply \( p \Vdash_p \neg \psi \). However, this follows implicitly from the proof of the Forcing Theorem 14.10, which is the core result of forcing.

The Forcing Theorem. In order to prove that ZFC holds in every generic extension of any model \( V \) of ZFC, we need a tool which allows us to decide within \( V \) whether a given first-order formula is true or false in a certain generic model. The following theorem is the required tool.

Theorem 14.10 (Forcing Theorem). Let \( \psi(x_1, \ldots, x_n) \) be a first-order formula with all free variables shown, i.e., \( \text{free}(\psi) \subseteq \{x_1, \ldots, x_n\} \). Let \( V \) be a model of ZFC, let \( \mathbb{P} = (P, \leq) \) be a forcing notion which belongs to \( V \), let \( \bar{x}_1, \ldots, \bar{x}_n \in V^\mathbb{P} \) be any \( \mathbb{P} \)-names, and let \( G \subseteq P \) be \( \mathbb{P} \)-generic over \( V \).

1. If \( p \in G \) and \( p \Vdash_{\mathbb{P}} \psi(\bar{x}_1, \ldots, \bar{x}_n) \), then \( V[G] \models \psi(\bar{x}_1[G], \ldots, \bar{x}_n[G]) \).
2. If \( V[G] \models \psi(\bar{x}_1[G], \ldots, \bar{x}_n[G]) \), then \( \exists p \in G \{ p \Vdash_{\mathbb{P}} \psi(\bar{x}_1, \ldots, \bar{x}_n) \} \).

Proof. The proof is by induction on the complexity of \( \psi(\bar{x}_1, \ldots, \bar{x}_n) \). So, we first prove (1) and (2) for atomic formulae \( \psi \).

\( \psi(\bar{x}_1, \bar{x}_2) \equiv (\bar{x}_1 = \bar{x}_2) \): When \( \psi(\bar{x}_1, \bar{x}_2) \) is \( \bar{x}_1 = \bar{x}_2 \), the proof is by transfinite induction on \( \text{rk}(\bar{x}_1, \bar{x}_2) := \max\{\text{rk}(\bar{x}_1), \text{rk}(\bar{x}_2)\} \), using clause (a) of Definition 14.8: If \( \text{rk}(\bar{x}_1, \bar{x}_2) = 0 \), then \( \bar{x}_1 = \bar{x}_2 = \emptyset \). Now, \( \emptyset \models G = \emptyset \), which implies (1), and for all \( p \in P \) we have \( p \Vdash_{\mathbb{P}} \emptyset = \emptyset \), which implies (2). For \( \text{rk}(\bar{x}_1, \bar{x}_2) > 0 \) we shall check (1) and (2) separately.

1. Assume that \( p \in G \) and \( p \Vdash_{\mathbb{P}} \bar{x}_1 = \bar{x}_2 \), and that (1) holds for all names \( y_1, y_2 \) with \( \text{rk}(y_1, y_2) < \text{rk}(\bar{x}_1, \bar{x}_2) \). We show \( \bar{x}_1[G] = \bar{x}_2[G] \) by proving that \( x_1[G] \subseteq x_2[G] \) using (a) of Definition 14.8(a); the proof of \( x_2[G] \subseteq x_1[G] \) using (b) is the same. Every element of \( x_1[G] \) is of the form \( y_1[G] \), where \( (y_1, s_1) \in x_1[G] \) for some \( s_1 \in G \). We must show that \( y_1[G] \in x_2[G] \). Since \( G \) is directed, there is an \( \bar{r} \in G \) with \( s_1 \leq r \geq p \). By Fact 14.9(a), \( r \Vdash_{\mathbb{P}} \bar{x}_1 = \bar{x}_2 \), and by Definition 14.8(a),(a) and Fact 14.7, there is a \( q \in G \) such that \( q \geq r \) (in particular \( q \geq s_1 \)) and

\[ \exists (y_2, s_2) \in x_2[q \geq s_2 \land q \Vdash_{\mathbb{P}} y_1 = y_2]. \]  

Fix \( (y_2, s_2) \in x_2 \) as in (3), then \( \text{rk}(y_1, y_2) < \text{rk}(\bar{x}_1, \bar{x}_2) \) and by our assumption we get \( y_1[G] = y_2[G] \). Further, since \( q \geq s_2 \) and \( G \) is downwards closed we have \( s_2 \in G \) which implies \( y_2[G] \in x_2[G] \); consequently we get \( y_1[G] \in x_2[G] \).

2. To check (2), assume \( \bar{x}_1[G] = \bar{x}_2[G] \), and that (2) holds for all names \( y_1, y_2 \) with \( \text{rk}(y_1, y_2) < \text{rk}(\bar{x}_1, \bar{x}_2) \). Let \( D_{\bar{x}_1, \bar{x}_2} \subseteq P \) be the set of all conditions \( r \in P \) such that \( r \Vdash_{\mathbb{P}} \bar{x}_1 = \bar{x}_2 \), or we are at least in one of the following two cases:
(α') there exists a name \( \langle y_1, s_1 \rangle \in x_1 \) such that \( r \geq s_1 \) and
\[
\forall \langle y_2, s_2 \rangle \in x_2 \forall q \in P'(q \geq s_2 \land q \forces x y_1 = y_2) \Rightarrow q \perp r,
\]
(β') there exists a name \( \langle y_2, s_2 \rangle \in x_2 \) such that \( r \geq s_1 \) and
\[
\forall \langle y_1, s_1 \rangle \in x_1 \forall q \in P'(q \geq s_1 \land q \forces x y_1 = y_2) \Rightarrow q \perp r.
\]
First we show that no condition \( r \in G \) can satisfy (α') or (β'): Indeed, if \( r \in G \) and \( \langle y_1, s_1 \rangle \in x_1 \) as in (α'), then \( s_1 \in G \) and therefore \( y_1[G] \in x_1[G] = x_2[G] \) (by our assumption). Now, fix \( \langle y_2, s_2 \rangle \in x_2 \) with \( s_2 \in G \) and \( y_1[G] = y_2[G] \). Since \( \text{rk}(y_1, y_2) < \text{rk}(x_1, x_2) \) there is a condition \( q_0 \in G \) such that \( q_0 \forces y_1 = y_2 \), and since \( G \) is directed there is a \( q \in G \) such that \( q_0 \leq q \geq s_2 \). By Fact 14.9(a) we have \( q \forces y_1 = y_2 \), and hence by (α') we get \( q \perp r \), which contradicts the fact that \( G \) is directed.

If there is no \( r \in G \) such that \( r \forces y_1 = y_2 \), then \( D_{1, 2} \cap G = \emptyset \). We would be done if we could show that \( D_{1, 2} \) is dense in \( P \) since this would contradict the fact that \( G \) meets every dense set in \( V \): Fix an arbitrary condition \( p \in P \). Either \( p \forces y_1 = y_2 \), or otherwise, (α) or (β) of Definition 14.8(a) fails. If (α) fails, then there are \( \langle y_1, s_1 \rangle \in x_1 \) and \( r \geq p \) such that \( r \geq s_1 \) and for all \( q \geq r \) we have:
\[
\forall \langle y_2, s_2 \rangle \in x_2 (\neg (q \forces y_1 = y_2) \land q \geq s_2)
\]
If \( \langle y_2, s_2 \rangle \in x_2 \), \( q \geq s_2 \), and \( q \forces y_1 = y_2 \), then \( q \perp r \), since a common extension \( q' \) of \( q \) and \( r \) would contradict (γ). Thus, \( r \geq p \) and \( r \) satisfies (α'), in particular \( r \in D_{1, 2} \). Likewise, if (β) fails then there is a condition \( r \geq p \) which satisfies (β').

\( \psi(x_1, x_2) \equiv (x_1 \in x_2) \): When \( \psi(x_1, x_2) \) is \( x_1 \in x_2 \) we check again (1) and (2) separately.

(1): Assume that there is a condition \( p \in G \) such that \( p \forces y_1 = x_2 \). Then, by Definition 14.8(b), the set
\[
D_p = \{ q \in P : \exists (y, s) \in x_2(q \geq s \land q \forces p (y = x_1) \}
\]
is dense above \( p \). Fix a condition \( q \in G \cap D_p \) and a \( \dot{P} \)-name \( \langle y, s \rangle \in x_2 \) such that \( q \geq s \) and \( q \forces p y = x_1 \). Since \( s \in G \) and \( \langle y, s \rangle \in x_2 \) we get \( y[G] \in x_2[G] \), and since \( q \in G \) and \( q \forces p y = x_1 \), by (1) applied to \( y = x_1 \) we also get \( y[G] = x_1[G] \). Thus, we have \( y[G] \in x_2[G] \) as well as \( y[G] = x_1[G] \), which obviously implies that \( x_1[G] \in x_2[G] \).

(2): Assume now \( x_1[G] \in x_2[G] \). By definition of \( x_2[G] \) there is a name \( \langle y, s \rangle \in x_2 \) such that \( s \in G \) and \( y[G] = x_1[G] \). By (1) for \( y[G] = x_1[G] \), there is an \( r \in G \) such that
\[
r \forces y = x_1.
\]
Finally, let \( p \in G \) be such that \( s \leq p \geq r \). Then
\[
\forall q \geq p (q \geq s \land q \forces y \in x_2),
\]
and consequently \( p \forces x_1 \in x_2 \).
This concludes the proof of (1) and (2) for atomic formulae. The proofs for non-atomic formulae are much easier than the preceding proofs, but even though it is enough to prove (1) and (2) for formulae $\psi$ of the form $\neg \varphi$, $\varphi_1 \land \varphi_2$, and $\exists x \varphi(x)$, there are still six cases to be checked.

$\psi(x_1, \ldots, x_n) \equiv \neg \varphi$: Let $\psi(x_1, \ldots, x_n)$ be a negated formula, i.e., of the form $\neg \varphi$ for some formula $\varphi$.

1. We assume (2) for $\varphi$ and conclude (1) for $\neg \varphi$. Assume $p \in G$ and $p \Vdash \neg \varphi$. We have to show that $V[G] \Vdash \neg \varphi$. If $V[G] \Vdash \varphi$, then by (2) for $\varphi$ there is a $q \in G$ such that $q \Vdash \neg \varphi$. Since $G$ is directed, there is an $r \in G$ such that $q \leq r \geq p$ and by Fact 14.9(a) we would have $r \Vdash \varphi$, contradicting the definition of $p \Vdash \neg \varphi$.

2. We assume (1) for $\varphi$ and conclude (2) for $\neg \varphi$. Assume that $V[G] \Vdash \neg \varphi$. We have to show that there is a condition $p \in G$ such that $p \Vdash \neg \varphi$. Consider the set $\Delta_\varphi := \{ r \in P : (r \Vdash \varphi) \lor (r \Vdash \neg \varphi) \}$. By Fact 14.9(b), $\Delta_\varphi$ is open dense in $P$ and therefore $\Delta_\varphi \cap G \neq \emptyset$. Fix a condition $p \in \Delta_\varphi \cap G$. Then if $p \Vdash \varphi$, then we are done; and if $p \Vdash \neg \varphi$, then by (1) for $\varphi$ we have $V[G] \Vdash \varphi$, a contradiction.

$\psi(x_1, \ldots, x_n) \equiv \varphi_1 \land \varphi_2$: Let $\psi(x_1, \ldots, x_n)$ be of the form $\varphi_1 \land \varphi_2$ for some formulae $\varphi_1$ and $\varphi_2$.

1. We assume (1) for $\varphi_1$ and $\varphi_2$ and conclude (1) for $\varphi_1 \land \varphi_2$. Assume $p \in G$ and $p \Vdash \varphi_1 \land \varphi_2$. Then $p \Vdash \varphi_1$ and $p \Vdash \varphi_2$, hence, by (1) for $\varphi_1$ and $\varphi_2$ we have $V[G] \Vdash \varphi_1$ and $V[G] \Vdash \varphi_2$ which implies $V[G] \Vdash \varphi_1 \land \varphi_2$.

2. We assume (2) for $\varphi_1$ and $\varphi_2$ and conclude (2) for $\varphi_1 \land \varphi_2$. Assume $V[G] \Vdash \varphi_1 \land \varphi_2$. By (2) for $\varphi_1$ and $\varphi_2$ there are $p_1, p_2 \in G$ such that $p_1 \Vdash \varphi_1$ and $p_2 \Vdash \varphi_2$. Let $r \in G$ be such that $p_1 \leq r \geq p_2$. Then $r \Vdash \varphi_1$ and $r \Vdash \varphi_2$, hence, $r \Vdash \varphi_1 \land \varphi_2$.

$\psi(x_1, \ldots, x_n) \equiv \exists x \varphi(x)$: Let $\psi(x_1, \ldots, x_n)$ be an existential formula of the form $\exists x \varphi(x)$ for some formula $\varphi$.

1. We assume (1) for $\varphi(x)$ and conclude (1) for $\exists x \varphi(x)$. Assume $p \in G$ and $p \Vdash \exists x \varphi(x)$. Then the set

$$\{ r \in P : \exists x (r \Vdash \varphi(x)) \}$$

is dense above $p$. So, we find a $q \in G$ and a $P$-name $\dot{x}_0 \in V^P$ such that $q \Vdash \varphi(\dot{x}_0)$. By (1) for $\varphi(\dot{x}_0)$ we get $V[G] \Vdash \varphi(\dot{x}_0[G])$, and therefore $V[G] \Vdash \exists x \varphi(x)$.

2. We assume (2) for $\varphi(x) \in V[G]$ and conclude (2) for $\exists x \varphi(x)$. Assume $V[G] \Vdash \exists x \varphi(x)$. Then there exists an $x_0 \in V[G]$ such that $V[G] \Vdash \varphi(x_0)$ and let $\dot{x}_0$ be such that $V[G] \Vdash \dot{x}_0[G] = x_0$. By (2) for $\varphi(x_0[G])$ there is a $p \in G$ such that $p \Vdash \varphi(x_0)$. Then for all $r \geq p$ we have $r \Vdash \varphi(\dot{x}_0)$, which implies that $p \Vdash \exists x \varphi(x)$. $

One might be tempted to prove the following result (which is to some extent the converse of the Forcing Theorem 14.10): If for all $P$-generic filters

$\dot{x}$
\[ G \subseteq P \text{ containing a certain } \mathbb{P}\text{-condition } p \text{ we have } V[G] \models \psi \text{ (for a given sentence } \psi), \text{ then } p \Vdash \neg \psi. \text{ For the proof we notice first that } p \Vdash \neg \psi \text{ would imply that there exists a condition } q > p \text{ such that } q \Vdash \neg \psi. \text{ Now, if we could show that there exists a } \mathbb{P}\text{-generic filter } G \text{ containing } q \text{ we would have } V[G] \models \neg \psi, \text{ which contradicts our assumption. However, as we shall see below, the existence of a } \mathbb{P}\text{-generic filter } G \text{ (no matter if it contains } q \text{ or not) cannot be proved within } ZFC.\]

However, assume for the moment — as we shall later always do — that for any condition } q \text{ there exists a generic filter containing } q. \text{ As an application of the Forcing Theorem 14.10 we prove the following lemma, which is one of the standard results about forcing.}

**Lemma 14.11.** Let } \mathbb{P} = (P, \leq) \text{ be a forcing notion, let } G \text{ be } \mathbb{P}\text{-generic over } V, \text{ and let } p \in G.\]

(a) If } p \Vdash q \in y, \text{ then there exist a } \mathbb{P}\text{-name } \dot{x} \text{ with } \text{rk}(\dot{x}) < \text{rk}(y) \text{ and a } \mathbb{P}\text{-condition } q \geq p \text{ in } G \text{ such that } q \Vdash \dot{x} = \dot{x}.\]

(b) If } p \Vdash f \in B \wedge \dot{x}_0 \in A, \text{ then there is a } \mathbb{P}\text{-name } \langle y, r \rangle \in B \text{ with } r \in G \text{ and a condition } q \geq p \text{ in } G \text{ such that } q \Vdash f(\dot{x}_0) = y.\]

**Proof.** (a) Since } p \in G, \text{ } V[G] \models \dot{z}[G] \in y[G], \text{ and since } y[G] = \{ \dot{x}[G] : \dot{x} \in y \}, \text{ there is a name } \langle \dot{x}_0, r \rangle \in y \text{ with } r \in G \text{ such that } x_0[G] = \dot{z}[G]. \text{ In particular, } \text{rk}(\dot{x}_0) < \text{rk}(y). \text{ Now, since } V[G] \models \dot{x}_0[G] = \dot{z}[G], \text{ there is a condition } p' \in G \text{ such that } p' \Vdash \dot{z} = \dot{x}_0. \text{ Further, since } G \text{ is directed, there is a } q \in G \text{ such that } p \leq q \geq p'. \text{ Thus, } q \Vdash f(\dot{x}_0) = y.\]

(b) Since } p \in G, \text{ there is a set } z \in V[G] \text{ such that } V[G] \models z \in B[G] \wedge \langle \dot{x}_0[G], z \rangle \in f[G].\]

Let } \dot{z} \text{ be a } \mathbb{P}\text{-name in } V \text{ for } z \text{ (i.e., } \dot{z}[G] = z). \text{ By the proof of } (a) \text{ there is a } \mathbb{P}\text{-name } \langle y, r \rangle \in B \text{ with } r \in G \text{ and a } p' \in G \text{ such that } p' \Vdash \dot{z} = \dot{y} \wedge \dot{y} \in B. \text{ Since } G \text{ is directed, there is a } q \in G \text{ such that } p \leq q \geq p'. \text{ Thus, we have } q \Vdash \text{otp}(\dot{x}_0, y) \in f, \text{ or in other words, } q \Vdash f(\dot{x}_0) = y.\]

**The Generic Model Theorem.** With the Forcing Theorem 14.10 we would now be able to prove that generic extensions of models of ZFC are also models of ZFC (however, we omit most of the quite tedious proof).

**Theorem 14.12 (Generic Model Theorem).** Let } V \text{ be a transitive standard model of ZFC (i.e., a transitive model with the standard membership relation), let } \mathbb{P} = (P, \leq) \text{ be a forcing notion which belongs to } V, \text{ and let } G \subseteq P \text{ be } \mathbb{P}\text{-generic over } V. \text{ Then } V[G] \models ZFC. \text{ Moreover, the class } V \text{ is a subclass of } V[G], \text{ } G \in V[G], \text{ and every transitive standard model of ZFC containing } V \text{ as a subclass and } G \text{ as an element also contains } V[G] \text{ (i.e., } V[G] \text{ is the smallest standard model of ZFC containing } V \text{ as a subclass and } G \text{ as a set). Furthermore, } \Omega^V[G] = \Omega^V, \text{ i.e., every ordinal in } V[G] \text{ belongs to } V, \text{ and vice versa.}
Instead of the full Generic Model Theorem, let us just prove the following four partial results.

**Fact 14.13.** If \( V \models \text{ZFC} \) and \( G \) is \( \mathbb{P} \)-generic over \( V \), then \( V[G] \) satisfies the Axiom of Pairing.

**Proof.** Let \( G \) be an arbitrary \( \mathbb{P} \)-generic filter and let \( x \) and \( y \) be \( \mathbb{P} \)-names for some sets \( x \) and \( y \) in \( V[G] \) (i.e., \( x[G] = x \) and \( y[G] = y \) respectively). Because \( G \) is downwards closed we have \( 0 \in G \) and therefore we get

\[
up(x, y|G) = \{ x[G], y[G] \} = \{ x, y \}.
\]

Thus, if \( x \) and \( y \) belong to \( V[G] \), then also \( \{ x, y \} \) belongs to \( V[G] \). \( \dashv \)

**Proposition 14.14.** If \( V \models \text{ZFC} \) and \( G \) is \( \mathbb{P} \)-generic over \( V \), then \( V[G] \models \text{AC} \).

**Proof.** Let \( x \in V[G] \) be an arbitrary set. Since the Well-Ordering Principle implies AC, it is enough to prove that in \( V[G] \) there exists an injective function from \( x \) into \( \Omega \) (notice that the empty function is injective). Let \( x \) be a \( \mathbb{P} \)-name in \( V \) for \( x \) and let

\[
y = \{ y : \exists p \in P \ ( (y, p) \in x) \}.
\]

Obviously, \( y \) is a set of \( \mathbb{P} \)-names which belongs to \( V \). By the Axiom of Choice, which holds in \( V \), we can write \( y = \{ y_\alpha : \alpha \in \kappa \} \), where \( \kappa = |y| \) is a cardinal in \( V \). Now let

\[
\mathcal{R} = \{ \text{op}(\alpha, y_\alpha) : \alpha \in \kappa \} \times \{ 0 \}
\]

which is a \( \mathbb{P} \)-name in \( V \) for a set of ordered pairs in \( V[G] \). Since \( 0 \in G \), \( \mathcal{R}[G] \) induces a surjection from \( \{ \alpha \in \kappa : \exists p \in G \ ( (y_\alpha, p) \in x) \} \subseteq \kappa \) onto the set \( x = x[G] = \{ (y_\alpha[G]) : \exists p \in G \ ( (y_\alpha, p) \in x) \} \), and consequently the set \( x \in V[G] \) can be well-ordered. Hence, since \( x \) was arbitrary, \( V[G] \models \text{AC} \). \( \dashv \)

**Fact 14.15.** If \( V \models \text{ZFC} \) and \( G \) is \( \mathbb{P} \)-generic over \( V \), then \( G \in V[G] \) and \( V \) is a subclass of \( V[G] \).

**Proof.** Let \( G \) be an arbitrary \( \mathbb{P} \)-generic filter. By definition of \( G \), \( G[G] = G \), and hence, by definition of \( V[G] \), \( G \in V[G] \). Further, \( G \) is downwards closed and therefore contains \( 0 \) (the smallest element of \( P \). Hence, for each \( x \in V \) we have \( x[G] = x \) and consequently \( x \in V[G] \). \( \dashv \)

**Proposition 14.16.** Let \( V \models \text{ZFC} \), let \( \mathcal{P} \) be a forcing notion in \( V \), and let \( G \) be \( \mathbb{P} \)-generic over \( V \); then \( \Omega^{V[G]} = \Omega^V \).

**Proof.** Since \( V \subseteq V[G] \), we obviously have \( \Omega^V \subseteq \Omega^{V[G]} \). On the other hand, assume towards a contradiction that there exists an ordinal in \( V[G] \) which does not belong to \( V \). Since the class \( \Omega^{V[G]} \) is well-ordered in \( V[G] \), there is a smallest ordinal in \( V[G] \), say \( \gamma \), which does not belong to \( V \). Let \( \gamma \) be a \( \mathbb{P} \)-name for \( \gamma \), i.e., \( \gamma = \gamma[G] \). Then \( \{ x : \exists p ((x, p) \in \gamma) \} \) is a set in \( V \), hence, the collection of all ordinals \( \alpha \in \gamma \) is in fact a set in \( V \). This implies that \( \gamma \) belongs to \( V \) and contradicts our assumption. \( \dashv \)
Until now we did not show that generic filters exist, but let us postpone this topic until the end of this chapter and let us show first how a statement (e.g., "there are Ramsey ultrafilters") can be forced to become true in a certain generic model.

**Forcing notions which do not add reals.** In this section, we shall see that the forcing notion \( \mathbb{U} \) adds a Ramsey ultrafilter to the ground model \( V \). In fact we shall see that whenever \( G \) is \( \mathbb{U} \)-generic over \( V \), then \( G \) induces a filter over \( \omega \) such that for any colouring \( \pi : [\omega]^2 \to 2 \) in \( V \) there is an \( x \in G \) such that \( \pi|_{[\omega]^2} \) is constant. However, in order to make this approach work we have to show that forcing with \( \mathbb{U} \) does not add any new reals (i.e., subsets of \( \omega \) or functions \( [\omega]^2 \to 2 \)) to \( V \); if \( \mathbb{U} \) would add new reals to \( V \), there might be a colouring \( \rho : [\omega]^2 \to 2 \) in \( V[G] \) such that no set \( x \in G \) is homogeneous for \( \rho \), and consequently, \( \{ x \in [\omega]^{\omega} : \exists y \in G (y \subseteq x) \} \) would just be a filter in \( V[G] \).

So, let us first prove that whenever \( G \) is \( \mathbb{U} \)-generic over \( V \), then \( [\omega]^{\omega} \cap V = [\omega]^{\omega} \cap V[G] \), i.e., every subset of \( \omega \) which is in \( V[G] \) is also in \( V \), and vice versa.

A forcing notion \( \mathbb{P} = (P, \leq) \) is said to be \( \sigma \)-closed if whenever \( \langle p_n : n \in \omega \rangle \) is an increasing sequence of elements of \( P \) (i.e., \( m < k \to p_m \leq p_k \) ), then there exists a condition \( q \in P \) such that for all \( n \in \omega \), \( q \geq p_n \).

By the proof of the fact that \( \mathbb{P} \) is uncountable (cf. Theorem 8.1) we get that the forcing notion \( \mathbb{U} \) is \( \sigma \)-closed.

The next result shows that forcing with a \( \sigma \)-closed forcing notion does not add new reals to the ground model.

**Lemma 14.17.** Let \( \mathbb{P} = (P, \leq) \) be a \( \sigma \)-closed forcing notion, \( G \) a \( \mathbb{P} \)-generic filter over \( V \), \( X \) a set in \( V \), and \( f : \omega \to X \) a function in \( V[G] \), i.e., \( V[G] \models f \in \omega X \); then \( f \) belongs to \( V \).

**Proof.** Let \( f \in \omega X \) be a function in \( V[G] \) and let \( f \) be a \( \mathbb{P} \)-name for \( f \). Assume towards a contradiction that \( f[G] \notin V \). By the Forcing Theorem 14.10(2) there is a condition \( q \in P \) (in fact, \( q \in G \)) such that

\[
q \Vdash f \in \omega X \land f \notin \omega X.
\]

Notice the difference between \( \omega X \) (which is a \( \mathbb{P} \)-name for the set \( \omega X \in V[G] \)) and \( \omega X \) (which is the canonical \( \mathbb{P} \)-name for the set \( \omega X \in V \)). By Lemma 14.11(b), let \( p_0 \geq q \) be such that \( p_0 \Vdash f(0) = x_0 \) (for some \( x_0 \in X \)), and for \( n \in \omega \) let \( p_{n+1} \geq p_n \) be such that \( p_{n+1} \Vdash f(n+1) = x_{n+1} \) (for some \( x_{n+1} \in X \)). Notice that by Lemma 14.11(b), \( p_0 \) and \( p_{n+1} \) exist and that the construction can be carried out in \( V \). Finally, let \( p \in P \) be such that for all \( n \in \omega \), \( p \geq p_n \). Then, by Fact 14.9(a), for all \( n \in \omega \) there is an \( x_n \in X \) such that \( p \Vdash f(n) = x_n \). Thus,

\[
p \Vdash f \in \omega X,
\]

which is a contradiction to our assumption. \( \Diamond \)
Since \( \mathbb{U} \) is \( \sigma \)-closed and every real \( x \in [\omega]^\omega \) corresponds to a function \( f_x \in \omega^2 \) (stipulating \( f_x(n) = 1 \iff n \in x \)), by Lemma 14.17, ultrafilter forcing \( \mathbb{U} \) does not add any new reals to the ground model \( \mathbb{V} \). In other words, if \( G \) is \( \mathbb{U} \)-generic over \( \mathbb{V} \), then \( [\omega]^\omega \cap \mathbb{V} = [\omega]^\omega \cap \mathbb{V}[G] \). With this observation we are ready to prove the following result.

**Proposition 14.18.** If \( G \) is \( \mathbb{U} \)-generic over \( \mathbb{V} \). Then \( \bigcup G \) is a Ramsey ultrafilter in \( \mathbb{V}[G] \) which is different from all ultrafilters in \( \mathbb{V} \), i.e., ultrafilter forcing \( \mathbb{U} \) adds a new Ramsey ultrafilter to \( \mathbb{V} \). In particular, \( \mathbb{V}[G] \) contains a Ramsey ultrafilter.

**Proof.** Firstly we show that \( \bigcup G = \{ x \in [\omega]^\omega : [x]^\omega \subseteq G \} \) is an ultrafilter over \( \omega \) which is different from all ultrafilters in \( \mathbb{V} \). Since \( G \) is downwards closed, directed, and meets every maximal anti-chain in \( [\omega]^\omega \)/fin which belongs to \( \mathbb{V} \) (in particular all anti-chains of the form \( \{ \{ z \} : \omega \setminus z \} \) for co-infinite sets \( z \in [\omega]^\omega \)), and since forcing with \( \mathbb{U} \) does not add reals, \( \bigcup G \) is an ultrafilter over \( \omega \). Let now \( \mathbb{W} \in \mathbb{V} \) be an arbitrary ultrafilter over \( \omega \). Then

\[
D_\mathbb{W} = \{ [x]^\omega : x \notin \mathbb{W} \}
\]

is an open dense subset of \( [\omega]^\omega \)/fin. Thus, \( G \cap D_\mathbb{W} \neq \emptyset \) which implies \( \bigcup G \neq \mathbb{W} \), and since \( \mathbb{W} \) was arbitrary, the ultrafilter \( \bigcup G \) is different from all ultrafilters in \( \mathbb{V} \).

Secondly we show that \( \bigcup G \) is a Ramsey ultrafilter: Let \( \pi : [\omega]^2 \to 2 \) be an arbitrary colouring in \( \mathbb{V}[G] \). Since forcing with \( \mathbb{U} \) does not add reals, \( \pi \in \mathbb{V} \). Now the set

\[
D_\pi := \{ [x]^\omega : \pi|_{[x]^\omega} \text{ is constant} \}
\]

is an open dense subset of \( [\omega]^\omega \)/fin. Thus, \( G \cap D_\pi \neq \emptyset \) which implies that there exists an \( [x]^\omega \subseteq G \) such that \( \pi|_{[x]^\omega} \) is constant, and since \( \pi \) was arbitrary, \( \bigcup G \) is a Ramsey ultrafilter.

The preceding theorem is a typical example how to force the existence of a certain set whose existence cannot be proved in ZFC. By the same forcing construction as above we shall see in Chapter 24 that there may be a Ramsey ultrafilter even in the case when \( p < \kappa \).

**Forcing notions which do not collapse cardinals.** Now we consider the forcing notion \( \mathbb{C}_\kappa \) (for an arbitrary cardinal \( \kappa \)) and show that the forcing notion \( \mathbb{C}_\kappa \) adds \( \kappa \) reals to the ground model \( \mathbb{V} \). As a consequence we get that whenever \( G \) is \( \mathbb{C}_\kappa \)-generic over \( \mathbb{V} \), then \( \mathbb{V}[G] \models c \geq \kappa \) (where \( c \) denotes the cardinality of the continuum). In particular, for \( \kappa > \omega_1 \) we get \( \mathbb{V}[G] \models \neg \text{CH} \). However, in order to make this approach work we have to show that \( \kappa \) is the same cardinal in \( \mathbb{V}[G] \) as it is in \( \mathbb{V} \). Let us explain this problem in greater detail: Let \( \mathbb{P} \) be a forcing notion and let \( G \) be \( \mathbb{P} \)-generic over \( \mathbb{V} \). Further, let \( \kappa \) be an arbitrary infinite cardinal in \( \mathbb{V} \). By definition, \( \kappa \) is an ordinal such that there is no bijection between \( \kappa \) and any of its elements (recall that the elements
of ordinal are ordinals). Since $V$ and $V[G]$ contain the same ordinals, $\kappa$ is an ordinal number in $V[G]$. However, since $V[G]$ is an extension of $V$, there might be an injective function in $V[G]$ which maps $\kappa$ to one of its elements. In other words, the ordinal number $\kappa$, which is a cardinal in $V$, might become an ordinary ordinal in $V[G]$, i.e., we might have $V \models |\kappa| = \kappa$ but $V[G] \models |\kappa| \in \kappa$.

If this is the case, then we say that $\mathbb{P}$ collapses $\kappa$; otherwise, we say that $\mathbb{P}$ preserves $\kappa$. If $\mathbb{P}$ preserves all cardinal numbers, i.e., $|\kappa|^{V[G]} = \kappa$ whenever $|\kappa|^V = \kappa$, then we simply say that $\mathbb{P}$ preserves cardinalities. Notice that all finite cardinals are preserved by any forcing notion, and consequently also $\omega$ must be preserved, i.e., we always have $|\omega|^V = |\omega|^{V[G]} = \omega$. On the other hand, any uncountable cardinal number can be collapsed; moreover, any uncountable cardinal can be forced to become a countable ordinal.

Now, let us prove that the forcing notion $C_\kappa$ preserves cardinals, but first we prove a slightly more general result.

Recall that a forcing notion $\mathbb{P} = (P, \leq)$ is said to satisfy the countable chain condition, denoted ccc, if every anti-chain in $P$ is at most countable — in which case we usually just say “$\mathbb{P}$ satisfies ccc”. For example, by Corollary 13.3 we know that the forcing notion $C_\kappa$ satisfies ccc.

In order to show that a forcing notion which satisfies ccc does not collapse any cardinal, we shall show the slightly more general result that a forcing notion which preserves cofinalities also preserves cardinalities. A forcing notion $\mathbb{P}$ preserves cofinalities if whenever $G$ is $\mathbb{P}$-generic over $V$ and $\kappa$ is a cardinal in $V$, then $\text{cf}(\kappa)^V = \text{cf}(\kappa)^{V[G]}$.

**Lemma 14.19.** If $\mathbb{P}$ preserves cofinalities, then $\mathbb{P}$ preserves cardinalities.

**Proof.** Assume $\mathbb{P}$ preserves cofinalities and let $G$ be $\mathbb{P}$-generic over $V$.

Firstly, let $\kappa$ be a regular cardinal in $V$, i.e., $V \models \text{cf}(\kappa) = \kappa$. Then, since $\mathbb{P}$ preserves cofinalities, the ordinal $\text{cf}(\kappa)^V$ is equal to the ordinal $\text{cf}(\kappa)^{V[G]}$. Thus, $V[G] \models \kappa = \text{cf}(\kappa)$ which shows that the ordinal $\kappa$, which is a regular cardinal in $V$, is still a regular cardinal in $V[G]$.

Secondly, if $\lambda > \omega$ is a limit cardinal in $V$, then the set of cardinals $C = \{\kappa < \lambda : \kappa \text{ regular}\}$ is cofinal in $\lambda$ (recall that by Proposition 5.10 successor cardinals are regular), and since the cardinals in $C$ remain (regular) cardinals in $V[G]$, $C^V = C^{V[G]}$ and consequently $\lambda$ is a cardinal (in fact a limit cardinal) in $V[G]$ as well.

**Lemma 14.20.** If $\mathbb{P} = (P, \leq)$ is a forcing notion which satisfies ccc, then $\mathbb{P}$ preserves cofinalities as well as cardinalities.

**Proof.** Let $\mathbb{P} = (P, \leq)$ be a forcing notion which satisfies ccc and which belongs to some model $V$ of ZFC, and let $G$ be $\mathbb{P}$-generic over $V$. By Lemma 14.19 it is enough to prove that $\mathbb{P}$ preserves cofinalities. Let $\kappa$ be an infinite cardinal in $V$ and let $S$ be a $P$-name for a strictly increasing sequence of length $\lambda = \text{cf}(\kappa)$ in $V[G]$ which is cofinal in $\kappa$, i.e., we have $S[G] : \lambda \to \kappa$ with $\bigcup \{ S[G](\alpha) : \alpha \in \lambda \} = \kappa$. Thus, there is a $P$-condition $p \in G$ such that...
\[ p \models_{\mathbb{P}} \mathcal{S} \in \lambda \land \bigcup \{ \mathcal{S}(\alpha) : \alpha \in \lambda \} = \kappa. \]

Work for a moment in the ground model \( V \): For each \( \alpha \in \lambda \) let
\[ D_\alpha = \{ q \geq p : \exists y (q \models_{\mathbb{P}} \mathcal{S}(\alpha) = y) \}. \]

Then, by Fact 14.9.(b), \( D_\alpha \) is open dense above \( p \). For each \( \alpha \in \lambda \) define
\[ Y_\alpha = \{ \gamma \in \kappa : \exists q \in D_\alpha (q \models_{\mathbb{P}} \mathcal{S}(\alpha) = \gamma) \}. \]

Then, for every \( \alpha \in \lambda \), the set \( Y_\alpha \subseteq \kappa \) is in \( V \), and since \( \mathbb{P} \) satisfies ccc, \( |Y_\alpha| \leq \omega \). Indeed, if \( q_1 \models_{\mathbb{P}} \mathcal{S}(\alpha) = \gamma_1 \) and \( q_2 \models_{\mathbb{P}} \mathcal{S}(\alpha) = \gamma_2 \), where \( \gamma_1 \neq \gamma_2 \), then \( q_1 \not\equiv q_2 \).

Let us turn back to the model \( V[G] \): For every \( \alpha \in \lambda \) let \( A_\alpha \) be a maximal anti-chain in \( D_\alpha \). By Fact 14.6.(b) and Fact 14.7, \( G \) meets every set \( A_\alpha \), which implies that for every \( \alpha \in \lambda \), \( S[G](\alpha) \in Y_\alpha \). Let \( Y := \bigcup \{ Y_\alpha : \alpha \in \lambda \} \); then \( Y \subseteq \kappa \) is a set in \( V \) such that \( \bigcup Y = \kappa \). Since the cardinal \( \lambda \) is infinite we get \( |Y| \leq \lambda \cdot \omega = \lambda \), which implies that \( \text{cf}(\kappa)^V \leq \lambda \). Thus, since \( \lambda = \text{cf}(\kappa)^V \), we have \( \text{cf}(\kappa)^V = \text{cf}(\kappa)^{V[G]} \).

\[ \square \]

**Independence of CH: The Gentle Way**

Since \( C_\kappa \) satisfies ccc, in order to prove the following result we just have to show that forcing with \( C_\kappa \) adds \( \kappa \) different real numbers to the ground model \( V \), i.e., the continuum in \( V[G] \) is at least of cardinality \( \kappa \).

**Theorem 14.21.** If \( V \models \text{ZFC} \) and \( G \) is \( C_\kappa \)-generic over \( V \), then \( V[G] \models \kappa \geq \kappa \).

In particular, if \( \kappa > \omega_1 \), then \( V[G] \models \lnot \text{CH} \).

**Proof.** Let \( G \) be \( C_\kappa \)-generic over \( V \). Since \( C_\kappa \) satisfies ccc, by Lemma 14.20 it is enough to prove that with \( G \) one can construct \( \kappa \) different real numbers. To keep the notation short let \( C_\kappa := \text{Fn}(\kappa \times \omega, 2) \).

Firstly we show that \( \bigcup G \) is a function from \( \kappa \times \omega \) to 2. For \( \alpha \in \kappa \) and \( n \in \omega \) let
\[ D_{\alpha,n} = \{ p \in C_\kappa : \langle \alpha, n \rangle \in \text{dom}(p) \}. \]

Then for any \( \alpha \in \kappa \) and \( n \in \omega \), \( D_{\alpha,n} \) is an open dense subset of \( C_\kappa \) and therefore \( G \cap D_{\alpha,n} \neq \emptyset \). Thus, for every \( \alpha \in \kappa \) and for every \( n \in \omega \) there is a \( p \in G \) such that \( p \) is defined on \( \langle \alpha, n \rangle \), and since \( G \) is directed, \( \bigcup G \) is a function with \( \text{dom}(\bigcup G) = \kappa \times \omega \).

Secondly we show how to construct \( \kappa \) different real numbers from \( G \): For each \( \alpha \in \kappa \) define \( r_\alpha \in \omega \) by stipulating \( r_\alpha(n) := \bigcup G(\langle \alpha, n \rangle) \) (for all \( n \in \omega \)). Now, for \( \alpha, \beta \in \kappa \) let
\[ D_{\alpha,\beta} = \{ p \in C_\kappa : \exists n \in \omega (\langle \alpha, n \rangle, \langle \beta, n \rangle) \subseteq \text{dom}(p) \land p(\langle \alpha, n \rangle) \neq p(\langle \beta, n \rangle) \}. \]
Independence of CH: the gentle way

Then for any distinct ordinals \( \alpha, \beta \in \kappa \), \( D_{\alpha, \beta} \) is an open dense subset of \( C_\kappa \) and therefore \( G \cap D_{\alpha, \beta} \neq \emptyset \). Thus, for any distinct \( \alpha, \beta \in \kappa \) there is an \( n \in \omega \) and a \( p \in G \) such that \( p((\alpha, n)) \neq p((\beta, n)) \), and therefore \( r_\alpha(n) \neq r_\beta(n) \).

We can even show that \( G \) adds \( \kappa \) new reals to the ground model \( V \). To see this, let \( f : \omega \to 2 \) be an arbitrary function in \( V \), and for any \( \alpha \in \kappa \) let

\[
D_{f, \alpha} = \{ p \in C_\kappa : \exists n \in \omega ((\alpha, n) \in \text{dom}(p) \land p((\alpha, n)) \neq f(n)) \}.
\]

Since \( D_{f, \alpha} \) is obviously open dense in \( C_\kappa \), \( r_\alpha \neq f \), and since the function \( f \in V \) was arbitrary, for each \( \alpha \in \kappa \) we have \( r_\alpha \notin V \). \( \dashv \)

Now we show that for each ordinal \( \alpha \), the statement \( 2^{\omega_\alpha} = \omega_{\alpha+1} \) is consistent with ZFC. In particular, for \( \alpha = 0 \) we get the relative consistency of the Continuum Hypothesis: but first we have to introduce some notations.

Let \( \kappa \) be an infinite cardinal. We say that a forcing notion \( \mathbb{P} = (P, \leq) \) is \( \kappa \)-closed if whenever \( \gamma < \kappa \) and \( \{ p_\xi : \xi \in \gamma \} \) is an increasing sequence of elements of \( P \) (i.e., \( \xi_0 < \xi_1 \Rightarrow p_{\xi_0} \leq p_{\xi_1} \)), then there exists a condition \( q \in P \) such that for all \( \xi \in \gamma \), \( q \geq p_\xi \). In particular, \( \omega_1 \)-closed is the same as \( \omega \)-closed.

The following fact is just a generalisation of Lemma 14.17 and we leave the proof as an exercise to the reader.

**Fact 14.22.** Let \( \mathbb{P} = (P, \leq) \) be a \( \kappa \)-closed forcing notion, \( \lambda \) an ordinal in \( \kappa \), \( G \) a \( \mathbb{P} \)-generic filter over \( V \), \( X \) a set in \( V \), and \( f : \lambda \to X \) a function in \( V[G] \); then \( f \) belongs to \( V \).

For ordinals \( \alpha \) let \( K_\alpha \) be the set of all functions \( p \) from a subset of \( \omega_{\alpha+1} \) to \( \mathcal{P}(\omega_\alpha) \) such that \( |\text{dom}(p)| < \omega_{\alpha+1} \) (i.e., \( |\text{dom}(p)| \leq \omega_\alpha \)), and let \( K_\alpha := (K_\alpha, \subseteq) \). Since \( \omega_{\alpha+1} \) is an infinite successor cardinal, it is regular, and therefore \( K_\alpha \) is \( \omega_{\alpha+1} \)-closed. Thus, by Fact 14.22, for each ordinal \( \beta \), every function from \( \omega_\alpha \) to \( \beta \) in a \( K_\alpha \)-generic extension belongs to the ground model. As a consequence we get that the forcing notion \( K_\alpha \) preserves all cardinals \( \leq \omega_{\alpha+1} \) and does not add new subsets of \( \omega_\alpha \).

With the forcing notion \( K_\alpha \) we can now easily construct a generic model in which \( 2^{\omega_\alpha} = \omega_{\alpha+1} \).

**Theorem 14.23.** If \( V \models \text{ZFC} \) and \( G_\alpha \) is \( K_\alpha \)-generic over \( V \), then \( V[G_\alpha] \models 2^{\omega_\alpha} = \omega_{\alpha+1} \). In particular we get \( V[G_0] \models \text{CH} \).

**Proof.** We shall show that \( \bigcup G_\alpha \) is a surjective function from \( \omega_{\alpha+1} \) onto \( \mathcal{P}(\omega_\alpha) \). Work in \( V \). For \( \xi \in \omega_{\alpha+1} \) and \( x \in \mathcal{P}(\omega_\alpha) \) let

\[
D_{\xi, x} = \{ p \in K_\alpha : \xi \in \text{dom}(p) \land x \in \text{ran}(p) \}.
\]

Then for every \( \xi \in \omega_{\alpha+1} \) and every \( x \in \mathcal{P}(\omega_\alpha) \), \( D_{\xi, x} \) is an open dense subset of \( K_\alpha \) and therefore \( G_\alpha \cap D_{\xi, x} \neq \emptyset \). Thus, for all \( \xi \in \omega_{\alpha+1} \) and \( x \in \mathcal{P}(\omega_\alpha) \) there is a \( p \in G_\alpha \) such that \( \xi \in \text{dom}(p) \) and \( x \in \text{ran}(p) \), and since \( G_\alpha \) is directed, this implies that the set \( \bigcup G_\alpha \) in \( V[G] \) is indeed a surjective function from \( \omega_{\alpha+1} \) onto \( \mathcal{P}(\omega_\alpha) \). Hence, \( V[G_\alpha] \models |\mathcal{P}(\omega_\alpha)| \leq \omega_{\alpha+1} \), and since \( 2^{\omega_\alpha} \geq \omega_{\alpha+1} \) we finally get \( V[G_\alpha] \models 2^{\omega_\alpha} = \omega_{\alpha+1} \). \( \dashv \)
By the two preceding theorems it follows that there are models of ZFC in which the Continuum Hypothesis holds as well as some in which it fails, and as a consequence we get that CH is independent of ZFC. However, the construction of the corresponding generic models relied on the existence of the corresponding generic filters, and it is now time to discuss this issue.

**On the Existence of Generic Filters**

Let \( V \) be again a model of ZFC and let \( P = ( P, \leq ) \) be a forcing notion which belongs to \( V \). We know from Chapter 5 that if ZF is consistent, then so is ZFC and that there is a smallest standard model of ZFC containing the ordinals, namely Gödel's constructible universe \( L \). So, we can assume \( V = L \) (in fact we have no other choice because \( L \) is the only model of ZFC we know of). Now assume that the set \( G \subseteq P \) is \( P \)-generic over \( V \), where \( P \) belongs to \( V \) and \( V \) is a model of ZFC (e.g., \( V = L \)). We first show that \( G \) does not belong to the model \( V \).

**Fact 14.24.** If \( V \) is a model of ZFC, \( P = ( P, \leq ) \) a forcing notion in \( V \), and \( G \subseteq P \) is \( P \)-generic over \( V \), then the set \( G \) does not belong to \( V \).

**Proof.** Let \( D_G = P \setminus G \) and let \( p \in P \) be an arbitrary \( P \)-condition. Since \( P \) is a forcing notion, there are incompatible elements above \( p \), i.e., \( \exists q_1, q_2 \in P \ (p \leq q_1 \land p \leq q_2 \land q_1 \perp q_2) \). Now, since \( G \) is directed, at most one of these two elements belongs to \( G \), or in other words, at least one of these two elements belongs to \( D_G \). Therefore, \( D_G \) is dense in \( P \) and since \( G \) is downwards closed, \( D_G \) is also open. Hence, \( D_G \) is an open dense subset of \( P \). If \( G \) belongs to \( V \), then \( D_G \) belongs to \( V \) as well, but obviously \( G \cap D_G = \emptyset \) which implies that \( G \) is not \( P \)-generic over \( V \). \( \square \)

This leads to the following question: If \( P \)-generic filters do not belong to the ground model \( V \), why do we know that \( P \)-generic filters exist? Informally, people living in \( V \) may ask: Is there life beyond \( V \)?

Unfortunately, one cannot prove within ZFC that \( P \)-generic filters exist, but at least, this one can prove: Consider the constructible universe \( L \). All sets in \( L \) are constructible, and vice versa, all constructible sets are in \( L \). If we add the statement \textit{all sets are constructible}, denoted \( V = L \), as a kind of axiom to ZFC, then there exists just a single transitive standard model of ZFC + \( V = L \) containing all the ordinals, namely \( L \) (at the same time we get that \( V = L \) is consistent with ZFC). Thus, as a consequence of \( V = L \) we get that there are no \( P \)-generic filters whatsoever.

Let us now explain how to get around this difficulty: Firstly construct a small (i.e., countable) model \( M \) of a large enough fragment of ZFC inside \( V \), and then extend \( M \) within \( V \) to a suitable generic model \( M[G] \). For example to show that \( \neg \text{CH} \) is consistent with ZFC, by the Compactness Theorem 3.7
it is enough to show that whenever $\Phi$ is a finite set of axioms of ZFC, then $\Phi + \neg CH$ has a model. Let $\Phi \subseteq ZFC$ be an arbitrary but fixed finite set of axioms. Now, take a countable set $M \in V$ such that $M$ can be extended in $V$ to a set model $M[G]$ (still in $V$) such that $M[G] \models \Phi$ but also $M[G] \models \neg CH$. Because $\Phi$ was arbitrary, this shows that $\neg CH$ is consistent with ZFC.

In the next chapter we show how to construct countable models for arbitrary finite fragments of ZFC and in Chapter 16 we finally show how to get proper independence proofs. However, in later chapters we shall skip this quite tedious construction and just work with the — in fact equivalent — approach presented here.

Notes

The creation of forcing. The notion of forcing and of generic sets were introduced by Paul Cohen [1] in 1963 to prove that $\neg AC$ is consistent with ZF and that $\neg CH$ is consistent with ZFC, and since Gödel’s constructible universe $L$ is a model of $ZF + AC + CH$, this implies that AC and CH are even independent of ZF and ZFC respectively. Cohen’s original approach and notation were modified for example by Scott, who defined essentially the forcing relationship given in Definition 14.8 and introduced the corresponding forcing symbol “$\Vdash$” (this definition of forcing and the corresponding symbol were first published in Feferman [6, p. 328f]). Notice the similarity between “$\models$” and “$\Vdash$”, and compare the Forcing Theorem 14.10 with Gödel’s Completeness Theorem 3.4. For a description of how Cohen had come to forcing we refer the reader to Cohen [5], and a history of the origins and the early development of forcing can be found in Moore [9] and Kanamori [7] (but see also Cohen [1, 2, 3, 4]).

The approach taken here. The way we introduced forcing was motivated by Kunen [8, Chapter VII, §§2-5], from where for example Definition 14.8 as well as the proof of the Forcing Theorem 14.10 were taken, and where one can also find a complete proof of the Generic Model Theorem 14.12 (cf [8, Chapter VII, Theorem 4.2]). However, Kunen considers generic extensions of countable transitive models of finite fragments of ZFC (whereas we considered generic extensions of models of full ZFC). This way he gets model-theoretic theorems whereas we just get results in the metatheory.

References

Models of finite fragments of Set Theory

In this chapter we summarise the model-theoretic facts which will be used in the next chapter in which the independence of the Continuum Hypothesis will be proved. Most of the following statements are classical results and are stated without proper proofs (for which we refer the reader to standard textbooks in axiomatic Set Theory like Jech [4] or Kunen [5]).

Basic Model-Theoretical Facts

Let $\mathcal{L}$ be an arbitrary but fixed language. Two $\mathcal{L}$-structures $M$ and $N$ with domain $A$ and $B$ respectively are called isomorphic if there is a bijection $f : A \to B$ between $A$ and $B$ such that:

1. $f(c^M) = c^N$ (for each constant symbol $c \in \mathcal{L}$)
2. $R^M(a_1, \ldots, a_n) \iff R^N(f(a_1), \ldots, f(a_n))$ (for each $n$-ary relation symbol $R \in \mathcal{L}$)
3. $f(F^M(a_1, \ldots, a_n)) = F^N(f(a_1), \ldots, f(a_n))$ (for each $n$-ary function symbol $F \in \mathcal{L}$)

If the $\mathcal{L}$-structures $M$ and $N$ are isomorphic and $f : A \to B$ is the corresponding bijection, then for all $a_1, \ldots, a_n \in A$ and each formula $\varphi(x_1, \ldots, x_n)$ we have:

$$M \models \varphi(a_1, \ldots, a_n) \iff N \models \varphi(f(a_1), \ldots, f(a_n))$$

This shows that isomorphic $\mathcal{L}$-structures are essentially the same, except that their elements have different “names”, and therefore, isomorphic structures are usually identified. For example the dihedral group of order six and $S_3$ (i.e., the symmetric group of order six) are isomorphic; whereas $C_6$ (i.e., the cyclic group of order six) is not isomorphic to $S_3$ (e.g., consider $\varphi(x_1, x_2) \equiv x_1 \cdots x_2 = x_2 \cdot x_1$).
If $M$ and $N$ are $\mathcal{L}$-structures and $B \subseteq A$, then $N$ is said to be an **elementary substructure** of $M$, denoted $N \prec M$, if for every formula $\varphi(x_1, \ldots, x_n)$ and every $b_1, \ldots, b_n \in B$:

$$N \models \varphi(b_1, \ldots, b_n) \iff M \models \varphi(b_1, \ldots, b_n)$$

For example the linearly ordered set $(\mathbb{Q}, <)$ is an elementary substructure of $(\mathbb{R}, <)$. On the other hand, $(\mathbb{Z}, <)$ is not an elementary substructure of $(\mathbb{Q}, <)$, e.g., $\forall x \forall y (x < y \rightarrow \exists z (x < z < y))$ is false in $(\mathbb{Z}, <)$ but true in $(\mathbb{Q}, <)$.

The key point in construction of elementary substructures of a given structure $M$ with domain $A$ is the following fact: A structure $N$ with domain $B \subseteq A$ is an elementary substructure of $M$ if and only if for every formula $\varphi(u, x_1, \ldots, x_n)$ and all $b_1, \ldots, b_n \in B$:

$$\exists a \in A : M \models \varphi(a, b_1, \ldots, b_n) \iff \exists b \in B : M \models \varphi(b, b_1, \ldots, b_n)$$

Notice that the implication from the right to the left is obviously true (since $B \subseteq A$). Equivalently we get that $N \prec M$ if for every formula $\varphi(u, x_1, \ldots, x_n)$ and all $b_1, \ldots, b_n \in B$:

$$\forall a \in A : M \models \varphi(a, b_1, \ldots, b_n) \iff \forall b \in B : M \models \varphi(b, b_1, \ldots, b_n)$$

Notice that in this case, the implication from the left to the right is obviously true.

The following theorem is somewhat similar to Corollary 15.5 below, even though it goes beyond ZFC (see Related Result 86). However, it is not used later, but it is a nice consequence of the characterisation of elementary submodels given above.

**Theorem 15.1 (Löwenheim-Skolem Theorem).** Every infinite model for a countable language has a countable elementary submodel. In particular, every model of ZFC has a countable elementary submodel.

### The Reflection Principle

Instead of aiming for a set model of all of ZFC, we can restrict our attention to **finite fragments** of ZFC (i.e., to finite sets of axioms of ZFC), denoted by ZFC$^\ast$.

We will see that for every finite fragment ZFC$^\ast$ of ZFC, there is a set which is a model of ZFC$^\ast$, but before we can state this result we have to give some further notions from model theory.

Let $V \models \text{ZFC}$, let $M \in V$ be any set, and let $M = (M, \in)$ be an $\in$-structure with domain $M$. An $\in$-structure $M = (M, \in)$, where $M \in V$ is a set, is called a **set model**. Notice that this definition of *model* is slightly
different to the one given in Chapter 3, where we defined models with respect to a set of formulae. For any formula $\varphi$ we define $\varphi^M$, the relativisation of $\varphi$ to $M$, by induction on the complexity of the formula $\varphi$:

- $(x = y)^M = x = y$.
- $(x \in y)^M = x \in y$.
- $(\psi_1 \wedge \psi_2)^M = \psi_1^M \wedge \psi_2^M$.
- $(\neg \psi)^M = \neg (\psi^M)$.
- $(\exists x \psi)^M = \exists x (x \in M \land \psi^M)$.

In other words, $\varphi^M$ is the formula obtained from $\varphi$ by replacing the quantifiers "$\exists x$" by "$\exists x \in M$". If $\varphi(x_1, \ldots, x_n)$ is a formula and $x_1, \ldots, x_n \in M$, then $\varphi^M(x_1, \ldots, x_n)$ is the same as $\varphi(x_1, \ldots, x_n)$ except that the bound variables of $\varphi$ range over $M$. (For $x_1, \ldots, x_n$ not all in $M$, the interpretation of $\varphi^M(x_1, \ldots, x_n)$ is irrelevant.) Notice that in the definition of $\varphi^M$, the interpretation of the non-logical symbol "$\in$" remains unchanged. Further, notice that also the sets themselves remain unchanged (which will not be the case for example when we apply Mostowski’s Collapsing Theorem 15.4).

For a formula $\varphi$ and a set model $M$, $M \models \varphi$ means $\varphi^M$ (where the free variables take arbitrary values in $M$). Similarly, for a set of formulae $\Phi$, $M \models \Phi$ means $M \models \varphi$ for each formula $\varphi \in \Phi$. If $M = (M, \in)$ and for all formulae $\varphi \in \Phi$ we have $M \models \varphi \iff V \models \varphi$,

then we say that $M$ reflects $\Phi$.

The following theorem shows that if ZFC is consistent, then any finite fragment of ZFC has a set model.

**Theorem 15.2 (Reflection Principle).** Assume that ZFC has a model, say $V$, let $M_0 \in V$ be an arbitrary set, and let $\text{ZFC}^* \subseteq \text{ZFC}$ be an arbitrarily large finite fragment of ZFC. Then we have:

(a) There is a set $M \supseteq M_0$ in $V$ such that $M$ reflects $\text{ZFC}^*$. In other words, there is a set $M \supseteq M_0$ such that for $M = (M, \in)$ we have

$$M \models \text{ZFC}^*.$$  

(b) There is even a transitive set $M \supseteq M_0$ that reflects $\text{ZFC}^*$ (recall that a set $x$ is transitive if $z \in y \in x$ implies $z \in x$).

(c) Moreover, there is a limit ordinal $\lambda$ such that $V_\lambda \supseteq M_0$ and the set $V_\lambda$ reflects $\text{ZFC}^*$.

(d) There is an $M \supseteq M_0$ such that $M$ reflects $\text{ZFC}^*$ and $|M| \leq \max \{|M_0|, \omega\}$. In particular, for $M_0 = \{\emptyset\}$, there is a countable set $M$ that reflects $\text{ZFC}^*$.

The crucial point in the proof of the Reflection Principle 15.2 is to show that for any existential formula $\exists x \varphi(x, y)$ and any set $M_0$ there exists a set
Let $V$ be a model of ZFC and let \( \varphi(x, y_1, \ldots, y_n) \) be a formula with \( \{x, y_1, \ldots, y_n\} \subseteq \text{free}(\varphi) \). For each non-empty set $M_0$ there is a set $M \supseteq M_0$ (where $M \in V$) such that for all $c_1, \ldots, c_n \in M$ we have:

$$V \models \exists x \varphi(x, c_1, \ldots, c_n) \rightarrow \exists a \in M \varphi(a, c_1, \ldots, c_n)$$

Moreover, we can construct $M' \supseteq M_0$ such that $|M'| \leq \max(|M_0|, \omega)$, in particular, if $M_0$ is countable, then $M'$ is countable as well.

Proof. Let $V \models \text{ZFC}$ and let $M_0$ be any non-empty set, e.g., $M_0 = \{ \emptyset \}$. Firstly, define in $V$ the class function $H : V^n \to V$ as follows:

If $V \models \exists x \varphi(x, u_1, \ldots, u_n)$ for some $u_1, \ldots, u_n \in V$, then let

$$H(u_1, \ldots, u_n) = \bigcap \{ V_a : a \in \Omega \land \exists x \in V_a \varphi(x, u_1, \ldots, u_n) \},$$

otherwise, $H(u_1, \ldots, u_n) := \{ \emptyset \}$.

Now, we construct the set $M \supseteq M_0$ by induction: For $i \in \omega$ let

$$M_{i+1} = M_i \cup \bigcup \{ H(c_1, \ldots, c_n) : c_1, \ldots, c_n \in M_i \}$$

and let

$$M = \bigcup_{i \in \omega} M_i.$$ 

If $c_1, \ldots, c_n \in M$, then there is an $i \in \omega$ such that $c_1, \ldots, c_n \in M_i$, and consequently, if $V \models \exists x \varphi(x, c_1, \ldots, c_n)$, then there is an $a \in M$ such that $V \models \varphi(a, c_1, \ldots, c_n)$.

By AC, fix a well-ordering $<$ of $M$, and define the partial function $h(c_1, \ldots, c_n) : M^n \to M$ as follows: If $H(c_1, \ldots, c_n) = \{ \emptyset \}$, then let $h(c_1, \ldots, c_n) := \emptyset$; otherwise, let $a \in M$ be the $\langle \cdot \rangle$-minimal element of $H(c_1, \ldots, c_n) \subseteq M$ and let $h(c_1, \ldots, c_n) := a$. We construct the set $M' \supseteq M_0$ again by induction: For $i \in \omega$ let

$$M'_{i+1} = M'_i \cup \{ h(c_1, \ldots, c_n) : c_1, \ldots, c_n \in M'_i \}$$

and let

$$M' = \bigcup_{i \in \omega} M'_i.$$ 

For all $i \in \omega$ we have $|M'_{i+1}| \leq |\text{seq}(M'_i)| = \max(|M'_i|, \omega)$, and therefore, $|M'| \leq \max(|M_0|, \omega)$. 

\[ \square \]
The Reflection Principle

**Proof of Theorem 15.2 (Sketch).** Let ZFC* be an arbitrary finite fragment of ZFC. Let \( \varphi_1, \ldots, \varphi_l \) be the finite list of all subformulae of formulae contained in ZFC*. We may assume that the formulae \( \varphi_1, \ldots, \varphi_l \) are written in the set-theoretic language \( \{\in\} \) and that no universal quantifier occurs in these formulae (i.e., replace \( \forall x \) by \( \neg \exists x \neg \)).

Applying the proof of Lemma 15.3 to all these formulae simultaneously, yields a set \( M \) such that for each \( i \) with \( 1 \leq i \leq l \) we have:

\[
V \models \exists x \varphi_i \rightarrow \exists x \in M \varphi_i
\]

A formula \( \varphi(x_1, \ldots, x_n) \) is said to be **absolute** for \( M = (M, \in) \) and \( V \), if for all \( a_1, \ldots, a_n \in M \) we have \( V \models \varphi(a_1, \ldots, a_n) \iff M \models \varphi(a_1, \ldots, a_n)^M \).

The proof is now by induction on the complexity of the formulae \( \varphi_1, \ldots, \varphi_l \): Let \( i, j, k \) be such that \( 1 \leq i, j, k \leq l \). If \( \varphi_i \) is atomic, i.e., \( \varphi_i \) is equivalent to \( x = y \) or \( x \in y \), then \( \varphi_i \) is obviously absolute for \( M \) and \( V \). If \( \varphi_i \) is of the form \( \varphi_j \land \varphi_k \), \( \varphi_j \lor \varphi_k \), \( \varphi_j \rightarrow \varphi_k \), where \( \varphi_j \) and \( \varphi_k \) are absolute for \( M \) and \( V \), then \( \varphi_i \) is absolute for \( M \) and \( V \) too. Finally, if \( \varphi_i \equiv \exists x \varphi_j \), then by construction of \( M \), \( \varphi_i \) is absolute for \( M \) and \( V \).

Hence, \( M \supseteq M_0 \), and the model \( M = (M, \in) \) has the desired properties.

\[\Box\]

The Reflection Principle 15.2 can be considered as a kind of ZFC-version of the Löwenheim-Skolem Theorem 15.1, and even though it is weaker than that theorem, it has many interesting consequences and important applications, especially in consistency proofs.

Some remarks:

1. If we compare (b) with (d) we see that we may require that the set \( M \) is transitive or that \( |M| \leq \max\{|M_0|, \omega\} \), but in general not both.

   For example let ZFC* be rich enough to define \( \omega \) as the smallest uncountable ordinal and assume that \( M = (M, \in) \) reflects ZFC*. If \( M \) is countable, then \( M \) cannot be transitive; and if \( M \) is transitive, then \( M \) must be uncountable.

2. As a consequence of the Reflection Principle 15.2 and of Gödel’s Second Incompleteness Theorem 3.9, it follows that ZFC is not finitely axiomatisable (i.e., there is no way to replace the two axiom schemata by just finitely many single axioms).

   On the other hand, by the Reflection Principle 15.2 we get that for each finite fragment ZFC* of ZFC, there is a proof in ZFC that ZFC* has a set model, whereas by Gödel’s Second Incompleteness Theorem 3.9 the existence of a model of ZFC is not provable within ZFC.

3. Let ZFC* be a finite fragment of ZFC and assume that ZFC* \( \vdash \varphi \) (for some sentence \( \varphi \)). Further, assume that \( M \) reflects ZFC* and let \( M = (M, \in) \).

   Then, in the model-theoretic sense, \( M \models ZFC^* \), and consequently, \( M \models \varphi \).

   As we will see later, this is the first step in order to show that a given
sentence \( \varphi \) is consistent with \( \text{ZFC} \). By the \textbf{Compactness Theorem 3.7} it is enough to show that whenever \( \Phi \subseteq \text{ZFC} \) is a finite fragment of \( \text{ZFC} \), then \( \Phi + \varphi \) has a model. Let \( \Phi \) be an arbitrary but fixed finite set of axioms of \( \text{ZFC} \). Now, let \( M \in \mathcal{V} \) be a set model of \( \Phi \), where \( \mathcal{V} \) is a certain finite fragment of \( \text{ZFC} \) which makes sure that the model \( M \) can be extended to a set model \( M[X] \) such that \( M[X] \models \Phi + \varphi \). Thus, since \( \Phi \) was arbitrary, this shows that \( \varphi \) is consistent with \( \text{ZFC} \). (This method is used and explained again in Chapter 16.)

**Countable Transitive Models of Finite Fragments of \( \text{ZFC} \)**

As mentioned above, a set model \( M = (M, \in) \) of a finite fragment of \( \text{ZFC} \) can be taken to be countable or transitive, but in general not both. However, as a consequence of \textbf{Mostowski's Collapsing Theorem 15.4} we can get also a transitive set model which is isomorphic to \( M \). This is done by reinterpreting the elements of \( M \) and as a result we get a model which is countable and transitive, but which is not a submodel of \( M \). Before we can state \textbf{Mostowski's Collapsing Theorem 15.4}, we have to introduce some notions.

Let \( M \) be an arbitrary set. For a binary relation \( E \subseteq M \times M \) on \( M \) and each \( x \in M \) let

\[
\text{ext}_E(x) = \{ z \in M : z \; E \; x \}
\]

be the \textit{extension} of \( x \).

A binary relation \( E \) on \( M \) is said to be \textit{well-founded} if every non-empty subset of \( M \) has an \( E \)-minimal element (i.e., for each non-empty \( A \subseteq M \) there is an \( x_0 \in A \) such that \( \text{ext}_E(x_0) \cap A = \emptyset \)).

A well-founded binary relation \( E \) on \( M \) is \textit{extensional} if for all \( x, y \in M \) we have

\[
\text{ext}_E(x) = \text{ext}_E(y) \rightarrow x = y.
\]

In other words, \( E \) is extensional iff \((M, E)\) satisfies the Axiom of Extensionality (with respect to the binary relation \( E \)).

The following result shows that for every structure \((M, E)\) which satisfies the Axiom of Extensionality, there exists a transitive set \( N \) such that \((M, E)\) and \((N, \in)\) are isomorphic.

**Theorem 15.4 (Mostowski's Collapsing Theorem).** If \( E \) is a well-founded and extensional binary relation on a set \( M \), then there exists a unique transitive set \( N \) and an isomorphism \( \pi \) between \((M, E)\) and \((N, \in)\), i.e., \( \pi : M \rightarrow N \) is a bijection and for all \( x, y \in M \), \( y \; E \; x \leftrightarrow \pi(y) \in \pi(x) \).

**Proof (Sketch).** Let \( x_0 \in M \) be an \( E \)-minimal element of \( M \). Since \( E \) is extensional, \( x_0 \) is unique. Define \( \pi(x_0) := \emptyset \) and let \( A_0 = \{ x_0 \} \). If, for some \( \alpha \in \Omega \), \( A_\alpha \) is already defined and \( M \setminus A_\alpha \neq \emptyset \), then let \( X_\alpha \) be the set of all \( E \)-minimal elements of \( M \setminus A_\alpha \), let \( A_{\alpha+1} := A_\alpha \cup X_\alpha \), and for each \( x \in X_\alpha \)
define $\pi(x) := \{\pi(y) : y \in x\}$. Now, $M = \bigcup_{\alpha \in \lambda} A_\alpha$ (for some $\lambda \in \Omega$) and we define $N := \pi[M]$. We leave it as an exercise to the reader to show that $\pi$ and $(N, \in)$ have the required properties and that $(N, \in)$ is unique. \[ \square \]

It is worth mentioning that not just the set $N$, but also the isomorphism $\pi$ is unique. We also would like to mention that Mostowski's Collapsing Theorem 15.4 is a ZFC result and that $\pi$ is just a mapping between two sets.

As an immediate consequence of Mostowski's Collapsing Theorem 15.4 we get

**Corollary 15.5.** Let $V$ be a model of ZFC and let $M = (M, \in)$ be a countable set model in $V$. If $\text{ZFC}^*$ is a finite fragment of ZFC containing the Axiom of Extensionality and $M \models \text{ZFC}^*$, then there is a countable transitive set $N$ in $V$ such that $N = (N, \in)$ is a set model in $V$ which is isomorphic to $M$ (in particular, $N \models \text{ZFC}^*$).

**Proof.** Let $M = (M, \in)$ be a countable set model of ZFC*. Because $M$ is a set, the relation “$\in$” is obviously a well-founded and extensional binary relation on $M$. Thus, by Mostowski’s Collapsing Theorem 15.4, there is a transitive set $N$ such that $M = (M, \in)$ and $N = (N, \in)$ are isomorphic, and since $\pi : M \rightarrow N$ is a bijection, $N$ is countable. \[ \square \]

Let $\text{ZFC}^*$ be any finite fragment of ZFC and let $V$ be a model of ZFC. Then, by the Reflection Principle 15.2(d), there is a countable set $M$ in $V$ that reflects $\text{ZFC}^*$ and for $M = (M, \in)$ we have $M \models \text{ZFC}^*$. Thus, by Corollary 15.5, there is a countable transitive set $N$ that reflects $\text{ZFC}^*$. In other words, for any finite fragment $\text{ZFC}^* \subseteq \text{ZFC}$ there is a countable transitive model $N$ in $V$ such that $N \models \text{ZFC}^*$.

Let us briefly discuss the preceding constructions: We start with a model $V$ of ZFC and an arbitrary large but finite set of axioms $\text{ZFC}^* \subseteq \text{ZFC}$. By the Reflection Principle 15.2(d) there is a countable set $M$ in $V$ such that $M = (M, \in)$ is a model of $\text{ZFC}^*$. By applying Mostowski’s Collapsing Theorem 15.4 to $(M, \in)$ we obtain a countable transitive model $N = (N, \in)$ in $V$ such that the models $N = (N, \in)$ and $M$ are isomorphic, and consequently, $N$ is a model of $\text{ZFC}^*$.

It is worth mentioning that the model $M = (M, \in)$ is a genuine submodel of $V$ and therefore contains the real sets of $V$. For example if

$M \models \text{“$\lambda$ is the least uncountable ordinal”}$

then $\lambda = \omega_1$, i.e., $\omega_1 \in M$. However, since the set $M$ is countable in $V$, there are countable ordinals in $V$ which do not belong to the set $M$, and therefore not to the model $M$ (which implies that $M$ is not transitive). In other words,

$V \models \lambda = \omega_1 \land \omega_1 \in M \land |\lambda \cap M| = \omega.$

On the one hand, the model $N = (N, \in)$ is in general not a submodel of $V$ and just contains a kind of copies of countably many set of $V$. For example if
\[ N \models \text{"}\lambda \text{ is the least uncountable ordinal"} \]

then \( \lambda \), which corresponds to \( \omega_1 \) in \( N \), is just a countable ordinal in \( V \). However, since \( N \) is transitive, every ordinal in \( V \) which belongs to \( \lambda \) also belongs to the set \( N \), and therefore to the model \( N \). In other words,

\[ V \models \lambda \in \omega_1 \land \lambda \in N \land \lambda \cap N = \lambda. \]

The relationships between the three models \( V, M, \) and \( N \) are illustrated by the following figure:

As we shall see in the next chapter, countable transitive models of finite fragments of \( \text{ZFC} \) play a key role in consistency and independence proofs.

Notes

For concepts of model theory and model-theoretical terminology we refer the reader to Hodges [3] or to Chang and Keisler [1]. However, the preceding results (including proofs) can also be found in Jech [4, Chapter 12].

The Löwenheim-Skolem Theorem 15.1 was already discussed in the notes of Chapter 3; the Reflection Principle 15.2 was introduced by Montague [7] (see also Lévy [6]); and the transitive collapse was defined by Mostowski [8].

Related Results

82. *A model of \( \text{ZF} - \text{Inf} \) and the consistency of PA.* \( V_\omega \models \text{ZF} - \text{Inf} \), where Inf denotes the Axiom of Infinity, and moreover, we even have Con(PA) \( \iff \) Con(\( \text{ZF} - \text{Inf} \)) (see Jech [4, Exercise 12.9] and Kunen [5, Chapter IV, Exercise 30]).

83. *Models of \( \text{Z} \).* Let \( Z \) be \( \text{ZF} \) without the Axiom Schema of Replacement. For every limit ordinal \( \lambda > \omega \) we have \( V_\lambda \models Z \) (see Jech [4, Exercise 12.7] or Kunen [5, Chapter IV, Exercise 6]).
For every infinite regular cardinal $\kappa$, let $H_\kappa := \{ x : \text{TC}(x) < \kappa \}$. The elements of $H_\kappa$ are said to be hereditarily of cardinality $< \kappa$. In particular, $H_\omega$, which coincides with $V_\omega$, is the set of hereditarily finite sets and $H_{\omega_1}$ is the set of hereditarily countable sets.

84. **Models of ZFC – P.** If AC holds in $V$, then for all cardinals $\kappa > \omega$ we have $H_\kappa \models Z - P$, where $P$ denotes the Axiom of Power Set. Moreover, for regular cardinals $\kappa > \omega$ we even have $H_\kappa \models ZFC - P$ (see Kunen [5, Chapter IV, Exercise 7] and Kunen [5, Chapter IV, Theorem 6.5]).

An uncountable regular cardinal $\kappa$ is said to be inaccessible if for all $\lambda < \kappa$, $2^\lambda < \kappa$. The inaccessible cardinals owe their name to the fact that they cannot be obtained (or accessed) from smaller cardinals by the usual set-theoretical operations. To some extent, an inaccessible cardinal is to smaller cardinals what $\omega$ is to finite cardinals and what is reflected by the fact that $H_\omega \models ZFC - \text{Inf}$ (cf. Jech [4, Exercise 12.9]). Notice that by Cantor’s Theorem 3.25, every inaccessible cardinal is a regular limit cardinal. One cannot prove in ZFC that inaccessible cardinals exist; moreover, one cannot even prove that uncountable regular limit cardinals exist (see Kunen [5, Chapter VI, Corollary 4.13] but also Hausdorff’s remark [2, p. 131]).

85. **Models of ZFC.** If $\kappa$ is inaccessible, then $H_\kappa \models ZFC$ (cf. Kunen [5, Chapter IV, Theorem 6.6]). Let us show that if ZFC is consistent, then ZFC $\not\models \text{Inacc}$, where Inacc denotes the axiom “$\exists \kappa (\kappa \text{ is inaccessible})$”. Since $H_\kappa \models ZFC$ (if $\kappa$ is inaccessible), it is provable from ZFC $+ \text{Inacc}$ that ZFC has a model which is equivalent to saying that ZFC is consistent. Now, if ZFC $\not\models \text{Inacc}$, then we consequently get that ZFC proves its own consistency, which is impossible by Gödel’s Second Incompleteness Theorem 3.9 (unless ZFC is inconsistent).

86. **The Löwenheim-Skolem Theorem.** Even though the Löwenheim-Skolem Theorem 15.1 for ZFC — which says that every model of ZFC has a countable elementary submodel — is somewhat similar to Corollary 15.5, it can neither be formulated in First-Order Logic nor can it be proved in ZFC: Firstly, notice that ZFC consists of infinitely many axioms. Thus, we cannot write these axioms as a single formula as we have done above in order to prove the Reflection Principle 15.2. Furthermore, even in the case when we would work in higher order Logic, if every model $V$ of ZFC would have a countable elementary submodel $V'$, then the set of ordinals in $V'$ (i.e., $\Omega^V \cap V'$) would be countable in $V$ (but not in $V'$, of course). Now, in $V$ we can build the sequence $a_0 := \bigcup \Omega^V \cap V'$, $a_1 := \bigcup (\Omega^{V'} \cap V''$), and so on. This would result in an infinite, strictly decreasing sequence $a_0 \not\geq a_1 \not\geq \ldots$ of ordinals in $V$, which is a contradiction to the Axiom of Foundation.

**References**


Consistency and Independence Proofs: The Proper Way

We have seen in Chapter 14 how we could extend models of ZFC to models in which for example CH fails — supposed we have suitable generic filters at hand. On the other hand, we have also seen in Chapter 14 that there is no way to prove that generic filters exist.

However, in order to show that for example CH is independent of ZFC we have to show that ZFC + CH as well as ZFC + ¬CH has a model. In other words we are not interested in the generic filters themselves, but rather in the sentences which are true in the corresponding generic models; on the other hand, if there are no generic filters, then there are also no generic models.

The trick to avoid generic filters (over models of ZFC) is to carry out the whole forcing construction within a given model V of ZFC — or alternatively in ZFC: In V we first construct a countable model N of a suitable finite fragment of ZFC. Then we define a kind of “mini-forcing” P which belongs to the model N and show that there is a set G in V which is P-generic over N. From the point of view of N, N[G] is a proper generic extension of N, and since G is a set in V, also N[G] belongs to V. This shows that certain generic extensions exist, in particular generic extensions of countable models of finite fragments of ZFC.

What we gain with this approach is that the whole construction takes place in the model V, but the price we pay is that neither N nor N[G] is a proper generic extension of N, and since G is a set in V, also N[G] belongs to V. This shows that certain generic extensions exist, in particular generic extensions of countable models of finite fragments of ZFC.

What we gain with this approach is that the whole construction takes place in the model V, but the price we pay is that neither N nor N[G] is a model of ZFC; but now it is time to describe the proper way for obtaining consistency and independence results in greater detail:

0. The goal: Suppose we would like to show that a given sentence ϕ is consistent with ZFC, i.e., we have to show that Con(ZFC) implies Con(ZFC + ϕ). By Gödel’s Completeness Theorem 3.4 this is equivalent to showing that ZFC + ϕ has a model whenever there is a model V of ZFC.
1. **Getting started:** By the **Compactness Theorem 3.7**, \( \text{ZFC} + \varphi \) is consistent if and only if for every finite set of axioms \( \Phi \) of \( \text{ZFC} \), \( \Phi + \varphi \) is consistent, i.e., \( \Phi + \varphi \) has a model. Below, we show how to construct a model of \( \Phi_0 + \varphi \), where \( \Phi_0 \) is an arbitrary but fixed finite set of axioms of \( \text{ZFC} \).

2. **A suitable forcing notion \( \mathbb{P} \):** In the model \( V \) define a forcing notion \( \mathbb{P} = (P, \leq) \) which has the property that there is a condition \( p_0 \in P \) such that \( p_0 \Vdash \varphi \). For example if \( \varphi \) is \( \neg \text{CH} \), then by the methods presented in Chapter 14, \( \mathcal{C}_\omega \) would have the required properties.

3. **Choosing a suitable finite set of axioms:** Let \( \text{ZFC}^* \subseteq \text{ZFC} \) be a finite fragment of \( \text{ZFC} \) such that:
   (a) Each axiom of \( \Phi_0 \) belongs to \( \text{ZFC}^* \).
   (b) \( \text{ZFC}^* \) is strong enough to define the forcing notion \( \mathbb{P} \), the existence of the condition \( p_0 \), as well as some properties of \( \mathbb{P} \) like satisfying \( \text{ccc} \), being \( \sigma \)-closed, \( \text{et cetera} \).
   (c) \( \text{ZFC}^* \) is strong enough to prove that every sentence in \( \Phi_0 \) is forced to be true in any \( \mathbb{P} \)-generic extension of \( V \).
   (d) \( \text{ZFC}^* \) is strong enough to prove that various concepts like “finite”, “partial ordering and dense sets”, \( \text{et cetera} \), are absolute for all countable transitive models.

   The properties (b)-(d) of \( \text{ZFC}^* \) are necessary to prove **Theorem 16.1**; however, we will omit most of the quite tedious and technical proof of that theorem.

4. **The corresponding countable transitive model \( N \):** Let \( M_0 = \{ p_0, P, R_\leq \} \), where \( R_\leq = \{ (p, q) \in P \times P : p \leq q \} \). By the **Reflection Principle 15.2** there is a countable set \( M \supseteq M_0 \) in \( V \) such that \( M \) reflects \( \text{ZFC}^* \), i.e., for \( M = (M, \in) \) we have \( M \models \text{ZFC}^* \). By **Corollary 15.5** and **Mostowski’s Collapsing Theorem 15.4**, there is a countable transitive model \( N = (N, \in) \) in \( V \) such that \( N \models \text{ZFC}^* \), and in addition there is a bijection \( \pi : M \to N \) such that for all \( x, y \in M \), \( y \in x \iff \pi(y) \in \pi(x) \). Define \( P^N := \pi[P] \) and \( \leq^N := \pi[R_\leq] \). Notice that for all \( p, q \in P^N \), \( N \models p \leq^N q \) iff \( \pi^{-1}(p) \leq \pi^{-1}(q) \).

5. **Relativisation of \( \mathbb{P} \)-generic filters to \( N \):** For a set \( G \subseteq P^N \) let
   \[
   N[G] = \{ x[G] : x \text{ is a } \mathbb{P} \text{-name in } N \}.
   \]

   A set \( G \subseteq P^N \) is \( P^N \)-generic over \( N \) if it meets every open dense subset \( D \subseteq P^N \) which is in \( N \).

6. **Relativisation of the Generic Model Theorem:** There is even a relativisation of the **Generic Model Theorem 14.12** which states as follows.

**Theorem 16.1.** Let \( V \) be a model of \( \text{ZFC} \), let \( \mathbb{P} = (P, \leq) \) be a forcing notion in \( V \) and let \( p_0 \) be an arbitrary condition in \( P \). Furthermore, let \( \Phi_0 \) and \( \text{ZFC}^* \) be as above and let \( N = (N, \in) \) be a countable transitive
model in $V$ such that $N \models \text{ZFC}^\ast$. Then there is a set $G \subseteq P^N$ in $V$ which contains $p_0$ and which is $\mathbb{P}^N$-generic over $N$. Moreover, $N[G] = (N[G], \in)$ is a countable transitive model in $V$ and $N[G] \models \Phi_0$.

Proof (Sketch). Firstly, let us show that there exists a set $G \subseteq P^N$ in $V$ which is $\mathbb{P}^N$-generic over $N$ and contains $p_0$; because the model $N$ is countable in $V$, from the point of view of $V$, the model $N$ contains just countably many sets which are open dense subsets of $P^N$. Let $\{D_n : n \in \omega\}$ be this countable set. Since $D_0$ is dense, we can take a condition $q_0 \in D_0$ such that $q_0 \geq p_0$; and in general, for $n \in \omega$ take $q_{n+1} \in D_{n+1}$ such that $q_{n+1} \geq q_n$. Finally let

$$G = \{p \in P^N : \exists n \in \omega (p \leq q_n)\}.$$ 

Then $G \subseteq P^N$ is a set in $V$ which contains $p_0$ and meets every open dense subset of $P^N$ which belongs to $N$, and hence, $G$ is $\mathbb{P}^N$-generic over $N$. Notice that even though each $q_n$ belongs to the model $N$, the sequence $\{q_n : n \in \omega\}$ —and consequently the set $G$— does not belong to $N$. Notice also that since $N$ is countable in $V$, there are only countably many names in $N$ and consequently $N[G]$ is countable in $V$.

Secondly, let us show that $N[G] \models \Phi_0$: By the choice of $\text{ZFC}^\ast$ (in step 3), we can show in $N$, by using the technique introduced in Chapter 14, that whenever $G$ is $\mathbb{P}^N$-generic over $N$ and contains $p_0$, then $N[G] \models \Phi_0$.

7. The final step: In step 2 we assumed that $V[G] \models \varphi$ whenever $G$ is $\mathbb{P}$-generic over $V$ and contains $p_0$. Thus, by Theorem 16.1, we get that $N[G] \models \varphi$

whenever $G$ is $\mathbb{P}^N$-generic over $N$ and $p_0 \in G$. On the other hand, by the choice of the set of axioms $\text{ZFC}^\ast$ and since $N \models \text{ZFC}^\ast$, we get $N[G] \models \Phi_0$, hence,

$$N[G] \models \Phi_0 + \varphi$$

which shows that $\Phi_0 + \varphi$ is consistent.

8. Conclusion: Since the finite set of axioms $\Phi_0$ we have chosen in step 1 was arbitrary, $\Phi + \varphi$ is consistent for every finite set of axioms $\Phi$ of $\text{ZFC}$, and consequently we get that $\varphi$ is consistent with $\text{ZFC}$. This is what we were aiming for and what is summarised by the following result.

Proposition 16.2. Let $\varphi$ be an arbitrary sentence in the language of Set Theory. If there is a forcing notion $\mathbb{P} = (P, \leq)$ and a condition $p \in P$ such that $p \Vdash \varphi$, then $\varphi$ is consistent with $\text{ZFC}$. 


The model-theoretic part of the above construction is illustrated by the following figure:

![Diagram](image)

The most inelegant part in the proof of the consistency of $\varphi$ is surely step 3, where we have to find a finite set of axioms $\text{ZFC}^* \subseteq \text{ZFC}$ which is strong enough to prove that whenever $N \models \text{ZFC}^*$ and $G$ is $\mathbb{P}$-generic over $N$, then $N[G] \models \Phi_0$. On the other hand, for a consistency proof it is not necessary to display explicitly the axioms in $\text{ZFC}^*$; it is sufficient to know that such a finite set of axioms exists.

The crucial point in the proof of the consistency of $\varphi$ is step 2, where we have to find (or define) a forcing notion $\mathbb{P}$ such that there is a $\mathbb{P}$-name $p_0$ which forces $\varphi$. In fact it will turn out that $p_0$ is always equal to 0, in which case we say that $\mathbb{P}$ forces $\varphi$, i.e., $\varphi$ is true in all $\mathbb{P}$-generic extensions of $V$. For example $\mathbb{K}_0$ and $\mathbb{C}_{\omega_2}$ (both defined in Chapter 14) force $\text{CH}$ and $\neg \text{CH}$ respectively.

Now, let us turn our attention to independence results: Firstly recall that a sentence $\varphi$ is independent of $\text{ZFC}$ if $\varphi$ as well as $\neg \varphi$ is consistent with $\text{ZFC}$. So, in order to show that a sentence $\varphi$ is independent of $\text{ZFC}$ we would have to go twice through the procedure described above. However, since the only crucial point in the proof is step 2, all what we have to do is to find two suitable forcing notions:

In order to show that a given set-theoretic sentence $\varphi$ is independent of $\text{ZFC}$, we have to show that there are two forcing notions such that one forces $\varphi$ and the other one forces $\neg \varphi$.

As a first example let us consider the case when $\varphi$ is $\text{CH}$.

**Theorem 16.3.** $\text{CH}$ is independent of $\text{ZFC}$.

**Proof.** On the one hand, by **Theorem 14.21** we get that whenever $G$ is $\mathbb{C}_\kappa$-generic over $V$ and $\kappa > \omega_1$, then $V[G] \models \neg \text{CH}$, and therefore we get that
The cardinality of the continuum

Until now we just have seen that for each infinite cardinal \( \kappa \) there is a model in which \( \kappa \geq \varepsilon \), but we did not give any estimate how large \( \varepsilon \) actually is in such a model. Of course, since \( \varepsilon^\omega = \varepsilon \), \( \varepsilon = \kappa \) implies that \( \kappa \) must also satisfy \( \kappa^\omega = \kappa \). Surprisingly, this is the only demand for \( \kappa \) to make it possible to force that \( \varepsilon = \kappa \).

**Theorem 16.4.** For every cardinal \( \kappa \) which satisfies \( \kappa^\omega = \kappa \) we have:

\[
\text{Con}(\text{ZFC}) \Rightarrow \text{Con}(\text{ZFC} + \varepsilon = \kappa)
\]

**Proof.** Let \( V \models \text{ZFC} \) and let \( \kappa \) be a cardinal in \( V \) which satisfies \( \kappa^\omega = \kappa \). Consider the forcing notion \( C_\kappa = (\text{Fn}(\kappa \times \omega, 2), \subseteq) \). For convenience, we write \( C_\kappa \) instead of \( \text{Fn}(\kappa \times \omega, 2) \). If \( G \) is \( C_\kappa \)-generic over \( V \), then \( V[G] \models \varepsilon \geq \kappa \) (cf. Theorem 14.21). Thus, it remains to show that \( V[G] \models \kappa \leq \kappa \).

Firstly we investigate \( C_\kappa \)-names for subsets of \( \omega \). Let \( \check{x} \) be an arbitrary \( C_\kappa \)-name for a subset of \( \omega \). For each \( n \in \omega \) let

\[
\Delta_{n\in\omega} = \{ p \in C_\kappa : (p \force C_\kappa \eta \in \check{x}) \lor (p \force C_\kappa \eta \notin \check{x}) \}.
\]

By Fact 14.9(b), for each \( n \in \omega \) the set \( \Delta_{n\in\omega} \) is open dense in \( C_\kappa \). For each \( n \in \omega \) choose a maximal anti-chain \( A_n \) in \( \Delta_{n\in\omega} \) and define

\[
\check{x} = \{ (p, \check{\eta}) : p \in A_n \land p \force C_\kappa \eta \in \check{x} \}.
\]

A name for a subset of \( \omega \) of the form like \( \check{x} \) is called a nice name (i.e., nice names are a special kind of names for subsets of \( \omega \)). Now we show that \( 0 \force C_\kappa \check{x} = \check{x} \) by showing that for each \( n \in \omega \) the set

\[
D_n = \{ q \in C_\kappa : q \force C_\kappa \eta \in \check{x} \iff n \in \check{x} \}
\]

is dense in \( C_\kappa \). Fix \( n \in \omega \) and let \( p \) be an arbitrary \( C_\kappa \)-condition. Since \( \Delta_{n\in\omega} \) is dense in \( C_\kappa \), there is a \( p_0 \geq p \) such that \( p_0 \in \Delta_{n\in\omega} \), and since \( A_n \) is a maximal anti-chain in \( \Delta_{n\in\omega} \), there is a \( q_0 \in A_n \) such that \( p_0 \) and \( q_0 \) are compatible. Thus, there is a \( q \in C_\kappa \) such that \( p_0 \leq q \geq q_0 \). By construction we get

\[
q \force C_\kappa \eta \in \check{x} \iff n \in \check{x},
\]

and since \( p \leq q \) and \( p \) was arbitrary this shows that \( D_n \) is dense in \( C_\kappa \). In particular we get that for every \( C_\kappa \)-name \( \check{x} \) for a subset of \( \omega \) there exists a nice name \( \check{x} \) such that \( 0 \force C_\kappa \check{x} = \check{x} \).
Secondly we compute the cardinality of the set of nice names: Since $\kappa$ is infinite, $|[\kappa \times \omega \times 2]|^{<\omega} = \kappa$ (cf. Corollary 5.8), and consequently $|C_\kappa| = \kappa$ (we leave the details as an exercise to the reader). Recall that $C_\kappa$ satisfies $\text{ccc}$, i.e., every anti-chain in $C_\kappa$ is at most countable. Now, every nice name is the countable union of at most countable sets of ordered pairs, where each ordered pair is of the form $(n,p)$ for some $n \in \omega$ and $p \in C_\kappa$. Thus, there are at most

$$((\omega \cdot \kappa)^\omega)^\omega = \kappa^{\omega \omega} = \kappa^\omega = \kappa$$

nice names for subsets of $\omega$. Now, because each set $x \subseteq \omega$ which is in $V[G]$ has a $C_\kappa$-name in $V$, and because every $C_\kappa$-name for a subset of $\omega$ corresponds to a nice name, there are at most $\kappa$ subsets of $\omega$ in $V[G]$. Hence, $V[G] \models \exists \epsilon \leq \kappa$ and we finally get $V[G] \models \epsilon = \kappa$.

Notes

Approaches to forcing. There are different ways of presenting the forcing technique, and even though they all yield precisely the same consistency proofs, they can be quite different in their metamathematical conception. The approach to forcing presented in this chapter is essentially taken from Kunen [4, Chapter VII]. Another approach — taken for example by Jech in [3, Chapter 14] and in [2, Part I, Section 3] — uses Boolean-valued models. For a discussion of different approaches, as well as for some historical background, we refer the reader to Kunen [4, Chapter VII, §9].

Related Results

87. The $\kappa$-chain condition. Let $\kappa$ be a regular cardinal. We say that a forcing notion $\mathbb{P} = (P, \leq)$ satisfies the $\kappa$-chain condition, denoted $\kappa$-cc, if every anti-chain in $P$ has cardinality $< \kappa$ (i.e., strictly less than $\kappa$). In particular, $\omega_1$-cc is equivalent to $\text{ccc}$.

One can show that if a forcing notion $\mathbb{P}$ satisfies the $\kappa$-cc, then forcing with $\mathbb{P}$ preserves all cardinals $\geq \kappa$ (see for example Kunen [4, Chapter VII, Lemma 6.9] or Jech [2, Part I, Section 2]).

88. On the consistency of $2^{\omega_\alpha} > \omega_{n+1}$. With essentially the same construction as in the proof of Theorem 16.4, but replacing the $\text{ccc}$ forcing notion by a similar one satisfying the $\omega_{n+1}$-chain condition, one can show that $2^{\omega_\alpha} = \kappa$ is consistent with ZFC whenever $\text{cf}(\kappa) > \omega_n$. Notice that by Corollary 5.12, the condition $\text{cf}(\kappa) > \omega_n$ is necessary. A more general result is obtained using Easton forcing (see Easton [1] or Chapter 18, Related Result 100).

References

References

Models in which AC fails

In Chapter 7 we have constructed models of Set Theory in which the Axiom of Choice failed. However, these models were models of Set Theory with atoms, denoted \( \mathcal{ZFA} \), where atoms are objects which do not have any elements but are distinct from the empty set. In this chapter we shall demonstrate how one can construct models of Zermelo-Fraenkel Set Theory (i.e., models of ZF) in which AC fails. Moreover, we shall also see how we can embed arbitrary large fragments of permutation models (i.e., models of \( \mathcal{ZFA} \)) into models of ZF.

**Symmetric Submodels of Generic Extensions**

Let \( \mathbf{V} \) be a model of ZFC and let \( \mathcal{P} = (P, \leq) \) be a forcing notion which is defined in \( \mathbf{V} \) with smallest element 0. A mapping \( \alpha : P \to P \) is called an **automorphism** of \( \mathcal{P} \) if \( \alpha \) is a one-to-one mapping from \( P \) onto \( P \) such that for all \( p, q \in P \):

\[
\alpha p \leq \alpha q \iff p \leq q.
\]

In particular we get \( \alpha 0 = 0 \). If \( \alpha \) is an automorphism of \( \mathcal{P} \), then we define, by induction on \( \text{rk}(x) \), an automorphism of the class of \( \mathcal{P} \)-names \( \mathbf{V}^\mathcal{P} \) by stipulating

\[
\alpha x = \{ (\alpha y, \alpha p) : (y, p) \in x \}.
\]

Notice that in particular we have \( \alpha 0 = 0 \). Moreover, if \( x = \{ (y, 0) : y \in x \} \) is the canonical \( \mathcal{P} \)-name for a set \( x \in \mathbf{V} \) and \( \alpha \) is an arbitrary automorphism of \( \mathcal{P} \), then \( \alpha x = x \). Furthermore, with respect to the forcing relationship “\( \Vdash \)”, we have

\[
p \Vdash \varphi(x_1, \ldots, x_n) \iff \alpha p \Vdash \varphi(\alpha x_1, \ldots, \alpha x_n)
\]

where \( \varphi(x_1, \ldots, x_n) \) is a first-order formula with all free variables shown and \( x_1, \ldots, x_n \in \mathbf{V}^\mathcal{P} \) are arbitrary \( \mathcal{P} \)-names.
Let now \( \mathcal{G} \) be an arbitrary but fixed group of automorphisms of \( \mathbb{P} \). In other words, let \( \mathcal{G} \) be an arbitrary subgroup of the automorphism group of \( \mathbb{P} \).

For each \( \mathbb{P} \)-name \( x \) we define the symmetry group \( \text{sym}_\mathcal{G}(x) \subseteq \mathcal{G} \) of \( x \) by stipulating

\[
\text{sym}_\mathcal{G}(x) = \{ \alpha \in \mathcal{G} : \alpha x = x \}.
\]

In particular, if \( x \) is the canonical \( \mathbb{P} \)-name for a set \( x \in V \), then \( \text{sym}_\mathcal{G}(x) = \mathcal{G} \). Further, if \( \beta \in \text{sym}_\mathcal{G}(x) \) and \( \alpha \) is an arbitrary automorphisms of \( \mathbb{P} \), then \((\alpha \beta \alpha^{-1})(\alpha x) = \alpha x\); and therefore

\[
\text{sym}_\mathcal{G}(\alpha x) = \alpha \text{sym}_\mathcal{G}(x) \alpha^{-1},
\]

which shows that \( \beta \in \text{sym}_\mathcal{G}(x) \) iff \( \alpha \beta \alpha^{-1} \in \text{sym}_\mathcal{G}(\alpha x) \).

A set \( \mathcal{F} \) of subgroups of \( \mathcal{G} \) is a normal filter on \( \mathcal{G} \) if for all subgroups \( H, K \) of \( \mathcal{G} \) we have:

- \( \mathcal{G} \subseteq \mathcal{F} \)
- if \( H \in \mathcal{F} \) and \( H \subseteq K \), then \( K \in \mathcal{F} \)
- if \( H \in \mathcal{F} \) and \( K \in \mathcal{F} \), then \( H \cap K \in \mathcal{F} \)
- if \( \alpha \in \mathcal{G} \) and \( H \in \mathcal{F} \), then \( \alpha H \alpha^{-1} \in \mathcal{F} \)

Let \( \mathcal{F} \) be an arbitrary but fixed normal filter on \( \mathcal{G} \). Then \( x \in \check{V}^\mathcal{F} \) is said to be symmetric if \( \text{sym}_\mathcal{G}(x) \subseteq \mathcal{F} \). In particular, canonical \( \mathbb{P} \)-names \( x \) for sets \( x \in V \) are symmetric (since \( \text{sym}_\mathcal{G}(x) = \mathcal{G} \) and \( \mathcal{G} \in \mathcal{F} \)), and if \( x \) is symmetric and \( \alpha \in \mathcal{G} \), then also \( \alpha x \) is symmetric (since \( \text{sym}_\mathcal{G}(x) \subseteq \mathcal{F} \) iff \( \text{sym}_\mathcal{G}(\alpha x) \subseteq \mathcal{F} \)).

The class \( \text{HS} \) of hereditarily symmetric names is defined by induction on \( \check{\text{rk}}(x) \):

\[
x \in \text{HS} \iff x \text{ is symmetric and } \{ y : \exists p \in P(\langle y, p \rangle \in x) \} \subseteq \text{HS}.
\]

Since for all \( x \in V \) and each automorphism \( \alpha \) of \( \mathbb{P} \) we have \( \alpha x = x \), all canonical names for sets in \( V \) are in \( \text{HS} \). Furthermore, if a \( \mathbb{P} \)-name \( x \) is hereditarily symmetric and \( \alpha \in \mathcal{G} \), then also \( \alpha x \) is hereditarily symmetric. Thus, for all \( \alpha \in \mathcal{G} \) we have \( \alpha x \in \text{HS} \) iff \( x \in \text{HS} \).

Now, for any \( G \subseteq P \) which is \( \mathbb{P} \)-generic over \( V \) define

\[
\check{V} = \{ \check{x}[G] : x \in \text{HS} \}.
\]

In other words, \( \check{V} \) is the subclass of \( \check{V}[G] \) which contains all elements of \( V[G] \) that have a hereditarily symmetric \( \mathbb{P} \)-name. Since \( \mathbb{P} \)-names for \( \mathbb{P} \)-generic filters are in general not symmetric, the set \( G \), which belongs to \( V[G] \), is in general not a member of \( \check{V} \). However, \( \check{V} \) is a transitive model of \( \text{ZF} \) which is called symmetric submodel of \( V[G] \).

**Proposition 17.1.** Every symmetric submodel \( \check{V} \) of \( V[G] \) is a transitive model of \( \text{ZF} \) which contains \( V \), i.e., \( V \subseteq \check{V} \subseteq V[G] \) and \( \check{V} \models \text{ZF} \).
Symmetric submodels of generic extensions

Proof (Sketch). Like for the Generic Model Theorem 14.12, we shall prove just a few facts; the remaining parts of the proof are left as an exercise to the reader.

The heredity of the class $\mathbf{HS}$ implies that the class $\mathbf{V}$ is transitive, and by the definition of $\mathbf{V}$ we get $\mathbf{V} \subseteq \mathbf{V}[G]$. Further, since $x \in \mathbf{HS}$ for every $x \in \mathbf{V}$, we get $\mathbf{V} \subseteq \mathbf{V}$.

As a consequence of the transitivity of $\mathbf{V}$ we get that $\mathbf{V}$ satisfies the Axiom of Extensionality as well as the Axiom of Foundation.

To see that the Axiom of Empty Set and the Axiom of Infinity are valid in $\mathbf{V}$, just notice that the canonical $\mathbb{P}$-names for $\emptyset$ and $\omega$ respectively are hereditarily symmetric.

For the Axiom of Pairing, let $x_0$ and $x_1$ be arbitrary sets in $\mathbf{V}$ and let $x_0, x_1 \in \mathbf{HS}$ be $\mathbb{P}$-names for $x_0$ and $x_1$ respectively. Let $y := \{(x_0, 0), (x_1, 0)\}$. Then $y[G] = \{x_0, x_1\}$, and since $y \in \mathbf{HS}$ we get $\{x_0, x_1\} \in \mathbf{V}$.

For the Axiom Schema of Separation, let $\varphi(x, y_1, \ldots, y_n)$ be a first-order formula with $\text{free}(\varphi) \subseteq \{x, y_1, \ldots, y_n\}$. Let $u, a_1, \ldots, a_n$ be sets in $\mathbf{V}$ and let $\bar{u}, g_1, \ldots, g_n$ be the corresponding hereditarily symmetric $\mathbb{P}$-names for these sets. We have to find a hereditarily symmetric $\mathbb{P}$-name for the set

$$w = \{ v \in u : \varphi(v, a_1, \ldots, a_n) \}.$$  

For this, let $\bar{u} := \{(v, p) : \exists q \in P \{ q \leq p \land (v, q) \in u \}\}$ and let

$$w = \{ (v, p) \in \bar{u} : \bar{u} \mathrel{\upharpoonright} \varphi(v, g_1, \ldots, g_n) \}.$$  

Obviously we have $w[G] = w$ and it remains to show that $w \in \mathbf{HS}$. Since $u \in \mathbf{HS}$, also $\bar{u} \in \mathbf{HS}$, and it is enough to show that $\text{sym}_\varphi(w) \in \mathcal{F}$. Let $I := \text{sym}_\varphi(\bar{u}) \cap \text{sym}_\varphi(a_1) \cap \cdots \cap \text{sym}_\varphi(a_n)$. Then $I$, as the intersection of finitely many groups in $\mathcal{F}$, belongs to $\mathcal{F}$. For any $\alpha \in I$ we have $\alpha \bar{u} = \bar{u}$ and for every $1 \leq i \leq n$ we have $\alpha a_i = a_i$. Further we have

$$\alpha w = \{ (v, p) : \gamma \leq p \land \varphi(v, a_1, \ldots, a_n) \}$$  

Thus, $I \subseteq \text{sym}_\varphi(w) \in \mathcal{F}$ and we finally have $w \in \mathbf{HS}$.

As we shall see in the following examples, $\mathbf{V}$ does in general not satisfy the Axiom of Choice. Thus, in general we have $\mathbf{V} \not\models \mathbf{ZFC}$, even though $\mathbf{V}$ as well as $\mathbf{V}[G]$ are models of $\mathbf{ZFC}$. 

$\square$
Examples of Symmetric Models

A model in which the reals cannot be well-ordered

In this section we shall construct a symmetric model $\hat{V}$ in which there exists an infinite set $A$ of real numbers (i.e., $A \subseteq [\omega]^\omega$) such that $A$ is Dedekind-finite in $\hat{V}$, i.e., there is no injection in $\hat{V}$ which maps $\omega$ into $A$.

Consider the forcing notion $\mathbb{C}_\omega = (\text{Fn}(\omega \times \omega, 2), \subseteq)$ consisting of finite partial functions from $\omega \times \omega$ to $\{0, 1\}$. To keep the notation short let $C_\omega := \text{Fn}(\omega \times \omega, 2)$. Recall that the smallest element of $C_\omega$ is 0 and for $p, q \in C_\omega$, $p$ is stronger than $q$ iff the function $p$ extends $q$.

Before we construct the symmetric model $\hat{V}$, let us define a $\mathbb{C}_\omega$-name $\dot{A}$ for a set of reals. For each $n \in \omega$ define the $\mathbb{C}_\omega$-name $\dot{g}_n$ by stipulating

$$g_n = \{ (k, p) : k \in \omega \land p \in C_\omega \land p(n, k) = 1 \}$$

and let

$$\dot{A} = \{ \langle \dot{g}_n, 0 \rangle : n \in \omega \} .$$

First we show that $\dot{A}[G]$ is an infinite set in $V[G]$ whenever $G$ is $\mathbb{C}_\omega$-generic over some model $V$ of ZFC. For this, let $G \subseteq C_\omega$ be an arbitrary $\mathbb{C}_\omega$-generic filter over $V$. Then we obviously have $\dot{A}[G] = \{ \dot{g}_n[G] : n \in \omega \}$. Since for any integers $n, l \in \omega$ the set

$$\{ p \in C_\omega : \exists k \in \omega (k \geq l \land \langle k, p \rangle \in \dot{g}_n) \}$$

is open dense in $C_\omega$ we get $V[G] \models \dot{g}_n[G] \in [\omega]^\omega$. Furthermore, for any distinct integers $n, m \in \omega$, also

$$\{ p \in C_\omega : \exists k \in \omega (\langle n, k \rangle \in \text{dom}(p) \land \langle m, k \rangle \in \text{dom}(p) \land p(\langle n, k \rangle) \neq p(\langle m, k \rangle)) \}$$

is open dense in $C_\omega$ and therefore

$$V[G] \models "\dot{A}[G] \text{ is infinite}".$$  

Now we construct a symmetric submodel $\hat{V}$ of $V[G]$ in which $\dot{A}[G]$ is Dedekind-finite. If $\pi$ is a permutation of $\omega$ (i.e., $\pi$ is a one-to-one mapping from $\omega$ onto $\omega$), then $\pi$ induces an automorphism $\alpha_\pi$ of $\mathbb{C}_\omega$ by stipulating

$$\alpha_\pi p = \{ \langle \pi n, k \rangle, i \rangle : \langle n, k, i \rangle \in p \} ,$$

i.e., $\alpha_\pi p(\langle n, k \rangle) = p(\langle n, k \rangle)$.

Let $\mathcal{G}$ be the group of all automorphisms of $\mathbb{C}_\omega$ that are induced by permutations of $\omega$, i.e.,

$$\mathcal{G} = \{ \alpha_\pi : \pi \text{ is a permutation of } \omega \} .$$
A model in which the reals cannot be well-ordered

For every finite set $E \in \text{fin}(\omega)$ let

$$\text{fix}_E(E) = \{ \alpha_n \in \mathcal{G} : \pi n = n \text{ for each } n \in E \}.$$  

Let $\mathcal{F}$ be the filter on $\mathcal{G}$ generated by the subgroups $\{ \text{fix}_E(E) : E \in \text{fin}(\omega) \}$, i.e., a subgroup $H \subseteq \mathcal{G}$ belongs to $\mathcal{F}$ iff there is an $E \in \text{fin}(\omega)$ such that $\text{fix}_E(E) \subseteq H$. Then $\mathcal{F}$ is a normal filter (notice for example that $\alpha_\omega \text{fix}_E(E) \alpha_\omega^{-1} = \text{fix}_E(E)$ or see Chapter 7).

Finally, let $\mathbf{HS}$ be the class of all hereditarily symmetric $\mathcal{C}_\omega$-names and let $\mathbf{V}$ be the corresponding symmetric submodel of $\mathbf{V}[G]$.

In order to see that the set $\mathcal{A}[G]$ belongs to $\mathbf{V}$ we have to verify that $\mathcal{A} \subseteq \mathbf{HS}$. Firstly notice that each automorphism $\alpha_\pi$ corresponds to a permutation of the set $\{ g_n : n \in \omega \}$. In fact, for each $n \in \omega$ we have

$$\alpha_\pi g_n = \{ (\alpha_\pi k, \alpha_\pi p) : (k, p) \in g_n \}
= \{ (k, \alpha_\pi p) : \alpha_\pi p((\pi n, k)) = 1 \}
= \{ (k, q) : q((\pi n, k)) = 1 \} = g_{\pi n}.$$  

In particular, $\alpha_\pi g_n = g_n$ iff $\pi n = n$. Thus, for each $n \in \omega$, $\text{fix}_E(\{ n \}) = \text{sym}_E(g_n)$, and since $\{ k : \exists p \in C_\omega((k, p) \in g_n) \} \subseteq \mathbf{HS}$, each $g_n$ belongs to $\mathbf{HS}$. Furthermore, for each $\alpha_\pi \in \mathcal{G}$ we have

$$\alpha_\pi \mathcal{A} = \{ (\alpha_\pi g_n, \alpha_\pi \emptyset) : (g_n, \emptyset) \in \mathcal{A} \}
= \{ (g_{\pi n}, \emptyset) : (g_n, \emptyset) \in \mathcal{A} \} = \mathcal{A},$$

which shows that $\text{sym}_E(\mathcal{A}) = \mathcal{G}$. Thus, $\mathcal{A} \subseteq \mathbf{HS}$ which implies that $\mathcal{A}[G]$ belongs to $\mathbf{V}$. In fact, by $\langle \infty \rangle$, $\mathcal{A}[G]$ is an infinite set of reals which belongs to the model $\hat{\mathbf{V}}$, i.e.,

$$\hat{\mathbf{V}} \models \langle \mathcal{A}[G] \subseteq [\omega]^\omega \text{ and } \mathcal{A}[G] \text{ is infinite} \rangle.$$  

On the other hand we shall see that

$$\hat{\mathbf{V}} \models \langle \mathcal{A}[G] \text{ is } \text{D-finite} \rangle.$$  

Assume towards a contradiction that the function $f : \omega \leftrightarrow \mathcal{A}[G]$ is an injection which belongs to the model $\hat{\mathbf{V}}$. Then there is a hereditarily symmetric $\mathcal{C}_\omega$-name $\mathcal{f} \in \mathbf{HS}$ for $f$ and a condition $p \in C_\omega$ such that

$$p \mathcal{F} \mathcal{C}_\omega \mathcal{f} : \omega \leftrightarrow \mathcal{A}.$$  

Let the finite set $E_0 \in \text{fin}(\omega)$ be such that $\text{fix}_E(E_0) \subseteq \text{sym}_E(f)$. Since $f$ is an injective function with $\text{dom}(f) = \omega$, there is an $n_0 \in \omega \setminus E_0$, a $k \in \omega$, and a condition $p_0 \geq p$ such that

$$p_0 \mathcal{F} \mathcal{C}_\omega \mathcal{f}(k) = g_{n_0}.$$
Let now $\pi$ be a permutation of $\omega$ such that $\alpha_\pi \in \text{fix}_G(E_0)$, $\pi \eta \neq \eta$, but $\alpha_\pi \rho_0$ and $\rho_0$ are compatible (i.e., there is an $r \in C_\omega$ such that $\alpha_\pi \rho_0 \leq r \geq \rho_0$). Then the corresponding automorphism $\alpha_\pi \in \mathcal{G}$ belongs to $\text{sym}_G(f)$, in particular $\alpha_\pi f = f$. Recall that $\alpha_\pi k = k$ (for all $k \in \omega$). If $r \in C_\omega$ is such that $\alpha_\pi \rho_0 \leq r \geq \rho_0$, then we have

$$r \models_{C_\omega} f(k) = g_{\rho_0},$$

because $r \geq \rho_0$, as well as

$$r \models_{C_\omega} f(k) = g_{\pi \rho_0},$$

because $r \geq \alpha_\pi \rho_0$. Hence, $r \models_{C_\omega} g_{\rho_0} = g_{\pi \rho_0}$, but this contradicts the fact that $\rho_0 \neq \pi \rho_0 \rightarrow g_{\rho_0}[G] \neq g_{\pi \rho_0}[G]$. Obviously, this shows that there is no hereditarily symmetric name for an injection $f : \omega \leftrightarrow A[G]$, in other words, $\hat{\mathcal{V}} = \text{“}A[G]\text{”}$ is D-finite”.

Conclusion: Starting from a model $\mathcal{V}$ of ZFC we constructed a symmetric model $\hat{\mathcal{V}}$ of ZF in which there exists an infinite but D-finite set of reals. Thus, there is a model of ZF in which the reals cannot be well-ordered. In particular, the Well-Ordering Principle is not provable in ZF.

A model in which every ultrafilter over $\omega$ is principal

The following construction of a symmetric model $\hat{\mathcal{V}}$ in which every ultrafilter over $\omega$ is principal is essentially the same as in the example above, except that the set $\{g_n[G] : n \in \omega\}$ will not belong to the model $\hat{\mathcal{V}}$. Thus, let $\mathcal{V}$ be a model of ZFC and consider again the forcing notion $\mathcal{C}_\omega = (\text{Fn}(\omega \times \omega, 2), \subseteq)$.

For each $n \in \omega$ let $g_n = \{(k, p) : k \in \omega \land p \in C_\omega \land p(n, k) = 1\}$, and let $G \subseteq C_\omega$ be $\mathcal{C}_\omega$-generic over $\mathcal{V}$; then $\mathcal{V}[G] \models g_n[G] \in [\omega]^{\omega}$.

For every $X \subseteq \omega \times \omega$ we define an automorphism $\alpha_X$ of $\mathcal{C}_\omega$ by stipulating

$$\alpha_X p : \text{dom}(p) \mapsto \{0, 1\}$$

$$(n, m) \mapsto \begin{cases} p(n, m) & \text{if } (n, m) \notin X, \\ 1 - p(n, m) & \text{if } (n, m) \in X. \end{cases}$$

Let $\mathcal{G}$ be the group of all automorphisms $\alpha_X$, where $X \subseteq \omega \times \omega$, and let $\mathcal{F}$ be the normal filter on $\mathcal{G}$ generated by $\{\text{fix}_G(E \times \omega) : E \in \text{fin}(\omega)\}$, where

$$\text{fix}_G(E \times \omega) = \{\alpha_X : X \cap (E \times \omega) = \emptyset\}.$$
A model with a paradoxical decomposition of the real line

principal, i.e., \( U \) contains a finite set. Let \( U \in \mathbf{HS} \) be a name for \( U \) and let \( p \in G \) be such that

\[ p \models \alpha \text{ is an ultrafilter over } \omega. \]

Let \( E_0 \in \text{fin}(\omega) \) be such that \( \text{fix}_q(E_0 \times \omega) \subseteq \text{sym}_q(\mathbb{U}) \) and fix an natural number \( l \notin E_0 \). Then there is a \( q \geq p \) such that \( q \in \Delta_{\alpha_q, G} \), i.e., \( q \in G \) and \( q \) decides whether or not \( \alpha \in U \). Let us assume that \( q \models \alpha \models \mathbb{U} \) (the case when \( q \models \alpha \models \mathbb{U} \) is similar). Let \( m_0 \) be such that for all integers \( m \geq m_0 \) we have \( \langle l, m \rangle \notin \text{dom}(q) \) and let

\[ X_0 = \{ \langle l, m \rangle : m \geq m_0 \} \subseteq \omega \times \omega. \]

Let \( U := U[G], a_i := q_i[G], \) and for \( b_i := \alpha_{X_i} a_i \) let \( b_i := \beta_i[G] \). Then, for each \( m \geq m_0, m \in a_i \leftrightarrow m \notin b_i \), which implies that \( (\omega \setminus a_i) \cap (\omega \setminus b_i) \) is finite. Notice that since \( q \models \alpha \models \mathbb{U}, \alpha_{X_i} q \models \alpha \models \mathbb{U} \). By definition of \( X_0 \) we further have \( \alpha_{X_0} \models \text{fix}_q(E_0 \times \omega) \subseteq \text{sym}_q(\mathbb{U}) \) and therefore \( \alpha_{X_0} \models \mathbb{U} \), and since \( \alpha_{X_0} q = q \) and \( \alpha_{X_0} a_i = b_i \) we have \( q \models \alpha \models \mathbb{U} \). Thus, since \( q \in G \), we get that neither \( a_i \) nor \( b_i \) belongs to \( U \). Because \( U \) is an ultrafilter, \( \omega \setminus a_i \) as well as \( \omega \setminus b_i \) belongs to \( U \), and therefore \( (\omega \setminus a_i) \cap (\omega \setminus b_i) \in U \). Hence, \( U \) contains a finite set, or in other words, \( U \) is principal.

**Conclusion:** Starting from a model \( V \) of ZFC we constructed a symmetric model \( V \) of ZF in which every ultrafilter over \( \omega \) is principal. Thus, there is a model of ZF in which for example the Fréchet ideal cannot be extended to a prime ideal. In particular we get that the Prime Ideal Theorem is not provable in ZF.

**A model with a paradoxical decomposition of the real line**

Below, we shall construct a model of ZF in which the real line \( R \) can be partitioned into a family \( \mathcal{R} \), such that \( |\mathcal{R}| > |R| \). (Recall that \( \mathcal{R} \) is a partition of \( R \) if \( \mathcal{R} \subseteq \mathcal{P}(R) \) such that \( \bigcup \mathcal{R} = R \) and for any distinct \( x, y \in \mathcal{R} \), \( x \cap y = \emptyset \).)

By Corollary 4.13 it is enough to construct a model in which the set of reals \( \mathcal{P}(\omega) \) is a countable union of countable sets.

In order to construct a symmetric model in which \( \mathcal{P}(\omega) \) is a countable union of countable sets we start with a model \( V \) of ZFC such that for each \( n \in \omega, V \models z^\omega_n = \omega_{n+1} \). Such a model is for example Gödel’s constructible universe \( L \). Alternatively, such a model is also obtained by an iterated application of Theorem 14.23, or more precisely, by iterating the forcing notions of Theorem 14.23 using the iteration technique given in Chapter 18 (see also Related Result 100 of that chapter).

Now, let

\[ P = \{ p \in \text{Fn}(\omega \times \omega, \omega_n) : \forall (n, m) \in \text{dom}(p) \ (p(n, m) \in \omega_n) \}. \]

Then \( \mathbb{P} := (P, \subseteq) \) is a forcing notion.
Let \( G \subseteq P \) be \( \mathbb{P} \)-generic over \( V \). We construct a symmetric submodel \( \hat{V} \) of \( V[G] \) such that in \( \hat{V} \), the set of reals is a countable union of countable sets. For this, let \( \mathcal{F} \) be the group of all permutations \( \pi \) of \( \omega \times \omega \) such that

\[
\pi(n, i) = (m, j) \rightarrow n = m.
\]

Now, for each \( \pi \in \mathcal{F} \) and every \( n \in \omega \) let \( \pi_n \) be the permutation of \( \omega \) such that for every \( i \in \omega \),

\[
\pi(n, i) = (n, \pi_n(i)).
\]

Every \( \pi \in \mathcal{F} \) induces an automorphism \( \alpha_\pi \) of \( \mathbb{P} \) by stipulating

\[
\alpha_\pi p = \left\{ \left( (n, \pi_n i), \alpha \right) : \left( (n, i), \alpha \right) \in p \right\}.
\]

For every \( n \in \omega \), let \( H_n \) be the group of all \( \pi \in \mathcal{F} \) such that for all \( k \in n \), the corresponding permutation \( \pi_k \) is the identity, and let \( \mathcal{F} \) be the filter on \( \mathcal{F} \) generated by the subgroups \( \{ H_n : n \in \omega \} \). We leave it as an exercise to the reader to verify that \( \mathcal{F} \) is a normal filter. Finally, let \( \hat{V} \) be the symmetric submodel of \( V[G] \) which is determined by \( \mathcal{F} \).

Now, we show that there are countably many countable sets of reals \( R_n \) in \( \hat{V} \) such that \( \hat{V} \models \mathbb{P}(\omega) = \bigcup_{n \in \omega} R_n \). Firstly we construct canonical names for reals in \( \hat{V} \): Let \( x \in \text{HS} \) be a name for a real (i.e., for a subset of \( \omega \)), or more precisely, let \( x \subseteq \left\{ (k, p) : k \in \text{HS} \land p \in P \right\} \) be such that for each \( (k, p) \in \hat{x} \),

\[
p \vdash P k \in \omega \quad \text{(notice that we also have } p \vdash \neg k \in \omega)\]

Since \( x \in \text{HS} \) there is an \( n_0 \in \omega \) such that \( H_{n_0} \subseteq \text{sym}_G(x) \), which implies that for all \( \alpha_\pi \in H_{n_0} \) we have

\[
x = \left\{ (k, p) : (k, p) \in \hat{x} \right\} = \left\{ (\alpha_\pi k, \alpha_\pi p) : (k, p) \in \hat{x} \right\} = \alpha_\pi x.
\]

With respect to \( x \), the canonical name \( \bar{x} \in \text{HS} \) is defined as follows:

\[
\bar{x} = \left\{ (m, q) : \exists (k, p) \in \hat{x} \exists r \geq p \quad q = r|_{n_0 \times \omega} \land r \vdash m = k \right\}
\]

**Claim:** \( V[G] \models \bar{x}[G] = \bar{x}[G] \).

**Proof of Claim.** First we show that \( \bar{x}[G] \subseteq \bar{x}[G] \): Let \( (m, q) \) be an arbitrary but fixed element of \( \bar{x} \) such that \( q \in G \). In particular, \( m[G] \in \bar{x}[G] \). We show that \( m[G] \in \bar{x}[G] \). By definition of \( x \), there is a \( (k, p) \in \hat{x} \) and a condition \( r_0 \geq p \) such that \( q = r_0|_{n_0 \times \omega} \) and \( r_0 \vdash m = k \) and \( k \in \hat{x} \). Now, for every condition \( r' \geq q \) we can find an automorphism \( \alpha_\pi \in H_{n_0} \) and a condition \( r \) such that \( r' \leq r \geq \alpha_\pi r_0 \) which implies that \( r \vdash \bar{m} = \alpha_\pi k \land \alpha_\pi \bar{x} \in \bar{x} \) (recall that \( \alpha_\pi \bar{x} = \bar{x} \) and that for all \( \pi \in \mathcal{F} \), \( \alpha_\pi \bar{m} = m \)). Since \( \alpha_\pi \bar{x} \in H_{n_0} \) we get \( \alpha_\pi \bar{r}|_{n_0 \times \omega} = \bar{r}|_{n_0 \times \omega} = q \) and therefore the set \( \{ r \geq q : r \vdash \bar{m} \in \bar{x} \} \) is dense above \( q \). Thus, \( m[G] \in \bar{x}[G] \), and since \( (m, q) \in \bar{x} \) was arbitrary (with the property that \( q \in G \), we get \( V[G] \models \bar{x}[G] \subseteq \bar{x}[G] \).

Now we show that \( \bar{x} \subseteq \bar{x} \): If \( V[G] \models \bar{m} \in \bar{x}[G] \), then there exist an \( r \in G \) and a name \( (k, p) \in \hat{x} \) such that \( r \geq p \) and \( r \vdash m = k \in \hat{x} \), which implies that \( (m, r)|_{n_0 \times \omega} \in \bar{x} \) and shows that \( V[G] \models \bar{x}[G] \subseteq \bar{x}[G] \).
A model with a paradoxical decomposition of the real line

Thus, each real \( x \in \hat{V} \) (i.e., each subset of \( \omega \) in \( \hat{V} \)) has a canonical name \( \dot{x} \) which is a subset of \( \{ (m, q) : m \in \omega \land q \in P_n \} \), where \( n_0 \in \omega \) and \( P_n := \{ p \in P : \forall (n, m) \in \text{dom}(p) \} \). If \( x \) is a canonical name for a real \( x \in V \) with \( Q_x \subseteq P_n \), where \( Q_x = \{ q \in P : \exists m ((m,q) \in x) \} \), then \( \text{sym}_x(x) \supseteq H_n \) and since \( m \in \text{HS} \) for any \( m \in \omega, x \in \text{HS} \). Moreover, for every \( \alpha \in \mathcal{G} \), if \( \dot{x} \) is a canonical name for a real then also \( \dot{\alpha}x \) is a canonical name for a real. To see this, let \( x \in \text{HS} \) be a name for some real \( x \in \hat{V} \), let \( \dot{x} \) be the canonical name for \( x \) which corresponds to \( x \), and let \( \dot{\alpha} \in \mathcal{G} \). Then \( \dot{\alpha}x \) is a hereditarily symmetric name for a real in \( \hat{V} \) with corresponding canonical name \( \dot{\alpha}x \).

Now, for each \( n \in \omega \) let

\[
R_n = \{ (x, \emptyset) : x \text{ is a canonical name for a real } x \text{ with } Q_x \subseteq P_n \}.
\]

Notice that \( R_n \) is in \( V \) and that for each \( n \in \omega \) and all \( \alpha \in \mathcal{G} \) we have \( \dot{\alpha}R_n = R_n \), which shows that \( \text{sym}_G(R_n) = \emptyset \), and since \( \text{sym}_G(x) \supseteq H_n \) for all \( x \in R_n \), we even have \( R_n \subseteq \text{HS} \), i.e., \( R_n[G] \in \hat{V} \). Moreover, also the function which maps each \( n \in \omega \) to \( R_n[G] \) belongs to \( \hat{V} \) (notice that the name \( \{ (\text{op}(n, R_n), \emptyset) : n \in \omega \} \) is hereditarily symmetric). Further, the set \( \bigcup \{ R_n[G] : n \in \omega \} \) contains all reals in \( \hat{V} \). So, in order to prove that the set of reals in \( \hat{V} \) can be written as a countable union of countable sets, it is enough to prove that each \( R_n[G] \) is countable in \( \hat{V} \), which is done in two steps:

Firstly recall that \( V = \{ \lambda \in \omega_n : n \in \omega \} \) for each \( n \in \omega \). Now, by counting (in the ground model \( V \)) the canonical names which belong to \( R_n \) we get that for each \( n \in \omega \), \( |R_n| = (\omega_{n+1})^\omega \).

Secondly, for each \( n \in \omega \) define

\[
f_n = \{ (\text{op}(k, q), p) : p \in P_{n+1} \land (n, k) \in \text{dom}(p) \land p((n, k)) = \alpha \}.\]

Then, for every \( n \in \omega \), \( f_n \) is a name for a function from \( \omega \) to \( \omega_n \), \( \text{sym}_G(f_n) \supseteq H_{n+1} \), and \( f_n \in \text{HS} \), hence \( f_n[G] \in \hat{V} \). Moreover, \( f_n[G] : \omega \to \omega_n^\omega \) is surjective which implies that \( \omega_n^\omega \) is countable in \( \hat{V} \). Now, since \( |R_n| = (\omega_{n+1})^\omega \) (for each \( n \in \omega \)), each \( R_n[G] \) is countable in \( \hat{V} \) whereas \( \bigcup \{ R_n[G] : n \in \omega \} = \mathcal{P}(\omega)^\omega \) is uncountable in \( \hat{V} \).

**Conclusion:** Starting from a model \( V \) of \( \text{ZFC} + \forall n \in \omega (2^{\omega_n} = \omega_{n+1}) \) we constructed a symmetric model \( \hat{V} \) of \( \text{ZF} \) in which the set of reals is a countable union of countable sets. In particular, this shows that without some form of AC we cannot prove that countable unions of countable sets is countable. Furthermore, we get that in the absence of AC it might be possible that the real line \( \mathbb{R} \) can be partitioned into a family \( \mathcal{A} \), such that \( |\mathcal{A}| > |\mathbb{R}| \). Moreover, by Fact 4.3 we know that \( |[0,1]|^2 = |\mathbb{R}| \) is provable in \( \text{ZF} \) only, and therefore we get that in the absence of AC, it might be possible to decompose a square into more parts than there are points on the square.
Simulating Permutation Models by Symmetric Models

The following theorem provides a method which enables us to embed an arbitrarily large fragment of a given permutation model (i.e., a model of ZFA) into a well-founded model of ZF. In particular, if \( \varphi \) is a statement which holds in a given permutation model and whose validity depends only on a certain fragment of that model, then there is a well-founded model of ZF in which \( \varphi \) holds as well. For example assume that there are two sets \( R \) and \( S \) in some permutation model \( \mathcal{V} \) of ZFA such that \( \mathcal{V} \models |R| < |S| \land |S| \leq^* |R| \), i.e., there is an injection from \( R \) into \( S \), a surjection from \( R \) onto \( S \), but no bijection between the two sets (cf. Theorem 4.21 and Proposition 7.14). Notice that the surjection from \( R \) onto \( S \) induces a partition \( \mathcal{R} \) of \( R \) of cardinality \( |S| \), i.e., \( |\mathcal{R}| > |R| \). Now, the validity of the sentence \( \exists R \exists S (|R| < |S| \land |S| \leq^* |R|) \), which holds in \( \mathcal{V} \), depends only on a certain fragment of that model, and thus, by the following theorem, there is a well-founded model of ZF in which we find sets \( \tilde{R} \) and \( \tilde{S} \) such that \( |\tilde{R}| < |\tilde{S}| \land |\tilde{S}| \leq^* |\tilde{R}| \).

**Theorem 17.2 (Jech-Sochor Embedding Theorem).** Let \( \mathcal{V} \models \text{ZFA} \) be a permutation model in which AC holds in the kernel of \( \mathcal{V} \). Furthermore, let \( A \) be the set of all atoms of \( \mathcal{V} \), let \( \gamma \) be an arbitrary but fixed ordinal number, and let \( \mathcal{V}_\gamma := \mathcal{P}^\gamma(A) \cap \mathcal{V} \). Then there exist a symmetric model \( \mathcal{V}' \) (i.e., a model of ZF) and an embedding \( x \mapsto \hat{x} \) of \( \mathcal{V} \) into \( \mathcal{V}' \) whose restriction to \( \mathcal{V}_\gamma \) is an \( \varepsilon \)-isomorphism between the sets \( \mathcal{V} \), and \( \mathcal{P}^\gamma(A)^{\mathcal{V}} \), where \( f : S \to T \) is an \( \varepsilon \)-isomorphism between \( S \) and \( T \) if \( f \) is a bijection and for all \( x, y \in S \), \( x \in y \iff f(x) \in f(y) \). In other words, one can simulate arbitrarily large fragments of permutation models by symmetric models, which is visualised by the following figure:

![Diagram](image)

**Proof.** Let \( \mathcal{M} \) be a model of ZFA + AC and let \( \mathcal{V} := \mathcal{P}^\infty(\emptyset) \subseteq \mathcal{M} \) be the kernel of \( \mathcal{M} \); then \( \mathcal{V} \models \text{ZFC} \). Let \( A_0 \) be the set of all atoms of \( \mathcal{M} \). We consider
a group \( G_0 \) of permutations of \( A_0 \) and a normal filter \( \mathcal{F}_0 \) on \( G_0 \), and let \( V \subseteq \mathcal{M} \) be the permutation model (i.e., a model of ZFA) given by \( G_0 \) and \( \mathcal{F}_0 \). Further, let \( \gamma \) be an arbitrary but fixed ordinal number and let \( V_\gamma := \mathcal{P}^\gamma(A) \cap V \).

In order to construct a symmetric submodel of a generic extension, we have to work in a ground model of ZFC. So, we shall work in the model \( V \) and first construct a generic extension \( V[G] \) of \( V \): Let \( A \) be a set in \( V \) such that \( \mathcal{M} \models |A| = |A_0| \) and fix in \( \mathcal{M} \) a bijection \( \iota : A_0 \to A \). Let \( \kappa \) be a regular cardinal (in \( V \)) such that \( \kappa > |\mathcal{P}^\gamma(A)| \). The set \( P \) of forcing conditions consists of functions \( p : \text{dom}(p) \to \{0, 1\} \) such that \( \text{dom}(p) \subseteq (A \times \kappa) \times \kappa \) and \(|\text{dom}(p)| < \kappa \). As usual let \( p \leq q \iff p \subseteq q \). Then, by the choice of \( \kappa \), \( P = (P, \leq) \) is a \( \kappa \)-closed forcing notion. Below, for \( p \in P \) and \( \langle \langle a, \xi, \eta \rangle \rangle \in \text{dom}(p) \) we shall write \( p(a, \xi, \eta) \) instead of \( p(\langle \langle a, \xi, \eta \rangle \rangle) \). For each \( a \in A_0 \) and each \( \xi \in \kappa \) let

\[ x_{a, \xi} = \{ (y, p) : p(a, \xi, \eta) = 1 \}, \]

and for each \( a \in A_0 \) define

\[ q = \{ (x_{a, \xi}, \emptyset) : \xi \in \kappa \} \]

and let \( A = \{ q : a \in A_0 \} \). Having now defined \( q \) for each \( a \in A_0 \), by transfinite recursion we define \( x \) for each \( x \in \mathcal{M} \) by stipulating

\[ x = \{ (q, 0) : \mathcal{M} \models y \in x \}. \]

**Claim 1.** If \( G \) is \( \mathbb{P} \)-generic over \( V \), then for all \( x, y \in \mathcal{M} \):

\[ \mathcal{M} \models y \in x \iff V[G] \models y[G] \in x[G] \]

\[ \mathcal{M} \models y = x \iff V[G] \models y[G] = x[G] \]

**Proof of Claim 1.** Notice first that \( x_{a, \xi}[G] \neq x_{a', \xi'}[G] \) whenever \( \langle a, \xi \rangle \neq \langle a', \xi' \rangle \), that \( x_{a, \xi}[G] \neq x[G] \) whenever \( x \in V \), and that for all \( x \in \mathcal{M} \) and \( a \in A_0 \), \( x[G] \notin q[G] \). Consequently we have \( q[G] \neq q'[G] \) whenever \( a \neq a' \) are atoms and that the atoms do not contain any elements of the form \( x[G] \). Further, for all \( a \in A_0 \), all \( \xi \in \kappa \), and every \( x \in \mathcal{M} \), we have \( x[G] \neq x_{a, \xi}[G] \). To see this, notice that on the one hand, for all \( x \in V \) we have \( x[G] = x[G] \) and therefore \( x[G] \neq x_{a, \xi}[G] \); on the other hand, if \( x \in \mathcal{M} \setminus V \) then \( \text{TC}(x) \) (i.e., the transitive closure of \( x \)) contains an atom \( a_0 \in A_0 \), and hence, \( x_{a_0, \xi}[G] \in \text{TC}(x[G]) \) (for every \( \xi \in \kappa \)), whereas for example \( x_{a_0, \xi}[G] \notin \text{TC}(x[G]) \).

Now we can prove the claim simultaneously for \( "\in" \) and \( "=" \) by induction on rank, where, for a set \( x \), \( \text{rk}_\mathcal{M}(x) \) is the least \( \alpha \in \Omega \) such that \( x \in \mathcal{P}^\alpha(A_0) \). Notice that \( \text{rk}_\mathcal{M}(\emptyset) = 1 \), whereas \( \text{rk}_\mathcal{M}(a) = 0 \) for all atoms \( a \in A_0 \). Assume that the claim is valid for \( y \in z \) and \( y = z \) whenever \( \text{rk}_\mathcal{M}(z) < \text{rk}_\mathcal{M}(x) \); we shall show that the claim is also valid for \( y \in x \) and \( y = x \).

**Case (\( "\in" \)):** If \( \mathcal{M} \models y \in x \), then \( V[G] \models y[G] \in x[G] \) follows by definition of \( x \). Conversely, if \( V[G] \models y[G] \in x[G] \), then \( x \) can neither be the name for
an atom nor for the empty set, since otherwise we would have $p \models y \in x$ (for
some $p \in P$), which is obviously impossible. Hence, $V[G] \models y[G] = z[G]$ for
some $z \in x$ (i.e., $z \in x$), and we have $\mathcal{M} \models y = z$ by the induction hypothesis,
thus $\mathcal{M} \models y \in x$.

(=): Obviously, if $\mathcal{M} \models y = x$, then $V[G] \models y[G] = x[G]$. Conversely, if
$\mathcal{M} \models y \neq x$, then either both $x$ and $y$ are atoms or the empty set and then
$V[G] \models y[G] \neq x[G]$; or for example $x$ contains some $z$ which is not in $y$,
and then, by the $\mathcal{M}$ already proceed, $V[G] \models z[G] \in x[G] \setminus y[G]$, hence,

Notice that the proof of Claim 1 does not depend on the particular $P$-generic
filter $G$.

The next step is to construct a symmetric submodel $\tilde{V}$ of $V[G]$ which
reflects to some extent the model $V$. We define a group $G$ of automorphisms
of $P$ and a normal filter $\mathcal{F}$ on $G$ as follows. For every permutation $\sigma$ of $A_0$,
let $\tilde{\sigma}$ be the set of all permutations $\pi$ of $\tilde{A} \times \kappa$ such that for all $a \in A_0$
and all $\xi \in \kappa$:

$$\pi(a, \xi) = \langle \sigma(a), \xi' \rangle$$

for some $\xi' \in \kappa$.

One can visualise the set $\tilde{A} \times \kappa$ as a set $\tilde{A}$ of pairwise disjoint blocks, each block
consisting of $\kappa$ elements. Every permutation $\sigma$ of $A_0$ induces a permutation
$\sigma'$ of the blocks and every $\pi \in \tilde{\sigma}$ permutes the elements of $\tilde{A} \times \kappa$ in such a
way that $\pi$ acts on the blocks exactly as $\sigma'$ does.

Let

$$G = \bigcup \{ \tilde{\sigma} : \sigma \in G_0 \}$$

and for every subgroup $H$ of $G_0$ let $H = \bigcup \{ \tilde{\sigma} : \sigma \in H \}$. Since every permutation $\pi$ of $\tilde{A} \times \kappa$ corresponds to an automorphism of $P$ by stipulating

$$\pi_p (\pi(a, \xi), \eta) := p(\tilde{a}, \xi, \eta),$$

we consider $G$ as well as its subgroups as groups of automorphisms of $P$. For
every finite $E \in \text{fin}(\tilde{A} \times \kappa)$ let

$$\text{fix}_G(E) = \{ \pi \in G : \pi x = x \text{ for each } x \in E \}.$$  

We let $\mathcal{F}$ be the filter on $G$ generated by

$$\{ \tilde{H} : H \in \mathcal{F}_0 \} \cup \{ \text{fix}_G : E \in \text{fin}(\tilde{A} \times \kappa) \}.$$  

We leave it as an exercise to the reader to check that $\mathcal{F}$ is a normal filter.

Now, let $\mathbf{HS}$ be the class of all hereditarily symmetric names (with respect
to $G$ and $\mathcal{F}$), let $G$ be $P$-generic over $V$, and let $\tilde{V} = \{ x[G] : x \in \mathbf{HS} \}$ be the
corresponding symmetric submodel of $V[G]$. As an immediate consequence of
the definition of $\mathcal{F}$ we have:

- $x_{a\xi}[G] \in \tilde{V}$ for all $a \in A_0$ and $\xi \in \kappa$, because $\text{sym}_G(x_{a\xi}) = \text{fix}_G (\{ (a, \xi) \})$. 
Simulating permutation models by symmetric models

- \( \sigma[G] \in \hat{V} \) for all \( a \in A_0 \), because \( \text{sym}_G(a) = \text{sym}_{\sigma[G]}(a) \), i.e., for every \( \sigma \in \text{sym}_G(a) \), \( \hat{\sigma} \subseteq \text{sym}_{\sigma[G]}(a) \).
- \( A[G] \in \check{V} \), because \( \text{sym}_G(A) = \mathcal{G} \).

Below, we shall write \( \hat{x} \) for \( x[G] \). So, in particular we have \( a \in \hat{V} \) and \( \hat{A} \in \check{V} \), i.e., the "atoms" (more precisely, the surrogates of atoms introduced by the forcing) as well as the set of all "atoms" belongs to the model \( \check{V} \).

The next task is to show that \( x \in \check{V} \) iff \( \hat{x} \in \hat{V} \), which is done in the following two steps.

**Claim 2.** For all \( x \in \mathcal{M} : x \in \mathcal{V} \iff x \in \text{HS} \).

**Proof of Claim 2.** It suffices to show that

\[ \text{sym}_{\sigma[G]}(x) \in \check{F}_0 \iff \text{sym}_G(x) \in \hat{F}. \]

If \( \sigma \in \mathcal{G}_0 \) and \( \pi \in \sigma \), then \( \sigma x \) is the canonical name for \( \sigma x \), and therefore \( \text{sym}_G(x) = \text{sym}_{\sigma[G]}(x) \). Thus, if \( \text{sym}_{\sigma[G]}(x) \in \check{F}_0 \), then \( \text{sym}_G(x) \in \hat{F} \). On the other hand, if \( \text{sym}_G(x) \in \hat{F} \), then \( \text{sym}_{\sigma[G]}(x) \subseteq H \cap \text{fix}_G(E) \) for some \( H \in \mathcal{F}_0 \) and a finite set \( E \in \text{fin}(A \times \kappa) \). Let \( E\{A_0 = \{a \in A_0 : \exists x((a, \xi) \in E)\} \). Then \( \text{sym}_{\sigma[G]}(x) \subseteq H \cap \text{fix}_G(E|A_0) \), and since \( \mathcal{F}_0 \) is a normal filter on \( \mathcal{G}_0 \) we have \( \text{fix}_G(E|A_0) \in \mathcal{F}_0 \) and hence \( \text{sym}_{\sigma[G]}(x) \in \mathcal{F}_0 \).

**Claim 3.** For all \( x \in \mathcal{M} : x \in \mathcal{V} \iff \hat{x} \in \hat{V} \).

**Proof of Claim 3.** By **Claim 2**, it suffices to show that if \( \hat{x} \in \hat{V} \), then \( x \in \mathcal{V} \).

Assume towards a contradiction that there exists an \( x \in \mathcal{M} \) such that \( \hat{x} \in \hat{V} \) and \( x \notin \mathcal{V} \), but for all \( y \in x, y \in \mathcal{V} \). Thus \( x \notin \mathcal{V} \), and since \( \hat{x} \in \hat{V} \), there exist a name \( \hat{z} \in \text{HS} \) and a condition \( p_0 \in G \) such that \( p_0 \vdash \pi \hat{z} = x \). In other words, \( x \notin \text{HS} \) but there exists a name \( \hat{z} \in \text{HS} \) such that \( \hat{x} = \hat{z}[G] \), and consequently \( \hat{x} \in \hat{V} \). Since we have \( \text{sym}_G(\hat{z}) \in \hat{F} \), there is a group \( H_0 \in \mathcal{F}_0 \) and a finite set \( E_0 \in \text{fin}(A \times \kappa) \) such that \( \text{sym}_G(\hat{z}) \supseteq H_0 \cap \text{fix}_G(E_0) \).

Assume there are permutations \( \sigma \in \mathcal{G}_0 \) and \( \pi \in \sigma \) such that

(a) \( \pi \in H_0 \cap \text{fix}_G(E_0) \),
(b) \( \sigma \neq \pi \), and
(c) \( \pi p_0 \) and \( p_0 \) are compatible.

Then we have \( \pi \hat{z} = \hat{z} \) by (a), \( p_0 \vdash \pi \hat{z} \neq x \) by (b) and **Claim 1**, and since \( \pi p_0 \vdash \pi \hat{z} = \pi x \), by (c) there is a \( q_0 \in P \) such that \( \pi p_0 \leq q_0 \geq p_0 \) and

\[ q_0 \vdash (\hat{z} = x) \land (x \neq \pi x) \land (\pi x = \hat{z}), \]

a contradiction. To see that permutations \( \sigma \) and \( \pi \) with the above properties exist, notice first that since \( x \) is not symmetric (i.e., \( x \notin \mathcal{V} \)), there exists a \( \sigma \in H_0 \cap \text{fix}_G(E_0|A_0) \) such that \( \sigma \neq \pi \). Since \( |\text{dom}(p)| < \kappa \), there is a \( \delta \in \kappa \) such that

\[ \{ (a, \xi) : a \in A_0 \land \delta \leq \xi \in \kappa \} \cap (\text{dom}(p) \cup E_0) = \emptyset, \]

and we define \( \pi \in \sigma \) as follows.
• If $a \in E_0|A_0$, then for all $\xi \in \kappa$:
  \[ \pi(a, \xi) = \langle a, \xi \rangle. \]

• If $a \not\in E_0|A_0$ and $\xi \in \delta$, then
  \[ \pi(a, \xi) = \langle \nu(a), \delta + \xi \rangle. \]

• If $a \not\in E_0|A_0$ and $\xi \in \delta + 1 \in \kappa$, then
  \[ \pi(a, \delta + \xi) = \langle \nu(a), \delta + \xi \rangle. \]

By definition it follows that $\pi \in \hat{H}_0 \cap \text{fix} g(E_0)$ and that $\pi p_0$ and $p_0$ are compatible.

The final step in the proof of Theorem 17.2 is to show that the embedding $x \mapsto \dot{x}$ is a bijection between $V_\gamma$ and $\mathcal{P}^\gamma(\dot{A})^\dot{V}$.

Claim 4. $\{ \dot{x} : x \in V_\gamma \} = \mathcal{P}^\gamma(\dot{A})^\dot{V}$

Proof of Claim 4. By Claim 3, the left-hand side is included in the right-hand side; thus, it suffices to show that $\mathcal{P}^\gamma(\dot{A})^\dot{V} \subseteq \{ \dot{x} : x \in V_\gamma \}$, which will be done by transfinite recursion: Let $x \in V_\gamma$, and let $y \in \dot{V}$ be such that $\dot{V} \models y \in \dot{x}$. We have to show that $y = \dot{z}$ for some $z \in V$. Let $y$ be a $\mathbb{P}$-name for $y$. Since $\mathbb{P}$ is $\kappa$-closed and $\kappa > |x|$ (since $\kappa > |\mathcal{P}^\gamma(\dot{A})|$), there is a $p \in G$ which decides $y \in y$ for all $u \in x$, more formally, $p \in G \cap \bigcap_{u \in x} \Delta_u \in y$. Hence, $y \dot{=} \dot{z}$, where $z = \{ u \in x : p \Vdash u \in y \}$, and since $\dot{z} \in \dot{V}$, by Claim 3 we get $z \in V$. Finally, by Claim 4 we get that the embedding $x \mapsto \dot{x}$ of $V$ into $\dot{V}$ is such that $\{ \dot{x} : x \in V_\gamma \} = \mathcal{P}^\gamma(\dot{A})^\dot{V}$, and for all $x, y \in V_\gamma$ we have $V \models y \in x$ iff $\dot{V} \models \dot{y} \in \dot{x}$, which shows that $V_\gamma$ and $\mathcal{P}^\gamma(\dot{A})^\dot{V}$ are indeed $\in$-isomorphic, i.e., the embedding $x \mapsto \dot{x}$ restricted to $V_\gamma$ is an $\in$-isomorphism between $V_\gamma$ and $\mathcal{P}^\gamma(\dot{A})^\dot{V}$.

Corollary 17.3. Let $\nu$ be an ordinal and let $\varphi$ be a sentence of the form $\exists X \psi(X, \nu)$, where the only quantifiers we allow in $\psi$ are the restricted quantifiers $\exists \gamma \in \mathcal{P}^\nu(X)$ and $\forall u \in \mathcal{P}^\nu(X)$. If $V \models \text{ZFA}$ is a permutation model in which AC holds in the kernel and $V \models \varphi$, then there exists a symmetric model $\dot{V} \models \text{ZF}$ such that $\dot{V} \models \varphi$.

Proof. Let $X \in V$ be such that $V \models \psi(X, \nu)$ and let $\gamma \in \Omega$ be such that $\mathcal{P}^\nu(X) \subseteq \mathcal{P}^\nu(\dot{A})$, where $\dot{A}$ is the set of atoms of $\dot{V}$. By the Jech-Sochor Embedding Theorem 17.2 there exists a symmetric model $\dot{V}$ of ZF such that $V_\gamma$ and $\mathcal{P}^\gamma(\dot{A})$ are $\in$-isomorphic. Now, by the choice of $\gamma$ and since $V \models \psi(X, \nu)$ we have $(V_\gamma, \in) \models \psi(X, \nu)$, and therefore $(V_\gamma, \in) \models \varphi$. Hence, $(\mathcal{P}^\gamma(\dot{A}), \in) \models \varphi$ which shows that $\dot{V} \models \varphi$. 

---
Applications: Most of the results of Chapter 7 — obtained by permutation models — can now be transferred to proper models of ZF. For example the existence of a set $X$, such that $|X^2| < |\mathcal{P}(X)|^2$ is consistent with ZF (cf. Proposition 7.18), or in other words, ZF $\not\models \forall X \left( |X^2| \not< |\mathcal{P}(X)|^2 \right)$. Similarly we can show that ZF $\not\models \forall X \left( |\text{seq}(X)| \not< |\text{fin}(X)| \right)$ (cf. Proposition 7.17).

Notes

Symmetric submodels of generic extensions. The idea of using symmetry arguments to construct models in which the Axiom of Choice fails goes back to Fraenkel [6]. Cohen incorporated the symmetry arguments into his method and constructed for example the model given above in which the reals are not well-orderable. The formulation of Cohen’s method in terms of symmetric submodels of generic extensions is due to Scott and Jech (cf. Jech [11, Chapter 15]).

Three examples of symmetric models. The first model (i.e., the one in which the reals are not well-orderable) is due to Cohen (cf. [3, Chapter IV, §9]) and is sometimes called the basic Cohen model (cf. Jech [9, Chapter 5, §3]); the second model we presented (i.e., the one in which every ultrafilter over $\omega$ is principal) is due to Feferman [4]; and the third model (i.e., the one in which the set of reals is a countable union of countable sets) is due to Feferman and Levy [5]. However, the constructions can also be found in Jech [11, Chapter 15], and in greater detail in Jech [10, Chapter 3, Section 21] and [9, Chapter 10, §1] respectively.

Simulating permutation models by symmetric models. The Jech-Schönh Embedding Theorem 17.2 is due to Jech and Schönh [12, 13], where numerous applications of the theorem are given in the second paper [13] (see also Jech [9, Theorem 6.1] and [11, Chapter 13]). The limits of the Jech-Schönh Embedding Theorem 17.2 are discussed in Related Result 93.

Related Results

89. Choice principles in the basic Cohen model. We have seen that in the basic Cohen model — the model in which the reals cannot be well-ordered — there is an infinite set of reals which does not contain a countable infinite subset and thus, the Axiom of Choice fails in that model. On the other hand, the following choice principles are still valid in the basic Cohen model:

- If $X$ is infinite, then $\mathcal{P}(X)$ is transfinite, i.e., $\aleph_0 \leq |\mathcal{P}(X)|$ (see Jech [9, p. 81, Problem 20]).
- For every family $\mathcal{F}$ of sets, each containing at least two elements, there is a function $F$ such that for each set $S \in \mathcal{F}$, $\emptyset \neq F(S) \subseteq S$ (see Jech [9, p. 82, Problem 21]).
- Every family of non-empty well-orderable sets has a choice function (see Jech [9, p. 82, Problem 22] and compare with Chapter 7 | Related Result 48).

90. A model in which every ultrafilter is principal. Blass constructed in [1] a model — similar to Feferman’s model given above — in which every ultrafilter (and not just ultrafilters over $\omega$) is principal.
91. $\omega_1$ can be singular. It is provable in ZF that there exists a surjection from the reals onto $\omega_1$ (cf. Theorem 4.11). Hence, in the model in which the set of reals is a countable union of countable sets, $\omega_1$ is a limit of a countable sequence of countable ordinals, and therefore $\omega_1$ is singular in that model (compare with Proposition 5.10 where it is shown that in the presence of AC, successor cardinals are always regular).

92. $\omega_1$ can be even measurable. An uncountable aleph $\kappa$ is called a measurable cardinal if there exists a non-principal ultrafilter $\mathcal{U}$ over $\kappa$ which is $\kappa$-complete, i.e., if $\alpha \in \kappa$ and $\{x_\xi : \xi \in \alpha\} \subseteq \mathcal{U}$, then

$$\bigcap \{x_\xi : \xi \in \alpha\} \in \mathcal{U}.$$  

In the presence of AC, measurable cardinals are extremely large, even much larger than inaccessible cardinals, on which Hausdorff [7, p. 131] wrote that already the smallest of those cardinals — if they exist — is of an exorbitant magnitude. However, under the assumption that there is a measurable cardinal in the ground model, Jech constructed in [8] a symmetric model of ZF in which $\omega_1$ is measurable (see also Jech [9, Chapter 12, §1]).

93. Nontransferable statements. Not every statement which hold in a permutation model (i.e., in a model of ZFA) can be transferred into ZF. There are even statements which imply AC in ZF but are weaker than AC in ZFA. For example, Multiple Choice and Kurepa’s Principle are such statements (see Theorem 5.4 and Jech [9, Theorem 9.2]).

94. Bases in vector spaces and the Axiom of Choice. In Chapter 5 we have seen that the Axiom of Choice follows in ZF from the assertion that every vector space has a basis (cf. Theorem 5.4). However, it is still open whether the Axiom of Choice is deducible in ZFA from the assertion that every vector space has a basis, or at least from the assertion that in every vector space every independent set is included in a basis.

95. Inaccessible cardinals in ZF. In [2], Blass, Dimitriou, and Löwe introduce and investigate definitions for inaccessible cardinals (see page 315) in the absence of AC. They produce four possible definitions that are equivalent in ZFC but not in ZF, and provide a complete implication diagram (in ZF) for these four different concepts.

References


References


Combining Forcing Notions

In this chapter we shall investigate how one can combine various forcing notions. For this we first consider just two (not necessarily distinct) forcing notions, say $\mathbb{P} = (P, \leq_P)$ and $\mathbb{Q} = (Q, \leq_Q)$.

The simplest way to combine $\mathbb{P}$ and $\mathbb{Q}$ is to form the disjoint union of $\mathbb{P}$ and $\mathbb{Q}$ (where conditions of $\mathbb{P}$ are incomparable with those of $\mathbb{Q}$). Obviously, a generic filter of the disjoint union is either $\mathbb{P}$-generic or $\mathbb{Q}$-generic, and therefore, this construction is useless for independence proofs.

Another way to combine $\mathbb{P}$ and $\mathbb{Q}$ is to build the product $\mathbb{P} \times \mathbb{Q} = (P \times Q, \leq_{P \times Q})$. Since the forcing notion $\mathbb{P} \times \mathbb{Q}$ belongs to $V$, forcing with $\mathbb{P} \times \mathbb{Q}$ is in fact just a one-step extension of $V$. Products of forcing notions will be investigated in the first part of this chapter, where the focus will be on products of Cohen forcing notions.

A more sophisticated way to combine $\mathbb{P}$ and $\mathbb{Q}$ is to iterate $\mathbb{P}$ and $\mathbb{Q}$, i.e., we first force with $\mathbb{P}$ and then—in the $\mathbb{P}$-generic extension—by $\mathbb{Q}$. In this case, the forcing notion $\mathbb{Q}$ does not necessarily belong to $V$. To see this, let $G$ be $\mathbb{P}$-generic over $V$ and let $\mathbb{Q} = (\text{Fn}(G, 2), \subseteq)$. Obviously, the forcing notion $\mathbb{Q}$ does not belong to $V$. However, since $\mathbb{Q}$ belongs to $V[G]$, there is a $\mathbb{P}$-name $\check{Q}$ in $V$ such that $\mathbb{Q}[G] = \check{Q}$. Two-step iterations of this type are denoted by $\mathbb{P} * \mathbb{Q}$.

In the second part of this chapter we shall see how to transform a two-step iteration into a one-step forcing extension. Furthermore, we shall see different ways to define general iterations of forcing notions.

From now on, a forcing notion is just a partially ordered set $\mathbb{P} = (P, \leq)$ with a smallest element; in particular, we no longer require that there are incompatible conditions above each $p \in P$. 
Products

General Products of Forcing Notions

Before we investigate products of Cohen forcing notions — which will be the most frequently used product of forcing notions — we consider first the general case.

For two forcing notions $\mathbb{P}_0 = (P_0, \leq_0, 0_0)$ and $\mathbb{P}_1 = (P_1, \leq_1, 0_1)$, the product forcing notion

$$\mathbb{P}_0 \times \mathbb{P}_1 = (P_0 \times P_1, \leq, 0)$$

is defined by stipulating $0 := (0_0, 0_1)$ and

$$\langle p_0, p_1 \rangle \leq \langle q_0, q_1 \rangle \iff p_0 \leq q_0 \wedge p_1 \leq q_1 .$$

We leave it as an exercise to the reader to show that $\mathbb{P}_0 \times \mathbb{P}_1 = (P_0 \times P_1, \leq, 0)$ is indeed a forcing notion.

In general, if $\kappa$ is a non-zero cardinal number and $\langle P_\alpha : \alpha \in \kappa \rangle$ is a sequence of forcing notions, where for all $\alpha \in \kappa$, $P_\alpha = (P_\alpha, \leq_\alpha, 0_\alpha)$, then we define the product forcing notion

$$\prod_{\alpha \in \kappa} P_\alpha = \left( \prod_{\alpha \in \kappa} P_\alpha, \leq, 0 \right)$$

by stipulating $0 := \langle 0_\alpha : \alpha \in \kappa \rangle$ and

$$\langle p_\alpha : \alpha \in \kappa \rangle \leq \langle q_\alpha : \alpha \in \kappa \rangle \iff \forall \alpha \in \kappa (p_\alpha \leq_\alpha q_\alpha) .$$

Let us now have a closer look at the product $\prod_{\alpha \in \kappa} P_\alpha$ for some $\kappa \geq 2$. If $G$ is $\prod_{\alpha \in \kappa} P_\alpha$-generic over $V$, then $G \subseteq \prod_{\alpha \in \kappa} P_\alpha$. Thus, each $p \in G$ is of the form $p = \langle p(\alpha) : \alpha \in \kappa \rangle$. For each $\alpha \in \kappa$ let $G(\alpha) := \{ p(\alpha) : p \in G \}$; in particular, $G \subseteq \prod_{\alpha \in \kappa} G(\alpha)$. Obviously, for each $\alpha \in \kappa$, $G(\alpha)$ is $P_\alpha$-generic over $V$. Moreover, we have $G = \prod_{\alpha \in \kappa} G(\alpha)$, which implies that $V[G] = V[\prod_{\alpha \in \kappa} G(\alpha)] = V[\{G(\alpha) : \alpha \in \kappa\}]$ (the details are left as an exercise to the reader). In fact, we can prove even more:

**Lemma 18.1.** Let $\kappa$ be a cardinal, let $\prod_{\alpha \in \kappa} P_\alpha$ be a product of forcing notions $P_\alpha = (P_\alpha, \leq_\alpha, 0_\alpha)$, and let $G$ be $\prod_{\alpha \in \kappa} P_\alpha$-generic over $V$. Then, for each $\gamma \in \kappa$, $G(\gamma)$ is $P_\gamma$-generic over $V[\{G(\alpha) : \alpha \in \kappa \setminus \{ \gamma \}\}]$.

**Proof.** The cases when $\kappa = 0$ or $\kappa = 1$ are trivial. For the other cases, notice first that it is enough to prove the result just in the case when $\kappa = 2$, for we can always consider the product $P \times Q$ where $P := P_\gamma$ and $Q := \prod_{\alpha \in \kappa \setminus \{ \gamma \}} P_\alpha$.

So, let $G(0)$ be $P$-generic over $V$, where $P = (P, \leq, 0_P)$. We have to show that $G(1)$ is $Q$-generic over $V[G(0)]$, where $Q = (Q, \leq, 0_Q)$. Let $D \subseteq Q$ be an open dense set which belongs to the model $V[G(0)]$ — notice that $D$
does not necessarily belong to \( V \). In \( V \) there exist a \( \mathbb{P} \)-name \( D \) for \( D \) and a \( \mathbb{P} \)-condition \( p_0 \in G(0) \) such that
\[
V \models p_0 \Vdash \text{"} D \text{ is an open dense subset of } Q.\text{"}
\]
In other words, for every \( r \in Q \) there exists a \( \mathbb{P} \)-name \( q \) for a condition in \( Q \) such that \( p_0 \Vdash q \models r \land q \in D \). Now, let
\[
D'_1 = \{(p,q) \in P \times Q : p \geq p_0 \land p \Vdash q \in D \} \subseteq P \times Q.
\]
We leave it as an exercise to the reader to show that \( D'_1 \) is dense above \( \langle p_0, 0_Q \rangle \). Since \( p_0 \in G(0) \) and \( G(1) \) is \( \mathbb{Q} \)-generic over \( V \), by Fact 14.7 there are conditions \( p' \in P \) and \( q' \in Q \) such that \( \langle p', q' \rangle \in D'_1 \cap (G(0) \times G(1)) \). In particular we have \( p' \in G(0) \) and \( p' \Vdash q' \in D \), which implies that \( V[G(0)] \models q' \in D[G(0)] \). Finally, since \( q' \in G(1) \) and \( D[G(0)] = D \), we get \( q' \in D \cap G(1) \), i.e., \( D \cap G(1) \) is non-empty.

We now introduce the notion of support of a condition — a notion which we shall meet again in the definition of iterated forcing.

Let \( p = \langle p(\alpha) : \alpha \in \kappa \rangle \) be a \( \prod_{\alpha \in \kappa} \mathbb{P}_\alpha \)-condition, i.e., for each \( \alpha \in \kappa \) we have \( p(\alpha) \in P_\alpha \), where \( \mathbb{P}_\alpha = (P_\alpha, \leq_\alpha, 0_\alpha) \). Then the set \( \{ \alpha \in \kappa : p(\alpha) \neq 0_\alpha \} \) is called the support of \( p \) and is denoted by \( \text{supp}(p) \). Notice that for any \( \prod_{\alpha \in \kappa} \mathbb{P}_\alpha \)-conditions \( p \) and \( q \), \( p \leq q \) implies \( \text{supp}(p) \subseteq \text{supp}(q) \). A finite \textbf{support product} of forcing notions is a product of forcing notions consisting of those conditions that have finite support.

\textbf{Products of Cohen Forcing}

In this section we shall consider the product of countably many Cohen forcing notions is essentially the same as Cohen forcing.

For this, let us first consider Cohen forcing \( \mathbb{C} = (\text{Fn}(\omega, 2), \subseteq) \), as it was defined in Chapter 14. If \( G \) is \( \mathbb{C} \)-generic over some ground model \( V \), then \( c := \bigcup G \) is a function in \( V[G] \) from \( \omega \) to \( \{0, 1\} \) (i.e., \( c \in \omega^\omega \)) which has the property that the set \( \{ p \in \text{Fn}(\omega, 2) : p \subseteq c \} \) is \( \mathbb{C} \)-generic over \( V \). A real \( c \in \omega^\omega \) (in some model \( V \)) with this property is called a \textbf{Cohen real} over \( V \). Obviously, every \( \mathbb{C} \)-generic filter over \( V \) corresponds to a Cohen real, and vice versa, every Cohen real over \( V \) corresponds to a \( \mathbb{C} \)-generic filter over \( V \).

Sometimes it is convenient to consider a Cohen real, defined as an element of \( \omega^\omega \), as a function from \( \omega \) to \( \omega \). Of course, there exist natural mappings between the sets \( \omega^\omega \) and \( \omega^\omega \). However, there is a more elegant way to get Cohen reals \( c \in \omega^\omega \): Consider again Cohen forcing \( \mathbb{C} = (\text{Fn}(\omega, 2), \subseteq) \), and for the moment let \( \mathbb{C} := (\bigcup_{n \in \omega} ^n \mathbb{C} \subseteq), \mathbb{C}(\omega) := (\text{Fn}(\omega, \omega), \subseteq) \), and \( \mathbb{C}(\omega) := (\bigcup_{n \in \omega} ^n \mathbb{C} \subseteq) \).

We shall show that the forcing notions \( \mathbb{C}, \mathbb{C}(\omega) \), and \( \mathbb{C}(\omega) \) are all equivalent to Cohen forcing \( \mathbb{C} \), i.e., no matter whether we force (over some ground model \( V \)) with \( \mathbb{C} \) or with one of \( \mathbb{C}, \mathbb{C}(\omega) \), or \( \mathbb{C}(\omega) \), we always get the same generic extension.
Proposition 18.2. \( C \approx \bar{C} \approx C(\omega) \approx \bar{C}(\omega) \).

Proof. In order to prove that two forcing notions \( P = (P, \leq) \) and \( Q = (Q, \leq) \) are equivalent, it is enough to show that there exists a dense embedding \( h : P \to Q \) (see Fact 14.3).

\( C \approx \bar{C} \) and \( C(\omega) \approx \bar{C}(\omega) \): The identities \( \iota_1 : \bigcup_{\omega \in \omega} 2 \to \text{Fn}(\omega, 2) \) and \( \iota_2 : \bigcup_{\omega \in \omega} \omega \to \text{Fn}(\omega, \omega) \) are obviously dense embeddings.

\( \bar{C}(\omega) \approx \bar{C} \): We shall define a dense embedding \( h : \bigcup_{\omega \in \omega} \omega \to \bigcup_{\omega \in \omega} 2 \).
For this, take an arbitrary function \( p : n_0 \to \omega \). If \( n_0 = 0 \), then \( h(p) \) is the empty function. Otherwise, by induction on \( n_0 \) we first define integers \( b_k \) such that for all \( k \in n_0 \) we have

\[
b_k = \begin{cases} p(0) & \text{if } k = 0, \\ b_{k-1} + p(k) + 1 & \text{if } k > 0. \end{cases}
\]

Let \( x_p := \{ b_k : k \in n_0 \} \) and define the function \( h(p) : b_{n_0} + 1 \to 2 \) by stipulating

\[
h(p)(j) = \begin{cases} 1 & \text{if } j \in x_p, \\ 0 & \text{if } j \notin x_p. \end{cases}
\]

Notice that we always have \( h(p)(b_{n_0}) = 1 \). On the other hand, if the function \( q : k_0 + 1 \to 2 \) is such that \( q(k_0) = 1 \), then there exists a \( p : l \to \omega \), where \( l = \{ m \in k_0 + 1 : q(m) = 1 \} \), such that \( h(p) = q \). In fact, \( h(p) \) is the sequence of \( p(0) \) zeros, a single 1, \( p(1) \) zeros, a single 1, et cetera. We leave it as an exercise to the reader to verify that \( h \) is indeed a dense embedding.

Since the forcing notions \( C, \bar{C}, C(\omega), \bar{C}(\omega) \), are all equivalent, we shall not distinguish between these four forcing notions, and in order to simplify the terminology, each of these four forcing notions is called Cohen forcing and is denoted by \( C \).

Let us now consider products of Cohen forcing: For any ordinal \( \lambda \in \Omega \) let \( C_\lambda = (\text{Fn}(\omega \times \lambda, 2), \subseteq) \) and let \( C^\lambda \) denote the finite support product of \( \lambda \) copies of Cohen forcing \( C = (\text{Fn}(\omega, 2), \subseteq) \). We shall show that for any ordinal \( \lambda \), \( C_\lambda \approx C^\lambda \), and in addition, if \( \lambda \) is a non-zero countable ordinal, then both forcing notions are equivalent to Cohen forcing \( C \).

Proposition 18.3. For every ordinal \( \lambda \) we have \( C_\lambda \approx C_{\lambda \mid \lambda} \approx C^\lambda \approx C^\lambda \), and for every non-zero countable ordinal \( \gamma \) we have \( C \approx C_\gamma \approx C^\gamma \).

Proof. It is sufficient to show that for every non-zero countable ordinal \( \gamma \) we have \( C \approx C_\gamma \), and that for every ordinal \( \lambda \) we have \( C_\lambda \approx C_{\lambda \mid \lambda} \approx C^\lambda \), and \( C_\lambda \approx C^\lambda \).

\( C \approx C_\gamma \): Let \( \xi : \omega \times \gamma \to \omega \) be a bijection and let \( h : \text{Fn}(\omega \times \gamma, 2) \to \text{Fn}(\omega, 2) \) be such that for each \( p \in \text{Fn}(\omega \times \gamma, 2) \), \( \text{dom}(h(p)) = \xi(\text{dom}(p)) \) and for all \( j \in \xi(\text{dom}(p)) \) we have \( h(p)(j) = p(\xi^{-1}(j)) \). Then \( h \) is obviously a dense embedding; in fact, \( h \) is even an isomorphism.
A model in which \( a < c \)

\( C^\lambda \approx C_\lambda \): Since \( C^\lambda \) is a finite support product, for every \( C^\lambda \)-condition \( p = (p(\beta) : \beta \in \lambda) \), the set \( \text{supp}(p) = \{ \beta \in \lambda : p(\beta) \neq 0 \} \) is finite. Now, for every \( C^\lambda \)-condition \( p \) let \( h(p) \in \text{Fn}(\omega \times \lambda, 2) \) be such that

\[
\text{dom}(h(p)) = \{ (\beta, n) \in \text{supp}(p) \times \omega : n \in \text{dom}(p(\beta)) \}
\]

and \( h(p)(\beta, n) = p(\beta)(n) \). Then \( h \) is obviously a dense embedding; in fact, it is even an isomorphism.

Finally, let \( \zeta : \lambda \to [\lambda] \) be a bijection. Then \( \zeta \) induces a bijection between \( \omega \times \lambda \) and \( \omega \times [\lambda] \), as well as a bijection between the set of \( C^\lambda \)-conditions and the set of \( C^{[\lambda]} \)-conditions, which shows that \( C_\lambda \approx C_{[\lambda]} \) and that \( C^\lambda \approx C^{[\lambda]} \). \( \square \)

As an immediate consequence of PROPOSITION 18.3 we get that for every non-zero countable ordinal \( \lambda \), each \( C^\lambda \)-generic filter can be encoded by a single Cohen real. Roughly speaking, adding one Cohen real is the same as adding countably many Cohen reals. Since this is one of the main features of Cohen forcing, we state it in a more formal way.

**FACT 18.4.** If \( G \) is \( C^\lambda \)-generic over \( V \) and \( G' \) is \( C_\lambda \)-generic over \( V \), where \( \lambda \) is a non-zero countable ordinal, then there are Cohen reals \( c \) and \( c' \) over \( V \) such that \( V[G] = V[c] \) and \( V[G'] = V[c'] \).

**A Model in which \( a < c \)**

As a first application of a product of Cohen forcing we shall construct a model of \( \text{ZFC} \) in which \( c \) is large and \( a \) is small. Recall that \( a \) is the least cardinality of an infinite, maximal almost disjoint family (called mad family), where a family \( \mathcal{F} \subseteq [\omega]^\omega \) is almost disjoint if any two distinct elements of \( \mathcal{F} \) have finite intersection (see Chapter 8).

**PROPOSITION 18.5.** \( \omega_1 = a < c \) is consistent with \( \text{ZFC} \).

**Proof.** Let \( V \) be a model of \( \text{ZFC} + \text{CH} \), let \( \kappa \geq \omega_2 \) be a cardinal, and let \( G \) be \( C_\kappa \)-generic over \( V \) (by PROPOSITION 18.3 we could equally well work with the finite support product \( C^\kappa \)). By THEOREM 14.21 we know that \( V[G] \models c \geq \kappa \).

Thus, it remains to show that \( V[G] \) contains a mad family of size \( \omega_1 \). Firstly, we shall construct a family \( \mathcal{A}_0 \subseteq [\omega]^\omega \) of size \( \omega_1 \) in \( V \) such that whenever \( g \) is \( C_\kappa \)-generic over \( V \), then \( V[g] \models \mathcal{A}_0 \text{ is mad} \). Then we shall show that \( \mathcal{A}_0 \) — which is obviously an almost disjoint family in \( V[G] \) — is still maximal in \( V[G] \).

**Construction of \( \mathcal{A}_0 \) in \( V \):** Consider Cohen forcing \( \mathbb{C} = (\text{Fn}(\omega, 2), \subseteq) \). Within \( V \), let \( \{ (p_\xi, x_\xi) : \omega \leq \xi \in \omega_1 \} \) be an enumeration of all pairs \( (p, x) \) such that \( p \in \text{Fn}(\omega, 2) \) and \( x \) is a nice name for a subset of \( \omega \), i.e., for all \( (\eta, q_1), (\eta, q_2) \in x \), either \( q_1 = q_2 \) or \( q_1 \perp q_2 \) (see the proof of THEOREM 16.4).

Notice that since \( V \models \text{CH} \), there are just \( \omega_1 \) nice names in \( V \) for subsets of \( \omega \). The set \( \mathcal{A}_0 = \{ A_\xi \in [\omega]^\omega : \xi \in \omega_1 \} \) is constructed as follows: Let
\( \{ A_n \in \omega^\omega : n \in \omega \} \) be any family of pairwise disjoint infinite subsets of \( \omega \).

Let \( \omega \leq \xi \in \omega_1 \) and assume that we have already defined \( A_\eta \) for all \( \eta \in \xi \).

Then, choose \( A_\xi \in \omega^\omega \) such that the following conditions are satisfied:

1. For all \( \eta \in \xi \), \( A_\eta \cap A_\xi \) is finite.
2. If \( p_\xi \Vdash \exists x \in [\omega]^{\omega} \land \forall \eta \in \xi (p_\xi \Vdash |x_\xi \cap A_\eta| < \omega) \), \( (\ast) \)

then the set \( \{ r \geq p_\xi : r \Vdash (A_\xi \cap x_\xi) = \omega \} \) is dense above \( p_\xi \).

To see that \( A_\xi \) may be chosen that way, notice that whenever \( (\ast) \) fails, then we just have to take care of (1) and we simply apply the fact that \( \xi \) is countable and therefore the almost disjoint family \( \{ A_\eta : \eta \in \xi \} \) cannot be maximal. On the other hand, if \( (\ast) \) holds, then whenever \( g \) is \( C \)-generic over \( V \) and \( p_\xi \in g \) we have

\[ V[g] \vDash \exists x_\xi \in [\omega]^\omega \land \forall \eta \in \xi (|x_\xi \cap A_\eta| < \omega) . \]

In other words, \( x_\xi \in [\omega]^\omega \) witnesses that the almost disjoint family \( \{ A_\eta : \eta \in \xi \} \) is not maximal in \( V[g] \).

Now, we construct \( A_\xi \), satisfying (1), such that \( V[g] \vDash \exists x_\xi \in [\omega]^\omega \land \forall \eta \in \xi (|x_\xi \cap A_\eta| < \omega) \).

For this, let \( \{ B_i : i \in \omega \} \) be an enumeration of the set \( \{ A_\eta : \eta \in \xi \} \) and let \( \{ (m_i, q_i) : i \in \omega \} \) be an enumeration of \( \omega \times \{ q : q \geq p_\xi \} \). By (\ast), for each \( i \in \omega \) we obviously have

\[ q_i \Vdash \exists x_\xi \in \omega^\omega \land \forall \eta \in \xi (|x_\xi \cap A_\eta| < \omega) . \]

Thus, we find a \( C \)-condition \( r_i \geq q_i \) as well as an integer \( m_i \geq n_i \) such that \( m_i \notin \{ B_0 \cup \ldots \cup B_i \} \) and \( r_i \Vdash m_i \in x_\xi \), and define \( A_\xi := \{ m_i : i \in \omega \} \).

What have we achieved? By (\ast), for every \( q \geq p_\xi \), every \( n \in \omega \), and every finite set \( \{ \eta_0, \ldots, \eta_k \} \subseteq \xi \), there is a condition \( q' \geq q \) and an integer \( m \geq n \) such \( q' \Vdash m \in x_\xi \land m \notin \bigcup_{\eta \in \xi} A_\eta \). Thus, \( x_\xi \in [\omega]^\omega \) is not a witness for the statement "\( (A_\eta : \eta \in \xi + 1) \) is not a mad family in \( V[g] \)" , which implies that \( \mho \) is not a mad family in \( V[g] \). In other words, \( \mho \) is a mad family in \( V \) which remains mad after adding a single Cohen real.

In the next step we show that the same is true even if we add many Cohen reals.

\( \mho \) is mad in \( V[G] \): Consider now the forcing notion \( C_\kappa \). Let \( G \) be \( C_\kappa \)-generic over \( V \) and assume towards a contradiction that

\[ V[G] \vDash \exists x \in [\omega]^\omega \land \forall A_\xi \in \mho \land (|x \cap A_\xi| < \omega) . \]

Then there would be a \( C_\kappa \)-name \( x \) for a subset of \( \omega \) and a \( C_\kappa \)-condition \( p \) such that for all \( \xi \in \omega_1 \),

\[ p \Vdash \exists x \in [\omega]^\omega \land |x \cap A_\xi| < \omega . \]

By the facts proved earlier and since \( C_\kappa \) satisfies \( \text{ccc} \) and every \( C_\kappa \)-condition is finite, there is a countable set \( I_0 \subseteq \kappa \) such that, with respect to \( C_{I_0} = \)
Two-step iterations

(Fn(ω × I₀, 2), ⊆), there is a nice Cₗ₀-name x₀ for a subset of ω as well as a Cₗ₀-condition p₀ such that for all ξ ∈ ω₁,

\[ p₀ \Vdash C_\xi \left| x₀ \right| = \omega \land \left| x₀ \cap A_\xi \right| < \omega. \]

By Proposition 18.3, C ≅ Cₗ₀, and hence we can replace Cₗ₀ by C. Thus, there exists a pair \( \langle p_{ξ₀}, x_{ξ₀} \rangle \), consisting of a C-condition p_{ξ₀} and a nice name x_{ξ₀} for a subset of ω, such that for all ξ ∈ ω₁,

\[ p_{ξ₀} \Vdash C \left| x_{ξ₀} \right| = \omega \land \left| x_{ξ₀} \cap A_\xi \right| < \omega. \]

In particular, for A_{ξ₀} we would have

\[ p_{ξ₀} \Vdash C \left| x_{ξ₀} \cap A_{ξ₀} \right| < \omega, \]

which contradicts the construction of A_{ξ₀}.

For a proof using iterated forcing (introduced below) see Related Result 99.

Iterations

Below, we shall develop some methods to add generic filters step by step. The simplest case, which we consider first, is when only two generic filters are added. This so-called two-step iteration is quite easy to understand, but because it involves most of the tools which are used to handle longer iterations, it is worthwhile to consider this case in greater detail. Nevertheless, the situation becomes more difficult when the length of the iteration is infinite — which will be discussed in a slightly less detailed way.

Two-Step Iterations

Let us start with an example: Let V be a model of ZFC. Assume we want to construct an infinite set \( H \subseteq \omega \) in some generic extension of V which is almost homogeneous for each colouring \( \pi : [\omega]^n \to r \) which belongs to V (where \( n \in \omega \) and \( r \) is a positive integer). Recall that an infinite set \( H \subseteq \omega \) is almost homogeneous for a colouring \( \pi : [\omega]^n \to r \), if there is a finite set \( K \in \text{fin}(\omega) \) such that \( [H \setminus K]^n \) is monochromatic. There are many different ways to obtain such a real \( H \). For example, if there is a Ramsey ultrafilter \( \mathcal{U} \) in V, then it would be enough to force the existence of a set \( H \in [\omega]^\omega \) which is almost contained in each \( x \in \mathcal{U} \). Why? Since \( \mathcal{U} \) is a Ramsey ultrafilter, for every colouring \( \pi : [\omega]^n \to r \) there is an \( x \in \mathcal{U} \) which is homogeneous for \( \pi \). Now, if \( H \) is almost contained in \( x \), then \( H \) is almost homogeneous for \( \pi \). However, if there is no Ramsey ultrafilter in V (see for example Proposition 25.11), we first have to force the existence of a Ramsey ultrafilter. In order to force a Ramsey ultrafilter we use the forcing notion \( \mathbb{U} = ([\omega]^\omega / \text{fin,} \leq) \) which was
introduced in Chapter 14. Let $G_0$ be $U$-generic over $V$ and let $\mathcal{U} = \bigcup G_0$. Then, by Proposition 14.18, $\mathcal{U}$ is a Ramsey ultrafilter in $V[G_0]$. Now, we force the existence of a set $H \in [\omega]^{\omega}$ which is almost contained in each $x \in \mathcal{U}$. In $V[G_0]$, consider the forcing notion $Q_\mathcal{U} = (Q_\mathcal{U}, \leq)$, where $Q_\mathcal{U}$ is the set of all ordered pairs $\langle s, E \rangle$ such that $s \in \text{fin}(\omega)$ and $E \in \text{fin}(\mathcal{U})$, and for all $\langle s, E \rangle, \langle t, F \rangle \in Q_\mathcal{U}$ we define

$$\langle s, E \rangle \leq \langle t, F \rangle \iff s \subseteq t \wedge E \subseteq F \wedge (t \setminus s) \subseteq \bigcap E.$$ 

If $G_1$ is $Q_\mathcal{U}$-generic over $V[G_0]$, then the set

$$H_0 = \bigcup \{ s \in \text{fin}(\omega) : \exists E \in \text{fin}(\mathcal{U}) (\langle s, E \rangle \in G_1) \},$$

which belongs to the model $V[G_0][G_1]$, is almost homogeneous for all colourings $\pi : [\omega]^{\omega} \to r$ which belong to $V$.

Notice that the forcing notion $Q_\mathcal{U}$ belongs to $V[G_0]$, so, there is a $U$-name $Q_\mathcal{U}$ in $V$ for $Q_\mathcal{U}$. Forcing first with $U$ over $V$, followed by forcing with $Q_\mathcal{U}$ over $V[G_0]$, is a two-step “process” which we shall denote by $U \ast Q_\mathcal{U}$. The goal is now to find a forcing notion $P$ in $V$ such that $P$ is equivalent to $U \ast Q_\mathcal{U}$, in other words, the goal is to write the two-step “process” $U \ast Q_\mathcal{U}$ as a single forcing extension over the ground model $V$.

More generally, we have the following situation: We start in some ground model $V$ of ZFC, where in $V$ we have a forcing notion $P = (P, \leq_P, 0_P)$. If $G$ is $P$-generic over $V$, then $V[G]$ is again a model of ZFC. Assume that $Q = (Q, \leq_Q, 0_Q)$ is a forcing notion in $V[G]$ (which is not necessarily in $V$) and that $H$ is $Q$-generic over $V[G]$. Then $V[G][H]$ is a model of ZFC, too.

Since $Q$ belongs to $V[G]$, there is a $P$-name $Q$ in $V$ for $Q$. So, by combining the conditions in $P$ with $P$-names for $Q$-conditions, it should be possible to write the so-called two-step iteration $P \ast Q$ as a single forcing notion $R$ which belongs to the ground model $V$. Furthermore, it would be interesting to know whether some combinatorial properties of $P$ and $Q$ are preserved in the two-step iteration. For example, if $P$ and $Q$ both satisfy ccc, does this imply that $R$ also satisfies ccc? Before we can answer this question (in the affirmative), we first have to show that $P \ast Q$ is indeed equivalent to a single forcing notion which belongs to $V$ — which is consequently denoted by $P \ast Q$.

Let $V$ be a model of ZFC and let $P = (P, \leq_P, 0)$ be a forcing notion in $V$ with smallest element $0$. Notice that by Fact 14.4 we may always assume that the smallest element of a forcing notion is $0$, i.e., $0 = 0$. A $P$-name in $V$ for a forcing notion $Q = (Q, \leq, 0)$ in the $P$-generic extension of $V$ is a triple of $P$-names $\langle Q, \leq, 0 \rangle$ which has the following properties:

(a) $\emptyset \Vdash_P \leq$ is a partial ordering of $Q$" (recall that a partial ordering is a binary relation which is transitive, reflexive, and anti-symmetric).

(b) If $p \Vdash_P q \in Q$ for some $P$-name $q$, then there is a $P$-condition $p'$ such that $p \leq_P p'$, and there are $P$-names $r_1$ and $r_2$ such that

$$p' \Vdash_P r_1 \in Q \wedge r_2 \in Q \wedge q \leq r_1 \wedge q \leq r_2 \wedge r_1 \not\leq r_2.$$
Two-step iterations 351

(c) \( \emptyset \vdash \emptyset \in Q \).

(d) If \( p \vdash g \in Q \), then \( p \vdash \emptyset \not\leq g \).

Now, we first define a forcing notion \( \mathbb{R} \) in \( \mathbb{V} \), which depends on \( \mathbb{P} \) & \( \mathbb{Q} \), and then we show that forcing with \( \mathbb{R} \) yields the same generic extension as the two-step iteration \( \mathbb{P} \times \mathbb{Q} \).

Let \( \mathbb{R} = (\mathbb{R}, \leq_{\mathbb{R}}, 0_{\mathbb{R}}) \) where

\[
R = \{ (p, q) : p \in P \land p \vdash g \in Q \} \quad \text{and} \quad 0_{\mathbb{R}} = (\emptyset, \emptyset),
\]

and for all \( \langle p_1, q_1 \rangle, \langle p_2, q_2 \rangle \in R \), let

\[
\langle p_1, q_1 \rangle \leq_{\mathbb{R}} \langle p_2, q_2 \rangle \iff p_1 \leq p_2 \land p_1 \vdash q_1 \leq q_2.
\]

Before we show that forcing with \( \mathbb{R} \) is equivalent to \( \mathbb{P} \times \mathbb{Q} \), we have to show that \( \mathbb{R} = (\mathbb{R}, \leq_{\mathbb{R}}, 0_{\mathbb{R}}) \) is a forcing notion with smallest element \( 0_{\mathbb{R}} \).

For this, we first show that the binary relation \( \leq_{\mathbb{R}} \) is a partial ordering, i.e., we show that \( \leq_{\mathbb{R}} \) is (1) reflexive, (2) transitive, and (3) has the property that

\[
\langle p_1, q_1 \rangle \leq_{\mathbb{R}} \langle p_2, q_2 \rangle \land \langle p_2, q_2 \rangle \leq_{\mathbb{R}} \langle p_1, q_1 \rangle \implies \langle p_1 = p_2 \rangle
\]

and that \( p_1 \vdash q_1 = q_2 \). For (1)-(3), let \( \langle p, q \rangle, \langle p_1, q_1 \rangle, \langle p_2, q_2 \rangle, \langle p_3, q_3 \rangle \), be arbitrary \( \mathbb{R} \)-conditions.

(1) \( \langle p, q \rangle \leq_{\mathbb{R}} \langle p, q \rangle \iff p \leq p \land p \vdash g \not\leq q \).

Since \( \leq_{\mathbb{P}} \) is a partial ordering, \( p \leq_{\mathbb{P}} p \), and by (a) we have \( p \vdash g \not\leq q \).

(2) \( \langle p_1, q_1 \rangle \leq_{\mathbb{R}} \langle p_2, q_2 \rangle \land \langle p_2, q_2 \rangle \leq_{\mathbb{R}} \langle p_3, q_3 \rangle \iff \)

\[
p_1 \leq p_2 \land p_2 \leq p_3 \land p_1 \vdash q_1 \not\leq q_2 \land p_3 \vdash q_2 \not\leq q_3
\]

which implies \( p_1 \leq p_3 \) since \( p_2 \leq p_3 \) we get \( p_3 \vdash q_1 \not\leq q_2 \land q_2 \not\leq q_3 \).

By (a) we get \( p_1 \vdash q_1 \not\leq q_3 \), and hence, \( \langle p_1, q_1 \rangle \leq_{\mathbb{R}} \langle p_3, q_3 \rangle \).

(3) \( \langle p_1, q_1 \rangle \leq_{\mathbb{R}} \langle p_2, q_2 \rangle \land \langle p_2, q_2 \rangle \leq_{\mathbb{R}} \langle p_1, q_1 \rangle \iff \)

\[
p_1 \leq p_2 \land p_2 \leq p_1 \land p_2 \vdash q_1 \not\leq q_2 \land p_1 \vdash q_2 \not\leq q_1
\]

which implies \( p_1 = p_2 \) since \( p_1 = p_2 \) we get \( p_2 \vdash q_1 \not\leq q_2 \land q_2 \not\leq q_1 \).

By (a), \( \not\leq \) is forced to be anti-symmetric, thus, \( p_1 \vdash q_1 = q_2 \).

Now, we show that \( 0_{\mathbb{R}} \) (i.e., \( \langle \emptyset, \emptyset \rangle \)) belongs to \( R \) and that \( 0_{\mathbb{R}} \) is the smallest element (with respect to the partial ordering \( \leq_{\mathbb{R}} \)):

- \( \langle \emptyset, \emptyset \rangle \in R \iff \emptyset \vdash \emptyset \in Q \), which is just (c).
- Let \( \langle p, q \rangle \) be an arbitrary \( \mathbb{R} \)-condition. Since \( \langle p, q \rangle \in R \) we have \( p \vdash q \in Q \), and further we have \( \langle \emptyset, \emptyset \rangle \leq_{\mathbb{R}} \langle p, q \rangle \iff p \vdash \emptyset \not\leq q \), which is in fact just (d).
Finally, we show that \( R = (R, \leq_R) \) is indeed a forcing notion: For this we have to show that there are incompatible conditions above each \( \langle p, q \rangle \in R \). Let \( p_1, p_2 \in P \) be such that \( p \leq_R p_1, p \leq_R p_2 \), and \( p_1 \perp_R p_2 \). Then \( \langle p, q \rangle \leq_R \langle p_1, q \rangle \), \( \langle p, q \rangle \leq_R \langle p_2, q \rangle \), and \( \langle p_1, q \rangle \perp_R \langle p_2, q \rangle \), as required.

It remains to show that forcing with \( R \) is equivalent to the two-step iteration \( \mathbb{P} * \mathbb{Q} \). We shall give a detailed proof of one direction and leave the other direction as an exercise to the reader.

**Proposition 18.6.** Let \( \mathbf{V} \) be a model of \( \text{ZFC} \) and let \( G \) be \( \mathbb{R} \)-generic over \( \mathbf{V} \). Then there are sets \( G_0 \) and \( G_1 \) in \( \mathbf{V}[G] \), such that \( G_0 \) is \( \mathbb{P} \)-generic over \( \mathbf{V} \) and \( G_1 \) is \( \mathbb{Q}[G_0] \)-generic.

**Proof.** In the model \( \mathbf{V}[G] \) we define

\[
G_0 = \{ p \in P : \exists q \in Q(\langle p, q \rangle \in G) \}
\]

and

\[
G_1 = \{ q[G_0] \in \mathbb{Q}[G_0] : \exists p \in G_0(\langle p, q \rangle \in G) \}.
\]

We first show that \( G_0 \) and \( G_1 \) are filters, i.e., \( G_0 \) and \( G_1 \) are both downwards closed and directed.

**\( G_0 \) is downwards closed and directed:** If \( p \in G_0 \), then there is a \( q \in Q \) such that \( \langle p, q \rangle \in G \), and for any \( p' \leq p \) we have \( \langle p', q \rangle \leq_R \langle p, q \rangle \), since \( G \) is downwards closed, this implies \( \langle p', q \rangle \in G \), and therefore \( p' \in G_0 \). Furthermore, if \( p_0 \) and \( p_1 \) belong to \( G_0 \), then we find \( \langle p_0, q_0 \rangle \) and \( \langle p_1, q_1 \rangle \) in \( G \), and since \( G \) is directed, there is an \( R \)-condition \( \langle p, q \rangle \in G \) such that \( \langle p_0, q_0 \rangle \leq \langle p, q \rangle \leq \langle p_1, q_1 \rangle \). Thus, \( p \in G_0 \) and \( p_0 \leq p \leq p_1 \).

**\( G_1 \) is downwards closed and directed:** If \( q_0[G_0] \in G_1 \), then there is a \( p_0 \in G_0 \) such that \( \langle p_0, q_0 \rangle \in G \). Assume that in \( \mathbf{V}[G_0] \), \( q_1[G_0] \leq q_0[G_0] \). We have to show that \( q_1[G_0] \in G_1 \). Firstly, there is a \( p' \in G_0 \) such that \( p' \not \leq_R q_1 \). Secondly, since \( G \) is directed, there is a \( \langle p_1, q_2 \rangle \in G \) such that \( \langle p', q \rangle \leq_R \langle p_1, q_2 \rangle \), in particular we get \( p_1 \not \leq_R q_2 \). Now, since \( p_1 \geq_R p' \), we also have \( p_1 \not \leq_R q_1 \). Thus, \( p' \not \leq_R q_1 \), which implies \( \langle p_1, q_2 \rangle \geq_R \langle p_1, q_1 \rangle \), and since \( G \) is downwards closed, \( \langle p_1, q_1 \rangle \in G \). Hence, \( q_1[G_0] \in G_1 \). Furthermore, if \( q_0[G_0] \) and \( q_1[G_0] \) belong to \( G_1 \), then we find \( \langle p_0, q_0 \rangle \) and \( \langle p_1, q_1 \rangle \) in \( G \), and since \( G \) is directed, there is an \( R \)-condition \( \langle p, q \rangle \in G \) — and therefore \( q[G_0] \in G_1 \) — such that \( \langle p_0, q_0 \rangle \leq_R \langle p, q \rangle \leq_R \langle p_1, q_1 \rangle \). Thus, \( p \not \leq_R q_0 \leq q \geq q_1 \), and since \( p \in G_0 \) we get \( q_0[G_0] \leq q[G_0] \geq q_1[G_0] \).

Now we show that \( G_0 \) and \( G_1 \) are generic, i.e., \( G_0 \) and \( G_1 \) meet every open dense set in \( \mathbf{V} \) and \( \mathbf{V}[G_0] \) respectively.

**\( G_0 \) is generic:** Let \( D_0 \subseteq P \) be an open dense subset of \( P \) and let

\[
D_0 = \{ \langle p, q \rangle \in R : p \in D_0 \}.
\]

Then \( D_0 \) is an open dense subset of \( R \), and since \( G \) is \( \mathbb{R} \)-generic over \( \mathbf{V} \), there is an \( R \)-condition \( \langle p, q \rangle \in G \) — and therefore \( p \in G_0 \) — such that \( p \) belongs to \( D_0 \). Hence, \( G_0 \cap D_0 \neq \emptyset \), which shows that \( G_0 \) is \( \mathbb{P} \)-generic over \( \mathbf{V} \).
$G_1$ is generic: Let $D_1$ be an arbitrary open dense subset of $Q[G_0]$. Then there is a $\mathbb{P}$-name $D_1$ for $D_1$ and a $\mathbb{P}$-condition $p_0 \in G_0$ such that

$$p_0 \Vdash "D_1 \text{ is open dense in } Q".$$  

With respect to $D_1$ define

$$D'_1 = \{ (p, q) \in R : p \Vdash q \in D_1 \}.$$  

Then $D'_1 \subseteq R$ is open dense above $(p_0, \emptyset)$, and since $(p_0, \emptyset) \in G$ (because $p_0 \in G_0$, we get that $G \cap D'_1 \not= \emptyset$, say $(p_1, q_1) \in G \cap D'_1$. Now, $(p_1, q_1) \in G$ implies that $p_1 \in G_0$ and that $q_1[G_0] \in G_1$. Furthermore, by definition of $D'_1$ we get $p_1 \Vdash q_1 \in D_1$, and therefore $q_1[G_0] \in D_1$. Hence, $q_1[G_0] \in G_1 \cap D_1$, which shows that $G_1$ is $\mathcal{Q}[G_0]$-generic over $V[G_0]$. 

In the next section we shall investigate general iterations, but before let us show that two-step iterations of $\mathbb{P}$ forcing notions satisfy ccc.

**Lemma 18.7.** If $\mathbb{P}$ satisfies ccc and

$$0_\mathbb{P} \Vdash \text{"$\mathcal{Q}$ satisfies ccc"}$$  

then also $\mathbb{P} * \mathcal{Q}$ satisfies ccc.

**Proof.** Let $\mathbb{P} = (P, \leq)$ and let $\mathcal{Q} = (Q, \subseteq)$. Assume towards a contradiction that in the ground model $V$ there are uncountably many pairwise incompatible $\mathbb{P} * \mathcal{Q}$-conditions $\{ (p_\xi, q_\xi) : \xi \in \omega_1 \}$. Let $x  = \{ (\xi, p_\xi) : \xi \in \omega_1 \}$; then $x$ is a $\mathbb{P}$-name for a subset of $\omega_1$, i.e., $0_\mathbb{P} \Vdash x \subseteq \omega_1$. Let $G$ be $\mathbb{P}$-generic over $V$. Then $x[G] = \{ \xi \in \omega_1 : p_\xi \in G \}$. We shall show that there is an ordinal $\beta \in \omega_1$ such that $0_\mathbb{P} \Vdash x \subseteq \beta$, but first we prove the following

**Claim 1.** In $V[G]$, the set $\{ q_\xi[G] : \xi \in \omega_1 \}$ is an anti-chain in $Q[G]$.

**Proof of Claim 1.** Assume towards a contradiction that there are distinct $\xi, \eta \in \omega_1$ such that $q_\xi[G]$ and $q_\eta[G]$ are compatible elements of $Q[G]$. This would imply that there is a $\mathbb{P}$-condition $p \in G$, as well as a $\mathbb{P}$-name $q$ for a $\mathcal{Q}[G]$-condition, such that

$$p \Vdash q \in \mathcal{Q} \land q_\xi \vDash q \land q_\eta \not\vDash q.$$  

In fact, by extending $p$ if necessary, we get a $\mathbb{P} * \mathcal{Q}$-condition $(p, q)$ which is stronger than both $(p_\xi, q_\xi)$ and $(p_\eta, q_\eta)$, contradicting our assumption that $\{ (p_\xi, q_\xi) : \xi \in \omega_1 \}$ is a set of pairwise incompatible $\mathbb{P} * \mathcal{Q}$-conditions. 

Since $0_\mathbb{P} \Vdash \text{"$\mathcal{Q}$ satisfies ccc"}$, and therefore preserves $\omega_1$ (by Lemma 14.20), we get that $V[G] \models |x[G]| < \omega_1$ whenever $G$ is $\mathbb{P}$-generic over $V$, hence, $0_\mathbb{P} \Vdash |x| < \omega_1$.

**Claim 2.** There is an ordinal $\beta \in \omega_1$ such that $0_\mathbb{P} \Vdash x \subseteq \beta$. 

Proof of Claim 2. In $\mathcal{V}$, let

$$E = \{ \alpha \in \omega_1 : \exists r \in P \forall \beta \in \alpha \langle r \Vdash x \subseteq \alpha \land x \not\subseteq \beta \rangle \}. $$

Further, for every $\alpha \in E$ choose a $\mathbb{P}$-condition $r_\alpha$ such that for all $\beta \in \alpha$, $r_\alpha \Vdash x \subseteq \alpha \land x \not\subseteq \beta$. The set $\{ r_\alpha : \alpha \in E \}$, which belongs to $\mathcal{V}$, is an anti-chain in $\mathbb{P}$, and since $\mathbb{P}$ satisfies $\text{ccc}$, $|E| < \omega_1$. Thus, there exists a $\beta \in \omega_1$ such that $E \subseteq \beta$, which implies that $0\# \Vdash x \subseteq \beta$. \text{Claim 2}

By definition of $x$, for all $\xi \in \omega_1$ we have $p_\# \Vdash x \subseteq \xi$. In particular we get $p_\# \Vdash x \subseteq x$, which is a contradiction to $0\# \Vdash x \subseteq \beta$. \text{-}\text{-}1

As a matter of fact we would like to mention that Lemma 18.7 does not have an analogue for products; in other words, the product of two $\text{ccc}$ forcing notions does not necessarily satisfy $\text{ccc}$ (see Related Result 98).

General Iterations

In the previous section we have constructed a two-step iteration $U \ast Q_{\omega}$ in such a way that whenever $G$ is $U \ast Q_{\omega}$-generic over $\mathcal{V}$, then there is an infinite set $H_0 \in [\omega]^{\omega_1} \cap \mathcal{V}[G]$ which is almost homogeneous for all colourings $\pi : [\omega]^n \rightarrow r$ which belong to the ground model $\mathcal{V}$. Obviously, such a set $H_0$ cannot belong to $\mathcal{V}$. Now, we can ask what happens if we iterate the forcing notion $U \ast Q_{\omega}$?

As we have seen, at each stage we obtain a new set $H \in [\omega]^{\omega_1}$ which is almost homogeneous for all "old" colourings $\pi : [\omega]^n \rightarrow r$. So, for example an $\omega_1$-stage iteration of $U \ast Q_{\omega}$, starting in a model $\mathcal{V}$ of $\text{ZFC}$ in which $\epsilon = \omega_2$, would generate a family $\{ H_\alpha : \alpha \in \omega_1 \}$ of size $\omega_1^\omega$, where each $H_\alpha$ is almost homogeneous with respect to all "old" colourings $\pi : [\omega]^n \rightarrow r$. Recall that for any integers $n, r \geq 2$ there exists a bijection between the set of colourings $\pi : [\omega]^n \rightarrow r$ and the set of real numbers, thus, every "old" colouring can be encoded by an "old" real (and vice versa). Now, if every colouring $\pi : [\omega]^n \rightarrow r$ (i.e., real number) appears at some stage $\alpha \in \omega_1$ in the iteration, and if the cardinal numbers $\omega_1^\mathcal{V}, \omega_2^\mathcal{V}, \epsilon^\mathcal{V}$ are the same as $\omega_1, \omega_2, \epsilon$ in the final generic extension, then we would get a model in which $\omega_1 = \text{hom} < \omega_2 = \epsilon$. But do we really get such a model?

To understand the previous example as well as iterations in general, we have to answer questions like:

1. Is every iteration of forcing notions equivalent to a single forcing notion?
2. How is the iteration defined at limit stages?
3. Does the iteration add reals at limit stages of uncountable cofinality?
4. Does the iteration preserve cardinals?

Below, we shall give a complete answer to Questions 1–3 and we shall give an answer to Question 4 with respect to forcing notions satisfying $\text{ccc}$; regarding the forcing notion $U \ast Q$, we refer the reader to Chapter 20 and Chapter 23 | Related Result 138.
General iterations

Let us now consider \( \alpha \)-stage iterations of forcing notions for arbitrary ordinals \( \alpha \) (recall that by Fact 14.4 we may always assume that the smallest element of a forcing notion is \( \emptyset \)).

For \( \alpha = 1 \) we get ordinary forcing, and for \( \alpha = 2 \) we get two-step iterations which we already discussed in the previous section.

For \( \alpha = 3 \) we start with an arbitrary forcing notion \( P_1 = (P_1, \leq) \) which belongs to some ground model \( V \). Let \( Q_1 \) be a \( P_1 \)-name for a forcing notion \( (Q_1, \leq) \) in the \( P_1 \)-generic extension of \( V \) and let \( P_2 := P_1 \ast Q_1 \). Further, let \( Q_2 \) be a \( P_2 \)-name for a forcing notion \( (Q_2, \leq) \) in the \( P_2 \)-generic extension of \( V \) and let \( P_3 := P_2 \ast Q_2 \). Then every \( P_3 \)-condition is of the form \( \langle q_0, q_1 \rangle \), where \( q_0 \in P_1 \), \( q_0 \Vdash_{P_1} q_1 \in Q_1, \) and \( \langle q_0, q_1 \rangle \Vdash_{P_2} q_2 \in Q_2 \).

To form an \( \alpha \)-stage iteration for \( 3 < \alpha < \omega \), we just repeat this procedure. Thus, for positive integers \( n \), every \( P_\alpha \)-condition is of the form \( \langle \ldots \langle q_0, q_1 \rangle, q_2 \ldots, q_{\alpha-2}, q_{\alpha-1} \rangle \), for which we shall write the typographically less cumbersome (and easier to read) \( n \)-tuple \( \langle q_0, q_1, \ldots, q_{\alpha-1} \rangle \). With this convention, for positive integers \( n \), \( P_\alpha \)-conditions are sequences of length \( n \).

For \( n = 0 \) let \( P_0 := (\{ \emptyset \}, \subseteq) \). When we define \( P_\sigma \)-names, we find that \( G = \{ \emptyset \} \) is the unique \( P_\sigma \)-generic filter over \( V \). In particular we get that a 0-stage extension of \( V \) is just \( V \).

The sequence of forcing notions \( P_0, P_1, \ldots, P_n \), where \( P_k = (P_k, \leq, \emptyset) \), has the property that if \( p = (q_0, q_1, \ldots, q_{n-1}) \in P_n \), then for all \( k \in n \), \( p|_k \in P_k \) and \( p|_k \Vdash_{P_k} q_k \in Q_k \), where \( Q_k \) is a \( P_k \)-name for a forcing notion \( (Q_k, \leq) \) in the \( P_k \)-generic extension of \( V \). In particular, \( P_1 = Q_0 \) is a \( P_\sigma \)-name for a forcing notion \( (Q_0, \leq) \) in the \( P_0 \)-generic extension of \( V \), which is just \( V \) itself. In other words, \( P_1 \) is a \( P_0 \)-name for forcing notion \( (P_1, \leq) \) which belongs to \( V \). Thus, every \( P_\alpha \)-condition is of the form \( \langle q_0, q_1, \ldots, q_{\alpha-1} \rangle \), where \( q_0 \in P_0 \), \( q_0 \Vdash_{P_0} q_1 \in Q_0 \), and so on, \( \langle q_0, q_1, \ldots, q_{\alpha-1} \rangle \) is a \( P_\alpha \)-name for \( Q_0 \)-condition. This completes the definition of \( \alpha \)-stage iterations for \( \alpha \in \omega \).

Similarly, we define \((\alpha + 1)\)-stage iterations for arbitrary ordinals \( \alpha \): If the \( \alpha \)-stage iteration \( P_\alpha = (Q_\beta : \beta < \alpha) \) is already defined and \( Q_\alpha \) is a \( P_\alpha \)-name for a forcing notion in the \( P_\alpha \)-generic extension, then \( P_{\alpha + 1} := P_\alpha \ast Q_\alpha \).

Let us now consider the case when \( \alpha \) is a limit ordinal. At first glance, the set of \( P_\alpha \)-conditions consists of all \( \alpha \)-sequences \( \langle q_\beta : \beta < \alpha \rangle \), but having a closer look we see that there is some freedom in defining the set of \( P_\alpha \)-conditions. For example we can require that \( q_\beta = \emptyset \) for all but finitely many \( \beta < \alpha \), which is called finite support iteration, or that \( q_\beta = \emptyset \) for all but countably many \( \beta < \alpha \), which is called countable support iteration.

For \( P_\alpha \)-conditions \( p = \langle q_\beta : \beta < \alpha \rangle \) we define
\[
\text{supp}(p) = \{ \beta < \alpha : q_\beta \neq \emptyset \},
\]
and like for products we call \( \text{supp}(p) \) the support of \( p \). For example, a countable support iteration \( P_\alpha \) consists of all \( P_\alpha \)-conditions \( p \) that have countable support, i.e., \( |\text{supp}(p)| \leq \omega \).

Because of the following result (which will be stated without proof), finite support iterations are often used in iterations of forcing notions satisfying ccc.
Proposition 18.8. Any finite support iteration of ccc forcing notions satisfies ccc. In other words, if $P_\alpha$ is a finite support iteration of $\langle Q_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have

$$0_\beta \Vdash \text{“} Q_\beta \text{ satisfies ccc”},$$

then also $P_\alpha$ satisfies ccc.

Before we give an example of a finite support iteration, let us first settle some notation: Let $P_\alpha = \langle Q_\gamma : \gamma \in \alpha \rangle$ be any $\alpha$-stage iteration and let $G$ be $P_\alpha$-generic over some model $V$. Then, for $\beta \in \alpha$, let

$$G(\beta) = \{ q_\gamma : \exists (p_\gamma : \gamma \in \alpha) \in G \ (q_\beta = p_\beta[G]) \}$$

and

$$G|_\beta = \{ q_\gamma : \gamma \in \beta : \exists (p_\gamma : \gamma \in \alpha) \in G \ \forall \gamma \in \beta \ (q_\beta = p_\beta[G]) \}.$$

In other words, $G|_\beta$ denotes the $P_\beta$-generic filter generated by $G$. In abuse of notation, for $P_\alpha = \langle Q_\gamma : \gamma \in \alpha \rangle$ we usually write $P_\alpha = \langle Q_\gamma : \gamma \in \alpha \rangle$, where for all $\gamma \in \alpha$, $Q_\gamma := Q_\gamma[G|\gamma]$. In other words, we usually consider an $\alpha$-stage iteration $P_\alpha$, starting in some model $V$, as an $\alpha$-sequence of forcing notions $Q_\gamma$ (not just $P_\gamma$-names for forcing notions), where for each $\gamma \in \alpha$, $Q_\gamma$ belongs to the $P_\gamma$-generic extension $V[G|\gamma]$. Consequently, for $\beta \in \alpha$ we also write $V[(G(\gamma) : \gamma \in \beta)]$ instead of $V[G|_\beta]$, having in mind that we add one generic filter after the other, rather than adding just the single generic filter $G|_\beta$.

We conclude this section by showing that in finite support or countable support iterations or products of certain forcing notions (e.g., ccc forcing notions), no new reals are added at limit stages of uncountable cofinality — a result which will be used quite often in the forthcoming chapters.

Lemma 18.9. Let $\lambda$ be an infinite limit ordinal of uncountable cofinality (i.e., $\text{cf}(\lambda) > \omega$), let $P_\lambda = \langle Q_\alpha : \alpha \in \lambda \rangle$ be any finite support or countable support iteration or product of arbitrary forcing notions $Q_\alpha$, and let $G$ be $P_\lambda$-generic over some model $V$ of ZFC. If $V[G] \models \text{cf}(\lambda) > \omega$, then no new reals are added at stage $\lambda$; more formally,

$$\omega \omega \cap V[G] = \bigcup_{\alpha \in \lambda} \omega \omega \cap V[G|\alpha].$$

Proof. Let $f$ be a $P_\lambda$-name for a function in $\omega \omega \cap V[G]$. For every $\beta \in \lambda$ define a $P_\beta$-name $g_\beta$ for a partial function from $\omega$ to $\omega$ by stipulating

$$g_\beta = \{ (\text{op}(n, m), p) : \supp(p) \subseteq \beta \land p \in G \},$$

where $\text{op}(n, m)$ is the canonical $P_\lambda$-name for the ordered pair $\langle n, m \rangle$ (which was defined in Chapter 14). Now, we show that there exist an $\alpha \in \lambda$ such
that $V[G]_{\alpha} \models f[G]_{\alpha} = g_{\alpha}[G]_{\alpha}$, i.e., the function $f[G]$ appears already in the model $V[G]_{\alpha}$: Let us work in the model $V[G]$. For every $n \in \omega$ we can choose a $p_n \in G$ which decides the value of $f(n)$, i.e., $(\text{op}(n,m),p_n) \in f$ for some $m \in \omega$. Using the fact that $V[G] \models \text{cf}(\lambda) > \omega$ and that the supports of the $p_n$’s are at most countable (i.e., finite or countably infinite), we get that in $V[G]$, $\bigcup_{n \in \omega} \text{supp}(p_n) \subseteq \lambda$. Thus, there is an $\alpha \in \lambda$ such that $\bigcup_{n \in \omega} \text{supp}(p_n) \subseteq \alpha$, and by construction we have $g_{\alpha}[G]_{\alpha} \in ^\omega \omega \cap V[G]_{\alpha}$ and $V[G] \models f[G] = g_{\alpha}[G]$.

A Model in which $i < \kappa$

In this section we shall construct — by a finite support iteration of ccc forcing notions — a model in which $i < \kappa$, where $i$ is the least cardinality of a maximal independent family; but first, let us recall a few notions: A set $\mathcal{I} \subseteq [\omega]^\omega$ is an independent family, denoted i.f., if for any $A, B \in \text{fin}(\mathcal{I})$ with $A \cap B = \emptyset$ we have $\bigcap A \setminus \bigcup B$ is infinite, where we stipulate $\bigcap \emptyset := \omega$ (see Chapter 8). Furthermore, for independent families $\mathcal{I}$, let $bc(\mathcal{I})$ be the set of all finite boolean combinations of distinct elements of $\mathcal{I}$, in other words,

$$bc(\mathcal{I}) = \left\{ \bigcap A \setminus \bigcup B : \{A,B\} \subseteq \text{fin}(\mathcal{I}) \land A \cap B = \emptyset \right\}.$$ 

Notice that $bc(\mathcal{I}) \subseteq [\omega]^\omega$ and that for $\mathcal{I} = \emptyset$ we have $bc(\mathcal{I}) = \{\omega\}$.

The following lemma — which is in fact a ZFC result — will be crucial in the construction of the forcing notion which will be used in the iteration below.

**Lemma 18.10.** Let $V$ be an arbitrary model of ZFC and let $\mathcal{I} \subseteq [\omega]^\omega$ be an arbitrary i.f. in $V$. Then there exists an ideal $I \subseteq \mathcal{P}(\omega)$ in $V$ such that

(a) $I \cap \text{bc}(\mathcal{I}) = \emptyset$, and
(b) for every $y \in [\omega]^\omega \cap V$ there exists an $x \in \text{bc}(\mathcal{I})$ such that $x \cap y$ or $x \setminus y$ belongs to $I$.

**Proof.** Let $\{y_\alpha \in [\omega]^\omega : \alpha \in \kappa\}$ be an arbitrary enumeration of $[\omega]^\omega$. With respect to this enumeration we construct the ideal $I$ by induction on $\kappa$. Firstly, let $I_0 := \text{fin}(\omega)$. Then $I_0$ is an ideal and $I_0 \cap \text{bc}(\mathcal{I}) = \emptyset$. Assume that we have already defined the ideal $I_\alpha$ for some $\alpha \in \kappa$. If there are $x \in \text{bc}(\mathcal{I})$ and $u \in I_\alpha$ such that

$$x \subseteq y_\alpha \cup u,$$

then $I_{\alpha+1} := I_\alpha$; otherwise, $I_{\alpha+1}$ is the ideal generated by $I_\alpha \cup \{y_\alpha\}$, i.e.,

$u \in I_{\alpha+1}$ iff there is an $A \in \text{fin}(I_\alpha \cup \{y_\alpha\})$ such that $u \subseteq A$. Further, for limit ordinals $\lambda \in \kappa$ let $I_\lambda := \bigcup_{\alpha < \lambda} I_\alpha$, and let

$$I = \bigcup_{\alpha \in \kappa} I_\alpha.$$
It remains to show that the ideal $I$ has the required properties (we leave it as an exercise to the reader to show that $I$ is indeed an ideal):

(a) Assume towards a contradiction that there is an $x \in \text{bc } (\mathcal{I}) \cap I$. Since $I_0 \cap \text{bc } (\mathcal{I}) = \emptyset$, there exists a least ordinal $\alpha \in \mathfrak{c}$ such that $x \in I_{\alpha+1}$. In particular, $x \notin I_\alpha$, which implies that $I_{\alpha+1} \neq I_\alpha$. Hence, $I_{\alpha+1}$ must be the ideal generated by $I_\alpha \cup \{y_\alpha\}$. Thus, by construction, there is no $u \in I_\alpha$ such that $x \subseteq y_\alpha \cup u$. In other words, for each $u \in I_\alpha$ we have $x \nsubseteq y_\alpha \cup u$, which contradicts the fact that $x \in I_{\alpha+1}$.

(b) Take any $y \in [\omega]^\omega$ and let $\alpha \in \mathfrak{c}$ be such that $y = y_\alpha$. If there are $x \in \text{bc } (\mathcal{I})$ and $u \in I_\alpha$ such that $x \in y_\alpha \cup u$, then $x \setminus y_\alpha \subseteq u$, and consequently $x \setminus y \in I$; otherwise, $y_\alpha \in I_{\alpha+1}$, which implies that $x \cap y_\alpha \in I_{\alpha+1}$, and consequently $x \cap y \in I$.

Now we are ready to construct a model in which $i < \kappa$.

**Proposition 18.11.** $i < \kappa$ is consistent with ZFC.

**Proof.** The proof will be given in two steps: In the first step, with respect to some i.f. $\mathcal{I}$ we shall construct a forcing notion $Q_I$ (where $\mathcal{I}$ and $I$ are as in Lemma 18.10), and will show that $Q_I$ adds a generic real $g \in [\omega]^\omega$ (over some model $V$) which has the following properties:

- $\mathcal{I} \cup \{g\}$ is an i.f. in $V[g]$.
- If $y \in [\omega]^\omega \cap V$ is such that $\mathcal{I} \cup \{y\}$ is independent and $y \notin \mathcal{I}$, then $\mathcal{I} \cup \{g, y\}$ is not independent.

In the second step, by a finite support iteration of length $\omega_1$ of forcing notions $Q_I$, we shall construct a generic model in which the set of generic reals, added by the forcing notions $Q_I$, is a maximal i.f. of size $\omega_1$.

1st Step: Let $V$ be an arbitrary model of ZFC and let $\mathcal{I} \subseteq [\omega]^\omega$ be an arbitrary countable i.f. in $V$. Furthermore, let $I \subseteq \mathcal{P}(\omega)$ be the ideal constructed in Lemma 18.10 with respect to $\mathcal{I}$, i.e., $I \cap \mathcal{I} = \emptyset$, and for every $y \in [\omega]^\omega \cap V$ there exists an $x \in \text{bc } (\mathcal{I})$ such that $x \cap y$ or $x \setminus y$ belongs to $I$. With respect to the ideal $I$ we define the forcing notion $Q_I = (Q_I, \leq)$ as follows: A $Q_I$-condition is an ordered pair $\langle s, E \rangle$ where $s \in \text{fin}(\omega)$ and $E \in \text{fin}(I)$, and for $Q_I$-conditions $\langle s, E \rangle$ and $\langle t, F \rangle$ we define

$$\langle s, E \rangle \leq \langle t, F \rangle \iff s \subseteq t \land E \subseteq F \land (t \setminus s) \cap \bigcup_{u \in E} u = \emptyset.$$ 

Notice that for any $E, F \in \text{fin}(I)$ and any $s \in \text{fin}(\omega)$, $\langle s, E \rangle$ and $\langle s, F \rangle$ are compatible, and since the set $\text{fin}(\omega)$ is countable, $Q_I$ satisfies $\omega \omega$.

Let $G$ be $Q_I$-generic over $V$ and let

$$g = \bigcup\left\{s \in \text{fin}(\omega) : \exists E \in \text{fin}(I) (\langle s, E \rangle \in G)\right\}.$$
We leave it as an exercise to the reader to show that \( g \in [\omega]^{\omega} \) and that \( V[g] = V[G] \). Thus, we can equally well work with \( g \) instead of \( G \), in other words, \( g \) is a \( Q_I \)-generic real over \( V \).

Now, we show that \( \mathcal{I} \cup \{ g \} \) is an i.f. in \( V[g] \) which is even maximal with respect to the reals \( y \) which belong to \( V \) — notice that this property of \( y \) does not depend on the particular ideal \( I \) which is involved in the construction of the forcing notion \( Q_I \).

**Claim.** If \( g \) is \( Q_I \)-generic over \( V \), then \( \mathcal{I} \cup \{ g \} \) is an independent family in \( V[g] \), but for all \( y \in [\omega]^{\omega} \cap V \) with \( y \notin \mathcal{I} \), \( \mathcal{I} \cup \{ g, y \} \) is not independent.

**Proof of Claim.** Firstly we show that \( \mathcal{I} \cup \{ g \} \) is an i.f. in \( V[g] \), i.e., we have to show that for every \( x \in \text{bc}(\mathcal{I}) \), both sets \( g \cap x \) and \( (\omega \setminus g) \cap x \) are infinite. For every \( x \in \text{bc}(\mathcal{I}) \) and every \( n \in \omega \) define
\[
A_{n,x} = \{ (s, E) \in Q_I : |x \cap s| > n \},
\]
\[
B_{n,x} = \{ (s, E) \in Q_I : |\bigcup E \cap x| > n \}.
\]
We leave it as an exercise to the reader to show that for all \( x \in \text{bc}(\mathcal{I}) \) and \( n \in \omega \), \( A_{n,x} \) and \( B_{n,x} \) are open dense subsets of \( Q_I \), which implies that \( \mathcal{I} \cup \{ g \} \) is an i.f. in \( V[g] \).

Now, we show that for all \( y \in [\omega]^{\omega} \cap V \) with \( y \notin \mathcal{I} \), \( \mathcal{I} \cup \{ g, y \} \) is not independent: Let \( y \in [\omega]^{\omega} \cap V \) be an arbitrary real. If for all \( u \in I \) and \( x \in \text{bc}(\mathcal{I}) \) we have \( x \notin y \cup u \), then let
\[
C_y = \{ (s, E) \in Q_I : y \in E \},
\]
otherwise, there is a \( u_0 \in I \) and an \( x \in \text{bc}(\mathcal{I}) \) such that \( x \subseteq y \cup u_0 \) and we define
\[
C_y = \{ (s, E) \in Q_I : u_0 \in E \}.
\]
By the properties of the ideal \( I \) we get that \( C_y \) is an open dense subset of \( Q_I \) for all \( y \in [\omega]^{\omega} \). This implies that for each \( y \in [\omega]^{\omega} \) we find an \( x \in \text{bc}(\mathcal{I}) \) such that \( g \cap y \) is finite (in the case when \( y \in I \), or \( g \cap (x \setminus y) \) is finite (in the case when \( x \not\subseteq y \cup u \) for some \( u \in I \)). However, in both cases we get that \( \mathcal{I} \cup \{ g, y \} \) is not independent whenever \( y \in [\omega]^{\omega} \setminus \mathcal{I} \).

**2nd Step:** Now, we are ready to define the finite support iteration which will yield a generic model in which there exists a maximal independent family \( \mathcal{I} \) of cardinality \( \omega_1 \); Let \( V \) be an arbitrary model of \( \text{ZFC} \) in which \( \epsilon > \omega_1 \). We construct the i.f. \( \mathcal{I} \) by induction on \( \alpha \in \omega_1 \). Let \( \mathcal{I}_0 = \emptyset \) and assume that we have already constructed the i.f. \( \mathcal{I}_\alpha \) for some \( \alpha \in \omega_1 \). Furthermore, let \( I_\alpha \subseteq \mathcal{P}(\omega) \) be the ideal constructed in the proof of Lemma 18.10 with respect to the i.f. \( \mathcal{I}_\alpha \), and let \( g_\alpha \) be a \( Q_{I_\alpha} \)-generic real over \( V[\{g_\gamma : \gamma \in \alpha \}] \). Now, let \( \mathcal{I}_{\alpha+1} := \mathcal{I}_\alpha \cup \{ g_\alpha \} \); and for limit ordinals \( \lambda \in \omega_1 \) let \( \mathcal{I}_\lambda := \bigcup_{\beta \in \lambda} \mathcal{I}_\beta \). Notice that for each \( \alpha \in \omega_1 \), \( \mathcal{I}_\alpha = \{ g_\gamma : \gamma \in \alpha \} \) is a countable i.f. in \( V[\{g_\gamma : \gamma \in \alpha \}] \).

Let \( P_{\omega_1} = \langle Q_{I_\alpha} : \alpha \in \omega_1 \rangle \) be the finite support iteration of the forcing notions \( Q_{I_\alpha} \), let \( G = \langle g_\alpha : \alpha \in \omega_1 \rangle \), and let \( \mathcal{I} = \{ g_\alpha : \alpha \in \omega_1 \} \). Then \( G \) is \( P_{\omega_1} \)-generic over \( V \) and \( \mathcal{I} \) is an i.f. in \( V[G] \) of cardinality \( \omega_1 \). It remains to show
that $\mathcal{I}$ is maximal and that $V[G] \models \kappa > \omega_1$. Since $P_{\omega_1}$ is a finite support iteration of $\text{ccc}$ forcing notions (recall that $Q_I$ satisfies $\text{ccc}$), by Proposition 18.8 we get that also $P_{\omega_1}$ satisfies $\text{ccc}$, and therefore, by Lemma 14.20, all cardinals are preserved. In particular, since $V \models \kappa > \omega_1$, we get that $V[G] \models \kappa > \omega_1$.

Furthermore, by Lemma 18.9 we know that the iteration does not add new reals at stage $\omega_1$. Thus, for every real $y \in [\omega]^{\omega} \cap V[G]$ there exists an $\alpha \in \omega_1$ such that $y \in V\left([g_\gamma : \gamma \in \alpha]\right)$. Now, by the Claim we know that for each $y \in [\omega]^{\omega} \cap V\left([g_\gamma : \gamma \in \alpha]\right)$ which does not belong to $\mathcal{I}_\alpha$, $\mathcal{I}_\alpha \cup \{g_\alpha, y\}$ is not independent. Consequently, for each $y \in [\omega]^{\omega} \cap V[G]$ we get that $\mathcal{I} \cup \{y\}$ is not independent whenever $y \notin \mathcal{I}$. This shows that $\mathcal{I}$ is a maximal independent family in $V[G]$, and since $|\mathcal{I}| = \omega_1$ and $\omega_1 < \kappa$, we get that $\omega_1 = i < \kappa$ is consistent with ZFC.

Considering the diagram at the end of Chapter 8, we see that the independence number $i$ appears on the top of the diagram. However, as we have seen above, $i$ can be quite small compared to $\kappa$. In the next chapter we consider a cardinal characteristic on the bottom of the diagram, namely $p$, and show that $p$ can be equal to $\kappa$, even in the case when $\kappa > \omega_1$.

Notes

Products and iterations. For a more detailed introduction to products and iterations of forcing notions we refer the reader to Kunen [5, Chapter VIII], Baumgartner [1], and Goldstern [3]—where one can also find many more applications of these forcing tools. In particular, Proposition 18.5 is taken from Kunen [5, p. 256, Theorem 2.3] and the idea for the proof of Proposition 18.11 is taken from Kunen [5, p. 259, A2-12] (where the actual construction is due to Jörg Brendle).

Related Results

96. Iterating Cohen forcing. A special feature of Cohen forcing $C = (\text{Fn}(\omega, 2), \subseteq)$ is that the set $\text{Fn}(\omega, 2)$ is the same in every transitive model of ZFC. In particular, for any cardinal $\kappa$ we get that (finite/countable support) iterations of length $\kappa$ of Cohen forcing $C$ are equivalent to (finite/countable support) products of $\kappa$ copies of $C$ (cf. Lemma 21.9).

97. Products as two-step iterations. Let $P_0$ and $P_1$ be some forcing notions in some model $V$ of ZFC, let $G$ be $P_0 \times P_1$-generic over $V$, and let $G(0)$ and $G(1)$ be as above. Then $G(0)$ is $P_0$-generic over $V[G(1)]$ and $G(1)$ is $P_1$-generic over $V[G(0)]$ (see for example Kunen [5, Chapter VIII, Theorem 1.4] and compare with Lemma 18.1).

98. Products and the countable chain condition. It is consistent with ZFC that there are forcing notions $P$ and $Q$, both satisfying $\text{ccc}$, such that product $P \times Q$ does not satisfy $\text{ccc}$ (compare with Lemma 18.7). Examples of such forcing notions can be found in Kunen [5, Chapter VIII, p. 291 f].
99. The consistency of \( \varepsilon > \omega_1 \) reidentified. Let \( V \) be a model in which \( \varepsilon > \omega_1 \) and let \( \mathcal{A} \subseteq [\omega]^{\omega} \) be a countable almost disjoint family. With respect to \( \mathcal{A} \) we define the following forcing notion \( \mathbb{Q}_{\mathcal{A}} \): The conditions of \( \mathbb{Q}_{\mathcal{A}} \) are of the form \( (s, X) \), where \( s \) is a finite sequence of \( \omega \) and \( X \in [\omega]^{<\omega} \) and we define \( (s, X) \leq (s', X') \) if \( s \subseteq s' \), \( X \subseteq X' \), and \( (s' \setminus s) \cap X = \emptyset \). For \( \mathbb{B} = \{ B \in [\omega]^\omega : \forall A \in \mathcal{A}(\langle B \cap A \rangle \mathrel{<} \omega) \} \) we get that the generic real \( A \in [\omega]^\omega \), generated by the finite sets \( s \), is almost disjoint from every member of \( \mathcal{A} \) and has infinite intersection with each member of \( \mathcal{B} \) (cf. Kunen [5, Chapter II, Lemma 2.17]). Thus, \( \mathcal{A} \cup \{ A \} \) is a mad family for the old reals (i.e., every real \( x \in [\omega]^\omega \) in the ground model \( V \) has infinite intersection with either \( A \) or an element of \( \mathcal{A} \)). Furthermore, it is not hard to show that the forcing notion \( \mathbb{Q}_{\mathcal{A}} \) satisfies ccc (cf. Kunen [5, Chapter II, Lemma 2.14]). Now, let \( \mathcal{A}_0 \) be an arbitrary countable almost disjoint family in \( V \) and for non-zero ordinals \( \alpha \in \omega_1 \) define \( \mathcal{A}_\alpha \) by transfinite induction as follows: If \( \alpha \) is a limit ordinal, then \( \mathcal{A}_\alpha := \bigcup_{\beta<\alpha} \mathcal{A}_\beta \), and if \( \alpha = \beta + 1 \), then let \( \mathcal{A}_\alpha := \mathcal{A}_\beta \cup \{ A_\beta \} \), where \( A_\beta \in [\omega]^{\omega} \) is \( \mathbb{Q}_{\mathcal{A}_\beta} \)-generic over \( V[\langle A_\beta : \gamma \in \beta \rangle] \). Finally, by the facts mentioned above we get that the finite support iteration \( \langle \mathbb{Q}_{\mathcal{A}_\alpha} : \alpha \in \omega_1 \rangle \), starting in \( V \), yields a model in which we have still \( \varepsilon > \omega_1 \) and in which there exists a mad family of size \( \omega_1 \), namely \( \mathcal{A}_0 \cup \{ A_\alpha : \alpha \in \omega_1 \} \).

100. Easton forcing. With so-called Easton forcing, which is a product forcing notion, one can modify the powers of infinitely many regular cardinals at once. In fact, one can show that cardinal exponentiation on the regular cardinals can be anything not "obviously false". For example one can force that \( \forall \alpha \in \omega (z^{\alpha} = \omega_{\alpha+1}) \), but one cannot force that \( z^{\omega} = \omega_{\omega+1} \) (since \( cf(z^{\omega}) > \omega \)).

For Easton forcing see Easton [2] or Kunen [5, Chapter VIII, §4].

101. Preservation of \( \kappa \)-chain condition. In Chapter 16 Related Result 87 we generalised the notion of ccc by saying that a forcing notion \( \mathbb{P} = (P, \leq) \) satisfies the \( \kappa \)-chain condition if every anti-chain in \( P \) has cardinality \( < \kappa \). Now, if \( \kappa \) is a regular uncountable cardinal and \( \mathbb{P}_\kappa = (\langle Q_\beta : \beta \in \kappa \rangle : \beta \in \kappa) \) is a finite support iteration, where for each \( \beta \in \kappa \) we have \( 0_\beta \vdash \langle Q_\beta : \beta \in \alpha \rangle \) satisfies the \( \kappa \)-chain condition", then \( \mathbb{P}_\kappa \) satisfies the \( \kappa \)-chain condition too (see for example Kunen [5, Chapter VIII, Lemma 5.12] or Jech [4, Part II, Theorem 2.7]).

References

Models in which $p = c$

In this chapter we shall consider models of ZFC in which $p = c$. Since $\omega_1 \leq p$ (by Theorem 8.1) and $p \leq c$, we have $p = c$ in all models in which $c = \omega_1$, but of course, these are not the models we are interested in.

By Theorem 13.6 we know that MA($\sigma$-centred) implies $p = \mathfrak{c}$, moreover, by Chapter 13 Related Result 79 we even have MA($\sigma$-centred) $\iff p = \mathfrak{c}$. On the other hand, in a model in which $\omega_1 < p = c$ we do not necessarily have MA (because MA($\sigma$-centred) is weaker than MA) and in fact it is slightly easier to force just $\omega_1 < p = c$ than to force MA + $\neg$CH. Thus, we shall first construct a model of $\omega_1 < p = c$, which — by Chapter 13 Related Result 79 — proves the consistency of MA($\sigma$-centred) + $\neg$CH with ZFC, and then we shall sketch the construction of a generic model in which we have MA + $\neg$CH. Finally, we shall consider the case when a single Cohen real $c$ is added to a model $\mathcal{V} \vDash \text{ZFC}$ in which MA + $\neg$CH holds. Even though full MA fails in $\mathcal{V}[c]$ (see Related Result 104), we shall see that $p = \mathfrak{c}$ still holds in $\mathcal{V}[c]$ — a result which will be used in Chapter 27.

A Model in which $p = c = \omega_2$

In this section, we shall construct a generic model in which $p = c = \omega_2$ — for the general case see Related Result 102.

**Proposition 19.1.** $p = c = \omega_2$ is consistent with ZFC.

**Proof.** We start with a model $\mathcal{V} \vDash \text{ZFC} + \text{CH}$ in which we have $\mathcal{V} \vDash 2^{\omega_1} = \omega_2$. In order to obtain such a model, use the techniques developed in Chapter 14 or see Chapter 18 Related Result 100.

In $\mathcal{V}$ we shall define a finite support iteration $\mathbb{P}_{\omega_2} = \langle Q_\xi : \xi \in \omega_2 \rangle$ of ccc forcing notions $Q_\xi$, such that in the $\mathbb{P}_{\omega_2}$-generic model $\mathcal{V}[G]$ we have $\mathcal{V} \vDash p = c$. Since for each $\xi \in \omega_2$ the forcing notion $Q_\xi$ will satisfy ccc, by Proposition 18.8 we get that also each $P_\xi$ will satisfy ccc, and therefore, by
Lemma 18.9 and the proof of Theorem 16.4, for any $\xi \in \omega_2$ we shall have $V[G][\xi] \models \mathfrak{c} = \omega_1 \land 2^{\omega_1} = \omega_2$. Furthermore, since for each $\xi \in \omega_2$ the forcing notion $Q_\xi$ will be of cardinality at most $\omega_1$, also $\mathbb{P}_\xi$ will be of cardinality at most $\omega_1$.

Like in the proof of Theorem 16.4, one can show that for any $\nu \in \omega_2$, there are $\omega_1$ nice $\mathbb{P}_\nu$-names for subsets of $\omega$, and because $V[G][\nu] \models 2^{\omega_1} = \omega_2$, for each $\nu \in \omega_2$ there exists a bijection $A_\nu : \omega_2 \rightarrow \mathcal{P}(\omega)\omega$ in $V[G][\nu]$. In particular, for all $\nu, \eta \in \omega_2$ we have $A_\nu(\eta) \subseteq [\omega]^{\omega_1}$, and since $\mathfrak{c} = \omega_1$ we get $|A_\nu(\eta)| \leq \omega_1$. Strictly speaking, we should work with some $\mathbb{P}_\nu$-name for $A_\nu$, not with the actual function, but for the sake of simplicity we shall omit this technical difficulty and leave it as an exercise to the reader.

Now we are ready to construct the ccc forcing notions $Q_\xi$: To start with, fix a bijection $g : \omega_2 \rightarrow \omega_2 \times \omega_2$ in $V$ (which will serve as a bookkeeping function) such that for every $\xi \in \omega_2$ we have

$$\langle g(\xi) = \langle \nu, \eta \rangle \rangle \mapsto \nu \leq \xi.$$ Let $\xi \in \omega_2$ be an arbitrary but fixed ordinal number and let $\langle \nu, \eta \rangle := g(\xi)$. Since $\nu \leq \xi$, $V[G][\nu] \subseteq V[G][\xi]$, and the set $A_\nu(\eta) \subseteq [\omega]^{\omega_1}$, originally defined in $V[G][\nu]$, also belongs to $V[G][\xi]$.

In order to define $Q_\xi = (Q_\xi, \leq)$ we work in $V[G][\nu]$ and consider the following two cases: If the family $A_\nu(\eta) \subseteq [\omega]^{\omega_1}$ has the strong finite intersection property sfiP (i.e., intersections of finitely many members of $A_\nu(\eta)$ are infinite), then we define

$$Q_\xi = \{ \langle s, E \rangle : s \in \text{fin}(\omega) \land E \in \text{fin}(A_\nu(\eta)) \},$$

and for $\langle s, E \rangle, \langle t, F \rangle \in Q_\xi$ we stipulate

$$\langle s, E \rangle \leq \langle t, F \rangle \iff s \subseteq t \land E \subseteq F \land (t \setminus s) \subseteq \bigcap F.$$ In the case when $A_\nu(\eta)$ does not have the sfiP, let $Q_\xi$ be the trivial forcing notion $\langle \emptyset, \subseteq \rangle$.

The forcing notion $Q_\xi$ (in the case when $Q_\xi$ is non-trivial) was already introduced in the proof of Theorem 13.6, where it was shown that $Q_\xi$ satisfies ccc and that the generic filter induces a pseudo-intersection of $A_\nu(\eta)$. Hence, we either have $V[G][\xi] = V[G][\xi]$, or the family $A_\nu(\eta)$ has a pseudo-intersection in $V[G][\xi + 1]$. In particular, the family $A_\nu(\eta)$, which is a family of cardinality at most $\omega_1$, is not a witness for $p = \omega_1$.

Let $G$ be $\mathbb{P}_{\omega_2}$-generic over $V$ and let $\mathcal{F} \subseteq [\omega]^{\omega_1}$ be an arbitrary family in $V[G]$ of cardinality $\omega_1$ which has the sfiP. Since for each $\xi \in \omega_2$, $Q_\xi$ satisfies ccc, by Proposition 18.8, also $\mathbb{P}_{\omega_2}$ satisfies ccc, and therefore, by Lemma 18.9, $V[G][\xi] \models \mathfrak{c} = \omega_2$.

Since $|\mathcal{F}| = \omega_1$, similar to Claim 2 in the proof of Proposition 24.12, there exists a $\nu \in \omega_2$ such that the family $\mathcal{F}$ belongs to $V[G][\nu]$. In particular, there is an $\eta \in \omega_2$ such that $V[G][\nu] \models \mathcal{F} = A_\nu(\eta)$. Hence, for $\xi = g^{-1}(\langle \nu, \eta \rangle)$,
there is a pseudo-intersection for $\mathcal{F}$ in $V[G|_{\xi+1}]$, and since $\mathcal{F}$ was arbitrary, we get $V[G] \models p \geq \omega_2$. Now, since $V[G] \models c = \omega_2$, we finally get $V[G] \models p = c = \omega_2$.

**On the Consistency of MA + \neg CH**

In this section we shall sketch the proof that MA $+$ $\neg$ CH is consistent with ZFC (for the general case see Related Result 103). The crucial point in the proof is the fact that every ccc forcing notion is equivalent to a forcing notion of cardinality strictly less than $c$, but let us recall first Martin’s Axiom:

**Martin’s Axiom (MA):** If $\mathbb{P} = (P, \leq)$ is a partially ordered set which satisfies ccc, and $\mathcal{D}$ is a set of less than $c$ open dense subsets of $P$, then there exists a $\mathcal{D}$-generic filter on $P$.

At first glance, we can build a model in which we have MA $+$ $\neg$ CH by starting in some model of ZFC $+$ $\neg$ CH, and then add a $\mathcal{D}$-generic filters for every partially ordered set $\mathbb{P} = (P, \leq)$ satisfying ccc. However, the collection of all partially ordered sets satisfying ccc is a proper class. So, we first have to show that it is enough to consider just the set of ccc partially ordered sets $\mathbb{P} = (P, \leq)$ satisfying $|P| < c$.

**Lemma 19.2.** The following statements are equivalent:

(a) MA.

(b) If $\mathbb{P} = (P, \leq)$ is a partially ordered set that satisfies ccc and $|P| < c$, and if $\mathcal{D}$ is a set of less than $c$ open dense subsets of $P$, then there exists a $\mathcal{D}$-generic filter on $P$.

**Proof.** Obviously it is enough to prove that (b) implies (a): Let $P$ be a ccc partially ordered set, and let $\mathcal{D}$ be a family of fewer than $c$ open dense subsets of $P$, i.e., $|\mathcal{D}| = \kappa$ for some $\kappa < c$. For each $D \in \mathcal{D}$, let $A_D \subseteq D$ be a maximal incompatible subset of $D$. Then, since $\mathbb{P}$ satisfies ccc, each $A_D$ is countable. Now, we can construct a set $Q \subseteq P$ of cardinality at most $\kappa$ such that $Q$ contains each $A_D$, and whenever $p, q \in Q$ are compatible in $P$, then they are also compatible in $Q$ (i.e., there is an $r \in Q$ such that $p \leq r \geq q$) — for the latter notice that $|\kappa|^2 = \kappa$. By construction of $Q$ we get that for each $D \in \mathcal{D}$, $A_D$ is a maximal anti-chain in $Q$. Finally, for each $D \in \mathcal{D}$ let $E_D = \{ q \in Q : \exists p \in A_D (q \geq p) \}$. Then each $E_D$ is open dense in $Q$.

Now, $(Q, \leq)$ is a partially ordered set which satisfies ccc and $|Q| \leq \kappa < c$. Thus, by (b), there is a filter $G$ on $Q$ that meets every open dense set $E_D$, and consequently, $G = \{ p \in P : \exists q \in G (p \leq q) \}$ is a $\mathcal{D}$-generic filter on $P$. 

}\end{proof}
Proposition 19.3. MA + $\kappa = \omega_2$ is consistent with ZFC.

Proof (Sketch). The proof is essentially the same as the proof of Proposition 19.1. We start again in a model $V$ of ZFC in which $\kappa = \omega_1$ and $2^{\omega_1} = \omega_2$, and extend $V$ by a finite support iteration $P_{\omega_2} = \langle Q_\xi : \xi \in \omega_2 \rangle$, where for each $\xi \in \omega_2$, $Q_\xi = (Q_\xi, \leq)$ satisfies ccc and $Q_\xi \subseteq \omega_1$. Since in the final model $V[G]$ we have $\kappa = \omega_2$, by Lemma 19.2 we can arrange the iteration so that every ccc forcing notion in $V[G]$ of size $< \omega_2$ is isomorphic to some forcing notion $Q_\xi$ (for some $\xi \in \omega_2$). A minor problem is that by adding new generic sets, we also might add new dense subsets to old partially ordered sets. This problem is solved by making sure that every ccc forcing notion $Q_\xi$ appears arbitrarily late in the iteration, which is done by a bookkeeping function similar to that used in the proof of Proposition 19.1.

$p = \kappa$ is Preserved under Adding a Cohen Real

The following result, which will be used in the proof of Proposition 27.9, shows that $p = \kappa$ is preserved under adding a Cohen real (cf. Related Result 104).

Theorem 19.4. If $V \models p = \kappa$ and $c$ is a Cohen real over $V$, then $V[c] \models p = \kappa$.

Proof. Throughout this proof, we shall consider the Cohen forcing notion $C = (\bigcup_{\alpha < \omega} 2^{\alpha}, \leq)$. Let $V$ be a model of ZFC and let $c \in 2^{\omega}$ be a Cohen real over $V$.

If $V \models \text{CH}$, then also $V[c] \models \text{CH}$ which implies $V[c] \models p = \kappa$. So, let us assume that $V \models \kappa > \omega_1$ and therefore, since Cohen forcing preserves cardinals, $V[c] \models \kappa > \omega_1$.

We have to show that every family $\{X_\alpha \in [\omega]^\omega : \alpha < \kappa \}$ in $V[c]$ which has the sfp has also a pseudo-intersection. To start with, fix a cardinal $\kappa$ with $\omega_1 < \kappa < \kappa$, and let $\{X_\alpha : \alpha \in \kappa \} \subseteq [\omega]^\omega$ be an arbitrary but fixed family in $V[c]$ which has the sfp. Furthermore, let

$$\{X_\alpha : \alpha \in \kappa\}$$

be a set of $\mathbb{C}$-names such that $\{X_\alpha[c] : \alpha \in \kappa\} = \{X_\alpha : \alpha \in \kappa\}$. Now, since $\{X_\alpha : \alpha \in \kappa\}$ has the sfp in $V[c]$, there exists a $\mathbb{C}$-condition $q$ such that for all $E \in \text{fin}(\kappa)$ we have

$$q \Vdash \bigcap \{X_\alpha : \alpha \in E\} = \omega,$$

where we define $\bigcap \emptyset := \omega$. For the sake of simplicity, let us assume that $q = 0$. The goal is now to construct a set $Y \subseteq V[c]$ which is a pseudo-intersection of $\{X_\alpha[c] : \alpha \in \kappa\}$. For this, we define (in $V$) the following $\sigma$-centred forcing notion $\mathbb{P} = (P, \leq)$:
$p = \mathfrak{c}$ is preserved under adding a Cohen real

The set of $\mathbb{P}$-conditions $P$ consists of pairs $\langle h, A \rangle$, where $A \in \text{fin}(\omega)$ and

$$ h : \bigcup \{ k^2 : k \in m \} \to \text{fin}(\omega) \quad \text{for some } m \in \omega. $$

For $\langle h, A \rangle, \langle l, B \rangle \in P$, let $\langle h, A \rangle \leq \langle l, B \rangle$ if and only if

- $h \subseteq l$, $A \subseteq B$, and
- for each $p \in \text{dom}(l) \setminus \text{dom}(h)$ we have $p \Vdash_{\mathbb{P}} l(p) \subseteq \bigcap \{ X_\alpha : \alpha \in A \}.$

We leave it as an exercise to the reader to show that $|P| = \kappa$ and that $\mathbb{P}$ is $\alpha$-centred — for the latter, notice that for any $\langle h, A \rangle, \langle h, B \rangle \in P$ we have $\langle h, A \rangle \leq \langle h, A \cup B \rangle \geq \langle h, B \rangle$. Now, for every $\alpha \in \kappa$ and $n \in \omega$ we define the set $D_{\alpha,n} \subseteq P$ by stipulating $\langle h, A \rangle \in D_{\alpha,n}$ if and only if

- $\alpha \in A$,
- $\text{dom}(h) = \{ k^2 : k \in m \}$ for some $m \geq n$,
- for each $p \in \mathbb{m}^2$, $|\bigcup_{i \in m} h(p_i)| \geq n$.

We leave it as an exercise to the reader to show that every set $D_{\alpha,n}$ is an open dense subset of $P$ and that $\left| \{ D_{\alpha,n} : \alpha \in \kappa \land n \in \omega \} \right| = \kappa$. The open dense sets $D_{\alpha,n}$ make sure that the set $Y$, constructed below, will be a pseudo-intersection of $\{ X_\alpha[c] : \alpha \in \kappa \}$, in particular, $Y$ will be infinite. At the moment, just notice the following fact: If $\langle h, A \rangle \in D_{\alpha,n}$ and $\langle h, A \rangle \leq \langle l, B \rangle$, where $\text{dom}(l) = \{ k^2 : k \in m \}$, then for each $p \in \mathbb{m}^2$ we have $|\bigcup_{i \in m} l(p_i)| \geq n$, and for each $p \in \text{dom}(l) \setminus \text{dom}(h)$ we have $p \Vdash_{\mathbb{P}} l(p) \subseteq X_\alpha$.

The crucial point is now to show that there exists a filter $G \subseteq P$ in $\mathcal{V}$ which meets every set $D_{\alpha,n}$.

**Claim.** Let $\mathcal{D} = \{ D_{\alpha,n} : n \in \omega \land \alpha \in \kappa \}$. Then there exists in $\mathcal{V}$ a $\mathcal{D}$-generic filter $G$ on $P$, i.e., there exists a directed and downwards closed set $G \subseteq P$ which meets every open dense subset of $P$ which belongs to $\mathcal{D}$.

**Proof of Claim.** The following proof is essentially the proof of the fact that $p = \mathfrak{c}$ is equivalent to $\text{MA}(\alpha$-centred) (see Chapter 13 | RELATED RESULT 79).

Firstly notice that for each $m \in \omega$ there are just countably many functions $h : \bigcup \{ k^2 : k \in m \} \to \text{fin}(\omega)$. For each $m \in \omega$ fix an enumeration $\{ h_{m,i} : i \in \omega \}$ of all these countably many functions and let $\eta : \omega \times \omega \to \omega$ be a bijection. For each $n \in \omega$ we define the set $P_n \subseteq P$ by stipulating

$$ P_n = \{ \langle h_{m,i}, A \rangle \in P : \eta(m, i) = n \}. $$

Notice that $\bigcup_{n \in \omega} P_n = P$ and that each $P_n$ consists of pairwise compatible $\mathbb{P}$-conditions.

Secondly, for each $\mathbb{P}$-condition $p = \langle h, A \rangle \in P$ and for every open dense set $D \in \mathcal{D}$ let

$$ [p, D] = \{ n \in \omega : \exists q \in P_n(q \in D \land q \geq p) \}. $$
Notice that \([p, D] \in \omega^\omega\). Furthermore, for all \(k, r \in \omega\), any \(\mathbb{P}\)-conditions \(\langle h, A_0 \rangle, \ldots, \langle h, A_k \rangle \in \mathbb{P}_r\), and any open dense sets \(D_0, \ldots, D_k \in \mathcal{D}\), we get that \(\bigcap_{i \leq k} [\langle h, A_i \rangle, D_i]\) is infinite. This implies that for each \(r \in \omega\), the family \(\mathcal{F}_r = \{ [p, D] : p \in \mathbb{P}_r \land D \in \mathcal{D} \}\) has the sfrp. Now, since \(V \vDash p = c\) and \(|\mathcal{F}_r| = |\mathbb{P}_r \times \mathcal{D}| \leq \kappa \times \kappa = \kappa < c\), we have \(V \vDash |\mathcal{F}_r| < p\). Hence, in \(V\) there exists a pseudo-intersection \(I_r\) of \(\mathcal{F}_r\). In other words, for every \(r \in \omega\) there is an \(I_r \in \omega^\omega\) such that for all \(p \in \mathbb{P}_r\) and \(D \in \mathcal{D}\), \(I_r \setminus [p, D]\) is finite.

In the following step we encode the elements of the sets \(I_r\) by finite sequences: Let \(\text{seq}(\omega)\) be the set of all finite sequences which can be formed with elements of \(\omega\). For \(s \in \text{seq}(\omega)\) and \(i \in \omega\), \(\bar{s}i\) denotes the concatenation of the sequences \(s\) and \(\langle i \rangle\).

Now, define the function \(\nu : \text{seq}(\omega) \to \omega\) by stipulating

- \(\nu(\emptyset) = 0\), and
- for all \(s \in \text{seq}(\omega)\): \(\nu(\bar{s}i) : i \in \omega\) enumerates \(I_{\nu(s)}\) in ascending order.

In particular, \(\nu(\langle i \rangle) : i \in \omega\) = \(I_0\), where for all \(i, i' \in \omega\), \(i < i'\) implies \(\nu(\langle i \rangle) < \nu(\langle i' \rangle)\).

Furthermore, for every \(D \in \mathcal{D}\) and every \(s \in \text{seq}(\omega)\) we choose a \(\mathbb{P}\)-condition \(p_D^s \in P_{I_{\nu(s)}}\) such that for all \(i \in \omega\,

\[ \nu(\bar{s}i) \in [p_{D^s}, D] \rightarrow (p_{D^s} \leq p_D) \land (p_{D^s} \in D). \]  

(*)

Notice that for any \(D \in \mathcal{D}\) and \(s \in \text{seq}(\omega)\), \(I_{\nu(s)} \setminus [p_{D^s}, D]\) is finite. Thus, for each \(D \in \mathcal{D}\) and each \(s \in \text{seq}(\omega)\) there is a least integer \(g_D(s) \in \omega\) such that for every \(i \geq g_D(s)\) we have \(\nu(\bar{s}i) \in [p_{D^s}, D]\). So, for every \(D \in \mathcal{D}\), we obtain a function \(g_D : \text{seq}(\omega) \to \omega\). Then, the family \(\mathcal{E} = \{ g_D : D \in \mathcal{D}\}\) is a family of size \(\kappa\) of functions from the countable set \(\text{seq}(\omega)\) to \(\omega\).

Now we show that \(\mathcal{E}\) is bounded: For this, recall first that for the bounding number \(b\) we have \(p \leq b \leq c\) (see Chapter 8). Since in \(V\) we have \(p = c\), in particular \(V \vDash b = c\), and since \(|\mathcal{E}| = \kappa < c\), \(V \vDash |\mathcal{E}| < b\). Thus, \(\mathcal{E}\) is bounded in \(V\), i.e., in \(V\) there exists a function \(g : \text{seq}(\omega) \to \omega\) such that for each \(D \in \mathcal{D}\),

\[ g_D(s) < g(s) \]  

for all but finitely many \(s \in \text{seq}(\omega)\).

By induction on \(n \in \omega\), define the function \(f \in \omega^\omega\) such that for all \(n \in \omega\,

\[ f(n) := g(f|_n). \]  

Then, by definition of \(f\) and the property of \(g\), for each \(D \in \mathcal{D}\),

\[ g_D(f|_n) < f(n) \]  

for all but finitely many \(n \in \omega\).

In other words, for every \(D \in \mathcal{D}\) there exists an integer \(m_D \in \omega\) such that for all \(n \geq m_D\), \(f(n) > g_D(f|_n)\).

We are now ready to define the \(\mathcal{D}\)-generic set \(G \subseteq P\), but before we do so, let us summarise a few facts which we have achieved so far: Let \(D \in \mathcal{D}\) and \(n \geq m_D\) be arbitrary, and let \(s := f|_n\) and \(i := f(n)\).
$p = c$ is preserved under adding a Cohen real

(0) $f(n) = g(f[n]) = g(s)$, i.e., $i = g(s)$, and $f(n + 1) = g(f[n + 1]) = g(s^i)$.

(1) Since $n \geq m_D$, we get $g(f[n]) > g_D(f[n])$, i.e., $g(s) > g_D(s)$, and therefore $i > g_D(s)$.

(2) Since $i > g_D(s)$, we get $\nu(s^i) \in [p^*_D, D]$, i.e.,

$$\nu(f[n + 1]) \in [p^*_D, D].$$

(3) Thus, by (2) we get $p^*_D \leq p^*_D$ and $p^*_D \in D$, i.e.,

$$p^*_D \leq p^*_D + 1 \quad \text{and} \quad p^*_D \in D.$$ Now, let $G \subseteq P$ be defined by

$$G = \{q \in P : \exists D \in \mathcal{D} \exists n \in \omega \ (n \geq m_D \land q \leq p^*_D)\}.$$ It remains to check that $G$ has the required properties, i.e., $G$ is a filter which meets every $D \in \mathcal{D}$.

$G$ is a filter: By definition, $G$ is downwards closed. To see that $G$ is directed, take any $q, q' \in G$ and, for some $D, D' \in \mathcal{D}$ and $n, n' \in \omega$, let $p^*_D, p^*_D \in G$ be such that $q \leq p^*_D$ and $q' \leq p^*_D$. Without loss of generality we may assume that $n \geq n'$. Then $p^*_D \geq p^*_D$. Now, $p^*_D$ and $p^*_D$ both belong to $P_{\nu(f[n])}$ and are therefore compatible. Thus, there exists an $r \in P_{\nu(f[n])}$ such that $p^*_D \leq r \geq p^*_D$, and consequently we have $q \leq r \geq q'$ where $r \in G$.

$G$ is $\mathcal{D}$-generic: By (3), for each $D \in \mathcal{D}$ and every $n \geq m_D$ we have $p^*_D \in D \cap G$, and hence, $G \cap D \neq \emptyset$.

With the $\mathcal{D}$-generic filter $G \subseteq P$ constructed above we define the function

$$H = \bigcup \{h : \exists \nu(h, A) \in G\}.$$ By construction, the function $H : \bigcup_{n \in \omega} \omega^2 \to \omega$ has the following property: If $\alpha \in \kappa$ and $\langle h, A \rangle \in G$ with $\alpha \in A$, then for every $p \in \bigcup_{n \in \omega} \omega^2 \setminus \text{dom}(h)$ we have

$$p \not\leq H(p) \subseteq X_\alpha.$$ In particular, if $c$ is a Cohen real over $V$, then for $Y := \bigcup_{n \in \omega} H(c[n])$, which is a set in $V[c]$, we have

$$V[c] \models \forall \alpha \in \kappa \ Y \subseteq^+ X_\alpha[c].$$ We leave it as an exercise to the reader to show that $V[c] \models |Y| = \omega$ (for this, recall the definition of the open dense sets $D_{\alpha,n}$). Thus, in $V[c]$, the arbitrarily chosen family $\{X_\alpha[c] : \alpha \in \kappa \land \alpha \leq c\}$ has a pseudo-intersection, which shows that $V[c] \models V[c] = c$. 


Notes

The consistency of $\text{MA} + \lnot \text{CH}$. A complete proof for the consistency of $\text{MA} + \lnot \text{CH}$ with ZFC can be found for example in Kunen [5, Chapter VIII, §6] (see also Martin and Solovay [6]).

On $p = \kappa$ after adding one Cohen real. Theorem 19.4 is due to Roitman [7], but the proof given here follows the proof of Bartoszyński and Judah [1, Theorem 3.3.8], where the proof of the Claim, originally proved by Bell [2], is taken from Fremlin [3, 14C].

Related Results

102. On the consistency of $p = \kappa$. Let $V$ be a model of ZFC + GCH and assume that in $V$, $\kappa$ is an uncountable regular cardinal such that $|\kappa|^{<\kappa} = \kappa$. Then, by a slight modification of the proof of Proposition 19.1, we get a generic extension of $V$ in which $p = \kappa$.

103. On the consistency of $\text{MA} + \varepsilon = \kappa$. As in Related Result 102, let $V$ be a model of ZFC + GCH and assume that in $V$, $\kappa$ is an uncountable regular cardinal such that $|\kappa|^{<\kappa} = \kappa$. Then there exists a ccc forcing notion $P$ in $V$, such that in the $P$-generic extension $V[G]$ we have $\text{MA} + \varepsilon = \kappa$ (for a proof see Kunen [5, Chapter VIII, Theorem 6.3]).

104. Martin’s Axiom and Cohen reals. By Chapter 13 | Related Result 79, which asserts $p = \omega \iff \text{MA}(\sigma\text{-centred})$, we get that $V \vDash \text{MA}(\sigma\text{-centred})$ if and only if $V \vDash p = \omega$. Hence, Theorem 19.4 implies that $\text{MA}(\sigma\text{-centred})$ is preserved under Cohen forcing, i.e., if $V \vDash \text{MA}(\sigma\text{-centred})$ and $c$ is a Cohen real over $V$, then $V[c] \vDash \text{MA}(\sigma\text{-centred})$. However, this is not the case for full MA. In fact one can show that if $V \vDash \lnot \text{CH}$ and $c$ is a Cohen real over $V$, then $V[c] \vDash \lnot \text{MA}$. The proof uses the fact that if $V \vDash \text{MA}(\omega_1)$, then there is no Sulsin tree in $V$ (see for example Jech [4, Theorem 16.16]). On the other hand, one can show that whenever $c$ is a Cohen real over $V$, then $V[c]$ contains a Sulsin tree (see Shelah [8, §4], Todorcevic [9], or Bartoszyński and Judah [1, Section 3.3.1]).

References


Part III

Combinatorics of Forcing Extensions
...the parts sing one after another in so-called fugue (fuga) or consequence (consequenza), which some also call reditta. All mean the same thing: a certain repetition of some notes or of an entire melody contained in one part by another part, after an interval of time. The second part sings the same note values or different ones, and the same intervals of whole tones, semitones, or similar ones.

There are two type of fugues or consequences namely strict and free.

In free writing, the imitating voice duplicates the other in fugue or consequence only up to a point; beyond that point it is free to proceed independently.

**Gi奥seffo Zarlino**

*Le Istituzioni Harmoniche*, 1558
Properties of Forcing Extensions

In this chapter we shall introduce some combinatorial properties of forcing notions which will accompany us throughout the remainder of this book. Furthermore, these properties will be the main tool in order to investigate various combinatorial properties of generic models of ZFC.

However, before we start with some definitions, let us modify our notation concerning names in the forcing language: Let \( \mathbb{P} \) be a forcing notion and let \( G \) be \( \mathbb{P} \)-generic over some ground model \( \mathbb{V} \).

- Instead of canonical \( \mathbb{P} \)-names for sets in \( \mathbb{V} \) like \( \emptyset \), \( 27 \), \( \omega \), et cetera, we just write \( \emptyset \), \( 27 \), \( \omega \), et cetera.
- If \( f \) is a \( \mathbb{P} \)-name for a function in \( \mathbb{V}[G] \) with domain \( A \in \mathbb{V} \) and \( a \in A \), then we write \( f(a) \) instead of \( f\langle a \rangle \).

For example, if \( \mathbb{P} = \mathbb{C} \) and \( c \) is the canonical name for a Cohen real \( c \in \check{\omega} \), then, for \( k \in \omega \), \( c(k) = \{ \langle m, p \rangle : p \in \bigcup_{n \in \omega}^\omega \omega \land k \in \text{dom}(p) \land p(k) = m \} \) denotes the canonical \( \mathbb{C} \)-name for the integer \( c(k) \) — properly denoted by \( c(k) \).

Dominating, Splitting, Bounded, and Unbounded Reals

First we recall some notions defined in Chapter 8: For two functions \( f, g \in \check{\omega} \) we say that \( g \) is dominated by \( f \), denoted \( g \prec^* f \), if there is an \( n \in \omega \) such that for all \( k \geq n \) we have \( g(k) < f(k) \). For two sets \( x, y \in [\omega]_\omega \) we say that \( x \) splits \( y \) if \( y \cap x \) as well as \( y \setminus x \) is infinite.

Now let \( \mathbb{V} \) be any model of ZFC and let \( \mathbb{V}[G] \) be a generic extension (i.e., \( G \) is \( \mathbb{P} \)-generic over \( \mathbb{V} \) with respect to some forcing notion \( \mathbb{P} \)). Let \( f \in \check{\omega} \) be a function in the model \( \mathbb{V}[G] \). Then \( f \) is called a dominating real (over \( \mathbb{V} \)) if each function \( g \in \check{\omega} \cap \mathbb{V} \) is dominated by \( f \), and \( f \) is called an unbounded real (over \( \mathbb{V} \)) if it is not dominated by any function \( g \in \check{\omega} \cap \mathbb{V} \). Furthermore, a set \( x \in [\omega]_\omega \) in \( \mathbb{V}[G] \) is called a splitting real (over \( \mathbb{V} \)) if it splits each set
y ∈ [ω]ω in the ground model V. Notice that we identify functions f ∈ ωω with real numbers.

**Fact 20.1.** If V[G] contains a dominating real, then it also contains a splitting real.

**Proof.** We can just follow the proof of Theorem 8.4: Whenever a function f ∈ ωω belongs V[G], then also the set

$$\sigma_f = \bigcup \{ [f^{2n}(0), f^{2n+1}(0)) : n ∈ ω \}$$

belongs to V[G], where \([a, b) = \{ k ∈ ω : a ≤ k < b \} \) and \(f^{n+1}(0) = f(f^n(0))\) with \(f^0(0) := 0\). Now let \(f ∈ ωω\) be a dominating real. Without loss of generality we may assume that \(f\) is strictly increasing and that \(f(0) > 0\). Fix any \(x ∈ [ω]ω ∩ V\) and let \(g_x : ω → x\) be the (unique) strictly increasing bijection between \(ω\) and \(x\). Since \(f\) is dominating we have \(g_x <^* f\), which implies that there is an \(n_0 ∈ ω\) such that for all \(k ≥ n_0\) we have \(g_x(k) < f(k)\). For each \(k ∈ ω\) we have \(k ≤ f^k(0)\) as well as \(k ≤ g_x(k)\). Moreover, for \(k ≥ n_0\) we have

$$f^k(0) ≤ g_x(f^k(0)) < f(f^k(0)) = f^{k+1}(0)$$

and therefore \(g_x(f^k(0)) ∈ [f^k(0), f^{k+1}(0))\). Thus, for all \(k ≥ n_0\) we have \(g_x(f^k(0)) ∈ \sigma_f\) iff \(k\) is even, which shows that both \(x \cap \sigma_f \cap x\) and \(x \setminus \sigma_f\) are infinite. Hence, since \(x ∈ [ω]ω\) was arbitrary, \(\sigma_f\) is a splitting real.

It is worth mentioning that the converse of Fact 20.1 does not hold, i.e., we cannot construct a dominating real from a splitting real (cf. Lemma 21.2 and Lemma 21.3).

A forcing notion \(P\) is said to add dominating (unbounded, splitting) reals if every \(P\)-generic extension of V contains a dominating (unbounded, splitting) real. More formally, let \(V ⊨ ZFC\) and let \(P ∈ V\) be a forcing notion. Then we say that

\[\begin{align*}
P \text{ adds dominating reals } & \iff 0 \Vdash \exists f ∈ ωω \forall g ∈ ωω (g <^* f), \\
P \text{ adds unbounded reals } & \iff 0 \Vdash \exists f ∈ ωω \exists g ∈ ωω (f <^g g), \\
\end{align*}\]

and

\[\begin{align*}
P \text{ adds splitting reals } & \iff 0 \Vdash \exists x ⊆ ω \forall y ∈ [ω]ω (|y ∩ x| = |y \setminus x| = ω). \\
\end{align*}\]

Notice that in this context, i.e., in statements being forced, ωω and [ω]ω stand for the canonical names for sets in the ground model, whereas for example \(0\) is a \(P\)-name for the set \(ω\) in the \(P\)-generic extension.

A forcing notion \(P\) is called \(ω\)-bounding if there are no unbounded reals in \(P\)-generic extensions. In other words, if \(P\) is \(ω\)-bounding and \(G\) is \(P\)-generic over \(V\), then every function \(f ∈ ωω\) in \(V[G]\) is dominated by some
function from the ground model $V$. Obviously, a forcing notion which adds a dominating real also adds unbounded reals and therefore cannot be $\omega$-bounding, and by Fact 20.1, such a forcing notion also adds splitting reals. On the other hand, none of these implications is reversible. An example of a forcing notion which is $\omega$-bounding but adds splitting reals is Silver forcing (investigated in Chapter 22), and Cohen forcing, discussed in the next chapter, is an example of a forcing notion which adds unbounded and splitting reals but does not add dominating reals. Furthermore, Miller forcing (discussed in Chapter 23) adds unbounded reals but does not add splitting reals, and Mathias forcing (discussed in Chapter 24) adds dominating reals but does not add Cohen reals.

**The Laver Property and Not Adding Cohen Reals**

In the following chapters we shall investigate different forcing notions like Cohen forcing, Silver forcing, Mathias forcing, et cetera. In fact, we shall investigate what kind of new reals (e.g., dominating reals or Cohen reals) are added by (an iteration of) a given forcing notion. In particular, we have to decide whether an iteration of a given forcing notion adds Cohen reals. Our main tool to solve this problem will be the following combinatorial property.

**Laver Property:** Let $\mathcal{F}$ be the set of all functions $S : \omega \to \text{fin}(\omega)$ such that for every $n \in \omega$, $|S(n)| \leq 2^n$. A forcing notion $\mathbb{P}$ has the **Laver property** if and only if for every function $f \in \omega^\omega \cap V$ in the ground model and every $\mathbb{P}$-name $g$ for a function in $\omega^\omega$ such that $0 \Vdash \forall n \in \omega(g(n) \leq f(n))$, we have $0 \Vdash \exists S \in \mathcal{F} \cap V \forall n \in \omega(g(n) \in S(n))$.

Roughly speaking, if a forcing notion has the Laver property, then for every function $g \in \omega^\omega$ in the generic extension which is bounded by a function from the ground model, and for every $n \in \omega$, the value $g(n)$ belongs to some finite set of size $2^n$ and the sequence of these finite sets is in the ground model.

Now we show that a forcing notion which has the Laver property does not add Cohen reals.

**Proposition 20.2.** If the forcing notion $\mathbb{P}$ has the Laver property, then $\mathbb{P}$ does not add Cohen reals.

**Proof.** Suppose that $\mathbb{P}$ has the Laver property. Let $\{I_n : n \in \omega\}$ be a partition of $\omega$ (in the ground model $V$) such that for all $n \in \omega$, $|I_n| = 2n$ and $\max(I_n) < \min(I_{n+1})$. Let $h$ be a $\mathbb{P}$-name for an arbitrary element of $\omega^2$, i.e., $0 \Vdash h \in \omega^2$.

We show that $h$ is not the name for a Cohen real, i.e., $h$ is not the name for a real which corresponds to a $\mathbb{C}$-generic filter over $V$, where $\mathbb{C} = \bigcup_{n \in \omega} \mathbb{C}_n \subseteq \subseteq$.

For every $n \in \omega$, let $H(n) := h|_{I_n}$. Then $H(n) : I_n \to 2$, and since $|I_n| = 2n$, $H(n)$ amounts to an element of $2^{2n}$. Thus, we can encode $H(n)$
by a $\mathbb{P}$-name for an integer in $2^{2^n}$; let $\eta(H(n))$ be that code and let $g(n) := \eta(H(n))$. Thus, $0 \Vdash P \forall n \in \omega(g(n) \leq 2^{2^n})$, and since $P$ has the Laver property, $0 \Vdash \exists S \in \mathcal{F} \cap V \forall n \in \omega(g(n) \in S(n))$. In the ground model $V$, let $p_0$ be a $\mathbb{P}$-condition such that for some $S \in \mathcal{F} \cap V$ we have $p_0 \Vdash \forall n \in \omega(g(n) \in S(n))$.

Further, let

$$D = \left\{ s \in \bigcup_{n \in \omega}^{n \in \omega} n^2 : \exists k (I_k \subseteq \text{dom}(s) \land \eta(s|I_k) \notin S(k)) \right\}.$$ 

Then $D$ is an open dense subset of $\bigcup_{n \in \omega}^{n \in \omega} n^2$. Indeed, for any $m \in \omega$ and any $t \in m^2$ there exists $k > m$ such that $I_k \cap \text{dom}(t) = \emptyset$, and we find an $s \in \bigcup_{n \in \omega}^{n \in \omega} n^2$ such that $t \subseteq s$, $I_k \subseteq \text{dom}(s)$, and $\eta(s|I_k) \notin S(k)$ — here we use that for any positive integer $k$, $|S(k)| \leq 2^k < 2^{2^k} = |I_k^2|$.

Now, for every $n \in \omega$ define $A_n = \{ x \in \omega^2 : \eta(x|I_n) \in S(n) \} \subseteq \omega^2$ and let $A = \bigcap_{n \in \omega} A_n$. Since $p_0 \Vdash \forall n \in \omega(g(n) \in S(n))$, we have $p_0 \Vdash h \in A$, and consequently we get that $p_0 \Vdash \forall k \in \omega(h|I_k \notin D)$. Hence, $h$ is not a $\mathbb{P}$-name for a Cohen real over $V$, which completes the proof.

So, we know that if a forcing $P$ has the Laver property, then forcing with $P$ does not add Cohen reals; but what can we say about products or iterations of $P$? On the one hand, it is possible that $P \times P$ adds Cohen reals, even though $P$ has the Laver property (see for example Chapter 24). On the other hand, in the next section we shall see that the Laver property is preserved under countable support iteration of proper forcing notions. More precisely, if $P$ is a forcing notion which is proper (see below) and has the Laver property, then any countable support iteration of $P$ has the Laver property, and therefore does not add Cohen reals.

### Proper Forcing Notions and Preservation Theorems

#### The Notion of Properness

By Proposition 18.8 we know that finite support iterations of ccc forcing notions satisfy ccc. In other words, ccc is preserved under finite support iteration of ccc forcing notions. Below, we shall present a generalisation of that result, but before we have to introduce some preliminary definitions: For every infinite regular cardinal $\chi$ let

$$H_\chi = \{ x \in V_\chi : |\text{TC}(x)| < \chi \}.$$ 

For example the sets in $H_\omega$ are the hereditarily finite sets and the sets in $H(\omega_1)$ are the hereditarily countable sets. Notice that each $H_\chi$ is transitive and that $x \in H_\chi$ iff $|\text{TC}(x)| < \chi$, i.e., $H_\chi$ contains all sets which are hereditarily of cardinality $< \chi$. It is worth mentioning that for every regular uncountable cardinal $\chi$, $H_\chi$ is a model of ZFC minus the Axiom of Power Set (cf. Chapter 15 | Related Result 84).
For the following discussion, let $\chi$ be a “large enough” regular cardinal, where “large enough” means that for all forcing notions $\mathbb{P} = (P, \leq)$ we shall consider in the forthcoming chapters we have $\mathcal{P}(P) \in H_\chi$, i.e., the power set of $P$ is hereditarily of size $< \chi$. If we assume that GCH holds in the ground model, then $\chi = \omega_3$ would be sufficient, but to be on the safe side we let

$$\chi = \beth_\omega,$$

where the so-called beth function $\beth_\alpha$ is defined by induction on $\alpha \in \Omega$, stipulating $\beth_0 := \omega$, $\beth_{\alpha+1} := 2^{\beth_\alpha}$, and for limit ordinals $\alpha$, $\beth_\alpha := \bigcup \{ \beth_\beta : \beta < \alpha \}$.

Let $\mathcal{N} = (N, \in)$ be an elementary submodel of $(H_\chi, \in)$, i.e., $(N, \in) \prec (H_\chi, \in)$. Furthermore, let $\mathbb{P} = (P, \leq)$ be a forcing notion such that $(P, \leq) \in \mathcal{N}$. Since $\mathcal{N}$ is an elementary submodel of $(H_\chi, \in)$, for all $p, q \in P \cap N$ we have $N \models p \perp q$ implies $V \models p \perp q$, i.e., if $p$ and $q$ are incompatible in $\mathcal{N}$, then they are also incompatible in the ground model $V$. We say that $G \subseteq P$ is $\mathcal{N}$-generic for $\mathbb{P}$ if $G$ has the following property.

Whenever $D \in N$ and $\mathcal{N} \models \text{“} D \subseteq P \text{ is an open dense subset of } P \text{”},$

$G \cap N \cap D \neq \emptyset$.

Notice that $G$ is $\mathcal{N}$-generic iff $G \cap N$ is $\mathcal{N}$-generic. By Fact 14.6, we can replace “open dense” for example by “maximal anti-chain”. Furthermore, we say that a condition $q \in P$, which is not necessarily in $\mathcal{N}$, is $\mathcal{N}$-generic if

$\mathcal{N} \models q \Vdash \text{“} G \text{ is } \mathcal{N}\text{-generic} \text{”,}$

where $G$ is the canonical $\mathbb{P}$-name for the $\mathbb{P}$-generic filter over the ground model $V$. Notice that if $q$ is $\mathcal{N}$-generic and $q' \geq q$, then $q'$ is $\mathcal{N}$-generic too.

Now, a forcing notion $\mathbb{P} = (P, \leq)$ is called proper, if for all countable elementary submodels $\mathcal{N} = (N, \in) \prec (H_\chi, \in)$ which contain $\mathbb{P}$, and for all conditions $p \in P \cap N$, there exists a condition $q \geq p$ (in $V$) which is $\mathcal{N}$-generic.

As a first example let us show that any forcing notion $\mathbb{P} = (P, \leq)$ which satisfies ccc is proper: Firstly, for any countable set $A \in N$ we have $A \subseteq N$. For this, notice that since $(N, \in) \prec (H_\chi, \in)$, $A$ must be the range of a function $f : \omega \to \bigcup N$ which belongs to $\mathcal{N}$, and since for all $n \in \omega$, $n \in N$, we also have $f(n) \in N$ for all $n \in \omega$, which shows that $A \subseteq N$. Now, let $A \in \mathcal{N}$ be a maximal anti-chain in $P$. Then, since $\mathbb{P}$ satisfies ccc, $A$ is countable and we have $A \subseteq N$. Further, $\emptyset \Vdash A \cap G \neq \emptyset$, and therefore, $\emptyset \Vdash A \cap \mathcal{N} \cap G = A \cap G \neq \emptyset$.

As a second example let us show that any forcing notion $\mathbb{P}(P, \leq)$ which is $\sigma$-closed is proper: Since the model $\mathcal{N}$ is countable, there are just countably many open dense subsets of $P$ which belong to $\mathcal{N}$, say $\{D_n : n \in \omega \}$. Let $p \in P \cap N$ and let $\langle q_n : n \in \omega \rangle$ be such that $q_0 \geq p$ and for each $n \in \omega$, $q_{n+1} \geq q_n$, and $q_n \in D_n$. Now, since $\mathbb{P}$ is $\sigma$-closed, we find a condition $q$ such that for all $n \in \omega$, $q \geq q_n$. Obviously, $q \geq p$ and $q$ is $\mathcal{N}$-generic.
Let us finish this section by introducing a property of forcing notions which is slightly stronger than properness, but which is often easier to verify than properness (e.g., for the forcing notions introduced in the forthcoming chapters).

**Axiom A**: A forcing notion $\mathbb{P} = (P, \leq)$ is said to satisfy Axiom A if there exists a sequence $\{\leq_n : n \in \omega\}$ of orderings on $P$ (not necessarily transitive) which has the following properties:

1. For all $p, q \in P$, if $q \leq_{n+1} p$ then $q \leq_n p$ and $q \leq p$.
2. If $(p_n \in P : n \in \omega)$ is a sequence of conditions such that $p_n \leq_{n+1} p_{n+1}$, then there exists a $q \in P$ such that for all $n \in \omega$, $p_n \leq_n q$.
3. If $A \subseteq P$ is an anti-chain, then for each $p \in P$ and every $n \in \omega$ there is a $q \in P$ such that $p \leq_n q$ and $\{r \in A : r \text{ and } q \text{ are compatible}\}$ is countable.

Examples of forcing notions satisfying Axiom A are forcing notions which are $\sigma$-closed or satisfy ccc. Furthermore, one can show that every forcing notion which satisfies Axiom A is proper, but not vice versa (for a proof and a counterexample see Baumgartner [4], Theorem 2.4 and Section 3 respectively).

**Preservation Theorems for Proper Forcing Notions**

Below, we state without proofs some preservation theorems for countable support iteration of proper forcing notions. These preservation theorems will be crucial in the following chapters, where we consider countable support iterations of length $\omega_2$ of various proper forcing notions — usually starting with a model in which CH holds.

The first of these preservation theorems states that proper forcing notions do not collapse $\omega_1$ and that properness is preserved under countable support iteration of proper forcing notions (for proofs see Goldstern [6, Corollary 3.14] and Shelah $\mathfrak{P}$, III. §3).

**Theorem 20.3.** (a) If $\mathbb{P}$ is proper and $\text{cf}(\delta) > \omega$, then $0_\mathbb{P} \Vdash \text{cf}(\delta) > \omega$. In particular, $\omega_1$ is not collapsed.

(b) If $\mathbb{P}_\alpha$ is a countable support iteration of $\langle \mathbb{P}_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have $0_{\mathbb{P}_\beta} \Vdash \mathbb{Q}_\beta$ is proper”, then $\mathbb{P}_\alpha$ is proper.

The following lemma is in fact just a consequence of Theorem 20.3.

**Lemma 20.4.** Let $\mathbb{P}_\alpha$ be a countable support iteration of $\langle \mathbb{Q}_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have $0_{\mathbb{Q}_\beta} \Vdash \mathbb{Q}_\beta$ is a proper forcing notion of size $\leq \mathfrak{c}$. If CH holds in the ground model and $\alpha \leq \omega_2$, then for all $\beta \in \alpha$, $0_{\mathbb{Q}_\beta} \Vdash \text{CH}$.
Since, by Lemma 18.9, no new reals appear at the limit stage $\omega_2$ one can prove the following theorem — a result which we shall use quite often in the forthcoming chapters.

**Theorem 20.5.** Let $P_{\omega_2}$ be a countable support iteration of $\langle Q_\beta : \beta \in \omega_2 \rangle$, where for each $\beta \in \omega_2$ we have

$0_\beta \Vdash " Q_\beta$ is a proper forcing notion of size $\leq \kappa$ which adds new reals". Further, let $V$ be a model of ZFC + CH and let $G$ be $P_{\omega_2}$-generic over $V$. Then we have

(a) $V[G] \models \kappa = \omega_2$, and

(b) for every set of reals $\mathcal{F} \subseteq [\omega]^\omega \cap V[G]$ of size $\leq \omega_1$ there is a $\beta \in \omega_2$ such that $\mathcal{F} \subseteq V[G_{|\beta}]$.

Now, let us say a few words concerning preservation of the Laver property and of $\omega$-boundedness: It can be shown that a countable support iteration of proper $\omega$-bounding forcing notions is $\omega$-bounding (for a proof see Section 5 and Application 1 of Goldstern [6]).

**Theorem 20.6.** If $P_\alpha$ is a countable support iteration of $\langle Q_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have $0_\beta \Vdash " Q_\beta$ is proper and $\omega$-bounding" , then $P_\alpha$ is $\omega$-bounding.

Further, one can show that the Laver property is preserved under countable support iteration of proper forcing notions which have the Laver property (for a proof see Section 5 and Application 4 of Goldstern [6]).

**Theorem 20.7.** If $P_\alpha$ is a countable support iteration of $\langle Q_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have $0_\beta \Vdash " Q_\beta$ is proper and has the Laver property" , then $P_\alpha$ has the Laver property.

Another property which is preserved under countable support iteration of proper forcing notions is preservation of $P$-points: A forcing notion $P$ is said to **preserve $P$-points** if for every $P$-point $\mathcal{U} \subseteq [\omega]^\omega$,

$0_\alpha \Vdash " \mathcal{U}$ generates an ultrafilter over $\omega"$,

i.e., for every set $x \in [\omega]^\omega$ in the $P$-generic extension there exists a $y \in \mathcal{U}$ such that either $y \subseteq x$ or $y \subseteq \omega \setminus x$. In particular, if the forcing notion $P$ is proper and CH holds in the ground model, then the ultrafilter in the $P$-generic extension which is generated by the $P$-point $\mathcal{U}$ is again a $P$-point.

One can show that preservation of $P$-points is preserved under countable support iteration of proper forcing notions (for a proof see Blass and Shelah [5] or Bartoszyński and Judah [2, Theorem 6.2.6]).

**Theorem 20.8.** If $P_\alpha$ is a countable support iteration of $\langle Q_\beta : \beta \in \alpha \rangle$, where for each $\beta \in \alpha$ we have $0_\beta \Vdash " Q_\beta$ is proper and preserves $P$-points" , then $P_\alpha$ preserves $P$-points.
There are many more preservation theorems for countable support iteration of proper forcing notions. However, what we presented here is all that we shall use in the forthcoming chapters.

Notes

The notion of properness, which is slightly more general than Axiom A (introduced by Baumgartner [3]), was discovered and investigated by Shelah [8, 9], who realised that properness is a property that is preserved under countable support iteration and that allows to prove several preservation theorems (see for example Shelah [9, VI. §§1–2], where one can find also proofs of the preservation theorems given above). For a brief introduction to proper forcing we refer the reader to Goldstern [6] and for applications of the Proper Forcing Axiom, which is a generalisation of Martin’s Axiom, see Baumgartner [4].

Related Results

105. Reals of minimal degree of constructibility. Let $\mathbb{P} = (P, \leq)$ be a forcing notion and let $g$ be a real in some $\mathbb{P}$-generic extension of $V$. Then $g$ is said to be of minimal degree of constructibility, or just minimal, if $g$ does not belong to $V$ and for every real $f$ in $V[g]$ we have either $f \in V$ or $g \in V[f]$, where $V[f]$ is the smallest model of ZFC containing $f$ and $V$. In the latter case we say that $f$ reconstructs $g$. For example no Cohen real is minimal. Indeed, if $c \in {}^\omega \omega$ is a Cohen real over $V$, then the real $c' \in {}^\omega \omega \cap V[c]$ defined by stipulating $c'(n) := c(2n)$ (for all $n \in \omega$) is also a Cohen real over $V$. Moreover, $c$ is even $\mathbb{P}$-generic over $V[c']$, which implies that $c$ does not belong to $V[c']$.

106. Alternative definitions of properness. The notion of properness can also be defined in terms of games or with stationary sets (see for example Jech [7, Part III] or Baumgartner [4, Section 2]).

107. Preservation of ultrafilters. In general, a forcing notion which adds reals does not preserve all ultrafilters. More precisely, for any forcing notion which adds a new real, say $r$, to the ground model $V$, there exists an ultrafilter $\mathcal{U}$ in $V$ which does not generate an ultrafilter in $V[r]$, (see Bartoszyński, Goldstern, Judah, and Shelah [1] or Bartoszyński and Judah [2, Theorem 6.2.2]). Further, one can show that any forcing notion which adds Cohen, dominating, or random reals, does not preserve $P$-points (see Bartoszyński and Judah [2, Theorem 7.2.22]).

References

References

Cohen Forcing revisited

Properties of Cohen Forcing

Since Cohen forcing is countable, it satisfies ccc, hence, Cohen forcing is proper. Furthermore, since forcing notions with the Laver property do not add Cohen reals, Cohen forcing obviously does not have the Laver property.

Not so obvious are the facts that Cohen forcing adds unbounded and splitting, but no dominating reals.

Cohen Forcing adds Unbounded but no Dominating Reals


Proof. Consider Cohen forcing \( \mathbb{C} = \left( \bigcup_{i \in \omega} \omega, \subseteq \right) \), which is — as we have seen in Chapter 18 — equivalent to the forcing notion \( \left( \bigcup_{i \in \omega} \omega, \subseteq \right) \). Let \( c \in \omega^\omega \) be \( \mathbb{C} \)-generic over some ground model \( V \) and let \( \mathcal{C} \) be the canonical \( \mathbb{C} \)-name for \( c \). We show that the function \( c \) is not dominated by any function \( g \in \omega^\omega \cap V \).

Firstly notice that for every \( \mathbb{C} \)-condition \( p \) we have

\[ p \Vdash \mathcal{C} \upharpoonright \text{dom}(p) = p \, . \]

Let \( g \in \omega^\omega \) be any function in the ground model \( V \) (i.e., \( g \in \omega^\omega \cap V \)) and let \( n \in \omega \). Then there exist \( k \geq n \) and a \( \mathbb{C} \)-condition \( q \geq p \) such that \( k \in \text{dom}(q) \) and \( q(k) > g(k) \). This implies that for every \( n \in \omega \), the set of \( \mathbb{C} \)-conditions \( q \) such that

\[ q \Vdash \exists k \geq n \ (g(k) < c(k)) \]

is open dense in \( \bigcup_{i \in \omega} \omega \). Hence, there is no \( \mathbb{C} \)-condition which forces that \( c \) is dominated by some function \( g \in \omega^\omega \cap V \). Consequently, \( c \) is not dominated by any function from the ground model, or in other words, the function \( c \in \omega^\omega \) is unbounded.

\[ \square \]
Lemma 21.2. Cohen forcing does not add dominating reals.

Proof. Consider Cohen forcing $\mathbb{C} = (\text{Fn}(\omega, 2), \subseteq)$. Let $c \in \omega^\omega$ be $\mathbb{C}$-generic over some ground model $V$. Further, let $f \in \omega^\omega$ be an arbitrary but fixed function in $V[c]$ and let $f$ be a $\mathbb{C}$-name for $f$. In order to show that $f$ is not dominating we have to find a function $g \in \omega^\omega \cap V$ such that for every $n \in \omega$ there is a $k \geq n$ such that $g(k) \geq f(k)$. Let $\{p_k : k \in \omega\}$ be an enumeration of $\text{Fn}(\omega, 2)$, i.e., $\{p_k : k \in \omega\} = \text{Fn}(\omega, 2)$. For every $k \in \omega$ define

$$g(k) = \min \{ n : \exists q \geq p_k (q \forces_c f(k) = n) \}.$$

For every $\mathbb{C}$-condition $p$ and every $n \in \omega$ there is a $k \geq n$ such that $p_k \geq p$, and we find a $q \geq p_k$ such that $q \forces_c f(k) = g(k)$. Consequently, for every $n \in \omega$, the set of $\mathbb{C}$-conditions $q$ such that

$$q \forces_c \exists k \geq n (f(k) = g(k))$$

is open dense in $\text{Fn}(\omega, 2)$. Hence, $g \in \omega^\omega \cap V$ is not dominated by $f \in V[c]$, and since $f$ was arbitrary, this shows that there are no dominating functions in $V[c]$, or in other words, Cohen forcing does not add dominating reals. $\Box$

Cohen Forcing adds Splitting Reals


Proof. Consider Cohen forcing $\mathbb{C} = (\bigcup_{n \in \omega} \omega^2, \subseteq)$. We show that any real $c$ which is $\mathbb{C}$-generic over some ground model $V$ generates a splitting real: Let $\sigma_c := \{ k \in \omega : c(k) = 1 \}$ and let $\sigma_c$ be its canonical $\mathbb{C}$-name. Then for any infinite set $x \in [\omega]^\omega \cap V$ and any $n \in \omega$, the set of $\mathbb{C}$-conditions $p$ such that

$$p \forces_c |x \cap \sigma_c| > n \land |x \setminus \sigma_c| > n$$

is open dense, and therefore, $\sigma_c$ splits every real in the ground model $V$, or in other words, $\sigma_c$ is a splitting real. $\Box$

Cohen Reals and the Covering Number of Meagre Sets

Below, we shall give a topological characterisation of Cohen reals, but before we introduce a topology on $\omega^\omega$ and show how to encode “basic” meagre sets by reals.

For each finite sequence $s = \langle n_0, \ldots, n_k-1 \rangle$ of natural numbers, i.e., $s \in \text{seq}(\omega)$, define the basic open set

$$O_s = \{ f \in \omega^\omega : f|_k = s \}.$$
well as \( {}^\omega \omega \) are open. Notice that a set \( A \subseteq {}^\omega \omega \) is open iff for all \( x \in A \) there exists an \( s \in \text{seq}(\omega) \) such that \( x \in O_s \subseteq A \). Furthermore, a set \( A \subseteq {}^\omega \omega \) is called \textbf{closed} (in \( {}^\omega \omega \)) if \( {}^\omega \omega \setminus A \) is open. Evidently, arbitrary unions and finite intersections of open sets are open; or equivalently, arbitrary intersections and finite unions of closed sets are closed. On the other hand, an intersection of countably many open sets is not necessarily open, and a union of countably many closed sets is not necessarily closed (see below). Now, intersections of countably many open sets are called \textbf{G}_\delta \textbf{ sets}, and unions of countably many closed sets are called \textbf{F}_\sigma \textbf{ sets}. Notice that every open (closed) set is a \( G_\delta \) set (\( F_\sigma \) set), and that by De Morgan laws, each \( F_\sigma \) set is the complement of a \( G_\delta \) set and vice versa. For example the set \( C_0 \subseteq {}^\omega \omega \) of eventually constant functions (i.e., \( f \in C_0 \) iff there is an \( n \in \omega \) such that \( f|_{\omega \setminus n} \) is constant) is an \( F_\sigma \) set which is neither closed nor open.

A subset of \( {}^\omega \omega \) is \textbf{dense} (in \( {}^\omega \omega \)) if it meets every non-empty open subset of \( {}^\omega \omega \). For example \( C_0 \) is dense in \( {}^\omega \omega \). Notice that every dense subset of \( {}^\omega \omega \) must be infinite. On the other hand, \( A \subseteq {}^\omega \omega \) is called \textbf{nowhere dense} if \( {}^\omega \omega \setminus A \) contains an open dense set. Notice that every nowhere dense set is contained in a closed nowhere dense set (i.e., the closure of a nowhere dense set is nowhere dense).

Now, a subset of \( {}^\omega \omega \) is called \textbf{meagre} if it is contained in the union of countably many nowhere dense sets. For example \( C_0 \) is meagre. Since the closure of a nowhere dense set is nowhere dense, we get that every meagre set is contained in some meagre \( F_\sigma \) set, and that the complement of a meagre set contains a co-meagre \( G_\delta \) set. Moreover, we have the following result.

\textbf{Theorem 21.4 (Baire Category Theorem).} The intersection of countably many open dense sets is dense. In particular, the complement of meagre set is always dense.

\textit{Proof.} Let \( \langle D_n : n \in \omega \rangle \) be a sequence of open dense subsets of \( {}^\omega \omega \). We have to show that \( D = \bigcap_{n \in \omega} D_n \) is dense, i.e., we have to show that for each basic open set \( O_\alpha \), \( D \cap O_\alpha \neq \emptyset \). Let \( O_\alpha \) be an arbitrary but fixed basic open set. By induction on \( n \in \omega \) we construct a sequence \( t_0 \subseteq t_1 \subseteq \ldots \) of elements of \( \text{seq}(\omega) \) such that \( \bigcap_{n \in \omega} O_{t_n} \subseteq D \cap O_\alpha \). In fact, we just have to make sure that \( \bigcup_{n \in \omega} t_n \subseteq {}^\omega \omega \) and that for all \( n \in \omega \), \( O_{t_n} \subseteq D_n \). Since \( D_0 \) is open dense, there exists a \( t_0 \in \text{seq}(\omega) \) such that \( s_0 \subseteq t_0 \) and \( O_{t_0} \subseteq (D_0 \cap O_{t_0}) \). Assume that \( t_n \in \text{seq}(\omega) \) is already constructed for some \( n \in \omega \). Then, since \( D_{n+1} \) is open dense, there is a \( t_{n+1} \in \text{seq}(\omega) \) such that \( O_{t_{n+1}} \subseteq (D_{n+1} \cap O_{t_n}) \) and \( |t_{n+1}| \geq n + 1 \). Now, by construction, the sequence \( t_0 \subseteq t_1 \subseteq \ldots \) has the required properties.

By definition, subsets of meagre sets as well as countable unions of meagre sets are meagre. Thus, the collection of meagre subsets of \( {}^\omega \omega \), denoted by \( \mathcal{M} \), is an ideal on \( \mathcal{P}({}^\omega \omega) \). By the Baire Category Theorem 21.4, \( {}^\omega \omega \notin \mathcal{M} \) but for every \( f \in {}^\omega \omega \) we have \( \{ f \} \in {}^\omega \omega \), and therefore the set \( {}^\omega \omega \) can be covered by \( \mathcal{M} \) meagre sets. This observation leads to the following cardinal number.
Definition. The covering number of $\mathcal{M}$, denoted $\text{cov}(\mathcal{M})$, is the smallest number of sets in $\mathcal{M}$ with union $^\omega \omega$, more formally

$$\text{cov}(\mathcal{M}) = \min \{|\mathcal{C}| : \mathcal{C} \subseteq \mathcal{M} \land \bigcup \mathcal{C} = {^\omega \omega}\}.$$ 

Since countable unions of meagre sets are meagre, and since we can cover $^\omega \omega$ by $\mathfrak{c}$ meagre sets, we obviously have $\omega_1 \leq \text{cov}(\mathcal{M}) \leq \mathfrak{c}$. Moreover, we can show slightly more:

**Theorem 21.5.** $p \leq \text{cov}(\mathcal{M})$.

**Proof.** Let $\{A_\alpha : \alpha \in \kappa < p\}$ be any infinite family of cardinality $\kappa < p$ of meagre subsets of $^\omega \omega$. We have to show that $\bigcup_{\alpha \in \kappa} A_\alpha \neq {^\omega \omega}$, or equivalently, we have to show that for any family $D = \{D_\alpha : \alpha \in \kappa < p\}$ of open dense subsets of $^\omega \omega$ we have $\bigcap D \neq \emptyset$. Notice the similarity with the proof of the Baire Category Theorem 21.4. Let $\nu : \text{seq}(\omega) \to \omega$ be a bijection. For every $s \in \text{seq}(\omega)$ and every $\alpha \in \kappa$, let

$$I_{s,\alpha} = \{t \in \text{seq}(\omega) : s \subseteq t \land O_t \subseteq D_\alpha\}.$$ 

Since $D_\alpha$ is open dense, the set $y_{s,\alpha} := \{\nu(t) : t \in I_{s,\alpha}\}$ is an infinite subset of $\omega$.

For the moment, let $s$ be an arbitrary but fixed element of $\text{seq}(\omega)$. Then for any finitely many ordinals $\alpha_0, \ldots, \alpha_{k-1}$ in $\kappa$ we get that $\bigcap_{\alpha \in \kappa} y_{s,\alpha} \in [\omega]^\omega$.

Consider the family $\mathcal{F}_s = \{y_{s,\alpha} : \alpha \in \kappa\} \subseteq [\omega]^\omega$. Obviously, $\mathcal{F}_s$ has the strong finite intersection property, and since $\kappa < p$, $\mathcal{F}_s$ has a pseudo-intersection, say $x_s$. Thus, for every $\alpha \in \kappa$ there exist a $k \in \omega$ such that $x_s \setminus k \subseteq y_{s,\alpha}$.

Now, for each $\alpha \in \kappa$ define $h_\alpha : \text{seq}(\omega) \to \omega$ by stipulating $h_\alpha(s) := \min\{k \in \omega : x_s \setminus k \subseteq y_{s,\alpha}\}$, and let $g_\alpha \in {^\omega \omega}$ be such that for all $n \in \omega$, $g_\alpha(n) := h_\alpha(\nu^{-1}(n))$. Since $\kappa < p$ and $p \leq \mathfrak{b}$, there is a function $f \in {^\omega \omega}$ which dominates each $g_\alpha$. By construction,

$$U = \bigcup_{s \in \text{seq}(\omega)} \{O_t \subseteq {^\omega \omega} : \nu(t) \in x_s \setminus f(\nu(s))\}$$

is an open dense subset of $^\omega \omega$ which has the property that for each $\alpha \in \kappa$ there is a finite set $E_\alpha \in [\text{seq}(\omega)]^\omega$ such that $U_{E_\alpha} \subseteq D_\alpha$, where for $E \subseteq \text{seq}(\omega)$,

$$U_E = \bigcup_{s \in \text{seq}(\omega)} \{O_t \subseteq {^\omega \omega} : \nu(t) \in x_s \setminus f(\nu(s)) \land t \notin E\}.$$ 

Notice that for each $E \in [\text{seq}(\omega)]^\omega$, $U_E$ is open dense, and since there are only finitely many finite subsets of $\text{seq}(\omega)$, by the Baire Category Theorem 21.4 we get that

$$T = \bigcap \{U_E : E \in [\text{seq}(\omega)]^\omega\}$$

is dense, and since $T$ is contained in each $D_\alpha$ we have $T \subseteq \bigcap_{\alpha \in \kappa} D_\alpha$. \qed
With a product of Cohen forcing we shall construct a model in which $p < \text{cov}(\mathcal{M})$ (see Corollary 21.11). The crucial point in the construction will be the fact that Cohen reals over $V$ are not contained in any meagre $F_{\sigma}$ set which can be encoded (explained later) by a real $r \in \omega$ which is not the ground model $V$. For this, we have to show the relationship between Cohen reals and meagre sets and have to explain how to encode meagre $F_{\sigma}$ sets by real numbers; but first we give the relationship between Cohen reals and open dense subsets of $\omega$.

Consider Cohen forcing $C = \left( \bigcup_{n \in \omega} \omega^n, \subseteq \right)$. To every $\mathcal{C}$-condition $s$ we associate the open set $O_s \subseteq \omega$. Similarly, to every dense set $D \subseteq \bigcup_{n \in \omega} \omega^n$ we associate the set

$$\mathcal{D}(D) = \bigcup \{ O_s \subseteq \omega : s \in D \},$$

which is an open dense subset of $\omega$. On the other hand, if $O \subseteq \omega$ is an open dense subset of $\omega$, then the set

$$\mathcal{D}(O) = \{ s \in \bigcup_{n \in \omega} \omega^n : O_s \subseteq O \}$$

is an open dense subset of $\bigcup_{n \in \omega} \omega^n$. Notice that for every open dense set $O \subseteq \omega$ there is a dense set $D \subseteq \bigcup_{n \in \omega} \omega^n$ such that $O = \mathcal{D}(D)$. Hence, if $c \in \omega$ is a Cohen real over $V$, then in $V[c]$ we have

$$V[c] = \{ c \in \bigcap \{ \mathcal{D}(D) : D \text{ is dense in } \bigcup_{n \in \omega} \omega^n \wedge D \in V \} \}.$$ 

Considering the dense set $D \subseteq \bigcup_{n \in \omega} \omega^n$ as the code for the open dense set $\mathcal{D}(D) \subseteq \omega$, we get the following

**Fact 21.6.** A real $c \in \omega$ is a Cohen real over $V$ if and only if $c$ is contained in every open dense subset of $\omega$ whose code belongs to $V$.

In order to make the notion of codes more precise, we show how one can encode meagre $F_{\sigma}$ sets by real numbers $r \in \omega$. For this, take two bijections $h_1 : \omega \to \text{seq}(\omega)$ and $h_2 : \omega \times \omega \to \omega$, and for $r \in \omega$ let

$$\eta_r : \omega \times \omega \to \text{seq}(\omega), \quad \langle n, m \rangle \mapsto h_1 \left( h_2(n, m) \right).$$

For every $F_{\sigma}$ set $A = \bigcup_{m \in \omega} \bigcap_{n \in \omega} \omega \setminus O_{s_{n,m}}$, there is a real $r \in \omega$, called code of $A$, such that for all $n, m \in \omega$ we have $\eta_r(n, m) = s_{n,m}$. On the other hand, for every real $r \in \omega$ let $A_r \subseteq \omega$ be defined by

$$A_r = \{ f \in \omega : \exists m \in \omega \forall n \in \omega (\eta_r(n, m) \notin f) \}.$$ 

As a countable union of closed sets, $A_r$ is an $F_{\sigma}$ set. Thus, every real $r \in \omega$ encodes an $F_{\sigma}$ set, and vice versa, every $F_{\sigma}$ set can be encoded by a real
$r \in \omega$. Now, an $F_\omega$ set $A = \bigcup_{m \in \omega} \bigcap_{n \in \omega} \omega \setminus O_{n,m}$ is meagre iff $\bigcup_{n \in \omega} O_{n,m}$ is dense for each $m \in \omega$. So, for $r \in \omega$ we have

$A_r$ is meagre iff $\forall s \in \text{seq}(\omega) \forall m \in \omega \exists n \in \omega (s \subseteq \eta_r(n,m) \vee \eta_r(n,m) \subseteq s)$.

The way we have defined $A_r$, it does not only depend on the real $r$, but also on the model in which we construct $A_r$ from $r$ (notice that this fact also applies to the sets $O_s$). So, in order to distinguish the sets $A_r$ constructed in different models, for models $V$ of ZFC and $r \in \omega \cap V$ we write

$$A_r^V = \{ f \in \omega \cap V : \exists m \in \omega \forall n \in \omega (\eta_r(n,m) \not\subseteq f) \}.$$  

By Fact 21.6 we get that if $c \in \omega$ is a Cohen real over $V$, then $c$ is not contained in any meagre $F_\omega$ set $A_r^V$ with $r \in \omega \cap V$. Now, let $V$ and $V'$ be two transitive models of ZFC. Then, for every real $r \in \omega$ which belongs to both models $V$ and $V'$ we have

$$V \models A_r^V \text{ is meagre} \iff V' \models A_r^{V'} \text{ is meagre}.$$  

As a consequence we get the following characterisation of Cohen reals:

**Proposition 21.7.** Let $V$ be a model of ZFC, let $\mathbb{P}$ be a forcing notion in $V$, and let $G$ be $\mathbb{P}$-generic over $V$. Then the real $c \in \omega \cap V[G]$ is a Cohen real over $V$ if and only if $c$ does not belong to any meagre $F_\omega$ set $A_r^{V[G]}$ with code $r$ in $V$.

**Proof.** ($\Rightarrow$) If $c \in \omega \cap V[G]$ belongs to some meagre $F_\omega$ set $A_r^{V[G]}$ with code $r$ in $V$, then $c \in \bigcup_{m \in \omega} \bigcap_{n \in \omega} \omega \setminus O_{n,m}$. Thus, there is an $m_0 \in \omega$ such that $c$ does not belong to the open dense set $\bigcup_{n \in \omega} O_{n,m_0}$ . Now, consider Cohen forcing $\mathbb{C} = (\bigcup_{n \in \omega} \omega, \subseteq)$ and let $D := \{ \eta_r(n,m_0) : n \in \omega \}$ . Then $D$ is an open dense subset of $\bigcup_{n \in \omega} \omega$ which belongs to the model $V$. On the other hand we have $\{ c\mid n : n \in \omega \} \cap D = \emptyset$ which shows that $c$ is not a Cohen real over $V$.

($\Leftarrow$) Firstly, recall that every meagre set is contained in some meagre $F_\omega$ set and that $A_r^V$ is meagre iff $A_r^{V[G]}$ is meagre, and secondly, notice that $A_r^V \subseteq A_r^{V[G]}$ . Hence, a real $c \in \omega \cap V[G]$ which does not belong to any meagre $F_\omega$ set $A_r^{V[G]}$ with code $r$ in $V$ does belong to every open dense subset of $\omega \cap \omega$ whose code belongs to $V$, and therefore, by Fact 21.6, $c$ is a Cohen real over $V$.

**Corollary 21.8.** Let $\mathbb{P}$ be a forcing notion which does not add Cohen reals and let $G$ be $\mathbb{P}$-generic over $V$, where $V$ is a model of ZFC + $\mathbf{CH}$. Then $V[G] \models \text{cov} (\mathcal{H}) = \omega_1$, in particular, $V[G] \models p = \omega_1$.

**Proof.** In $V$, let $C = \{ r \in \omega : A_r \text{ is meagre} \}$. Then $|C| = \omega_1$ and we obviously have $\bigcup_{c \in C} A_r = \omega$. In other words, the set of meagre sets $\{ A_r :
A model in which \( a < \omega = \tau = \text{cov}(\mathcal{M}) \)

\( r \in C \) is of cardinality \( \omega_1 \) which covers \( \omega \). Now, since \( P \) does not add Cohen reals, in \( V[G] \) we have \( \omega_\omega \backslash \bigcup_{r \in C} A^V \subseteq \emptyset \). Hence, \( V[G] \models \bigcup_{r \in C} A^V[G] = \omega_\omega \), and since \( \text{cov}(\mathcal{M}) \) is uncountable we get \( V[G] \models \text{cov}(\mathcal{M}) = \omega_1 \). In particular, by Theorem 21.5, \( V[G] \models p = \omega_1 \).

We have seen that \( \text{cov}(\mathcal{M}) \leq \kappa \) and by Theorem 21.5 we know that \( p \leq \text{cov}(\mathcal{M}) \). So, \( \text{cov}(\mathcal{M}) \) is an uncountable cardinal number which is less than or equal to \( \kappa \). Below, we shall compare the covering number \( \text{cov}(\mathcal{M}) \) with other cardinal characteristics of the continuum and give a model of ZFC in which \( p < \text{cov}(\mathcal{M}) \).

**A Model in which \( a < \omega = \tau = \text{cov}(\mathcal{M}) \)**

The following lemma will be crucial in our proof that \( \omega_1 < \tau = \text{cov}(\mathcal{M}) = \kappa \) is consistent with ZFC (cf. Lemma 18.1 and Chapter 18|Related Result 97).

**Lemma 21.9.** Let \( \alpha \) be an ordinal number, let \( C^{\alpha+1} \) be the finite support product of \( \alpha + 1 \) copies of Cohen forcing \( \mathcal{C} = (\text{Fn}(\alpha, 2), \subseteq) \), and let \( G \) be \( C^{\alpha+1} \)-generic over some model \( V \) of ZFC. Then \( G(\alpha) \) is \( \mathcal{C} \)-generic over \( V[G] \), in particular, \( \bigcup G(\alpha) = \text{a Cohen real over } V[G] \).

**Proof.** Firstly notice that since \( \text{Fn}(\alpha, 2) \) contains only finite sets, for all transitive models \( V', V'' \) of ZFC we have \( \text{Fn}(\alpha, 2)^{V'} = \text{Fn}(\alpha, 2)^{V''} \), i.e., \( \text{Fn}(\alpha, 2) \) is the same in all transitive models of ZFC, and consequently we get \( C^{V'} = C^{V''} \).

In particular, \( C^{V[G]} = C^V \).

To simplify the notation, let us work with the forcing notion \( C_\alpha = (\text{Fn}(\alpha \times \omega, 2), \subseteq) \) instead of \( C^\alpha \) (recall that by Proposition 18.3, \( C_\alpha \approx C^\alpha \)). Now, in the model \( V[G] \), fix an arbitrary dense set \( D \subseteq \text{Fn}(\alpha, 2) \) and let \( D \) be a \( C_\alpha \)-name for \( D \). Further, let \( p_0 \in G \) be such that

\[ p_0 \Vdash \text{"}D\text{ is dense in } \text{Fn}(\alpha, 2)\text{"}, \]

and let

\[ E = \left\{ (q_0, q_1) \in \text{Fn}(\omega \times \alpha, 2) \times \text{Fn}(\omega, 2) : q_0 \geq p_0 \land q_0 \Vdash q_1 \in D \right\}. \]

We leave it as an exercise to the reader to show that \( E \), which is a subset of \( \text{Fn}(\omega \times \alpha, 2) \times \text{Fn}(\omega, 2) \), is dense above \( \langle p_0, 0 \rangle \). Thus, since \( \langle p_0, 0 \rangle \in (G(\alpha) \times G(\alpha)) \), there is some \( \langle q_0, q_1 \rangle \in (G(\alpha) \times G(\alpha)) \cap E \). So, \( q_0 \Vdash q_1 \in D \) where \( q_0 \in G(\alpha) \), and since \( q_1 \in G(\alpha) \) we get that \( q_1 \in D \) which shows that \( G(\alpha) \cap D \neq \emptyset \). Since \( D \subseteq \text{Fn}(\omega, 2) \) was chosen arbitrarily, we finally get that \( G(\alpha) \) is \( C \)-generic over \( V[G] \), or in other words, \( \bigcup G(\alpha) = \text{a Cohen real over } V[G] \).

**Proposition 21.10.** \( \omega_1 < \omega = \tau = \text{cov}(\mathcal{M}) = \kappa \) is consistent with ZFC.
Proof. Let $V$ be a model of $\text{ZFC} + \text{CH}$, let $\kappa \geq \omega_2$ be a regular cardinal, and let $G$ be $C^\kappa$-generic over $V$. Since $\kappa$ is regular and by Proposition 18.3 $CC^\kappa \cong C_\kappa$, by Theorem 14.21 we have $V[G] \cong \kappa. Therefore, it remains to show that $V[G]$ is a model in which $\emptyset = \tau = \text{cov}(\mathcal{M}) = c$.

By Lemma 18.9, for every real $x$ in $V[G]$ there is an $\alpha_x \in \kappa$ such that $x \in V[G|\alpha_x]$. Moreover, since $\kappa$ is regular, for every set of reals $X \in V[G]$ with $|X| < \kappa$ we get that $\bigcup \{\alpha_x : x \in X\} \in \kappa$.

Let $\mathcal{E}, \mathcal{E'} \subseteq [\omega]^{\omega} \cap V[G]$ and $\mathcal{F} \subseteq \omega^{\omega} \cap V[G]$ be three families in $V[G]$, each of cardinality strictly less than $\kappa$. Then there is an ordinal $\gamma \in \kappa$ such that all three families $\mathcal{E}, \mathcal{F}$, and $\mathcal{F}$, belong to $V[G|\gamma]$.

Since Cohen forcing adds splitting reals (by Lemma 21.3) and since $G(\gamma)$ is $\mathcal{C}$-generic over $V[G|\gamma]$ (by Lemma 21.9), in $V[G|\gamma+1]$ there is a real $s \in [\omega]^{\omega}$ which is a splitting real over $V[G|\gamma]$. Hence, the family $\mathcal{F}$, which belongs to $V[G|\gamma]$, is not a reaping family, and since $\mathcal{F}$ was arbitrary, we must have $V[G] \not\models \tau = c$ Similarly, $c$ Cohen forcing adds unboundable reals (by Lemma 21.1), in $V[G|\gamma+1]$ there is a function $f \in [\omega]^{\omega}$ which is unboundable over $V[G|\gamma]$. Hence, the family $\mathcal{F}$, which belongs to $V[G|\gamma]$, is not a dominating family, and since $\mathcal{F}$ was arbitrary, we must have $V[G] \not\models \mathfrak{d} = c$.

Assume now that $\mathcal{E}$ is a set of codes of meagre $\mathcal{F}_s$ sets, i.e., for every $r \in \mathcal{E}$, $A^V[G] \subseteq [\omega]^{\omega}$ is a meagre $\mathcal{F}_s$ set. Again, since $G(\gamma)$ is $\mathcal{C}$-generic over $V[G|\gamma]$, $\bigcup \{\mathcal{G}(\gamma) \in \bigcap_{r \in \mathcal{E}} [\omega \setminus A^V[G]|r]\}$. Hence, in $V[G]$ we get $\bigcup_{r \in \mathcal{E}} A^V[G] \neq \omega^{\omega}$, and since $\mathcal{E}$ was arbitrary, we must have $V[G] \not\models \text{cov}(\mathcal{M}) = c$. \quad \cdot$

As an immediate consequence of Proposition 18.5 and Proposition 21.10 (using the fact that $C_\kappa \cong C^\kappa$), we get the following consistency result.

**Corollary 21.11.** $\omega_1 = a < \emptyset = \tau = \text{cov}(\mathcal{M}) = c$ is consistent with ZFC.

In particular, since $p \leq a$, we get that $p < \text{cov}(\mathcal{M})$ is consistent with ZFC.

---

**A Model in which $s = b < \emptyset$**

The idea is to start with a model $V$ in which we have $\omega_1 < p = c$ (in particular, $V \models s = b = \emptyset = c$), and then add $\omega_1$ Cohen reals to $V$. It is not hard to verify that in the resulting model we have $\omega_1 = s = b$. Slightly more difficult to prove is the fact that we still have $\emptyset = c$, which is a consequence of the following result.

**Lemma 21.12.** Let $\mathbb{P} = (P, \leq)$ be a forcing notion and let $G$ be $\mathbb{P}$-generic over some model $V$ of ZFC. If $V \models |P| < b$, then for every function $f \in [\omega \cap V[G]$ we can construct a function $g_f \in [\omega \cap V]$ such that for all $h \in [\omega \cap V$ we have $h * f \rightarrow h * g_f$.

i.e., whenever the function $h$ is dominated by $f$ (in the model $V[G]$), it is also dominated by the function $g_f$ from the ground model $V$. In particular, if $V \models b > \omega_1$ and $G$ is $C^{\omega_1}$-generic over $V$, then $V[G] \models \emptyset \geq \emptyset^V$. 
Proof. Let \( f \in {}^\omega \omega \cap V[G] \) and let \( f \) be a \( P \)-name for \( f \) \( ( \text{in the ground model } V) \) such that \( 0 \Vdash_P f \in {}^\omega \omega \). For every \( P \)-condition \( p \in P \) define the function \( f_p \in {}^\omega \omega \cap V \) by stipulating
\[
f_p(n) = \min \{ k \in \omega : \exists q \geq p ( q \Vdash_P f(n) = k) \}.
\]
Consider the family \( \mathcal{F} = \{ f_p : p \in P \} \subseteq {}^\omega \omega \). Since \( |P| < b \), there exists a function \( g_f \in {}^\omega \omega \) \( ( \text{in the ground model } V) \) which dominates each member of \( \mathcal{F} \). Thus, whenever \( p \Vdash h <^* f \) we have \( h <^* f_p <^* g_f \), which shows that \( g_f \) dominates \( h \).

In order to see that \( V[G] \models \exists \delta \geq \delta^V \) whenever \( V \models b > \omega_1 \) and \( G \) is \( \mathbb{C}^{\omega_1} \)-generic over \( V \), recall that \( \mathbb{C}^{\omega_1} \approx \mathbb{C}_{\omega_1} \) and notice that \( |\text{Fn}(\omega \times \omega_1, \omega)| \leq \big| \text{fin}(\omega \times \omega_1 \times \omega) \big| = \omega_1 \).

The proof of the following result will be crucial in the proof of Proposition 27.9.

Proposition 21.13. \( \omega_1 = s = b < \delta = c \) is consistent with ZFC.

Proof. Let \( V \) be a model of ZFC + \( c = p > \omega_1 \) and let \( G = \{ c_\alpha : \alpha \in \omega_1 \} \) be \( \mathbb{C}^{\omega_1} \)-generic over \( V \), where we work with \( \mathbb{C} = \bigcup_{\alpha \in \omega_1} \mathbb{C}_\alpha \). We shall show that \( V[G] \models \omega_1 = s = b < \delta = c = \mathcal{V}^{\omega_1} \).

Since \( \mathbb{C}^{\omega_1} \) satisfies omc, all cardinals are preserved and we obviously have \( V[G] \models c = \omega \mathcal{V}^{\omega_1} > \omega_1 \). Furthermore, by Lemma 18.9, for all \( f \in {}^\omega \omega \) and \( x \in [\omega]^{\omega_0} \) which belong to \( V[G] \) there is a \( \gamma_0 \in \omega_1 \) such that \( f \) and \( x \) belong to \( V[(c_\alpha : \alpha \in \gamma_0)] \).

Since Cohen forcing adds unbounded reals \( (\text{by Lemma 21.1}) \) and since \( c_{\gamma_0} \) is \( \mathbb{C} \)-generic over \( V[G]_{\gamma_0} \) \( (\text{by Lemma 21.9}) \), \( c_{\gamma_0} \) \( \in {}^\omega \omega \) is not dominated by any function in \( V[(c_\alpha : \alpha \in \gamma_0)] \), in particular, \( c_{\gamma_0} \) is not bounded by \( f \). Thus, in \( V[G] \), the family \( \{ c_\alpha : \alpha \in \omega_1 \} \) is an unbounded family of cardinality \( \omega_1 \), which shows that \( V[G] \models \omega_1 = b \).

Similarly, let \( \sigma_{\gamma_0} \) be the splitting real over \( V[(c_\alpha : \alpha \in \gamma_0)] \) \( \text{we get from the Cohen real } c_{\gamma_0} \text{ using the construction in the proof of Lemma 21.3} \). Then \( \sigma_{\gamma_0} \) every infinite subset of \( \omega \), in particular, \( \sigma_{\gamma_0} \) splits \( x \). Thus, in \( V[G] \), the family \( \{ \sigma_\alpha : \alpha \in \omega_1 \} \) is a splitting family of cardinality \( \omega_1 \), which shows that \( V[G] \models \omega_1 = s \).

Finally, by Lemma 21.12 we have \( V[G] \vDash \delta = \delta^V > \omega_1 \) which shows that
\[
V[G] \models \omega_1 = s = b < \delta = c.
\]

Notes

The results presented in this chapter are all classical and most of them can be found in textbooks like Kunen [8] or Bartoszyński and Judah [3] (for example the model in which \( c > a = \omega_1 \) as well as the corresponding proofs are taken from Kunen [8, Chapter VIII, §2] and Lemma 21.12 is just Lemma 3.3.19 of Bartoszyński and Judah [3]).
Related Results

108. A combinatorial characterisation of $\text{cov}(\mathcal{M})$. Bartoszyński [2] (see also Bartoszyński and Judah [3, Theorem 2.4.1]) showed that $\text{cov}(\mathcal{M})$ is the cardinality of the smallest family $\mathcal{F} \subseteq \omega^\omega$ with the following property: For each $g \in \omega^\omega$ there is an $f \in \mathcal{F}$, such that for all but finitely many $n \in \omega$ we have $f(n) \neq g(n)$. For another characterisation of $\text{cov}(\mathcal{M})$ see Chapter 13. Related Result 80.

109. $p \leq \text{add}(\mathcal{M})$. The additivity of $\mathcal{M}$, denoted $\text{add}(\mathcal{M})$, is the smallest number of meagre sets such that the union is not meagre. Notice that we obviously have $\text{add}(\mathcal{M}) \leq \text{cov}(\mathcal{M})$. Piotrowski and Szymański showed in [12] that $p \leq \text{add}(\mathcal{M})$ which follows from the fact that $\text{add}(\mathcal{M}) = \min\{\text{cov}(\mathcal{M}), b\}$ (see Miller [10] and Truss [16], or Bartoszyński and Judah [3, Corollary 2.2.9]). For possible (i.e., consistent with ZFC) relations between $\text{add}(\mathcal{M})$ and $\text{cov}(\mathcal{M})$ and other cardinal characteristics of the continuum we refer the reader to Bartoszyński and Judah [3, Chapter 7].

110. Cohen-stable families of subsets of integers. Kurilic showed in [9] that adding a Cohen real destroys a splitting family $\mathcal{F}$ if and only if $\mathcal{F}$ is isomorphic to a splitting family on the set of rational numbers whose elements have nowhere dense boundaries. Consequently, $|\mathcal{F}| < \text{cov}(\mathcal{M})$ implies the Cohen-indestructibility of $\mathcal{F}$. Further, he showed that for a mad family in order to remain maximal in any Cohen extension, it is necessary and sufficient that every bijection from $\omega$ to the set of rational numbers must have a somewhere dense image on some member of the family.

A forcing notion, introduced by Solovay [13, 14], which is closely related to Cohen forcing $\mathbb{C}$ is the so-called random forcing, denoted $\mathbb{B}$, which is defined as follows: $\mathbb{B}$-conditions are closed sets $A \subseteq R$ of positive Lebesgue measure, and for two $\mathbb{B}$-conditions $A$ and $B$ let $A \leq B \iff A \subseteq B$. Further, if $G$ is $\mathbb{B}$-generic (over some model $V$), then $r = \bigcap G$ is called a random real.

111. Properties of random forcing. Obviously, random forcing satisfies $\text{ccc}$, and therefore, random forcing is proper. Furthermore, random forcing is $\omega$-bounding (see Jech [5, Part I, Lemma 3.3.(a)]), hence, random forcing does not add Cohen reals. For more properties of random forcing see Bartoszyński and Judah [3, Section 3.2] or Blass [4, Section 11.4].

112. Random reals versus Cohen reals. Let $c$ be a Cohen real over $V$ and let $r$ be a random real over $V[c]$. Then, in $V[c][r]$, there is a Cohen real (but no random real) over $V[r]$ (see Pawlikowski [11, Corollary 3.2]).

113. On partitions of $\omega$ into $\omega_1$ disjoint closed sets. If $\text{CH}$ holds, then the set of singletons $\{\{x\} : x \in \omega\}$ is obviously a partition of $\omega$ into $\omega_1$ disjoint closed sets. However, if $\text{CH}$ fails, then the existence of a partition of $\omega$ into $\omega_1$ disjoint closed sets is independent of ZFC:

Now, Stern [15, §1] showed that if $G$ is $\omega_2$-generic over $V$, where $V \models \text{GCH}$, then, in $V[G]$, there is no partition of $\omega$ into $\omega_1$ disjoint closed sets. On the other hand, Stern [15, §2] also showed that adding $\omega_2$ random reals to a model in which GCH holds, yields a model in which CH fails, but in which such a partition of $\omega$ still exists.
114. **On the existence of Ramsey ultrafilters.** It can be shown that $\text{cov}(\mathcal{M}) = \omega$ if and only if every filter generated by $\omega^2$ elements can be extended to a Ramsey ultrafilter (see Bartoszyński and Judah [3, Theorem 4.5.6]). In particular, adding $\omega_2$ Cohen reals to a model in which GCH holds, yields a model in which Ramsey ultrafilters exist. On the other hand, Kunen showed in [7] that adding $\omega_2$ random reals to a model in which GCH holds, yields a model in which there are no Ramsey ultrafilters (see also Jech [6, Theorem 91]).

115. **Random forcing and the ideal of Lebesgue measure zero sets.** Like the set of meagre sets $\mathcal{M}$, also the set $\mathcal{N}$ of Lebesgue measure zero sets forms an ideal. So, we can investigate $\text{add}(\mathcal{N})$ and $\text{cov}(\mathcal{N})$, and compare these cardinal characteristics with $\text{add}(\mathcal{M})$ and $\text{cov}(\mathcal{M})$.

For example, Bartoszyński showed in [1] that $\text{add}(\mathcal{N}) \leq \text{add}(\mathcal{M})$. Furthermore, by Theorem 20.6 (using the fact that random forcing is proper and "$\omega$-bounding") it follows that a countable support iteration of length $\omega_2$, starting in a model for CH, yields a model in which $\text{cov}(\mathcal{N}) > \text{cov}(\mathcal{M})$ (cf. Bartoszyński and Judah [3, Model 7.6.8]). For more results concerning random reals and the ideal $\mathcal{N}$ see Bartoszyński and Judah [3, Section 3.2].

**References**


Silver-Like Forcing Notions

On the one hand, we have seen that every forcing notion which adds dominating reals also adds splitting reals (see Fact 20.1). On the other hand, we have seen in the previous chapter that Cohen forcing is a forcing notion which adds splitting reals, but which does not add dominating reals. However, Cohen forcing adds unbounded reals and as an application we constructed a model in which \( s = b < \mathfrak{d} = \tau \). One might ask whether there exists a forcing notion which is even \( \omega \)-bounding but still adds splitting reals. In this chapter, we shall present such a forcing notion and as an application we shall construct a model in which \( s = b = \mathfrak{d} < \tau \).

Below, let \( \mathcal{E} \) be an arbitrary but fixed \( P \)-family (introduced in Chapter 10). For a set \( x \subseteq \omega \), let \( \mathcal{P} \) denote the set of all functions form \( x \) to \( \{0, 1\} \). **Silver-like forcing** with respect to \( \mathcal{E} \), denoted \( S_\mathcal{E} = (S_\mathcal{E}, \leq) \), is defined as follows:

\[
S_\mathcal{E} = \bigcup \{ \mathcal{P} : x^E \in \mathcal{E} \}
\]

where \( x^E := \omega \setminus x \), and for \( p, q \in S_\mathcal{E} \) we stipulate

\[
p \leq q \iff \text{dom}(p) \subseteq \text{dom}(q) \land \varphi_{|\text{dom}(p)} = p.
\]

If \( \mathcal{E} = [\omega]^{\omega} \), then we call \( S_\mathcal{E} \) just **Silver forcing**, and if \( \mathcal{E} \) is a \( P \)-point, then \( S_\mathcal{E} \) is usually called **Grigorieff forcing**.

As in the case of Cohen forcing we can identify every \( S_\mathcal{E} \)-generic filter with a real \( g \in \mathcal{P} \), called **Silver real**, which is in fact just the union of the functions which belong to the generic filter. More formally, if \( G \) is \( S_\mathcal{E} \)-generic over some model \( V \), then the corresponding Silver real \( g \in \mathcal{P} \) is defined by

\[
g = \bigcup \{ f \in S_\mathcal{E} : f \in G \}.
\]

On the other hand, from a Silver real one can always reconstruct the corresponding generic filter, and therefore, \( V[G] = V[g] \) (we leave the reconstruction as an exercise to the reader). Furthermore, Silver reals can be characterised as follows: A real \( g \in \mathcal{P} \) is a Silver real over a model \( V \) of \( \text{ZFC} \) iff for every open dense subset \( D \subseteq S_\mathcal{E} \) there is a \( p \in D \) such that \( g_{|\text{dom}(p)} = p \).
Properties of Silver-Like Forcing

Silver-Like Forcing is Proper and \( \omega^\omega \)-bounding

Before we show that Silver-like forcing \( S_\mathcal{E} \) is proper and \( \omega^\omega \)-bounding, let us introduce the following notation: For a condition \( p \in S_\mathcal{E} \) (i.e., \( p : x \rightarrow \{0, 1\} \)) where \( x' \in \mathcal{E} \) and a finite set \( t \subseteq \text{dom}(p) \) let

\[
\overline{p} - t = \{ q \in S_\mathcal{E} : \text{dom}(q) = \text{dom}(p) \land q_{|\text{dom}(q) \setminus t} = p_{|\text{dom}(p) \setminus t} \}.
\]

**Lemma 22.1.** Silver-like forcing \( S_\mathcal{E} \) is proper.

**Proof.** As described in Chapter 20, let \( \chi \) be a sufficiently large regular cardinal. We have to show that for all countable elementary submodels \( N = (N, \in) \prec (V, \in) \) which contain \( S_\mathcal{E} \), and for all conditions \( p \in S_\mathcal{E} \cap N \), there exists an \( S_\mathcal{E} \)-condition \( q \geq p \) in \( N \) which is \( N \)-generic (i.e., if \( g \in \omega^2 \) is a Silver real over \( V \) and \( q \leq g \), then \( g \) is also a Silver real over \( N \)).

So, let \( N = (N, \in) \) be an arbitrary countable elementary submodel of \((H_\chi, \in)\) and let \( p \in S_\mathcal{E} \cap N \) be an arbitrary \( S_\mathcal{E} \)-condition which belongs to \( N \). We shall construct in \( V \) an \( S_\mathcal{E} \)-condition \( q \geq p \) which is \( N \)-generic by using the fact that \( \mathcal{E} \) is a P-family. Firstly, let \( \{ D_n : n \in \omega \} \) be an enumeration (in \( V \)) of all open dense subsets of \( S_\mathcal{E} \) which belong to \( N \) and choose (in \( V \)) some well-ordering \( \prec \) on \( S_\mathcal{E} \cap N \). We construct the sought \( S_\mathcal{E} \)-condition \( q \geq p \) by running the game \( G^*_\mathcal{E} \): The MAIDEN starts the game by playing \( x_0 := \text{dom}(q_0) \mathcal{E} \), where \( q_0 \in N \) is the \( \prec \)-least condition such that \( q_0 \geq p \) and \( q_0 \in D_0 \), and DEATH responds with some finite set \( s_0 \subseteq x_0 \). Assume that for some \( n \in \omega \) we already have \( x_n, q_n, s_n \). Let \( t = \bigcup_{0 \leq i \leq n} s_i \) and \( y = x_n \setminus t \). Now, the MAIDEN plays \( x_{n+1} \subseteq y \) such that \( x_{n+1} = \text{dom}(q_{n+1}) \mathcal{E} \), where \( q_{n+1} \in N \) is the \( \prec \)-least condition such that \( q_{n+1} \geq q_n \) and \( q_{n+1} - t \subseteq D_{n+1} \), and DEATH responds with some finite set \( s_{n+1} \subseteq x_{n+1} \).

Since \( \mathcal{E} \) is a P-family, this strategy of the MAIDEN is not a winning strategy and DEATH can play so that \( x' = \bigcup_{n \in \omega} s_n \) belongs to \( \mathcal{E} \). For \( q' := \bigcup_{n \in \omega} q_n \) we have \( x' \subseteq \text{dom}(q') \), and thus, the function \( g := q'_{|x'} \) is an \( S_\mathcal{E} \)-condition.

In addition, if \( g \) is a Silver real over \( V \) such that \( q \leq g \), then, by construction of \( g \) and the properties of the \( q_n \)'s, for every \( n \in \omega \) we have \( g_{|\text{dom}(q_n)} \in D_n \), which shows that \( g \) is a Silver real over \( N \).

**Lemma 22.2.** Silver-like forcing \( S_\mathcal{E} \) is \( \omega^\omega \)-bounding.

**Proof.** Let \( G \) be \( S_\mathcal{E} \)-generic over \( V \), let \( f \in \omega^\omega \) be a function in \( V[G] \), and let \( f \) be an \( S_\mathcal{E} \)-name for \( f \). In order to show that \( f \) is bounded by some function in the ground model, it is enough to prove that for every \( S_\mathcal{E} \)-condition \( p \in S_\mathcal{E} \) there is a condition \( q_0 \geq p \) and a function \( g \in \omega^\omega \) in the ground model \( V \) such that \( q_0 \Vdash_{S_\mathcal{E}} \text{"} g \text{ dominates } f \text{"} \).

Firstly, choose some well-ordering \( \omega^\omega \) on \( S_\mathcal{E} \). We construct the condition \( q_0 \) by running the game \( G^*_\mathcal{E} \) where the MAIDEN plays according to the following...
strategy: Let \( m_0 \in \omega \) be the smallest integer for which there exists a condition \( r \geq p \) such that \( r \mathrel{\vDash_{\mathcal{S}_e}} f(0) < m_0 \) and let \( p_0 \) be the least such condition \( r \) with respect to the well-ordering \( \prec \). Then the \text{MAIDEN} plays \( x_0 = \text{dom}(p_0)^f \).

For positive integers \( i \in \omega \) let \( t_i = \bigcup_{k \in i} s_k \), where \( s_0, \ldots, s_{i-1} \) are the moves of \text{DEATH}, and let \( p_0 \leq \cdots \leq p_{i-1} \) be an increasing sequence of conditions. Further, let \( m_i \in \omega \) be the least number for which there exists a condition \( r \geq p_{i-1} \) with \( \text{dom}(r) \supseteq \text{dom}(p_{i-1}) \cup t_i \) such that for all \( q \in r - t_i \) we have \( r \mathrel{\vDash_{\mathcal{S}_e}} f(i) < m_i \), and again, let \( p_i \) be the least such condition \( r \) (with respect to \( \prec \)). Then the \text{MAIDEN} plays \( x_i = \text{dom}(p_i)^f \).

Since \( \mathcal{E} \) is a \( P \)-family, \text{DEATH} can play so that \( \bigcup_{i \in \omega} s_i \in \mathcal{E} \). Let \( h = \bigcup_{i \in \omega} p_i \); then \( h \in \mathcal{Z}_2 \) for some \( x \subseteq \omega \) (but \( h \) is not necessarily an \( \mathcal{S}_e \)-condition).

Now, let \( q_0 \in \mathcal{S}_e \) be such that \( \text{dom}(q_0) = \text{dom}(h) \setminus \bigcup_{i \in \omega} s_i \) and \( q_0 \equiv h \restriction_{\text{dom}(q_0)} \), and define the function \( g \in \omega^\omega \) by stipulating \( g(i) := m_i \) (for all \( i \in \omega \)). Then \( g \) belongs to the ground model \( \mathcal{V} \) and by construction we have

\[
q_0 \mathrel{\vDash_{\mathcal{S}_e}} \forall i \in \omega \left( f(i) < g(i) \right),
\]

which shows that \( q_0 \) forces that \( f \) is dominated by \( g \).

\[\square\]

\textbf{Silver-Like Forcing adds Splitting Reals}

\textbf{Lemma 22.3.} Silver-like forcing \( \mathcal{S}_e \) adds splitting reals.

\textbf{Proof.} Let \( g \in \mathcal{Z}_2 \) be a Silver real over \( \mathcal{V} \). We can identify \( g \) with the function \( f \in \omega^\omega \) by stipulating

\[
f(n) = k \iff g(k) = 1 \land |\{ m < k : g(m) = 1 \}| = n.
\]

Then the set

\[
\sigma_f = \bigcup \left\{ \left[ f(2n), f(2n+1) \right) : n \in \omega \right\}
\]

splits every real in the ground model, where \( [a, b) := \{ k \in \omega : a \leq k < b \} \). To see this, recall that \( \mathcal{E} \) is a free family, and notice that for each real \( x \in [\omega]^\omega \) in the ground model \( \mathcal{V} \) and for every \( n \in \omega \), the set

\[
D_{x,n} = \{ p \in \mathcal{S}_e : p \mathrel{\vDash_{\mathcal{S}_e}} \left( |x \cap \sigma_f| > n \land |x \setminus \sigma_f| > n \right) \}
\]

is open dense in \( \mathcal{S}_e \).

\[\square\]

\textbf{A Model in which } \mathcal{D} < \tau

\textbf{Proposition 22.4.} \( \omega_1 = \mathcal{D} < \tau = \alpha \) is consistent with \( \text{ZFC} \).

\textbf{Proof.} Let \( \mathcal{V} \) be a model of \( \text{ZFC} + \text{CH} \), let \( \mathbb{P}_{\omega_2} \) be an \( \omega_2 \)-stage, countable support iteration of Silver forcing (i.e., Silver-like forcing \( \mathcal{S}_e \) with \( \mathcal{E} = [\omega]^\omega \)),
and let $G$ be $\mathbb{P}_{\omega_2}$-generic over $V$. Since Silver forcing is of size $\varepsilon$, by Theorem 20.5.(a) we get $V[G] \models \varepsilon = \omega_2$. Furthermore, since Silver forcing is proper and $\omega$-bounding, by Theorem 20.6 we get that $\mathbb{P}_{\omega_2}$ is $\omega$-bounding, which implies that in $V[G]$, $\omega \cap V$ is a dominating family of size $\omega_1$ (recall that $V \models \text{CH}$), and therefore we have $V[G] \models \delta = \omega_1$. Finally, since Silver forcing adds splitting reals, by Theorem 20.5.(b) we get that no family $\mathcal{F} \subseteq [\omega]^{<\omega}$ of size $\omega_1$ can be a reaping family, thus, $V[G] \not\models \tau = \omega_2$. Hence, we get $V[G] \models \omega_1 = \delta < \tau = \omega_2 = \varepsilon$.

\section*{Notes}
Most of the results presented here can be found in Grigorieff [10] and Halbeisen [11] (see also Jech [12, p. 21f.] and Mathias [16]).

\section*{Related Results}

116. \textit{Silver-like forcing $S_\varepsilon$ is minimal.} Grigorieff proved in [10] that $S_\varepsilon$ is minimal whenever $\delta$ is a $P$-point and in Halbeisen [11] it is shown how Grigorieff's proof can be generalised to arbitrary $P$-families.

117. \textit{Silver-like forcing has the Laver property.} By similar arguments as in the proof of Lemma 22.2 one can show that Silver-like forcing has the Laver property.

118. \textit{$n$-Silver forcing.} For integers $n \geq 2$, the $n$-Silver forcing notion $S_n$ consists of functions $f : A \rightarrow n$, where $A \subset \omega$ and $\omega \setminus A$ is infinite. $S_n$ is ordered by inclusion, i.e., $f \leq g$ iff $g$ extends $f$. Notice that $S_2$ is the same as Silver forcing. If $G$ is $S_n$-generic, then the function $\bigcup G : \omega \rightarrow n$ is called an $S_n$-generic real. As a corollary of a more general result, it is shown in Rosłanowski and Steprāns [18] that no countable support iteration of $S_2$ adds an $S_n$-generic real.

119. \textit{Another model in which $\delta < \tau$.} A model in which $\omega_1 = a = \delta < \tau = \omega_2 = \varepsilon$ we get if we add $\omega_2$ random reals to a model $V$ of ZFC + CH (see for example Blass [2, Section 11.4]).

A forcing notion, introduced by Sacks [19], which is somewhat similar to Silver-like forcing, is the so-called \textit{Sacks forcing}, denoted $S$. To show the similarity to Silver-like forcing we shall define Sacks forcing in terms of perfect sets — but one can equally well define Sacks forcing in terms of trees. We say that a set $T \subseteq \omega^{<\omega}$ is \textbf{perfect} if for every $f \in T$ and every $n \in \omega$ there is a $g \in T$ and an integer $k \geq n$ such that $g|_n = f|_n$ and $f(k) \neq g(k)$. The set of $\mathcal{S}$-conditions consists of all perfect sets $T \subseteq \omega^{<\omega}$, and for any $\mathcal{S}$-conditions $S$ and $T$ we stipulate $S \subseteq T \iff T \subseteq S$. Furthermore, if $G$ is $\mathcal{S}$-generic then the real $\bigcap G \in \omega^{<\omega}$ is called a Sacks \textit{real}.

120. \textit{Properties of Sacks forcing.} One can show that Sacks forcing has the following properties:

- Sacks forcing is proper.
- Sacks forcing is $\omega$-bounding.
- Sacks forcing has the Laver property.
• Sacks forcing is minimal and every new real is a Sacks real.
• Sacks forcing collapses $\alpha$ to $\beta$.

For these and further properties of Sacks forcing, as well as for some applications of Sacks forcing to Ramsey Theory, see Sacks [39], Geschke and Quickert [8], Brendle [3, 4, 5], Simon [20], Blass [2, Section 11.3], Brendle and Lühr [7], and Brendle, Halbeisen, and Löwe [6].

121. Sacks forcing and splitting reals. Baumgartner and Laver [1] showed that Sacks forcing does not add splitting reals (see also Miller [17, Prop. 3.2]). Now, by applying the Weak Halpern-Läuchli Theorem 11.6, one can show that also finite products of Sacks forcing do not add splitting reals (see Miller [17, Remark, p. 149] and compare with Chapter 23 [Related Result 127]). Moreover, Laver [15] showed that even arbitrarily large countable support products of Sacks forcing do not add splitting reals.

122. Splitting families and Sacks forcing. Using the methods developed by Brendle and Yatake in [8], Kurilic investigated in [14] the stability of splitting families in several forcing extensions. For example, he proved that a splitting family is preserved by Sacks forcing if and only if it is preserved by some forcing notion which adds new reals (compare with Chapter 21 [Related Result 110]).

123. Sacks reals out of nowhere. Kellner and Shelah showed in [13] that there is a countable support iteration of length $\omega$ which does not add new reals at finite stages, but which adds a Sacks real at the limit stage $\omega$.

References

Miller Forcing

So far we have seen that Cohen forcing adds unbounded as well as splitting reals, but not dominating reals (see Chapter 21), and that Silver forcing adds splitting reals but not unbounded reals (see Chapter 22). Furthermore, it was mentioned that Sacks forcing adds neither splitting nor unbounded reals (see Chapter 22 | RELATED RESULT 120). In this chapter we shall introduce a forcing notion, called Miller forcing, which adds unbounded reals but no splitting reals. As an application of that forcing notion we shall construct a model in which \( t < \alpha \).

Before we introduce Miller forcing, let us first fix some terminology. We shall identify \( \text{seq}(\omega) \) (the set of finite sequences of \( \omega \)) with \( \bigcup_{n \in \omega} ^n \omega \). Consequently, for \( s \in \text{seq}(\omega) \) with \( |s| = n + 1 \) we can write \( s = \langle s(0), \ldots, s(n) \rangle \). Furthermore, for \( s, t \in \text{seq}(\omega) \) with \( |s| \leq |t| \) we write \( s \preceq t \) if \( t[[s]] = s \) (i.e., \( s \) is an initial segment of \( t \)). A set \( T \subseteq \text{seq}(\omega) \) is a tree, if it is closed under initial segments, i.e., \( t \in T \) and \( s \preceq t \) implies \( s \in T \). Elements of a tree are usually called nodes. Let \( T \subseteq \text{seq}(\omega) \) be a tree and let \( s \in T \) be a node of \( T \). Then the tree \( T_s \) is defined by

\[
T_s = \{ t \in T : t \preceq s \vee s \preceq t \}.
\]

Further, the set of immediate successors of \( s \) (with respect to \( T \)) is defined by

\[
\text{succ}_T(s) = \{ t \in T : \exists n \in \omega \ (t = \langle n \rangle) \},
\]

where \( s\langle n \rangle \) denotes the concatenation of the sequences \( s \) and \( \langle n \rangle \), and finally let

\[
\text{next}_T(s) = \{ n \in \omega : s\langle n \rangle \in T \}.
\]

A tree \( T \subseteq \text{seq}(\omega) \) is called superperfect, if for every \( t \in T \) there is an \( s \in T \) such that \( t \preceq s \) and \( |\text{succ}_T(s)| = \omega \), i.e., above every node \( t \) there is a node \( s \) with infinitely many immediate successors. If \( T \subseteq \text{seq}(\omega) \) is a superperfect tree, then let
split(T) = \{ s \in T : |\text{succ}_T(s)| = \omega \}.

Thus, a tree \( T \subseteq \text{seq}(\omega) \) is superperfect if and only if for each \( s \in T \) there exists a \( t \in \text{split}(T) \) — a so-called splitting node — such that \( s \preceq t \). For \( k \in \omega \) and \( T \subseteq \text{seq}(\omega) \), let

\[
\text{split}_k(T) = \{ s \in \text{split}(T) : |\{ t \in \text{split}(T) : t \preceq s \}| = k + 1 \},
\]

i.e., a splitting node \( s \in \text{split}(T) \) belongs to \( \text{split}_k(T) \) if and only if there are \( k \) splitting nodes below \( s \).

Now, Miller forcing, denoted by \( M = (M, \leq) \), also known as rational perfect set forcing, is defined as follows:

\[
M = \{ T \subseteq \text{seq}(\omega) : T \text{ is a superperfect tree} \},
\]

and for \( T, T' \in M \) we stipulate

\[
T \leq T' \iff T' \subseteq T.
\]

As in the case of Cohen and Silver forcing we can identify every \( M \)-generic filter with a real \( g \in \omega^\omega \), called Miller real, which is in fact the union of the intersection of the trees in the generic filter. More formally, if \( G \) is \( M \)-generic over some model \( V \), then the corresponding Miller real \( g \in \omega^\omega \) has the property that for each \( n \in \omega \) we have

\[
g|_n \in \bigcap \{ T \in \text{M} : T \in G \}.
\]

Since we can reconstruct the generic filter from the corresponding Miller real, we obviously have \( \mathcal{V}[G] = \mathcal{V}[g] \) (we leave the reconstruction as an exercise to the reader).

**Properties of Miller Forcing**

**Miller Forcing is Proper and adds Unbounded Reals**

**Lemma 23.1.** Miller forcing is proper.

**Proof.** As described in Chapter 20, let \( \chi \) be a sufficiently large regular cardinal. We have to show that for all countable elementary submodels \( N = (N, \in) \prec (H_\chi, \in) \) which contain \( M \), and for all conditions \( S \in M \cap N \), there exists an \( M \)-condition \( T \subseteq S \) (in \( V \)) which is \( N \)-generic.

So, let \( N = (N, \in) \) be an arbitrary countable elementary submodel of \( (H_\chi, \in) \) and let \( S \in M \cap N \) be an arbitrary \( M \)-condition which belongs to \( N \). We shall construct a superperfect tree \( T \subseteq S \) which meets every open dense subset of \( M \) which belongs to \( N \); in \( V \), let \( \{ D_n : n \in \omega \} \) be an enumeration of all open dense subsets of \( M \) which belong to \( N \). Firstly, choose a superperfect
tree $T^0 \subseteq S$ such that $T^0 \in (D_0 \cap N)$. Assume we have already constructed $T^0 \supseteq T^1 \supseteq \cdots \supseteq T^n$ such that for each $i \leq n$, $T^i \in (D_i \cap N)$. Let $\{s_j : j \in \omega\}$ be an enumeration of $\text{split}_{n+1}(T^n)$. For every $j \in \omega$ and for each $t \in \text{succ}_T(s_j)$ choose a superperfect tree $T^{j,t} \subseteq T^n$ such that $T^{j,t} \in (D_{n+1} \cap N)$ and let

$$T^{n+1} = \bigcup \{ T^{j,t} : j \in \omega \land t \in \text{succ}_T(s_j) \}.$$ 

Then $T^{n+1}$ is a superperfect tree and $T^{n+1} \subseteq T^n$. In addition, if $G$ is $M$-generic over $V$ and $T^{n+1} \in G$, then there exists a $T^{j,t} \subseteq T^{n+1}$ which belongs to $G$, and because $T^{j,t} \in D_{n+1}$, we get $G \cap D_{n+1} \neq \emptyset$. Now, let $T = \bigcap_{n \in \omega} T^n$. Then $T \subseteq S$ is a superperfect tree which is $N$-generic.

**Lemma 23.2.** Miller forcing adds unbounded reals.

**Proof.** In order to prove that Miller forcing adds unbounded reals, it is enough to show that whenever $g \in \omega^\omega$ is a Miller real over some model $V$, then $g$ is unbounded. Let $f \in \omega^\omega$ be an arbitrary function in $V$ and let

$$D_f = \{ T \in M : \forall s \in \text{split}(T) \forall n \in \text{next}_T(s) \left( f(|s|) < n \right) \}.$$ 

We leave it as an exercise to the reader to show that $D_f$ is open dense in $M$, which shows that $g \not\leq^* f$. Thus, $g$ is not dominated by $f$, and since $f$ was arbitrary, $g$ is unbounded. \qed

**Miller Forcing does not add Splitting Reals**

**Lemma 23.3.** Miller forcing does not add splitting reals.

**Proof.** Let $V$ be a model of ZFC, let $G$ be $M$-generic over $V$, and let $Y$ be an $M$-name for a subset of $\omega$ in $V[G]$, i.e., there is an $M$-condition $S \in M$ such that $S \Vdash_M Y \subseteq \omega$. We shall construct an $M$-condition $S' \subseteq S$ and an $X \in [\omega]^{\omega}$ (in $V$) such that $S' \Vdash_M (X \subseteq Y) \lor (X \cap Y = \emptyset)$, which shows that $Y$ is not a splitting real.

The construction of the superperfect tree $S'$ and the infinite set $X \in [\omega]^{\omega}$ is done in the following three steps.

**Claim 1.** There is an $M$-condition $T \subseteq S$ and a sequence $\langle Y_s : s \in \text{split}(T) \rangle$ (in $V$) of subsets of $\omega$, such that for every $s \in \text{split}(T)$, each $k \in \omega$, and for all but finitely many $n \in \text{next}_T(s)$ we have

$$T_{s,n} \Vdash_M Y \cap k = Y_s \cap k,$$

i.e., for every $k \in \omega$ there exists an $n_k \in \omega$ such that for all $n' \in \text{next}_T(s)$ with $n' \geq n_k$, $T_{s,n'} \Vdash_M Y \cap k = Y_s \cap k$.

**Proof of Claim 1.** We construct the condition $T$ by induction. In particular, the superperfect tree $T$ will be the intersection of superperfect trees $T^i$, where
\[ T^0 = S, \quad T^{i+1} = \bigcup_{s \in \text{split}_i(T^i)} \hat{T}'_s, \quad \text{and} \quad \hat{T}'_s \subseteq T'_s, \]

and where the superperfect trees \( \hat{T}'_s \) are constructed as follows: Fix an \( i \in \omega \) and a splitting node \( s \in \text{split}_i(T^i) \). For each \( n \in \text{next}_{T^i}(s) \), choose a superperfect tree \( \hat{T}'_s \subseteq T'_s \) such that, for some finite set \( b_n \in \text{fin}(\omega) \), we have

\[ \hat{T}'_s \models \forall Y \text{ such that } n = b_n. \]

For every \( k \in \omega \), let \( F_k = \{ b_n \cap k : n \in \text{next}_{T^i}(s) \} \). Notice that all sets \( F_k \) are finite, in fact, \( F_k \subseteq \mathcal{P}(k) \). Consider now the tree \( T \) with the infinite vertex set \( \{ (b, k) : k \in \omega \land b \in F_k \} \), where two vertices \( (b, k) \) and \( (b', k') \) are joined by an edge iff \( b \subseteq (b' \cap k) \) and \( k' = k + 1 \). Notice that \( T \) is an infinite, finitely branching tree. Hence, by König’s Lemma, \( T \) contains an infinite branch, say \( (0, 0), (a_1, 1), \ldots, (a_k, k), \ldots \). Let \( Y_s = \bigcup_{k \in \omega} a_k \) and define the strictly increasing sequence \( \langle n_j : j \in \omega \rangle \) of elements of \( \text{next}_{T^i}(s) \) so that for each \( k \in \omega \) and for all \( n_j \geq k \) we have

\[ \hat{T}'_s \models Y \cap k = a_k. \]

Hence, for each \( k \in \omega \) and for all but finitely many \( j \in \omega \) we have

\[ \hat{T}'_s \models Y \cap k = Y_s \cap k. \]

Now, let \( \hat{T}'_s = \bigcup_{j \in \omega} \hat{T}'_s \). Then, for each \( k \in \omega \) and for all but finitely many \( n \in \text{next}_{\hat{T}'_s}(s) \) we have

\[ \hat{T}'_s \models Y \cap n = Y_s \cap n. \]

Finally, let \( T^{i+1} = \bigcup \{ \hat{T}'_s : s \in \text{split}_i(T^i) \} \). Notice that for all \( j \leq i, \text{split}_i(T^{i+1}) = \text{split}_j(T^i) \); thus, \( T = \bigcap_{i \in \omega} T^i \) is a superperfect tree. By construction, for every \( s \in \text{split}(T) \), for each \( k \in \omega \), and for all but finitely many \( n \in \text{next}_{T}(s) \) we have

\[ T \models Y \cap k = Y_s \cap k, \]

where \( \langle Y_s : s \in \text{split}(T) \rangle \) is an infinite sequence of subsets of \( \omega \) which belongs to the ground model \( V \).

In the next step we prune the tree \( T \) so that the corresponding sets \( Y_s \) (or their complements) have the strong finite intersection property \text{split} \ (i.e., intersections of finitely many sets are infinite).

**Claim:** There exists a superperfect tree \( T' \subseteq T \) such that

1. \( \{ Y_s : s \in \text{split}(T') \} \) has the \text{split}; or
2. \( \{ \omega \setminus Y_s : s \in \text{split}(T') \} \) has the \text{split}.


Proof of Claim 2. Let \( \mathcal{U} \subseteq [\omega]^{\omega} \) be an arbitrary ultrafilter over \( \omega \). We partition the set \( \text{split}(T) \) according to whether the set \( Y_s \) belongs to \( \mathcal{U} \) or not. More precisely, let \( U = \{ s \in \text{split}(T) : Y_s \in \mathcal{U} \} \) and \( V = \{ s \in \text{split}(T) : (\omega \setminus Y_s) \notin \mathcal{U} \} \). Then \( U \cap V = \emptyset \) and \( U \cup V = \text{split}(T) \). We are in at least one of the following two cases:

- There exists an \( s \in \text{split}(T) \) such that \( \text{split}(T_s) \subseteq U \).
- For all \( s \in \text{split}(T) \) there exists a \( t \in \text{split}(T_s) \) with \( t \in V \).

In the former case, let \( T' = T_s \), and in the latter case, we can construct a superperfect tree \( T' \subseteq T \) such that \( \text{split}(T') \subseteq V \) — we leave the construction of \( T' \) as an exercise to the reader.

If \( \text{split}(T') \subseteq U \), then \( \{ Y_s : s \in \text{split}(T') \} \) has the slip, and if \( \text{split}(T') \subseteq V \), then \( \{ \omega \setminus Y_s : s \in \text{split}(T') \} \) has the slip.  

In the last step we construct a set \( X \in [\omega]^{\omega} \) which is not split by \( \mathcal{U} \).

Claim 3. Let \( T' \subseteq T \) be a superperfect tree such that

\[ \mathcal{B}_0 = \{ Y_s : s \in \text{split}(T') \} \quad \text{or} \quad \mathcal{B}_1 = \{ \omega \setminus Y_s : s \in \text{split}(T') \} \]

has the slp. Then there exists a sequence of superperfect trees \( (T^i : i \in \omega) \), where \( T^0 \subseteq T' \) and \( T^{i+1} \subseteq T^i \) (for all \( i \in \omega \)), as well as a sequence of natural numbers \( (m_i : i \in \omega) \), where \( m_i < m_{i+1} \) (for all \( i \in \omega \)), such that \( \bigcap_{i \in \omega} T^i \) is a superperfect tree and either

\[ \forall i \in \omega \left( T^i \models \forall m_i \in Y \right) \quad \text{or} \quad \forall i \in \omega \left( T^i \models \forall m_i \notin Y \right). \]

Proof of Claim 3. We just consider the case when \( \mathcal{B}_1 \) has the slp, in which case we shall later get \( X \cap Y = \emptyset \); the other case, in which would later get \( X \subseteq Y \), is handled analogously and is left as an exercise to the reader.

In order to get \( \bigcap_{i \in \omega} T^i \models M \), we shall construct an auxiliary sequence \( \langle F_i : i \in \omega \rangle \) of increasing finite subsets of \( \text{split}(T') \), i.e., for every \( i \in \omega \), \( F_i \subseteq F_{i+1} \) and \( F_i \in \text{fin}(\text{split}(T')) \). Moreover, we shall construct \( \langle F_i : i \in \omega \rangle \) such that \( \bigcup_{i \in \omega} F_i \) is infinite and \( \bigcap_{i \in \omega} F_i = \text{split}(\bigcap_{i \in \omega} T^i) \).

Let \( T^{-1} := T', m_{-1} := 0 \), and let \( F_{-1} = \{ s \} \) for some \( s \in \text{split}(T') \). Assume that for some \( i \in \omega \), we have already constructed a superperfect tree \( T^{-1} \models M, m_{i-1} \in \omega \), and \( F_{i-1} \in \text{fin}(\text{split}(T^{-1})) \). Choose a natural number \( m_i > m_{i-1} \) such that \( m_i \in \bigcap_{i \in \text{fin}(\text{split}(T^{-1}))} (\omega \setminus Y_s) \). This can be done since \( \mathcal{B}_1 \) has the slp, i.e., \( \bigcap_{i \in \text{fin}(\text{split}(T^{-1}))} (\omega \setminus Y_s) \) is infinite. Now, with respect to the finite set \( F_{i-1} \) define

\[ [F_{i-1}] = \{ t \in \text{seq}(\omega) : \exists s \in F_{i-1} \ (t \leq s) \} \, . \]

Then \( [F_{i-1}] \) is a finite subtree of \( T^{-1} \). Suppose that \( s_0 \in [F_{i-1}] \) is a terminal node of \( [F_{i-1}] \), i.e., for all \( n \in \omega \), \( s_0 \upharpoonright n \notin [F_{i-1}] \). By construction of \( Y_{s_0} \), for all but finitely many \( n \in \text{next}_{T^{-1}}(s_0) \) we have

\[ T^{-1}_{s_0} \models \forall m_i \in Y \cap (m_i + 1) = Y_{s_0} \cap (m_i + 1) . \]
Hence, since \( m_i \notin Y_{s_0} \), for all but finitely many \( n \in \text{next}_{T_{i-1}}(s_0) \) we have

\[
T_{s_0/n}^{i-1} \not

\Rightarrow m_i \notin Y.
\]

Now, we prune \( T_{s_0/n}^{i-1} \) by deleting the finitely many subtrees \( T_{s_0/n}^{i-1} \) with

\[
T_{s_0/n}^{i-1} \not

\Rightarrow m_i \notin Y.
\]

Furthermore, we do exactly the same for all other terminal nodes of the finite tree \( [F_{i-1}] \). Then, we do the same for all interior nodes of \( [F_{i-1}] \), except that we retain all subtrees \( T_{s_0/n}^{i-1} \) with \( s \notin [F_{i-1}] \).

The resulting tree \( T^i \) is superperfect and has the property that

\[
T^i \not

\Rightarrow m_i \notin Y.
\]

Notice that by construction, if \( s \in [F_{i-1}] \) is an interior node of \( [F_{i-1}] \) and \( s \notin [F_{i-1}] \) (for some \( n \in \omega \)), then \( s \notin T^i \). Now, choose a finite set \( F_i \) such that \( F_i \subseteq F_i \in \text{fin} \left( \text{split}(T^i) \right) \) which has the following property: For each \( s \in F_{i-1} \), for which there is an \( n_s \in \omega \) such that \( s \notin T^i \), there exists a \( t \in F_i \), \( F_{i-1} \) such that \( s \notin t \). We leave it as an exercise to the reader to verify that the resulting tree \( \bigcap_{i \in \omega} T^i \) is superperfect.

Now, let \( X := \{ m_i : i \in \omega \} \) and \( S' := \bigcap_{i \in \omega} T^i \). Then, in the case when \( \mathcal{B}_i \) has the sfp, we have

\[
S' \not

\Rightarrow X \cap Y = \emptyset,
\]

and otherwise we have

\[
S' \not

\Rightarrow X \subseteq Y.
\]

In other words, whenever \( G \) is \( \mathbb{M} \)-generic over \( V \), then \( Y[G] \) is not a splitting real over \( V \), and since \( Y \) was an \( M \)-name for an arbitrary subset of \( \omega \), this shows that Miller forcing does not add splitting reals.

As an immediate consequence we get

**Fact 23.4.** Miller forcing does not add dominating reals.

**Proof.** By Fact 20.1 we know that every forcing notion which adds dominating reals also adds splitting reals. Thus, since Miller forcing does not add splitting reals, it also does not add dominating reals.

**Miller Forcing Preserves P-Points**

By a similar construction as in the proof of Lemma 23.3 we can show that every \( P \)-point in the ground model generates an ultrafilter in the \( M \)-generic extension.

**Lemma 23.5.** Miller forcing preserves \( P \)-points.
Proof. Suppose that \( \mathcal{U} \subseteq [\omega]^\omega \) is a \( P \)-point in the ground model \( V \) and that \( G \) is \( \mathcal{M} \)-generic over \( V \). We have to show that \( \mathcal{U} \) generates an ultrafilter in \( V[G] \), i.e., for every \( Y \subseteq \omega \) in \( V[G] \) there exists an \( X \in \mathcal{U} \) in \( V \) such that either \( X \subseteq Y \) or \( X \cap Y = \emptyset \). For this, let \( Y \) be an \( \mathcal{M} \)-name for an arbitrary but fixed subset of \( \omega \) in \( V[G] \) (i.e., there is an \( \mathcal{M} \)-condition \( S \in \mathcal{M} \) such that \( S \mathrel{\ulcorner} Y \subseteq \omega \)). We shall construct an \( \mathcal{M} \)-condition \( S' \subseteq S \) and an \( X \in \mathcal{U} \) in \( V \) such that either \( S' \mathrel{\ulcorner} X \subseteq Y \) or \( S' \mathrel{\ulcorner} X \cap Y = \emptyset \). Since \( Y[G] \) is arbitrary, this would imply that the filter in \( V[G] \), generated by \( \mathcal{U} \), is an ultrafilter.

As in the proof of Lemma 23.3, we first construct an \( \mathcal{M} \)-condition \( T \subseteq S \) and a sequence \( \{ Y_s : s \in \text{split}(T) \} \) of subsets of \( \omega \), such that for every \( s \in \text{split}(T) \), for each \( k \in \omega \), and for all but finitely many \( n \in \text{next}_T(s) \), we have

\[
T_s \mathrel{\ulcorner} Y \cap k = Y_s \cap k.
\]

Now, we construct a superperfect tree \( T' \subseteq T \) such that either \( \{ Y_s : s \in \text{split}(T') \} \subseteq \mathcal{U} \) or \( \{ \omega \setminus Y_s : s \in \text{split}(T') \} \subseteq \mathcal{U} \). Since \( \mathcal{U} \) is a \( P \)-point, there exists an \( X' \in \mathcal{U} \) such that for all \( s \in \text{split}(T') \), either \( X' \subseteq^* Y_s \) or \( X' \subseteq^* (\omega \setminus Y_s) \).

Below we just consider the case when \( X' \subseteq^* Y_s \) and leave the other case as an exercise to the reader.

In the next step we build a sequence \( s_n \in \text{split}(T) \), such that both sets, \( \{s_n : n \in \omega \} \) and \( \{s_{n+1} : n \in \omega \} \), will be the splitting nodes of some \( \mathcal{M} \)-condition. At the same time we build a strictly increasing sequence of natural numbers \( \langle k_n : n \in \omega \rangle \), such that for all \( n \in \omega \), \( X' \setminus k_n \subseteq Y_{s_n} \).

The construction is by induction on \( n \): Firstly, let \( s_0 \in \text{split}_0(T) \), let \( s_1 = s_0 \), and let \( k_0 = 0 \). If necessary, modify \( X' \) such that \( X' \subseteq Y_{s_0} = Y_{s_1} \).

Assume that for some \( n \in \omega \), we have already constructed \( s_{2n}, s_{2n+1}, k_{2n}, \) and \( k_{2n+1} \). Let \( i, j \in \omega \) be such that

\[
n + 1 = \frac{(i+j)(i+j+1)}{2} + i.
\]

Notice that \( i \) and \( j \) are unique and that \( n+1 > i \). Now, we choose a new splitting node \( s_{2n+2} \in \text{split}(T) \), i.e., \( s_{2n+2} \notin \{ s_l : l \leq 2n+1 \} \), such that \( s_{2i} m_0 \preceq s_{2n+2} \) for some \( m_0 \in \text{next}_T(s_2i) \) with \( m_0 > k_{2n+1} \), and

\[
Y_{s_{2n+2}} \cap k_{2n+1} = Y_{s_{2i}} \cap k_{2n+1}.
\]

In order to see that such a splitting node \( s_{2n+2} \) exists, notice that \( 2n+2 > 2i \) and that for all but finitely many \( m \in \text{next}_T(s_{2i}) \),

\[
T'_{s_{2i}} \mathrel{\ulcorner} Y \cap k_{2n+1} = Y_{s_{2i}} \cap k_{2n+1}.
\]

Hence, there exists an \( m_0 \in \text{next}_T(s_{2i}) \) with \( m_0 > k_{2n+1} \), such that for all \( s_{2n+2} \gg s_0 m_0 \) we have \( Y_{s_{2n+2}} \cap k_{2n+1} = Y_{s_0} \cap k_{2n+1} \). Finally, we choose \( k_{2n+2} > k_{2n+1} \) large enough such that

\[
X' \setminus k_{2n+2} \subseteq Y_{s_{2n+2}}.
\]
The splitting node $s_{2n+3} \in \text{split}(T')$ (with $s_{2n+1} \sim m_0 \iff s_{2n+3}$) and the integer $k_{2n+3} > k_{2n+2}$ are chosen similarly.

Notice that by construction, for each node $s \in \{s_{2n} : n \in \omega\}$ there are infinitely many nodes $t \in \{s_{2n} : n \in \omega\}$ such that $s \sim t$, and the same holds for the set $\{s_{2n+1} : n \in \omega\}$. Thus, $\{s_{2n} : n \in \omega\}$ and also $\{s_{2n+1} : n \in \omega\}$ are the splitting nodes of superperfect subtrees of $T'$. Let $S_0, S_1 \subseteq T'$ be such that $\text{split}(S_0) = \{s_{2n} : n \in \omega\}$ and $\text{split}(S_1) = \{s_{2n+1} : n \in \omega\}$ respectively. Further, let

$$X_0 = X' \cap \bigcup \{[k_{2n}, k_{2n+1}) : n \in \omega\}$$

and

$$X_1 = X' \cap \bigcup \{[k_{2n+1}, k_{2n+2}) : n \in \omega\}$$

where $[k, k') = \{m \in \omega : k \leq m < k'\}$. Without loss of generality we may assume that $X_0 \subseteq S_0$. The goal is to show that $S_0 \not\equiv M X_0 \not\subseteq Y$, which is done in the following two claims:

**Claim 1.** For every $s \in \text{split}(S_0)$, $X_0 \subseteq Y_s$.

**Proof of Claim 1.** Firstly, notice that for every $s \in \text{split}(S_0)$ there is an $n \in \omega$ such that $s = s_{2n}$. We prove that $X_0 \subseteq Y_{s_{2n}}$ by induction on $n$. By the choice of $X'$ we have $X' \subseteq Y_{s_0}$; hence, $X_0 \subseteq Y_{s_0}$. If $n > 0$, then by the choice of $k_{2n}$ we have

$$X_0 \setminus k_{2n} \subseteq Y_{s_{2n}}$$

and by the definition of $X_0$ we have

$$X_0 \cap k_{2n} = X_0 \cap k_{2n-1}.$$

Therefore, we find an $i < n$ such that

$$Y_{s_{2n}} \cap k_{2n-1} = Y_{s_{2i}} \cap k_{2n-1}.$$

Now, by induction we have $X_0 \subseteq Y_{s_{2i}}$, thus, $(X_0 \cap k_{2n-1}) \subseteq Y_{s_{2i}} \cap k_{2n-1}$. Since $(X_0 \cap k_{2n}) = (X_0 \cap k_{2n-1})$ and $(X_0 \setminus k_{2n}) \subseteq Y_{s_{2n}}$, we finally get

$$X_0 = (X_0 \cap k_{2n}) \cup (X_0 \setminus k_{2n}) \subseteq (Y_{s_{2n}} \cap k_{2n-1}) \cup Y_{s_{2n}} = Y_{s_{2n}}.$$

**Claim 1:**

**Claim 2.** $S_0 \not\equiv M X_0 \not\subseteq Y$.

**Proof of Claim 2.** Assume towards a contradiction that there is an $M$-condition $\tilde{S} \subseteq S_0$ and an $m \in X_0$ such that

$$\tilde{S} \not\equiv M m \not\in Y'.$$

Let $s \in \text{split}_0(\tilde{S})$. By construction of $T$, and since $\tilde{S} \subseteq T$, for each $k \in \omega$ and for all but finitely many $n \in \text{next}_{\tilde{S}}(s)$ we have $\tilde{S}_{\sim n} \not\equiv M Y' \cap k = Y_s \cap k$. In particular, for $k = m + 1$ and for some $n_0 \in \text{next}_{\tilde{S}}(s)$ we have
A model in which \( \tau < \emptyset \)

\[
\tilde{S}_{x_0} M m \in Y \leftrightarrow m \in Y_
\]

Since \( X_0 \subseteq Y \) and \( m \in X_0 \), this implies

\[
\tilde{S}_{x_0} M m \in Y ,
\]

which contradicts our assumption that \( \tilde{S} M m \notin Y \).

Thus, in the case when for all \( s \in \text{split}(T') \), \( X' \subseteq Y \), there is an \( X \in \mathcal{Y} \) (where \( X \) is either \( X_0 \) or \( X_1 \)) and an \( M \)-condition \( S' \subseteq T' \) (where \( S' \) is either \( S_0 \) or \( S_1 \)) such that \( S' \vdash M X \subseteq Y \). In the other case (which was left to the reader), in which for all \( s \in \text{split}(T') \), \( X' \subseteq (\omega \setminus Y_s) \), there is an \( X \in \mathcal{Y} \) and an \( S' \subseteq T' \) such that \( S' \vdash M X \cap Y = \emptyset \). So, in both cases, \( \mathcal{Y} \) generates an ultrafilter in the \( M \)-generic extension, which is what we had to show.

A Model in which \( \tau < \emptyset \)

Below we show that after adding \( \omega_2 \) Miller reals to a model \( V \) of ZFC + CH, we get a model \( V[G] \) in which \( \tau = \omega_1 \) and \( \emptyset = \omega_2 \). The reason why \( V[G] \vDash \emptyset = \omega_2 \) is that Miller forcing adds unbounded reals, and the reason why \( V[G] \vDash \tau = \omega_1 \) is in fact a consequence of the following

**Fact 23.6.** If there exists an ultrafilter \( \mathcal{U} \) which is generated by some filter \( \mathcal{F} \subseteq [\omega]^\omega \) of cardinality \( \kappa \), then \( \tau < \kappa \).

**Proof.** Firstly notice that for all \( x \in [\omega]^{\omega} \), either \( x \in \mathcal{U} \) or \( \omega \setminus x \in \mathcal{U} \). Secondly, since \( \mathcal{F} \) generates \( \mathcal{U} \), for all \( x' \in \mathcal{U} \) there is a \( y \in \mathcal{F} \) such that \( y \subseteq x' \). This shows that \( \mathcal{F} \) is a reaping family. \( \dashv \)

**Proposition 23.7.** \( \omega_1 = \tau < \emptyset = \tau \) is consistent with ZFC.

**Proof.** Let \( \mathbb{P}_{\omega_2} \) be a countable support iteration of Miller forcing, let \( V \) be a model of \( \text{ZFC} + \text{CH} \), and let \( G \) be \( \mathbb{P}_{\omega_2} \)-generic over \( V \).

Since Miller forcing is of size \( \tau \), by Theorem 20.5.(a) we get \( V[G] \vDash \tau = \omega_2 \), and since Miller forcing adds unbounded reals, by Theorem 20.5.(b) we get that no family \( \mathcal{F} \subseteq [\omega]^{\omega} \) of size \( \omega_1 \) can be a dominating family. Hence, we get \( V[G] \vDash \emptyset = \omega_2 \).

Now we show that \( V[G] \vDash \tau = \omega_1 \). Firstly, notice that CH implies that every ultrafilter is of cardinality \( \omega_1 \), and recall that CH implies the existence of \( P \)-points. Thus, since \( V \models \text{CH} \), there are \( P \)-points in \( V \) of cardinality \( \omega_1 \). Since Miller forcing is proper and the iteration is a countable support iteration, by Theorem 20.8 we get that every \( P \)-point \( \mathcal{F} \) (of cardinality \( \omega_1 \)) in the ground model \( V \) generates an ultrafilter \( \mathcal{U} \subseteq [\omega]^{\omega} \) in \( V[G] \). Thus, by Fact 23.6, we have \( V[G] \vDash \tau = \omega_1 \). \( \dashv \)
Notes

All non-trivial results presented in this chapter are essentially due to Miller and can be found in [14]. In that paper, he introduced what is now called Miller forcing, but which he called rational perfect set forcing. Miller thought about this forcing notion when he worked on his paper [23], where he used a fusion argument which involved preserving a dynamically chosen countable set of points (see [13, Lemmata 8 & 9]). This led him to perfect sets in which the rationals in them are dense, and shortly after, he realised that this is equivalent to forcing with superperfect trees. Even though superperfect trees appeared first in papers of Kechris [10] and Louveau [12], Miller was the first who investigated the corresponding forcing notion.

Related Results

124. Characterising Miller reals. By the proof of Lemma 23.2 we know that every Miller real \( g \) is unbounded. On the other hand, one can show that every function \( f \in \omega^\omega \) in the \( \mathbb{M} \)-generic extension \( V[g] \) which is unbounded (i.e., not dominated by any function in \( V \)) is a Miller real (see Miller [14, Proposition 2]). Furthermore, one can show that Miller forcing is minimal (see Miller [14, p. 147]).

125. Miller forcing has the Laver property. One can show that Miller forcing has the Laver property (see Bartoszyński and Judah [1, Theorem 7.3.45]) and therefore does not add Cohen reals. Since the Laver property is preserved under countable support iterations, there are no Cohen reals in the model constructed in the proof of Proposition 23.7.

126. Miller forcing does not add Cohen, dominating, or random reals. Since every forcing notion which preserves \( P \)-points does not add Cohen, dominating, or random reals (see Chapter 20 | Related Result 107), Miller forcing adds neither Cohen, nor dominating, nor random reals.

127. \( \mathbb{M} \times \mathbb{M} \) adds splitting reals. Even though Miller forcing does not add splitting reals, a product of Miller forcing \( \mathbb{M} \times \mathbb{M} \) always adds splitting reals (see Miller [14, Remark p. 151]) and compare with Chapter 22 | Related Result 121).

128. Miller forcing satisfies Axiom A. Miller forcing is not just proper, it even satisfies the slightly stronger Axiom A (see Bartoszyński and Judah [1, p. 306]).

129. Miller forcing preserves \( \text{MA}(\sigma\text{-centred}) \). If \( V \models \text{MA}(\sigma\text{-centred}) \) and \( g \) is a Miller real over \( V \), then \( V[g] \models \text{MA}(\sigma\text{-centred}) \) (see Brendle [5]). Recall that by Chapter 13 | Related Result 79, \( \text{MA}(\sigma\text{-centred}) \iff \mathfrak{p} = \mathfrak{c} \), and compare this result with Theorem 19.4, which says that Cohen forcing preserves \( \mathfrak{p} = \mathfrak{c} \).

130. Cardinal characteristics in Miller’s model. In Miller’s model, which is the model constructed in the proof of Proposition 23.7, we also have \( \omega_1 = \mathfrak{a} = \mathfrak{c} \) (see for example Blass [2, Section 11.9]). Furthermore, the proof of Proposition 23.7 shows that in Miller’s model we even have \( u < \delta \) (see also Blass and Shelah [3]).

Another forcing notion with superperfect trees as conditions, which was introduced by Laver in [11], is the so-called Laver forcing, denoted \( L \in L \)-conditions are ordered
pairs \((s, T)\), where \(T \subseteq \text{seq}(\omega)\) is a superperfect tree, \(s \in T\), and for all \(t \in T\) we have either \(t \not\leq s\) or \(s \leq t \land t \in \text{split}(T)\) (i.e., \(T_s = T\) and every node \(t \geq s\) is a splitting node of \(T\)). For \(L\)-conditions \((s, T)\) and \((s', T')\) let \((s, T) \leq (s', T')\) \(\iff\) \(s \leq s' \land T' \subseteq T\). Furthermore, for ultrafilters \(\mathcal{U} \subseteq [\omega]^\omega\) we define \textbf{restricted Laver forcing}, denoted \(L_{\mathcal{U}}\), as follows: A pair \((s, T)\) is an \(L_{\mathcal{U}}\)-condition if it is an \(L\)-condition which has the property that for all \(t \in \text{split}(T)\) we have \(\text{next}_T(t) \in \mathcal{U}\).

131. \textbf{Laver forcing and Borel’s conjecture.} A set \(X \subseteq \mathbb{R}\) has \textbf{strong measure zero} if for every sequence of positive reals \(\{\varepsilon_n : n \in \omega\}\) there exists a sequence of intervals \(\{I_n = (n, \varepsilon_n) : n \in \omega\}\), such that for all \(n \in \omega\), \(\mu(I_n) \leq \varepsilon_n\), and \(X \subseteq \bigcup_{n \in \omega} I_n\). Furthermore, \textbf{Borel’s conjecture} is the statement that there are no uncountable strong measure zero sets (see Borel [4]). Now, Goldstern, Judah, and Shelah [6] showed that \(b = \omega_1\) implies that Borel’s conjecture fails. On the other hand, using Laver forcing, Laver showed in [11] that Borel’s conjecture is consistent with \(\text{ZFC} + \varepsilon = \omega_2\) (cf. Bartoszyński and Judah [1, Section 8.3]).

132. \textbf{Combinatorial properties of Laver forcing.} Laver forcing satisfies Axiom A (see Bartoszyński and Judah [1, Lemma 7.3.27]), and therefore, Laver forcing is proper. Since Laver forcing has the Laver property (see Bartoszyński and Judah [1, Theorem 7.3.29]), it does not add Cohen reals. However, Laver forcing adds dominating reals (see Bartoszyński and Judah [1, Lemma 7.3.28]), and therefore, Laver forcing adds splitting reals. Furthermore, one can show that Laver forcing is minimal (see Gray [8]).

133. \(L \times L\) adds Cohen reals. Even though Laver forcing does not add Cohen reals, by a similar argument as in the proof of \textbf{FACT} 24.9, one can show that a product of Laver forcing \(L \times L\) always adds Cohen reals.

134. \textbf{Two Laver reals added iteratively always force CH.} Brendle [5, Theorem 3.4] showed that Laver forcing collapses \(\delta\) to \(\omega_1\) and Goldstern, Repicky, Shelah, and Spinas [7, Theorem 2.7] showed that Laver forcing (as well as Miller forcing) collapses \(\omega\) to a cardinal \(\leq \delta\). Thus, two Laver reals added iteratively always force CH (cf. Chapter 24 | \textbf{RELATED RESULT} 139).

135. \textbf{On the consistency of \(s < b\).} An \(\omega_2\)-stage iteration with countable support of Laver forcing, starting in a model of \(\text{ZFC} + \text{CH}\), yields a model in which \(\omega_1 = s < b = \varepsilon\) (see Blass [2, Section 11.7]).

136. \textbf{Combinatorial properties of restricted Laver forcing \(L_{\mathcal{U}}\).} If \(\mathcal{U} \subseteq [\omega]^\omega\) is an ultrafilter, then restricted Laver forcing \(L_{\mathcal{U}}\) obviously satisfies \(\text{ccc}\). It is not hard to show that restricted Laver forcing \(L_{\mathcal{U}}\) adds dominating reals and therefore adds splitting reals. Furthermore, since restricted Laver forcing \(L_{\mathcal{U}}\) has pure decision (see Judah and Shelah [9, Theorem 1.7]), by a similar argument as in the proof of \textbf{COROLLARY} 24.8, one can show that \(L_{\mathcal{U}}\) has the Laver property.

137. \textbf{Restricted Laver forcing \(L_{\mathcal{U}}\) collapses \(\delta\) to \(\omega_1\).} Brendle [5, Corollary 3.10(a)] showed that restricted Laver forcing \(L_{\mathcal{U}}\) collapses \(\delta\) to \(\omega_1\) (cf. \textbf{RELATED RESULT} 134).

138. \textbf{On the consistency of \(\text{hom} < c\).} Judah and Shelah showed in [9, Theorem 1.16] that if a real \(r \in [\omega]^\omega\) is \(L_{\mathcal{U}}\)-generic over \(V\), then for each colouring \(\pi : [\omega]^2 \to 2\) in the ground model there exists an \(n \in \omega\) such that \(\pi|_{\omega \setminus n^2}\) is constant. Now,
let $P_{\omega_1} = (Q_\alpha : \alpha \in \omega_1)$ be an $\omega_1$-stage iteration with finite support, where for each $\alpha \in \omega_1$, $Q_\alpha$ is restricted Laver forcing $\mathbb{L}_\alpha$ (for some ultrafilter $\mathcal{U} \subseteq [\omega]_\omega$). Further, let $V$ be a model of ZFC in which $\kappa > \omega_1$ and let $G$ be $P_{\omega_1}$-generic over $V$. Then $V[G]$ is a model in which $\omega_1 = \text{hom} < \kappa$.

References

Mathias Forcing

In this chapter we investigate a forcing notion which is closely related to Ramsey’s Theorem 2.1 and to Ramsey ultrafilters (defined in Chapter 10). So, it is not surprising that also Ramsey families (also defined in Chapter 10) are involved.

With respect to an arbitrary but fixed Ramsey family $\mathcal{E}$ we define Mathias forcing $\mathbb{M}_\mathcal{E} = (\mathbb{M}_\mathcal{E}, \leq)$ as follows:

$$M_\mathcal{E} = \{(s, x) : s \in \text{fin}(\omega) \land x \in \mathcal{E} \land \max(s) < \min(x)\}$$

$$(s, x) \leq (t, y) \iff s \subseteq t \land y \subseteq x \land t \setminus s \subseteq x$$

If $\mathcal{E} = [\omega]^\omega$, then we write just $\mathbb{M}$ instead of $\mathbb{M}_\mathcal{E}$. The finite set $s$ of a Mathias condition $(s, x)$ is called the stem of the condition. Each $\mathbb{M}_\mathcal{E}$-generic filter $G$ corresponds to a generic real $m \in [\omega]^\omega$, called Mathias real, which is in fact just the union of the stems of the conditions which belong to the generic filter $G$, i.e., $m = \bigcup \{ s \in \text{fin}(\omega) : \exists x \in \mathcal{E}((s, x) \in G)\}$.

Properties of Mathias Forcing

Mathias Forcing adds Dominating Reals

**Lemma 24.1.** Mathias forcing $\mathbb{M}_\mathcal{E}$ adds dominating reals.

**Proof.** We show that a Mathias real is always dominating: Let $m$ be $\mathbb{M}_\mathcal{E}$-generic over the ground model $V$, let $p = (s, x)$ be an arbitrary $\mathbb{M}_\mathcal{E}$-condition, and let $g \in {}^\omega \cap V$ be an arbitrary function in $V$. It is enough to show that there exists an $\mathbb{M}_\mathcal{E}$-condition $q \geq p$ such that $q \Vdash_{\mathbb{M}_\mathcal{E}} "m \text{ dominates } g"$. In order to construct the condition $q$ we run the game $\mathcal{G}_\mathcal{E}$ where the Maiden plays according to the following strategy: The Maiden’s first move is

$$x_0 = x \setminus (g(n_0)^+)$$
where $n_0 = |s|$, and for $i \in \omega$ she plays

$$x_{i+1} = x_i \setminus \max \{ g(n_0 + i)^+, a_i^+ \},$$

where $a_i$ is the $i^{th}$ move of Death. Since $\mathcal{E}$ is a Ramsey family, this strategy is not a winning strategy for the Maiden and Death can play such that $y := \{ a_i : i \in \omega \} \in \mathcal{E}$. Now, by construction we get that $(s, y) \geq p$ and

$$(s, y) \models_{\mathcal{M}_E} \forall k \geq n_0 \left( m(k) > g(k) \right),$$

which shows that $m$ is a dominating real over $V$.

Together with Fact 20.1 we get

**Corollary 24.2.** Mathias forcing $\mathcal{M}_E$ adds splitting reals.

**Mathias Forcing is Proper and has the Laver Property**

Properness of Mathias forcing and that it has the Laver property follow quite easily from the fact that for every condition $(s, x)$ and every sentence $\varphi$ of the forcing language there is a $(s, y)$ which decides $\varphi$. This property of Mathias forcing is known as pure decision and is one of the main features of Mathias forcing.

**Theorem 24.3.** Let $(s, x)$ be an $\mathcal{M}_E$-condition and let $\varphi$ be a sentence of the forcing language. Then there is an $(s, y) \geq (s, x)$ — with the same stem as $(s, x)$ — such that either $(s, y) \models_{\mathcal{M}_E} \varphi$ or $(s, y) \models_{\mathcal{M}_E} \neg \varphi$ (i.e., $(s, y)$ decides the sentence $\varphi$).

Before we can prove the theorem, we have to introduce some terminology and prove some auxiliary results: For every $\mathcal{M}_E$-condition $(s, x) \in \mathcal{M}_E$ let

$$[s, x]^\omega = \{ z \in [\omega]^\omega : s \subseteq z \subseteq s \cup x \}.$$  

Notice that the sets $[s, x]^\omega$ agree with the sets of the base for the Ellentuck topology which was introduced in Chapter 9.

For a (fixed) open set $\mathcal{O} \subseteq \mathcal{M}_E$ let $\bar{\mathcal{O}} := \bigcup \{ [s, x]^\omega : (s, x) \in \mathcal{O} \}$. An $\mathcal{M}_E$-condition $(s, x)$ is called good (with respect to $\mathcal{O}$), if there is a condition $(s, y) \geq (s, x)$ such that $[s, y]^\omega \subseteq \bar{\mathcal{O}}$; otherwise it is called bad. Furthermore, the condition $(s, x)$ is called ugly if $(s \cup \{ a \}, x \setminus a^*)$ is bad for all $a \in x$. Notice that if $(s, x)$ is ugly, then $(s, x)$ is bad, too. Finally, $(s, x)$ is called completely ugly if $(s \cup \{ a_0, \ldots, a_n \}, x \setminus a_n^*)$ is bad for all $\{ a_0, \ldots, a_n \} \subseteq x$ with $a_0 < \ldots < a_n$.

**Lemma 24.4.** If an $\mathcal{M}_E$-condition $(s, x)$ is bad, then there is a condition $(s, y) \geq (s, x)$ which is ugly.
Mathias forcing is proper and has the Laver property

**Proof.** We run the game $G_{e}$ where the Maiden plays according to the following strategy: She starts the game by playing $x_{0} := x$, and then, for $i \in \omega$, she plays $x_{i+1} \subseteq (x_{i} \setminus a_{i}^{+})$ such that $[s \cup \{a_{i}\}, x_{i+1}]^{\omega} \subseteq \emptyset$ if possible, and $x_{i+1} = (x_{i} \setminus a_{i})^{+}$ otherwise. Strictly speaking we assume that $e$ is well-ordered and that $x_{i+1}$ is the first element of $e$ with the required properties. However, since this strategy is not a winning strategy for the Maiden, Death can play so that $z := \{a_{i} : i \in \omega\} \in e$. Now, let $y := \{a_{i} \in z : [s \cup \{a_{i}\}, x_{i+1}]^{\omega} \subseteq \emptyset\}$. Because $e$ is a free family, by Lemma 10.2 we get that $y$ or $z \setminus y$ belongs to $e$. If $y \in e$, then $[s, y]^{\omega} \subseteq \emptyset$ which would imply that $(s, x)$ is good, but this contradicts the premise of the lemma. Hence, $z \setminus y \in e$, which implies that $(s, z \setminus y)$ is ugly.

**Lemma 24.5.** If an $M_{e}$-condition $(s, x)$ is ugly, then there is a condition $(s, y) \geq (s, x)$ such that $(s, y)$ is completely ugly.

**Proof.** This follows by an iterative application of Lemma 24.4. In fact, for every $i \in \omega$, the Maiden can play a set $x_{i} \in e$ such that for each $t \subseteq \{a_{0}, \ldots, a_{i-1}\}$, either the condition $(s \cup t, x_{i})$ is ugly or $[s \cup t, x_{i}]^{\omega} \subseteq \emptyset$. Now Death can play such that $y := \{a_{i} : i \in \omega\} \in e$. Assume that there exists a finite set $t \subseteq y$ such that $(s \cup t, y \setminus \max(t)^{+})$ is good. Notice that since $(s, x)$ was assumed to be ugly, $t \neq \emptyset$. Now let $t_{0}$ be a smallest finite subset of $y$ such that $q_{0} = (s \cup t_{0}, y \setminus \max(t_{0}))$ is good and let $t_{0} = t_{0} \setminus \max(t_{0})$. Then by definition of $t_{0}$, the condition $q_{0} = (s \cup t_{0}, y \setminus \max(t_{0}))$ is not good, and hence, by the strategy of the Maiden, it must be ugly, but if $q_{0}$ is ugly, then $q_{0}$ is bad, which is a contradiction to our assumption. Thus, there is no finite set $t \subseteq y$ such that $(s \cup t, y \setminus \max(t)^{+})$ is good, which implies that all these conditions are ugly, and therefore $(s, y)$ is completely ugly.

Now we are ready to prove that Mathias forcing $M_{e}$ has pure decision:

**Proof of Theorem 24.3.** Let $(s, x)$ be an $M_{e}$-condition and let $\varphi$ be a sentence of the forcing language. With respect to $\varphi$ we define $O_{1} := \{q \in M_{e} : q \not\models_{M_{e}} \varphi\}$ and $O_{2} := \{q \in M_{e} : q \models_{M_{e}} \neg \varphi\}$. Clearly $O_{1}$ and $O_{2}$ are both open and $O_{1} \cup O_{2}$ is even dense in $M_{e}$. By Lemma 24.5 we know that for any $(s, x)$ there exists $(s, y) \geq (s, x)$ such that either $[s, y]^{\omega} \subseteq O_{1}$ or $[s, y]^{\omega} \cap O_{1} = \emptyset$. In the former case we have $(s, y) \models_{M_{e}} \varphi$ and we are done. In the latter case we find $(s, y') \geq (s, y)$ such that $[s, y']^{\omega} \subseteq O_{2}$. (Otherwise we would have $[s, y]^{\omega} \cap (O_{1} \cup O_{2}) = \emptyset$, which is impossible by the density of $O_{1} \cup O_{2}$.) Hence, $(s, y') \models_{M_{e}} \neg \varphi$.

As a consequence of Theorem 24.3 we can show that each infinite subset of a Mathias real is a Mathias real.

**Corollary 24.6.** If $m \in [\omega]^{\omega}$ is a Mathias real over $V$ and $m'$ is an infinite subset of $m$, then $m'$ is a Mathias real over $V$ too.
Proof. Let $D \subseteq M_\omega$ be an arbitrary open dense subset of $M_\omega$ which belongs to $V$ and let $D'$ be the set of all conditions $(s, z) \in M_\omega$ such that for all $t \subseteq s$, $[t, z]^\omega \subseteq D$. Notice that $D'$ belongs to $V$.

First we show that $D'$ is a dense (and by definition also open) subset of $M_\omega$: For this take an arbitrary condition $(s, x) \in D$ and let $\{t_i : 0 \leq i \leq h\}$ be an enumeration of all subsets of $s$. Because $D$ is open dense in $M_\omega$ we find a condition $(t_0, y_0)$ such that $y_0 \subseteq x$ and $[t_0, y_0]^\omega \subseteq D$. Moreover, for each $i < h$ we find a condition $(t_{i+1}, y_{i+1})$ such that $y_{i+1} \subseteq y_i$ and $[t_{i+1}, y_{i+1}]^\omega \subseteq D$.

Now, let $y := y_h$. Then $(s, y) \in D'$, which implies that $D'$ is dense in $M_\omega$.

Let $m \in [\omega]^\omega$ be a Mathias real over $V$ and let $m'$ be an infinite subset of $m$. Since $D'$ is an open dense subset of $M_\omega$ and $m$ is an $M_\omega$-generic real, there exists a condition $(s, x) \in D'$ such that $s \subseteq m \subseteq s \cup x$. For $t = m' \cap s$ we get $t \subseteq m' \subseteq t \cup x$, and by definition of $D'$ we have $[t, x]^\omega \subseteq D$. Thus, $m'$ meets the open dense set $D$, and since $D$ was arbitrary, this completes the proof.

As a consequence we get properness of Mathias forcing:

Corollary 24.7. Mathias forcing $M_\omega$ is proper.

Proof. Let $V$ be a model of ZFC. Further, let $N = (N, \in)$ be a countable elementary submodel of $(V, \in)$ which contains $M_\omega$, and let $(s, x) \in M_\omega \cap N$.

Since $N$ is countable (in $V$), there exists a Mathias real $m \in [s, x]^\omega \cap V$ over $N$. Notice that $(s, m \setminus s) \geq (s, x)$ and that $(s, m \setminus s)$ belongs to $V$. Now, by Corollary 24.6, every $m' \in [s, m \setminus s]^\omega$ is a Mathias real over $N$, and hence, the $M_\omega$-condition $(s, m \setminus s)$ is $N$-generic.

In Chapter 21 we have seen that Cohen forcing adds unboundedly reals, but not dominating reals. Now we shall show that Mathias forcing $M_\omega$ even though it adds dominating reals, it does not add Cohen reals (but see also Fact 24.9):

Corollary 24.8. Mathias forcing $M_\omega$ has the Laver property and therefore does not add Cohen reals.

Proof. Let $f \in {}^{\omega\omega} \cap V$ be an arbitrary function which belongs to $V$ and let $g$ be an $M_\omega$-name for a function in ${}^{\omega\omega}$ such that $0 \not\Vdash_{M_\omega} \forall n \in \omega( g(n) \leq f(n))$. Further, let $F$ be the set of all functions $S : \omega \to \text{fin}(\omega)$ such that for every $n \in \omega$, $|S(n)| \leq 2^n$. We have to show that $0 \not\Vdash_{M_\omega} \exists S \in F \cap V \forall n \in \omega( g(n) \in S(n))$. In other words, we have to show that for every $M_\omega$-condition $(s, x)$ there exists an $(s, y) \geq (s, x)$ and an $S \in F \cap V$ such that $(s, y) \Vdash_{M_\omega} \forall n \in \omega( g(n) \in S(n))$.

By Theorem 24.3, and since $g$ is bounded by $f(n)$, for every $M_\omega$-condition $(t, z)$ and for every $n \in \omega$ there exists a condition $(t', z') \geq (t, z)$ which decides $g(n)$, i.e., $(t, z) \Vdash_{M_\omega} g(n) = k$ for some $k \leq f(n)$. Let $(s, x)$ be any $M_\omega$-condition. We run the game $G_\omega$ where the MAIDEN plays according to the following strategy: She starts the game by playing $x_0 \subseteq x$ such that $(s, x_0)$
Mathias forcing is proper and has the Laver property

decides \( g(0) \), and we define \( S(0) := \{ k \leq f(0) : (s, x_0) \Vdash_{M_{\omega}} g(n) = k \} \).
Notice that \(|S(0)| = 1 = 2^0\). In general, for \( n \in \omega \), the Maiden plays \( x_{n+1} \subseteq (x_n \setminus a_n^0) \) such that for every \( \bar{a} \subseteq \{ a_0, \ldots, a_n \} \), \( (s \cup \bar{a}, x_{n+1}) \) decides \( g(n+1) \), and we define \( S(n+1) \) as the set of all \( k \leq f(n+1) \) such that, for some \( \bar{a} \subseteq \{ a_0, \ldots, a_n \} \), \( (s \cup \bar{a}, x_{n+1}) \Vdash_{M_{\omega}} g(n+1) = k \). Notice that \(|S(n+1)| \leq |\mathcal{P}(\{ a_0, \ldots, a_n \})| = 2^{n+1} \). Since this strategy is not a winning strategy for the Maiden, Death can play such that \( y := \{ a_n : n \in \omega \} \in \mathcal{D} \). Now, by construction, \( S \in F \cap V \) and for each \( n \in \omega \) we have \( (s, y) \Vdash_{M_{\omega}} g(n) \in S(n) \).
Thus, the set \( S \) and the \( M_{\omega} \)-condition \((s,y)\) have the required properties, which completes the proof. \( \square \)

Since Mathias forcing has the Laver property and is proper, a countable support iteration of Mathias forcing notions does not add Cohen reals. However, the next result shows that this is not true for a product of Mathias forcing (compare with Chapter 23 Related Result 127 and with Chapter 22 Related Result 121):

**Fact 24.9.** The product of any two Mathias forcing notions always adds Cohen reals.

**Proof.** Let \( G_1 \times G_2 \) be \( M_{\omega} \times M_{\omega} \)-generic over some model \( V \) of \( \text{ZFC} \) and let \( m_1 \) and \( m_2 \) be the corresponding Mathias reals (recall that \( m_1, m_2 \in [\omega]^\omega \)). Further, let \( \bar{m}_1, \bar{m}_2 \in \omega^\omega \) be the (unique) strictly increasing functions which map \( \omega \) onto \( m_1 \) and \( m_2 \) respectively (i.e., for \( i \in \{1, 2\} \), \( \bar{m}_i \) is strictly increasing and \( \bar{m}_i[\omega] = m_i \)). We shall show that \( c_{m_1, m_2} \in 2^2 \), defined by stipulating

\[
c_{m_1, m_2}(k) = \begin{cases} 0 & \text{if } \bar{m}_1(k) \leq \bar{m}_2(k), \\ 1 & \text{otherwise}, \end{cases}
\]

is a Cohen real over \( V \).

For \( s \in \text{fin}(\omega) \) we define \( \bar{s} \in [s]^\omega \) similarly; i.e., \( s = \{ s(k) : k \in |s| \} \) and for all \( k, l \in |s| \) with \( k < l \) we have \( s(k) < s(l) \). Further, for \( s, t \in \text{fin}(\omega) \) with \( |s| = |t| \) let \( \gamma_{s,t} \in [s]^\omega \) be such that

\[
\gamma_{s,t}(k) = \begin{cases} 0 & \text{if } s(k) \leq t(k), \\ 1 & \text{otherwise}. \end{cases}
\]

Now, let

\[
E = \{ \langle (s, x), (t, y) \rangle \in M_\omega \times M_\omega : |s| = |t| \}
\]

and consider the following function:

\[
\Gamma : E \to \bigcup_{n \in \omega} \omega^n,
\]

\[
\langle (s, x), (t, y) \rangle \mapsto \gamma_{s,t}
\]

Obviously, whenever \( D \subseteq \bigcup_{n \in \omega} \omega^n \) is open dense, then \( \Gamma^{-1}[D] = \{ p \in M_\omega \times M_\omega : \Gamma(p) \in E \} \) is dense in \( M_\omega \times M_\omega \), and since \( (m_1, m_2) \) is \( M_\omega \times M_\omega \)-generic over \( V \), we get that \( c_{m_1, m_2} \) is a Cohen real over \( V \). \( \square \)
A Model in which $p < h$

Before we construct a model in which $p < h$, we shall show that $M \cong U \ast M_\mathcal{U}$, where $U = ([\omega]^\omega/\text{fin}, \leq)$ (which was introduced in Chapter 14). To simplify the notation we write $\hat{\omega}$ instead of $[\omega]^\omega/\text{fin}$.

**Lemma 24.10.** $M \cong U \ast M_\mathcal{U}$, where $\mathcal{U}$ is the canonical $U$-name for the $U$-generic ultrafilter.

**Proof.** Firstly, recall that every $(U \ast M_\mathcal{U})$-condition is of the form $\langle \{x\}^\omega, (t, y) \rangle$, where

$$[x]^\omega \models U \langle (t, y) \rangle$$

is an $M_\mathcal{U}$-condition,

in particular, $[z]^\omega \models U y \in \mathcal{U}$. Furthermore, since $U$ does not add new reals, for every $U$-name $(t, y)$ for an $M_\mathcal{U}$-condition, and for every $U$-condition $[z]^\omega$, there is an $M$-condition $(s, x)$ in the ground model and a $U$-condition $[z']^\omega \geq [z]^\omega$ such that

$$[z']^\omega \models U (s, x) = (t, y).$$

With these facts one can show that the function $h : M \to \hat{\omega} \times M_\mathcal{U}$

$$h(s, x) \mapsto \langle \{x\}^\omega, (s, x) \rangle$$

is a dense embedding—we leave the details as an exercise to the reader. Hence, by Fact 14.3, we get that Mathias forcing $M$ is equivalent to the two-step iteration $U \ast M_\mathcal{U}$. $\dashv$

As a side-result of Lemma 24.10 we get that whenever $m \in [\omega]^\omega$ is a Mathias real over $V$, then the set $\mathcal{U} = \{ x \subseteq \omega : m \subseteq^* x \}$ is $U$-generic over $V$, in particular, $\mathcal{U}$ is a Ramsey ultrafilter in $V[\mathcal{U}]$. The following fact is just a reformulation of this observation.

**Fact 24.11.** If $m$ is a Mathias real over $V$, then $m$ is almost homogeneous for all colorings $\pi : [\omega]^2 \to 2$ which belong to $V$.

**Proposition 24.12.** $p = \text{cov}(M) < h$ is consistent with ZFC.

**Proof.** By Theorem 21.5, and since $\omega_1 < p$, it is enough to show that $\omega_1 = \text{cov}(M) < h = \omega_2$ is consistent with ZFC.

First we show that a $\omega_2$-iteration with countable support of Mathias forcing, starting from a model $V$ of ZFC + CH, yields a model in which $h = \omega_2$.

Let $\mathbb{P}_{\omega_2} = \langle Q_\alpha : \alpha \in \omega_2 \rangle$ be a countable support iteration of Mathias forcing, i.e., for all $\alpha \in \omega_2$ we have $0_\alpha \models \langle Q_\alpha : \alpha \in \omega_2 \rangle$ is Mathias forcing’. By Lemma 24.10 we may assume that for all $\alpha \in \omega_2$ we have $0_\alpha \models \langle Q_\alpha : \alpha \in \omega_2 \rangle$ is the two-step iteration $U \ast M_\mathcal{U}$.}
Let $V$ be a model of $ZFC + CH$ and let $G$ be $P_{\omega_2}$-generic over $V$. Since Mathias forcing is proper, by Theorem 20.5(a) we have $V[G] = V[\omega_2]$. In order to show that $V[G] = V[\omega_2]$ it is enough to show that in $V[G]$, the intersection of any family of size $\omega_1$ of open dense subsets of $\omega$ is non-empty.

**Claim 1.** If each family $\{D_\nu : \nu \in \omega_1\}$ of open dense subsets of $\omega$ which belongs to $V[G]$ has non-empty intersection, then $\beth > \omega_1$.

**Proof of Claim 1.** The proof is by contraposition. Assume that $\mathcal{A} = \{A_\nu : \nu \in \omega_1\}$ is a shattering family. For every $\nu \in \omega_1$ let

$$D_\nu = \{y \in [\omega]^{\omega_1} : \exists z \in A_\nu(y \subseteq^* z)\}.$$ 

Since $\mathcal{A}$ is shattering, for every $x \in [\omega]^{\omega_1}$ there is a $\nu_0 \in \omega_1$ such that $x$ has infinite intersection with at least two distinct members of $A_{\nu_0}$, which implies that $x \notin D_{\nu_0}$ and shows that $\bigcap \{D_\nu : \nu \in \omega_1\} = \emptyset$.

The following claim is a kind of reflection principle (cf. Theorem 15.2).

**Claim 2.** Let $\{D_\nu : \nu \in \omega_1\}$ be a family of open dense subsets of $\omega$ which belongs to $V[G]$. Then there is an $\alpha \in \omega_2$ such that for every $\nu \in \omega_1$ the set $D_\nu \cap V[G_\alpha]$ belongs to $V[G_\alpha]$ and is open dense in $\omega^{V[G_\alpha]}$.

**Proof of Claim 2.** It is enough to find an ordinal $\alpha \in \omega_2$ such that for every $\nu \in \omega_1$, $D_\nu \cap V[G_\alpha]$ belongs to $V[G_\alpha]$ and is dense in $\omega^{V[G_\alpha]}$—that $D_\nu \cap V[G_\alpha]$ is open in $\omega^{V[G_\alpha]}$ follows from the fact that $V[G_\alpha]$ is transitive.

Since Mathias forcing is proper and $V = CH$, by Lemma 20.4 we get that for each $\gamma \in \omega_2$, $V[G]\models CH$. For every $\gamma \in \omega_2$, let $\{x_\eta^\nu : \eta \in \omega_1\}$ be an enumeration of $[\omega]^{\omega_1} \cap V[G_\gamma]$. Since no new reals are added at limit stages of uncountable cofinality (see Lemma 18.9), for all $\eta, \nu \in \omega_2$ there is a least ordinal $\gamma_\nu > \gamma$, $\gamma_\nu^\nu \in \omega_2$, such that there is a set $y_\eta^\nu \in D_\nu \cap V[G_{\gamma_\nu}]$ with $y_\eta^\nu \subseteq^* x_\eta^\nu$. Let $\beta(\gamma) = \bigcup \{\gamma_\nu^\nu : (\eta, \nu) \in \omega_1 \times \omega_1\}$ and for $\xi \in \omega_1$ let

$$\beta(0) = \begin{cases} \bigcup_{\xi' < \xi} \beta(0) & \text{if } \xi \text{ is a limit ordinal}, \\ \beta(\beta(0)) & \text{if } \xi = \xi' + 1. \end{cases}$$

Then $\alpha = \bigcup \{\beta(0) : \xi \in \omega_1\}$, which is a limit ordinal below $\omega_2$ of cofinality $\omega_1$, has the required properties.

For every $\nu \in \omega_1$ let $D_\nu' = D_\nu \cap V[G_\alpha]$. Further, let $\mathcal{U}_\alpha$ be the $U$-generic Ramsey filter over $V[G_\alpha]$, determined by $G$. In the model $V[G_\alpha][\mathcal{U}_\alpha]$, $\mathcal{U}_\alpha$ meets every $D_\nu'$ (i.e., for every $\nu \in \omega_1$, $\mathcal{U}_\alpha \cap D_\nu' \neq \emptyset$). Now, for $m_\alpha$, the $M_{\mathcal{U}_\alpha}$-generic Mathias real over $V[G_\alpha][\mathcal{U}_\alpha]$ (i.e., the second component of the decomposition of Mathias forcing), we have $m_\alpha \in \bigcap \{D_\nu : \nu \in \omega_1\}$ which shows that $\bigcap \{D_\nu' : \nu \in \omega_1\}$ is non-empty. Thus, by Claim 1 and since $V[G] = \beth = \omega_2$, $V[G] = \omega_1$. It remains to show that $V[G] = \omega_1 = \text{cov}(M)$. For this, recall that Mathias forcing has the Laver property and therefore, by Proposition 20.2,
Mathias forcing does not add Cohen reals. Now, since the Laver property is preserved under countable support iteration of proper forcing notions (see Theorem 20.7), a countable support iteration of Mathias forcing does not add Cohen reals to the ground model. Hence, by Corollary 21.8 (which says that $\text{cov}(\mathcal{M})$ is preserved if no Cohen reals are added) we have $V[G] \models \omega_1 = \text{cov}(\mathcal{M})$.

Notes

Mathias forcing restricted to happy families (which are slightly more general than Ramsey families) was introduced and investigated by Mathias in [12]. However, most of the results presented in this chapter can be found in Halbeisen [5].

Related Results

139. Mathias forcing collapses $\mathcal{C}$ to $\mathcal{H}$ and $\mathcal{D}$ to $\omega_1$. The fact that Mathias forcing collapses $\mathcal{C}$ to $\mathcal{H}$ is just a consequence of Lemma 24.10 and the fact that ultrafilter forcing $\mathcal{U}$ collapses $\mathcal{C}$ to $\mathcal{H}$ (see Chapter 25 [Related Result 144]). Furthermore, Brendle [2, Corollary 3.10.(c)/(d)] showed that Mathias forcing collapses $\mathcal{D}$ to $\omega_1$, and since $\mathcal{H} \leq \mathcal{D}$, one gets that two Mathias reals added iteratively always force CH (cf. Chapter 23 [Related Result 134]).

140. Mathias forcing and Borel’s conjecture. By adding random reals to the model constructed in the proof of Proposition 24.12, Judah, Shelah, and Woodin [10] showed that Borel’s conjecture is consistent with $\mathcal{C}$ being arbitrarily large (cf. Chapter 23 [Related Result 131]), and see also Bartoszyński and Judah [1, Theorem 8.3.7]).

141. Restricted Mathias forcing which does not add dominating reals. Canjar showed in [3] that under the assumption $\mathcal{D} = \mathcal{C}$, there exists an ultrafilter $\mathcal{U}$ over $\omega$ such that $M_{\mathcal{U}}$ does not add dominating reals. Further, he showed that such an ultrafilter is necessarily a $P$-point.

142. Between Laver and Mathias forcing. If $\mathcal{U}$ is an ultrafilter, then restricted Mathias forcing $M_{\mathcal{U}}$ is equivalent to restricted Laver forcing $L_{\mathcal{U}}$ if and only if $\mathcal{U}$ is a Ramsey ultrafilter (see Judah and Shelah [8, Theorem 1.20]). On the other hand, if $\mathcal{U}$ is not a Ramsey ultrafilter, then $M_{\mathcal{U}}$ and $L_{\mathcal{U}}$ can be quite different (see Judah and Shelah [9]).

143. The Ramsey property of projective sets. The hierarchy of projective subsets of $[\omega]^\omega$ is defined as follows: Let $A \subseteq ([\omega]^\omega)^k$ be a $k$-dimensional set (for some positive integer $k$). Then $A$ is a $\Sigma^1_k$-set if $A$ the projection along $[\omega]^\omega$ of a closed set $C \subseteq ([\omega]^\omega)^{k+1}$, and $A$ is a $\Pi^1_k$-set if it is the complement of a $\Sigma^1_k$-set. In general, for integers $n \geq 1$, $A$ is a $\Sigma^1_{n+1}$-set if $A$ the projection along $[\omega]^\omega$ of a $(k+1)$-dimensional $\Pi^1_k$-set, and $A$ is a $\Pi^1_{n+1}$-set if it is the complement of a $\Sigma^1_{n+1}$-set. Furthermore, we say that $A$ is a $\Delta^1_n$-set if $A$ is a $\Sigma^1_n$-set as well as a $\Pi^1_n$-set. Below, $\Sigma^1_n$, $\Pi^1_n$, and $\Delta^1_n$, denote the collections of the corresponding subsets of $[\omega]^\omega$. The sets $A \subseteq [\omega]^\omega$ belonging to one of the collections $\Sigma^1_n$, $\Pi^1_n$, and $\Delta^1_n$. 


or $\Delta^1_n$, are called projective sets. With respect to inclusion, we get the following diagram:

$$
\Sigma^1_1 \supseteq \Sigma^1_2 \supseteq \Sigma^1_3 \\
\Delta^1_1 \supseteq \Delta^1_2 \supseteq \Delta^1_3 \\
\Pi^1_1 \supseteq \Pi^1_2 \supseteq \Pi^1_3
$$

If all $\Sigma^1_n$-sets $A \subseteq [\omega]^\omega$ have the Ramsey property (defined in Chapter 9), then we shall write $\Sigma^1_n(\mathcal{F})$; the notations $\Pi^1_n(\mathcal{F})$ and $\Delta^1_n(\mathcal{F})$ are defined accordingly. It is natural to ask whether all projective sets have the Ramsey property. Even though the answer to this question is not decidable in ZFC, one can show the following facts:

- For all $n \in \omega$: $\Sigma^1_n(\mathcal{F}) \iff \Pi^1_n(\mathcal{F})$ (trivial).
- $\Delta^1_1(\mathcal{F}) \iff \Sigma^1_2(\mathcal{F})$ (see Judah and Shelah [8, Theorem 2.7]).
- ZFC $\vdash \Sigma^1_1(\mathcal{F})$ (see Silver [13] or Ellentuck [4]).
- $\text{L} \not\models \Delta^1_2(\mathcal{F})$ (cf. Judah and Shelah [8, Lemma 2.2]).
- $\text{Con}(\text{ZFC}) \Rightarrow \text{Con}(\text{ZFC} + \Delta^1_3(\mathcal{F}))$ (see Judah [7, Theorem 0.8]).

Furthermore, Mathias showed in [11, Section 5] — using Mathias forcing — that if $\text{ZFC} + \text{"there is a strongly inaccessible cardinal"}$ is consistent (where $\kappa$ is strongly inaccessible if $\kappa$ is a regular limit cardinal and for all $\lambda < \kappa$, $\lambda^\lambda < \kappa$), then so is $\text{ZFC} + \text{"every projective set has the Ramsey property"}$. However, it is still open whether one can take “Mathias’ inaccessible” away, i.e., whether one can construct a model of ZFC in which all projective sets have the Ramsey property without assuming the existence of a strongly inaccessible cardinal (cf. Shelah [12]). Moreover, it is not even known whether $\Sigma^1_3(\mathcal{F})$ implies the existence of a strongly inaccessible cardinal. For partial results see Halbeisen and Judah [6, Theorem 5.3] and Brendle [2, Proposition 4.3].

References


On the Existence of Ramsey Ultrafilters

So far we have seen that \( p = \kappa \) implies the existence of Ramsey ultrafilters (see Proposition 10.9). In particular, if we assume CH, then Ramsey ultrafilters exist. Moreover, by Proposition 13.9 we know that MA(countable) implies the existence of \( \aleph_1 \) mutually non-isomorphic Ramsey ultrafilters. Furthermore, by Theorem 21.5 we know that \( p \leq \text{cov}(\mathcal{M}) \), and Chapter 13 Related Result 80 tells us that \( \text{MA(countable)} \) is equivalent to \( \text{cov}(\mathcal{M}) = \kappa \). Hence, \( \text{cov}(\mathcal{M}) = \kappa \) is a sufficient condition for the existence of Ramsey ultrafilters and it is natural to ask whether \( \text{cov}(\mathcal{M}) = \kappa \) is necessary, too. In the first section of this chapter we shall give a negative answer to this question by constructing a model of \( \text{ZFC} + \text{cov}(\mathcal{M}) < \kappa \) in which there is a Ramsey ultrafilter. Since in that model we have \( \mathfrak{h} = \kappa \) and \( \mathfrak{h} \) is related to the Ramsey property (cf. Chapter 9), one might think that perhaps \( \mathfrak{h} = \kappa \) implies the existence of a Ramsey ultrafilter; but this is not the case, as we shall see in the second section of this chapter.

There may be a Ramsey Ultrafilter and \( \text{cov}(\mathcal{M}) < \kappa \)

In the proof of Proposition 24.12 we have constructed a model \( V \) of ZFC, usually called Mathias’ model, in which \( \text{cov}(\mathcal{M}) < \kappa \). Furthermore, Proposition 14.18 states that if \( G \) is \( U \)-generic over \( V \), where \( U = (\omega^\omega/\text{fin}, \subseteq) \), then \( U \) is a Ramsey ultrafilter in \( V[G] \); in particular, ultrafilter forcing \( U \) adds a Ramsey ultrafilter to \( V \). Recall that \( \omega^\omega/\text{fin} = \{[x] : x < \omega]\} \) and \( [x] \leq [y] \iff y \subseteq^* x \). So, at first glance we just have to force with \( U \) over Mathias’ model. However, in order to get a model in which there exists a Ramsey ultrafilter and \( \text{cov}(\mathcal{M}) < \kappa \), it has to be shown that ultrafilter forcing \( U \) does not collapse \( \kappa \) to \( \text{cov}(\mathcal{M}) \) — for this, we first show that ultrafilter forcing \( U \) does not collapse \( \kappa \) to any cardinal below \( \mathfrak{h} \).

Lemma 25.1. If \( G \) is \( U \)-generic over \( V \), then \( V[G] \models \kappa \geq \mathfrak{h}^V \), in other words, ultrafilter forcing \( U \) does not collapse \( \kappa \) to any cardinal \( \kappa < \mathfrak{h}^V \).
Proof. Let $G$ be $\mathcal{U}$-generic over some model $V$ of ZFC. Since the forcing notion $\mathcal{U}$ is $\sigma$-closed (by the proof of Theorem 8.1), and since $\sigma$-closed forcing notions do not add reals (by Lemma 14.17), ultrafilter forcing $\mathcal{U}$ does not add any new reals to the ground model $V$. In particular, we have $V[G] \models \mathfrak{c} \leq \mathfrak{b}^V$. Thus, in order to show that $V[G] \models \mathfrak{c} \geq \mathfrak{b}^V$, it is enough to prove that in $V[G]$ there is no surjection from some $\kappa < \mathfrak{b}^V$ onto $\mathfrak{c}$ (which implies $\mathfrak{c} \leq \mathfrak{b}^V$).

Let $\kappa$ be a cardinal with $V \models \kappa < \mathfrak{b}$ and let $g \in V[G]$ be a function from $\kappa$ to $\mathfrak{c}$. In order to prove that $g$ fails to be surjective, it is enough to show that $g$ is in the ground model $V$—notice that this would imply $V \models \kappa \leq \kappa < \mathfrak{b}$, contradicting the fact that $\mathfrak{b} \leq \mathfrak{c}$. Let $g$ be a $\mathcal{U}$-name for $g$ and let $x_0 \in [\omega]^\omega$ be such that $[x_0] \mathcal{U}^V g : \kappa \rightarrow \mathfrak{c}$. For each $\alpha \in \kappa$ let

$$D_\alpha = \{[y] : [y \cap x_0] < \omega \vee (y \subseteq^* x_0 \wedge \exists \gamma \in \mathfrak{c} ([y] \mathcal{U}^V g(\alpha) = \gamma) \}.$$ 

Each $D_\alpha$ is open dense. Thus, for each $\alpha \in \kappa$ we can choose a mad family $\mathcal{A} \subseteq \bigcup D_\alpha$. Now, by Lemma 8.14 there is a mad family $\mathcal{A} \subseteq [\omega]^\omega$ such that

$$\forall \alpha \in \kappa \forall y \in \mathcal{A}_\alpha \exists x \in \mathcal{A}(x \subseteq^* y).$$

Furthermore, let $D = \{[y] : \exists x \in \mathcal{A}(y \subseteq^* x)\}$. Then $D$ is open dense and therefore $G \cap D$ is non-empty. For $[y_0] \in (G \cap D)$ we get $[y_0] \leq [x_0]$, in particular, $[y_0] \mathcal{U}^V g : \kappa \rightarrow \mathfrak{c}$. Moreover, by construction of $D$,

$$\forall \alpha \in \kappa \exists \gamma \in \mathfrak{c}([y_0] \mathcal{U}^V g(\alpha) = \gamma).$$

Let $g_0 : \kappa \rightarrow \mathfrak{c}$ be such that for all $\alpha \in \kappa$, $[y_0] \mathcal{U}^V g(\alpha) = g_0(\alpha)$. Then $g_0$ belongs to the ground model $V$ and in addition we have $[y_0] \mathcal{U}^V g = g_0$. Now, since $[y_0] \in G$, this shows that $g = g[G]$ belongs to $V$.

With this result, we easily can construct a model with a Ramsey ultrafilter in which $\text{cov}(\mathcal{M}) < \mathfrak{c}$.

**Proposition 25.2.** The existence of a Ramsey ultrafilter is consistent with ZFC + $\text{cov}(\mathcal{M}) < \mathfrak{c}$.

**Proof.** Let $V$ be Mathias’ model (i.e., the model constructed in the proof of Proposition 24.12), and let $G$ be $\mathcal{U}$-generic over $V$. Then we have

$$V \models \omega_1 = \text{cov}(\mathcal{M}) < \mathfrak{b} = \mathfrak{c} = \omega_2,$$

and by Lemma 25.1 we get $V[G] \models \mathfrak{b}^V = \mathfrak{c}$, in particular,

$$V[G] \models \text{cov}(\mathcal{M}) < \mathfrak{c}.$$

Finally, by Proposition 14.18 we get that $\bigcup G$ is a Ramsey ultrafilter in $V[G]$, and therefore, $V[G]$ is a model with a Ramsey ultrafilter in which $\text{cov}(\mathcal{M}) < \mathfrak{c}$.

$\blacksquare$
There may be no Ramsey Ultrafilter and \( h = c \)

The goal of this section is to show that there are no Ramsey ultrafilters in Mathias’ model— which is a model of \( h = c \). In fact we prove that not even rapid filters exist in that model. For this we first prove a few auxiliary results concerning \( \omega_2 \)-iterations of Mathias forcing. Then we recall the definition of rapid filters (cf. Chapter 10 | RELATED RESULT 70) and show that every Ramsey ultrafilter is a rapid filter; and finally we prove that there are no rapid filters in Mathias’ model.

Let us start by recalling some terminology of Mathias forcing \( M = ( M, \leq ) \) and by introducing some notation: Let \( ( s, x ) \) and \( ( t, y ) \) be two \( M \)-conditions. Recall that

\[
(s, x) \leq (t, y) \iff s \subseteq t \land y \subseteq x \land t \setminus s \subseteq x .
\]

Now, let us define

\[
(s, x) \leq^0 (t, y) \iff (s, x) \leq (t, y) \land s = t .
\]

In order to define “\( \leq^n \)” for positive integers \( n \in \omega \), we write sets \( x \in [\omega]^\omega \) in increasing order, i.e., \( x = \{ a_k : k \in \omega \} \) where \( k < k' \rightarrow a_k < a_{k'} \). By abuse of notation we shall just write \( x = \{ a_0 < a_1 < \cdots \} \). Now, for \( n \in \omega \) and \( x = \{ a_0 < a_1 < \cdots \} \) we define

\[
(s, x) \leq^n (t, y) \iff (s, x) \leq^0 (t, y) \land \forall k \in n (a_k \in y) .
\]

In this notation, the fact that Mathias forcing has pure decision (see Theorem 24.3) can be expressed as follows: Let \( p \in M \) be an \( M \)-condition and let \( \varphi \) be a sentence of the forcing language. Then there exists a \( q \in M \) with \( p \leq^0 q \) such that either \( q \forces^M \varphi \) or \( q \forces^M \neg \varphi \).

In order to get familiar with this notation we prove the following fact. Notice that this fact was already used implicitly in the previous chapter (e.g., in the proof of Corollary 24.8).

Fact 25.3. Let \( q \) be an \( M \)-name for a function \( g \in {}^\omega \omega \) and let \( n_0 \in \omega \) be a fixed integer. Further, let \( p \in M \) and \( k \in \omega \) be such that

\[
p \forces^M g(n_0) = k .
\]

Then there are \( q \in M \) and \( l_0 \in k \) such that \( p \leq^0 q \) and

\[
q \forces^M g(n_0) = l_0 .
\]

Proof. Since Mathias forcing has pure decision (see Theorem 24.3), there is a \( q_0 \in M \) with \( p \leq^0 q_0 \) such that

\[
q_0 \forces^M g(n_0) = 0 \quad \text{or} \quad q_0 \bigvee_{0 < l < k} g(n_0) = l ,
\]

where \( k = \{ \omega \setminus n_0 \} \).
where $\bigvee_{0 \leq i < k} \varphi_i$ is an abbreviation for $\varphi_1 \lor \cdots \lor \varphi_{k-1}$. In the latter case, by pure decision we find a $q_1 \in M$ with $q_0 \leq^0 q_1$ such that

$$q_1 \models_M g(n_0) = 1 \quad \text{or} \quad q_1 \models_{\mathcal{M}} \bigvee_{1 \leq i < k} g(n_0) = l.$$  

Proceeding this way, we finally find a $q \in M$ with $p \leq^0 q$ and an $l_0 \in k$ such that $q \models_M g(n_0) = l_0$.

To prove the following lemma, we just have to iterate this procedure.

**Lemma 25.4.** Let $q$ be an $\mathcal{M}$-name for a function $g \in \omega^\omega$ and let $n_0 \in \omega$ be a fixed integer. Further, let $p \in M$ and $k \in \omega$ be such that

$$p \models_M g(n_0) \in k.$$  

Then, for every $i \in \omega$, there are $q_i \in M$ and $I_i \subseteq k$ such that $p \leq^i q_i$, $|I_i| \leq i + 1$, and

$$q_i \models_{\mathcal{M}} \bigvee_{l \in I_i} g(n_0) = l.$$  

**Proof.** The proof is by induction on $i$: For $i = 0$, this is just Fact 25.3. So, let us assume that the lemma holds for some $i \in \omega$. In other words, there are $q_i \in M$ and $I_i \subseteq k$ such that $p \leq^i q_i$, $|I_i| \leq i + 1$, and $q_i \models_{\mathcal{M}} \bigvee_{l \in I_i} g(n_0) = l$.

Let $p = (s, x)$ and $q_i = (s, y_i)$, where $x = \{a_0 < a_1 < \cdots\}$ and $y_i = \{b_0 < b_1 < \cdots\}$ respectively. Notice that for all $j \in i$, $a_j = b_j$. If $a_i = b_i$, then, for $I_{i+1} := I_i$ and $q_{i+1} := q_i$, we get

$$q_{i+1} \models_{\mathcal{M}} \bigvee_{l \in I_{i+1}} g(n_0) = l.$$  

Otherwise, we have $a_i < b_i$ (since $p \leq^i q_i$), and by Fact 25.3, we find $y' \subseteq y \setminus a_i$ and $l_{i+1} \in k$ such that

$$(s \cup \{a_j : j \leq i\}, y') \models_M g(n_0) = l_{i+1}.$$  

Now, for $I_{i+1} := I_i \cup \{l_{i+1}\}$ and $q_{i+1} := (s \cup \{a_j : j \leq i\}, y')$ we get

$$q_{i+1} \models_{\mathcal{M}} \bigvee_{l \in I_{i+1}} g(n_0) = l,$$

where by construction, $p \leq^{i+1} q_{i+1}$ and $|I_{i+1}| \leq i + 2$. $

The next result uses the fact that Mathias forcing is proper (see Corollary 24.7).
Lemma 25.5. Let $V$ be a model of ZFC, let $\{q_k : k \in \omega\}$ be a countable set of $M$-names for ordinals, such that for some $p \in M$ we have

\[ p \Vdash_M \forall k \in \omega (q_k \in \omega_2). \]

Then, for every $i \in \omega$, there is a countable set $A \subseteq \omega_2$ in $V$, as well as a $q \in M$ with $p \leq q$, such that

\[ q \Vdash_M \forall k \in \omega (q_k \in A). \]

Proof. Let $N = (N, \in)$ be a countable elementary submodel of $(H_\infty, \in)$ which contains $M$, $\{q_k : k \in \omega\}$, and $p$, where $p = (s, x)$. Since $N$ is countable (in $V$), there exists a Mathias real $m_G \in [s, x]^\omega \cap V$ over $N$. Notice that $(s, m_G \setminus s) \geq (s, x)$ and that $(s, m_G \setminus s)$ belongs to $V$. By Corollary 24.6, every $m_G^G \in [s, m_G \setminus s]^\omega$ is a Mathias real over $N$, and hence, the $M$-condition $q = (s, m_G \setminus s)$ is $N$-generic. Now, for $A := N \cap \omega_2$, which is countable in $V$, we get that $q \Vdash_M \forall k \in \omega (q_k \in A)$, which proves the lemma in the case when $i = 0$. For $i > 0$, we can proceed as in the proof of Lemma 25.4 — the details are left as an exercise to the reader. 

In the following result we introduce what is called a fusion argument:

Fact 25.6. Let $\{p_n : n \in \omega\}$ be a sequence of $M$-conditions such that for all $n \in \omega$, $p_n \leq p_{n+1}$. Further assume that there is an $m_0 \in \omega$ such that for all $n \geq m_0$, $p_n \leq^n p_{n+1}$. Then there exists an $M$-condition $p_\omega$ such that for all $n \geq m_0$, $p_n \leq^n p_\omega$.

Proof. For $n \in \omega$, let $p_n = (s_n, x_n)$ where $x_n = \{x_n(0) < x_n(1) < \cdots\}$, and define

\[ p_\omega = (s_{m_0} \cup \{x_{m_0}(i) : i \in m_0\}, \{x_i(i-1) : m_0 \in i \in \omega\}). \]

We leave it as an exercise to the reader to show that $p_\omega$ has the required properties. 

Below we shall generalise the previous results to countable support iterations of Mathias forcing, but first let us introduce some notations: Let $V$ be a model of ZFC, let $\mathbb{P}_\omega = (\mathbb{Q}_\gamma : \gamma \in \omega_2)$ be the countable support iteration of length $\omega_2$ of Mathias forcing $\mathbb{M}$, and let $G = (G(\gamma) : \gamma \in \omega_2)$ be $\mathbb{P}_\omega$-generic over $V$. Furthermore, for $\beta \leq \omega_2$, $K \in \text{fin}(\beta)$, $\mathbb{P}_\beta$-conditions $p$ and $q$, and $n \in \omega$, define

\[ p \leq^K q \iff p \leq q \land \forall \gamma \in K \left( q|_\gamma \Vdash_{\mathbb{P}_\gamma} p(\gamma) \leq^n q(\gamma) \right). \]

The next result shows how fusion arguments work in countable support iterations of Mathias forcing.
Lemma 25.7. Let $\beta$ be an ordinal with $1 \leq \beta \leq \omega_2$ and let $\langle p_n : n \in \omega \rangle$ be a sequence of $\mathbb{P}_\beta$-conditions. Furthermore, let $\langle K_n : n \in \omega \rangle$ be an increasing chain of finite subsets of $\beta$ (i.e., $n < n' \rightarrow K_n \subseteq K_{n'}$) such that

$$\bigcup_{n \in \omega} K_n = \bigcup_{n \in \omega} \text{supp}(p_n) \quad \text{and} \quad \forall n \in \omega (p_n \leq p_{n+1}).$$

Then there is a $\mathbb{P}_\beta$-condition $p_\omega$ such that for each $n \in \omega$, $p_n \leq p_\omega$.

Proof. For every $\gamma < \beta$, $p_n(\gamma)$ is a $\mathbb{P}_\gamma$-name for an $M$-condition. Thus, $p_n(\gamma) = (s_n, x_n)$ where $x_n = \{x_n(0) < x_n(1) < \cdots \}$. For $\gamma \in \bigcup_{n \in \omega} K_n$, let $m_0 = \min\{n \in \omega : \gamma \in K_n\}$ and define

$$p_\omega(\gamma) = (s_{m_0} \cup \{x_{m_0}(i) : i \in m_0\}, \{x_{m_0}(i) : m_0 \equiv i \in \omega\}).$$

In the case when $\gamma \notin \bigcup_{n \in \omega} K_n$ define $p_\omega(\gamma) = 0_\gamma$. We leave it as an exercise to the reader to show that $p_\omega$ has the required properties. \qed

In order to state the next result, we have to introduce again some notation:

For ordinals $\alpha < \beta \leq \omega_2$ we say that $q$ is a $\mathbb{P}_\alpha$-condition iff there is a $\mathbb{P}_\omega$-condition $p = (p(\gamma) : \gamma \in \omega_2)$ such that $q = (p(\gamma) : \alpha \leq \gamma < \beta)$. In particular, $\mathbb{P}_{\alpha\beta}$-conditions are the same as $\mathbb{P}_\beta$-conditions.

Lemma 25.8. Let $\beta$ be an ordinal with $1 \leq \beta \leq \omega_2$ and let $p$ be a $\mathbb{P}_\beta$-condition. Furthermore, let $K = \{\alpha_1 < \cdots < \alpha_t\}$ be a finite subset of $\beta$ (i.e., $i \in \omega$) and let $n \in \omega$.

(a) Let $\{q_k : k \in \omega\}$ be a countable set of $\mathbb{P}_\beta$-names for ordinals such that

$$p \not\vdash_{\mathbb{P}_\beta} \forall k \in \omega (q_k \in \omega_2).$$

Then there is a countable set $A \subseteq \omega_2$ in $V$ and a $\mathbb{P}_\beta$-condition $p'$ with $p \leq^n p'$ such that

$$p' \not\vdash_{\mathbb{P}_\beta} \forall k \in \omega (q_k \in A).$$

(b) Let $\delta$ be an ordinal, where $\beta < \delta \leq \omega_2$, and assume that for some $\mathbb{P}_\beta$-name $\gamma$ we have

$$p \not\vdash_{\mathbb{P}_\beta} \text{"} \gamma \text{ is a $\mathbb{P}_\beta$-condition"}.$$

Then there is a $\mathbb{P}_\beta$-condition $p'$ with $p \leq^n p'$ and a $\mathbb{P}_\beta$-condition $q$ such that

$$p' \not\vdash_{\mathbb{P}_\beta} \gamma = q.$$

In particular, $p' \cup q$ is a $\mathbb{P}_\beta$-condition (which is in general not the case for $p' \cup \gamma$).

(c) Let $q$ be a $\mathbb{P}_\beta$-name for a function $g \in \mathbb{P}_\omega$ and let $n_0 \in \omega$ be a fixed integer. Further, assume that for some $k \in \omega$,

$$p \not\vdash_{\mathbb{P}_\beta} g(n_0) \in k.$$
Then there is an \( I \subseteq k \) with \( |I| \leq (n + 1)^i \) and a \( \mathbb{P}_\beta \)-condition \( p_0 \) with \( p \leq_K p_0 \) such that
\[
p_0 \forces_{\mathbb{P}_\beta} \bigvee_{i \in I} g(n_0) = l.
\]

Proof. (a) Firstly recall that since Mathias forcing is proper, also \( \mathbb{P}_\delta \), as a countable support iteration of proper forcing notions, is proper (see Theorem 20.3 (b)). Thus, let \( \mathcal{N} = (\mathbb{N}, \in) \) be a countable elementary submodel of \( (V_\kappa, \in) \) which contains \( \mathbb{P}_\delta \), \( \{q_k : k \in \omega \} \), \( p \), and \( q \). Now, by similar arguments as in the proof of Lemma 25.5 we can construct a \( \mathbb{P}_\beta \)-condition \( p' \) with the required properties — the details are left as an exercise to the reader.

(b) As a consequence of (a), there is a \( \mathbb{P}_\beta \)-condition \( p' \) with \( p \leq_K p' \) as well as a countable set \( A \subseteq [\beta, \delta] \) in \( V \) such that
\[
p' \forces_{\mathbb{P}_\beta} \text{supp}(q) \subseteq A.
\]

For \( \gamma \in [\beta, \delta) \setminus A \), let \( q(\gamma) := 0_\gamma \). Otherwise, for \( \gamma \in A \), let \( q(\gamma) := \varepsilon(\gamma) \). Then \( q \in \mathbb{P}_\beta \) and \( p' \forces_{\mathbb{P}_\beta} \gamma = q \), as required.

(c) The proof is by induction on \( \beta \), where \( 1 \leq \beta \leq \omega_2 \); thus, we have to consider the case when \( \beta = 1 \), which we have already done in Lemma 25.4, the case when \( \beta \) is a successor ordinal, and the case when \( \beta \) is a limit ordinal.

For \( \beta = \delta + 1 \), where \( 1 \leq \delta \), we just consider the case when \( \delta = \omega \), and leave the other case — which is similar to the case when \( \beta \) is a limit ordinal — as an exercise to the reader. For \( p(\delta) = (\bar{s}, \bar{x}) \), where \( \bar{x} = \{ x(0) < x(1) < \cdots \} \), and for every \( j \leq n \) let
\[
r_j = (\bar{s} \cup \{ \bar{x}(i) : i \in j \}, \{ \bar{x}(i) : j \leq i \in \omega \}).
\]

Notice that \( r_j \) is a \( \mathbb{P}_\beta \)-name for an \( M \) condition. In particular, if \( p|\delta \in G|\delta \) where \( G|\delta \) is \( \mathbb{P}_\delta \)-generic over \( V \), then \( V[G|\delta] \models \text{"} r_j \text{"} \text{is an } M \text{ condition} \). Since Lemma 25.4 holds in \( V[G|\delta] \), there is a \( \mathbb{P}_\beta \)-name \( r'_j \) for an \( M \)-condition such that
\[
p|\delta \forces_{\mathbb{P}_\beta} \left( r_j \leq r'_j \land \exists l \in k (r'_j \forces_M g(n_0) = l) \right).
\]

In particular, if \( p|\delta \in G|\delta \), then, for some \( l \in k \), \( V[G|\delta] \models r'_j[G|\delta] \forces_M g(n_0) = l \). Now, by induction on \( j \), where \( 0 \leq j \leq n \), we can construct \( \mathbb{P}_\beta \)-conditions \( q_j \), \( \mathbb{P}_\beta \)-names for \( M \)-conditions \( r'_j \), as well as subsets \( I_j \subseteq k \), which satisfy the following conditions:

- \( p|\delta \leq_K q_0 \leq_K q_1 \leq_K \cdots \leq_K q_n \).
- For each \( j \leq n \) we have \( |I_j| \leq (n + 1)^{j-1} \).
- For each \( j \leq n \), \( q_j \forces_{\mathbb{P}_\beta} \exists \bar{y} \in I_j \text{ (for this, encode } r'_j \text{ into } \mathbb{P}_\alpha \text{ by a function } g_{r'_j} \text{, stipulating } g_{r'_j}(n_0) = l \text{, and apply Lemma 25.4)} \).
- \( r'_n \) is such that \( q_n \cup r'_n \forces_{\mathbb{P}_\beta} g(n_0) \in \bigcup_{j \leq n} I_j \).
Then, for \( p_0 := q_n \cup r_n \) and \( I := \bigcup_{j \leq n} I_j \) we have \( p \leq_K p_0 \) and \( |I| \leq (n+1)^t \), and \( p_0 \models \bigvee_{I \in \mathcal{F}} g(n_0) = l \), as required.

Assume now that \( \beta \) is a limit ordinal and that the lemma is true for \( \alpha_i + 1 \) (notice that \( \alpha_i + 1 < \beta \)). Let \( \bar{r} \) be a \( \mathbb{P}_{\alpha_i + 1} \)-name for some \( \mathbb{P}_{\alpha_i, + 1} \)-condition such that

\[
p|\alpha_i + 1 \models \bar{p}_{\alpha_i + 1} \left( p|\alpha_i + 1, \beta \leq r \land \exists l \in k \left( r \models \bar{p}_{\alpha_i + 1, \beta} g(n_0) = l \right) \right).\]

Applying part (b) of the lemma to \( \alpha_i + 1 \), we get a \( \mathbb{P}_{\alpha_i + 1} \)-condition \( p' \) with \( p|\alpha_i + 1 \leq_K p' \) and a \( \mathbb{P}_{\alpha_i, + 1} \)-condition \( q \) such that

\[
p' \models \bar{p}_{\alpha_i + 1} \models q.
\]

By induction hypothesis, there is a \( \mathbb{P}_{\alpha_i + 1} \)-condition \( q' \) with \( p' \leq_K q' \) and an \( I \subseteq k \) with \( |I| \leq (n+1)^t \), such that

\[
q' \models \bar{p}_{\alpha_i + 1} \exists l \in I \left( q \models \bar{p}_{\alpha_i + 1, \beta} g(n_0) = l \right).
\]

Finally, let \( p_0 = q' \cup q \). Then \( p_0 \) has the required properties.

The next result, which will be crucial in the proof that there are no rapid filters in Mathias’ model, concludes our investigation of \( \omega_2 \)-stage countable support iterations of Mathias forcing.

**Lemma 25.9.** Let \( V \) be a model of \( \text{ZFC} \), let \( \mathbb{P}_{\omega_2} \) be the countable support iteration of length \( \omega_2 \) of Mathias forcing \( M \), and let \( G = \langle G(\gamma) : \gamma \in \omega_2 \rangle \) be \( \mathbb{P}_{\omega_2} \)-generic over \( V \). Furthermore, let \( \bar{f} \) be an \( M \)-name for the first Mathias real, more precisely, \( \bar{f} \) is the name for a strictly increasing function in \( \langle \omega \rangle \) such that

\[
0_{\omega_2} \models \bar{f}(i) \in \omega = \bigcup \left\{ s : \exists x \in [\omega]^s \left((s, x) \in G(0)\right) \right\}.
\]

If \( q \) is a \( \mathbb{P}_{\omega_2} \)-name for a strictly increasing function in \( \langle \omega \rangle \) such that for some \( \mathbb{P}_{\omega_2} \)-condition \( p \) we have

\[
p \models \forall i \in \omega (f(i) < g(i)),
\]

then there are infinite sets \( \mathcal{I}_0, \mathcal{I}_1 \subseteq \omega \) in \( V \), where \( \mathcal{I}_0 \cap \mathcal{I}_1 \) is finite, and \( \mathbb{P}_{\omega_2} \)-conditions \( p_0, p_1 \), where \( p_0 \leq p \leq p_1 \), such that

\[
p_0 \models \bar{f}(\omega) \subseteq \mathcal{I}_0 \text{ and } p_1 \models \bar{f}(\omega) \subseteq \mathcal{I}_1.
\]

**Proof.** Before we can start the proof, we have to introduce some notations: Firstly notice that if \( q \) is a \( \mathbb{P}_{\omega_2} \)-condition, then \( q(0) \) is an \( M \)-condition, i.e., \( q(0) = (s, x) \) where \( s \in \text{fin}(\omega) \) and \( x \in [\omega]^s \). We call \( s \) the stem of \( q(0) \) and write \( s = \text{stem} (q(0)) \). Let \( q \) be a \( \mathbb{P}_{\omega_2} \)-condition such that the stem of \( q(0) \) is empty, i.e., \( q(0) = (\emptyset, x) \) for some \( x \in [\omega]^\omega \). For every \( t \in \text{fin}(x) \) let
There may be no Ramsey ultrafilter and \( h = \epsilon \)

\( q(0) \colon \equiv (t, x \setminus \bar{t}^+), \text{ where } \bar{t}^+ = \max(t) + 1. \text{ Notice that } q(0)_t \text{ is an } M_0\text{-condition, stem } (q(0))_t = t, \text{ and } q(0) \leq q(0)_t. \)

Now, let us begin with the proof: Assume that for some \( \mathbb{P}_{\omega_2}\)-condition \( p \) we have

\[ \forall i \in \omega \left( f(i) < g(i) \right). \]

By induction on \( n \) we shall construct an infinite sequence \( \langle p_n : n \in \omega \rangle \) of \( \mathbb{P}_{\omega_2}\)-conditions such that \( p = p_0 \) and for every \( n \in \omega \) we have \( p_n \leq \mathcal{K}_n \quad p_{n+1}, \)

where the finite sets \( K_n \subseteq \omega_2 \) are such that \( 0 \in K_0, n < n' \rightarrow K_n \subseteq K_{n'}, \) and \( \bigcup_{n \in \omega} K_n = \bigcup_{n \in \omega} \supp(p_n) \) (the construction of the \( K_n \)'s with the required properties is left as an exercise to the reader).

For the sake of simplicity, let us assume that the stem of \( p(0) \) is empty (i.e., \( p = (0, x) \) for some \( x \in [\omega]^\omega \)), which implies that the stems of the \( p_n \)'s are empty, too. This way we even get infinite sets \( \mathcal{J}_0, \mathcal{J}_1 \subseteq \omega \) such that \( \mathcal{J}_0 \cap \mathcal{J}_1 = \emptyset. \) We leave it as an exercise to the reader to verify that the case when the stem of \( p(0) \) is non-empty yields infinite sets \( \mathcal{J}_0 \) and \( \mathcal{J}_1 \) such that the intersection \( \mathcal{J}_0 \cap \mathcal{J}_1 \) is still finite.

The goal is that for each \( n \in \omega \) and for each \( t = \{k_0 < \cdots < k_{n+1}\} \subseteq x_{n+1}, \) where \( p_{n+1}(0) = (\emptyset, x_{n+1}), \) we have

\[ p_{n+1}(0) t^\perp p_{n+1}|_{[1, \omega_2]} \vdash_{\mathbb{P}_{\omega_2}} g[\omega] \cap [k_n, k_{n+1}] \subseteq I_t, \]

where \( I_t \subseteq [k_n, k_{n+1}] \) is such that \( |I_t| \leq (n+1) \cdot (n+1)^{|K_n|}. \) The infinite sequence \( \langle p_n : n \in \omega \rangle \) is constructed as follows: Assume that we have already constructed \( p_n \) for some \( n \in \omega \) (recall that \( p_0 = p \)). So, \( p_n = (\emptyset, x_n) \) for some \( x_n \in [\omega]^\omega. \) Let \( t = \{k_0 < \cdots < k_{n+1}\} \subseteq x_{n+1} \) be an arbitrary but fixed subset of \( x_n \) of cardinality \( n+2 \) and let \( p_t := p_n(0)_t^\perp p_{n+1}|_{[1, \omega_2]} \). Then, for each \( i \leq n, \)

we obviously have

\[ p_t \vdash_{\mathbb{P}_{\omega_2}} g(i) \geq k_{n+1} \lor \bigvee_{i < k_{n+1}} g(i) = t. \]

Notice that since \( p_t \vdash_{\mathbb{P}_{\omega_2}} \forall i \leq n + 1 \land (f(i) = k_i), \) and since \( g \) is strictly increasing, \( p_t \vdash_{\mathbb{P}_{\omega_2}} \forall i > n \land (g(i) > k_{n+1}). \) Hence, by applying Lemma 25.8.(c) \( (n + 1)\)-times (for each \( i \leq n \)), we find a \( \mathbb{P}_{\omega_2}\)-condition \( q_i \) with \( p_t \leq \mathcal{K}_n \quad q_i, \) as well as a set \( I_t \subseteq [k_n, k_{n+1}], \) such that \( |I_t| \leq (n+1) \cdot (n+1)^{|K_n|} \) and

\[ q_i \vdash_{\mathbb{P}_{\omega_2}} g[\omega] \cap [k_n, k_{n+1}] \subseteq I_t. \]

Since \( t \) was arbitrary, for each \( t \in \text{fin}(x_n) \) of cardinality \( n+2 \) we find a \( q_t \) with \( p_t \leq \mathcal{K}_n \quad q_t \) such that \( q_t \vdash_{\mathbb{P}_{\omega_2}} g[\omega] \cap [k_n, k_{n+1}] \subseteq I_t, \) where \( I_t \) is as above. Moreover, by induction on \( \max(t) \) (similar to the proof of Claim below), we can construct a \( \mathbb{P}_{\omega_2}\)-condition \( p_{n+1} \) such that \( p_{n+1}(0) = (\emptyset, x_{n+1}) \) and \( p_n \leq \mathcal{K}_n \quad p_{n+1}, \) and for every finite set \( t = \{k_0 < \cdots < k_{n+1}\} \subseteq x_{n+1} \) of cardinality \( n+2 \) we have

\[ p_{n+1}(0)_t \vdash_{\mathcal{M}} p_{n+1}|_{[1, \omega_2]} = q_t|_{[1, \omega_2]} \]
and
\[ p_{n+1}(0) \models p_{n+1}|_{[1, \omega^2)} \models p_{\omega^2} g[\omega] \cap [k_n, k_{n+1}] \subseteq I_t . \]

Thus, \( p_{n+1} \) has the required properties, which completes the construction of the sequence \( \langle p_n : n \in \omega \rangle \).

By Lemma 25.7, let \( p_\omega \) be the fusion of the \( p_n \)'s. Since \( p \leq p_\omega \), the stem of \( p_\omega \) is empty, and therefore \( p_\omega = (\emptyset, z) \) for some \( z \in [\omega]^\omega \). By construction, for each \( t = \{ k_0 < \cdots < k_m \} \in \text{fin}(z) \), where \( m \in \omega \), we have
\[ p_\omega(0) \models p_\omega|_{[1, \omega^2)} \models p_{\omega^2} g[\omega] \cap [k_m, k_{m+1}] \subseteq I_t . \]

It remains to construct infinite sets \( \mathcal{I}_0, \mathcal{I}_1 \subseteq \omega \) in \( V \), where \( \mathcal{I}_0 \cap \mathcal{I}_1 \) is finite, and \( \mathcal{I}_2 \)-conditions \( \dot{\mathcal{I}}_0, \dot{\mathcal{I}}_1 \), where \( \dot{\mathcal{I}}_0 \geq p_\omega \leq \dot{\mathcal{I}}_1 \), such that \( \dot{\mathcal{I}}_0 \models \dot{p}_{\omega^2} g[\omega] \subseteq \mathcal{I}_0 \) and \( \dot{\mathcal{I}}_1 \models \dot{p}_{\omega^2} g[\omega] \subseteq \mathcal{I}_1 \). For this, we first prove the following

**Claim.** Let \( p_\omega(0) = (\emptyset, z) \) (for some \( z \in [\omega]^\omega \)), and for every \( x \in [\omega]^\omega \), let
\[ \mathcal{I}_x := \bigcup \{ I_t : t \in \text{fin}(x) \} \],
where \( I_t \) is as above. Then there are infinite sets \( \hat{x}, \hat{y} \in [z]^\omega \) such that \( \mathcal{I}_{\hat{x}} \cap \mathcal{I}_{\hat{y}} \) is finite. Moreover, since we assumed that \( \text{stem}(p_\omega(0)) = \emptyset \), we even get \( \mathcal{I}_{\hat{x}} \cap \mathcal{I}_{\hat{y}} = \emptyset \).

**Proof of Claim.** By construction, for every \( t = \{ k_0 < \cdots < k_{n+1} \} \in \text{fin}(z) \), \( I_t \subseteq [k_n, k_{n+1}] \) and \( |I_t| \leq (n+1)^{\text{fin}(z)} \). Notice that the size of \( I_t \) depends on \(|t|\) but not on the particular set \( t \). For every \( n \in \omega \), let \( F(n) := (n+1)^{\text{fin}(z)} \). Then, for every non-empty \( t \in \text{fin}(z) \) we have \( |I_t| \leq F(|t|) \) (notice that for every \( k_0 \in z \), \( I_{\{k_0\}} = \emptyset \)). For each non-empty set \( s = \{ k_0 < \cdots < k_n \} \in \text{fin}(z) \) let
\[ \text{succ}_s(z) = \{ t \in \text{fin}(z) : t = \{ k_0 < \cdots < k_n < k_{n+1} \} \} , \]
\( i.e., t \in \text{succ}_s(z) \) iff \( t = s \cup \{ k_{n+1} \} \) for some \( k_{n+1} \in z \) with \( k_{n+1} > k_n \).

Then, for each non-empty set \( s = \{ k_0 < \cdots < k_n \} \in \text{fin}(z) \) we get that
\[ \delta_s = \{ I_t : t \in \text{succ}_s(z) \} \]
is an infinite set of finite subsets of \([k_n, \omega]\), where the cardinality of the finite sets \( I_t \in \delta_s \) is bounded by \( F(|s|+1) \). By similar arguments as in the proof of the \( \Delta \)-System Lemma 13.2, for each non-empty set \( s = \{ k_0 < \cdots < k_n \} \in \text{fin}(z) \) we can construct an infinite set \( z' \in [z]^\omega \) and a finite set \( \Delta_s \subseteq [k_n, \omega] \), such that for any distinct \( t, t' \in \text{succ}_s(z) \) we have \( I_t \cap I_{t'} \subseteq \Delta_s \). In other words, for any distinct \( t, t' \in \text{succ}_s(z) \), \( I_t \setminus \Delta_s \) and \( I_{t'} \setminus \Delta_s \) are disjoint. Moreover, we can construct an infinite set \( z_0 \in [z]^\omega \), and for every non-empty \( s = \{ k_0 < \cdots < k_n \} \in \text{fin}(z_0) \) a finite set \( \Delta_s \subseteq [k_n, \omega] \), such that for any distinct \( t, t' \in \text{succ}_{z_0}(s) \) we have
\[ I_t \cap I_{t'} \subseteq \Delta_s . \]

Now, we are ready to construct the sets \( \hat{x} \) and \( \hat{y} \) in \([z]^\omega \) with the required properties: Firstly, let \( x_0 \) and \( y_0 \) be two disjoint infinite subsets of \( z_0 \). Let \( k_0 = \min(x_0) \) and let \( l_0 \in y_0 \) be such that \( l_0 > \max(\Delta_{\{k_0\}}) \). By \( \Delta \) we find sets \( x_1 \subseteq [x_0]^\omega \) and \( y_1 \subseteq [y_0]^\omega \) such that for all \( t \in \text{succ}_{x_1}(\{k_0\}) \) and all \( t' \in \text{succ}_{y_1}(\{l_0\}) \), \( I_t \cap I_{t'} = \emptyset \). Now, choose \( k_1 \in x_1 \) such that \( k_1 > k_0 \), and \( l_1 \in y_1 \) such that \( l_1 > \max\{ \max(\Delta_{\{k_1\}}), \max(\Delta_{\{k_0, k_1\}}) \} \). Again by \( \Delta \) we find sets
There may be no Ramsey ultrafilter and \( \mathfrak{h} = \mathfrak{c} \)

\[ x_2 \in [x_1]^{\omega} \text{ and } y_2 \in [y_1]^{\omega} \text{ such that for all } t \in \text{succ}_{x_2}({\{ k_1 \}}) \cup \text{succ}_{x_2}({\{ k_0, k_1 \}}) \text{ and all } t' \in \text{succ}_{y_2}({\{ l_1 \}}) \cup \text{succ}_{y_2}({\{ l_0, l_1 \}}), I_t \cap I_{t'} = \emptyset. \] 

Proceeding this way, we finally get \( \hat{x}, \hat{y} \in [z_0]^{\omega} \) such that for all \( t \in \text{fin}(\hat{x}) \) and all \( t' \in \text{fin}(\hat{y}) \) we have \( I_t \cap I_{t'} = \emptyset \), and hence, \( \mathcal{F}_x \cap \mathcal{F}_y = \emptyset \).

Now, let \( \hat{p}_0 := (\emptyset, \hat{x}) \upharpoonright [z_{12}]_\omega \) and \( \hat{p}_1 := (\emptyset, \hat{y}) \upharpoonright [z_{12}]_\omega \). Then \( \hat{p}_0 \geq p \geq \hat{p}_1 \), and by construction of \( \hat{x} \) and \( \hat{y} \) we have

\[
\hat{p}_0 \not\vdash_{P_{\omega_2}} g[\omega] \quad \text{ and } \quad \hat{p}_1 \not\vdash_{P_{\omega_2}} g[\omega] \subseteq \mathcal{F}_y,
\]

where \( \mathcal{F}_x \cap \mathcal{F}_y = \emptyset \), which completes the proof.

Before we show that every Ramsey ultrafilter is rapid, let us briefly recall the notion of rapid filters (given in Chapter 10 | Related Result 70), as well as the notion of \( Q \)-points (also given in Chapter 10):

A free filter \( \mathcal{F} \subseteq [\omega]^{\omega} \) is called a rapid filter if for each \( f \in \mathcal{F} \) there exists an \( x \in \mathcal{F} \) such that for all \( n \in \omega \), \( |x \cap f(n)| \leq n \). Furthermore, a free ultrafilter \( \mathcal{U} \subseteq [\omega]^{\omega} \) is a \( Q \)-point if for each partition of the \( \omega \) into finite pieces \( \{ I_n : n \in \omega \} \). (i.e., for each \( n \in \omega \), \( I_n \) is finite), there is an \( x \in \mathcal{U} \) such that for each \( n \in \omega \), \( |x \cap I_n| \leq 1 \). The following fact is just a consequence of these definitions.

**FACT 25.10.** Every \( Q \)-point is a rapid filter.

**Proof.** Let \( \mathcal{U} \subseteq [\omega]^{\omega} \) be a \( Q \)-point and let \( f \in \mathcal{U} \) be any strictly increasing function. Let \( I_0 := [0, f(0)) \), and for \( n \in \omega \) let \( I_{n+1} := [f(n), f(n+1)) \). Then \( \{ I_n : n \in \omega \} \) is obviously a partition of \( \omega \) into finite pieces. Since \( \mathcal{U} \) is a \( Q \)-point (in particular a free ultrafilter), there is an \( x \in \mathcal{U} \) such that \( x \cap f(0) = \emptyset \) and for each \( n \in \omega \), \( |x \cap I_n| \leq 1 \), i.e., for all \( n \in \omega \), \( |x \cap f(n)| \leq n \).

Thus, \( \mathcal{U} \) is a rapid filter.

By **FACT 10.10** we know that every Ramsey ultrafilter is a \( Q \)-point, and therefore, every Ramsey ultrafilter is rapid.

Now, we are ready to prove the main result of this section.

**PROPOSITION 25.11.** It is consistent with \( \text{ZFC + } \mathfrak{h} = \mathfrak{c} \) that there are no rapid filters. In particular, since every Ramsey ultrafilter is rapid, it is consistent with \( \text{ZFC + } \mathfrak{h} = \mathfrak{c} \) that there are no Ramsey ultrafilters.

**Proof.** Since \( \mathfrak{h} = \mathfrak{c} \) in Mathias’ model, it is obviously enough to prove that there are no rapid filters in Mathias’ model. So, let \( P_{\omega_1} = \langle Q, \gamma \in \omega_2 \rangle \) be the countable support iteration of length \( \omega_2 \) of Mathias forcing \( M \), starting in a model \( V \) of \( \text{ZFC + CH} \). Furthermore, let \( \mathcal{F} \) be a \( P_{\omega_1} \)-name for a filter in the \( P_{\omega_1} \)-generic extension of \( V \) (i.e., \( 0_{\omega_2} \not\vdash_{P_{\omega_2}} \mathcal{F} \text{ is a filter} \)) and let \( G \) be \( P_{\omega_1} \)-generic over \( V \). Then, similar to **CLAIM 2** in the proof of **PROPOSITION 24.12**, there is an \( \alpha < \omega_2 \) such that \( \mathcal{F}[G] \cap V[G[\alpha]] \in V[G[\alpha]] \). Let us work in the model \( V[G[\alpha]] \), i.e., we consider \( V[G[\alpha]] \) as the ground
model: In $\mathcal{V}[G\omega]$, let $f$ be an $M$-name in $\mathcal{V}[G\omega]$ for the next Mathias real, i.e., $f$ is the $M$-name for a strictly increasing function in $\omega$ such that

$$0_{\omega_2} \mathrel{\models} \mathcal{P}_{\omega_2} \{ f(n) : n \in \omega \} = \bigcup \{ s : \exists x \in [\omega]^\omega \left( (s, x) \in G(\alpha) \right) \}.$$  

Assume towards a contradiction that $\mathcal{F}$ is rapid. Then there is a $\mathcal{P}_{\omega_2}$-name $g$ for a strictly increasing function in $\omega$ and a $\mathcal{P}_{\omega_2}$-condition $p$, such that

$$p \mathrel{\models} \mathcal{F} \quad \forall n \in \omega (g(n) > f(n)) \land g[\omega] \in \mathcal{F}. \quad (\ast)$$

By Lemma 25.9 (with respect to the ground model $\mathcal{V}[G\omega]$), there are $\mathcal{P}_{\omega_2}$-conditions $p_0$ and $p_1$ with $p_0 \geq p \leq p_1$, and almost disjoint sets $\mathcal{I}_0, \mathcal{I}_1 \in [\omega]^\omega$ in $\mathcal{V}[G\omega]$, such that

$$p_0 \mathrel{\models} \mathcal{P}_{\omega_2} g[\omega] \subseteq \mathcal{I}_0 \quad \text{and} \quad p_1 \mathrel{\models} \mathcal{P}_{\omega_2} g[\omega] \subseteq \mathcal{I}_1.$$  

In particular, if $p_0 \mathrel{\models} \mathcal{P}_{\omega_2} g[\omega] \in \mathcal{V}[G\omega]$, then $p_1 \mathrel{\models} \mathcal{P}_{\omega_2} g[\omega] \notin \mathcal{F}[G\omega]$, and vice versa. Hence, $p \mathrel{\models} \mathcal{P}_{\omega_2} g[\omega] \in \mathcal{F}[G\omega]$, which is a contradiction to $(\ast)$. Thus, since $\mathcal{F}$ was arbitrary, there are no rapid filters in $\mathcal{V}[G]$. 

\textit{Notes}

Using results of Laver’s ([7, Lemmata 5 & 6]), Miller [8] showed that there are no rapid filters in Laver’s model (cf. Related Result 146). In the proof that there are no rapid filters in Mathias’ model given above, we essentially followed Miller’s proof by translating the corresponding results of Laver’s to iterations of Mathias forcing.

\textbf{Related Results}

144. Ultrafilter forcing $\mathcal{V}$ collapses $\kappa$ to $\mathcal{V}$. By Lemma 25.1 we already know that ultrafilter forcing $\mathcal{V}$ does not collapse $\kappa$ to any cardinal $\kappa < \mathcal{V}$, i.e., if $G$ is $\mathcal{V}$-generic over $\mathcal{V}$, then $\mathcal{V}[G] \models \kappa \geq \mathcal{V}$. Thus, in order to show that $\mathcal{V}[G] \models \kappa \leq \mathcal{V}$, it is enough to show that $\mathcal{V}[G] \models \kappa \leq \mathcal{V}$. In particular, it is enough to show that there is a surjection in $\mathcal{V}[G]$ which maps $\mathcal{V}$ onto $\kappa$. Let us work in the model $\mathcal{V}$. By the Base Matrix Lemma 2.11 of Bakar, Pelant, and Simon [1] (see Chapter 8 | Related Result 51), there exists a shattering family $\mathcal{H}_0 = \{ \mathcal{A}_x \subseteq [\omega]^\omega : x \in \mathcal{V} \}$ which has the property that for each $x \in [\omega]^\omega$ there is a $\xi \in \mathcal{V}$ and $A \in \mathcal{A}_x$ such that $A \subseteq x$. Now, for each $A \in [\mathcal{V}]^\omega$ let $\mathcal{A}_A \subseteq [A]^\omega$ be an almost disjoint family of cardinality $\kappa$ and let $h_A : \mathcal{A}_A \to A$ be a surjection. Furthermore, we define the $\mathcal{V}$-name $f$ for a function from some subset of $\mathcal{V}$ to $\kappa$ by stipulating

$$f = \left\{ \langle \xi, \gamma, [x] \rangle : \xi \in \mathcal{V} \land \gamma \in \kappa \land \exists A \in \mathcal{A}_x (x \in \mathcal{A}_A \land h_A(x) = \gamma) \right\}.$$  

In particular, if $\langle \xi, \gamma, [x] \rangle \in f$, then

$$[x] \mathrel{\models} f(\xi) = \gamma.$$
By the properties of $\mathcal{M}_0$, for every $y \in [\omega]^{\omega}$ there is a $\xi \in \mathfrak{h}$ and an $A \in \mathcal{M}_0$ such that $A \subseteq y$. Thus, there exists an $x \in \mathcal{E}_A$ (in particular, $x \subseteq^* y$), such that $h_A(x) = \gamma$. In other words, for every $y \in [\omega]^{\omega}$ and each $\gamma \in c$, there are $x \subseteq^* y$ and $\xi \in \mathfrak{h}$ such that $[x]^\mathfrak{h} \models f(\xi) = \gamma$. Hence,

$$D_\gamma = \{[x]^\mathfrak{h} : [x]^\mathfrak{h} \models \exists \xi \in c (f(\xi) = \gamma)\}$$

is an open dense subset of $[\omega]^{\omega}/\text{fin}$, and therefore, $f[G]$ is a surjection from some subset of $\mathfrak{h}$ onto $c$, which shows that $V[G] \models \mathfrak{c} \leq \mathfrak{h}^V$.

145. A model in which there are no Ramsey ultrafilters. The first model in which there are no Ramsey ultrafilters was constructed by Kunen [6] using measure algebras (see also Jech [4, Theorem 9.4]).

146. There are no rapid ultrafilters in Laver’s model. Miller [8] showed that there are no rapid ultrafilters in Laver’s model (i.e., the model we get after a countable support iteration of length $\omega_2$ of Laver forcing starting in a model of ZFC + CH). However, like in Mathias’ model, there are still $P$-points in Laver’s model (see Roitman [10]).

147. There are no $Q$-points in Miller’s model. According to Miller [9, p.156], there are no $Q$-points in Miller’s model (i.e., the model we get after a countable support iteration of length $\omega_2$ of Miller forcing starting in a model of ZFC + CH). On the other hand, since Miller forcing preserves $P$-points (by Lemma 23.5), there are still $P$-points in Miller’s model. Further notice that in Miller’s model we have $\mathfrak{d} = \mathfrak{c}$ (cf. Theorem 10.16).

148. Models without rapid ultrafilters and large continuum. We have seen that there exists a model of ZFC in which there are no rapid ultrafilters and $\mathfrak{c} = \omega_2$. It is natural to ask whether the continuum can be further increased without adding rapid ultrafilters; this is indeed the case: For any cardinal $\kappa$ there exists a model of ZFC in which there are no rapid ultrafilters and $\mathfrak{c} \geq \kappa$ (see Judah and Shelah [5, Theorem 2.6], or Bartoszynski and Judah [2, Theorem 4.6.7]).

149. Borel’s conjecture and the existence of Ramsey ultrafilters. Judah [3] showed that Borel’s conjecture holds in the model constructed in the proof of Proposition 25.2 (see Bartoszynski and Judah [2, Theorem 8.3.14]). Thus, Borel’s conjecture does not contradict the existence of a Ramsey ultrafilter (compare with Chapter 23 [Related Result 131 and Related Result 146]).

References


Combinatorial Properties of Sets of Partitions

In this chapter we shall investigate combinatorial properties of sets of partitions of $\omega$, where we try to combine as many topics or voices (to use a musical term) as possible. As explained in Chapter 11, partitions of $\omega$ are to some extent the dual form of subsets of $\omega$. Thus, we shall use the term “dual” to denote the partition forms of Mathias forcing, of Ramsey ultrafilters, of cardinal characteristics, et cetera. Firstly, we shall investigate combinatorial properties of a dual form of unrestricted Mathias forcing (which was introduced in Chapter 24). In particular, by using the Partition Ramsey Theorem 11.4, which is a dual form of Ramsey’s Theorem 2.1 (and which was the main result of Chapter 11), we shall prove that dual Mathias forcing has pure decision. Secondly, we shall dualise the shattering number $\mathcal{h}$ (introduced in Chapter 8 and further investigated in Chapter 9), and show how it can be increased by iterating dual Mathias forcing (cf. Proposition 24.12). Finally, we shall dualise the notion of Ramsey ultrafilters (introduced and investigated in Chapter 10), and show — using the methods developed in Part II and the previous chapter — that the existence of these dual Ramsey ultrafilters is consistent with ZFC + CH as well as with ZFC + ¬CH.

A Dual Form of Mathias Forcing

Firstly, let us recall some terminology — for more detailed definitions see Chapter 11: The set of all infinite partitions of $\omega$ is denoted by $(\omega)^\omega$, and $(\mathbb{N})$ denotes the set of all (finite) partitions of natural numbers. For $P \in (\mathbb{N})$ or $P \in (\omega)^\omega$, let $\text{Min}(P) := \{ \text{min}(p) : p \in P \}$ and $\text{Dom}(P) := \bigcup P$. For partitions $P$ and $Q$ (e.g., $P \in (\mathbb{N})$ and $Q \in (\omega)^\omega$) we write $P \sqsubseteq Q$ if $Q$ restricted to $\text{Dom}(P)$ is finer than $P$. Furthermore, for partitions $P$ and $Q$, let $P \cap Q$ ($P \cup Q$) denote the finest (coarsest) partition $R$ such that $\text{Dom}(R) = \text{Dom}(P) \cup \text{Dom}(Q)$ and $R$ is coarser (finer) than $P$ and $Q$. Let $S \in (\mathbb{N})$ and $X \in (\omega)^\omega$. If for each $s \in S$ there exists an $x \in X$ such that $x \cap \text{Dom}(S) = s$, then we write $S \leq X$. Similarly, for $S, T \in (\mathbb{N})$, where
Dom(S) ⊆ Dom(T), we write S ≼ T if for each s ∈ S there exists a t ∈ T
such that t ∩ Dom(S) = s. Finally, for S ∈ (N) and X ∈ (ω)ω with S ⊆ X, let
\[(S, X)^ω = \{ Y ∈ (ω)^ω : S ⊆ Y ⊆ X \} .\]
A set (S, X)ω, where S and X are as above, is called a dual Ellentuck
neighbourhood.
Now, we are ready to define a dual form of Mathias forcing (i.e., a form
of Mathias forcing in terms of partitions): Similar to Mathias forcing \( \mathcal{M} \)
introduced in Chapter 24, we define dual Mathias forcing \( \mathcal{M}^* = (M^*, \leq) \)
by stipulating:
\[M^* = \{ (S, X) : S ∈ (N) ∧ X ∈ (ω)^ω ∧ S \subseteq X \}\]
\[(S, X) ≤ (T, Y) ⇔ (T, Y)^ω ≤ (S, X)^ω\]
Notice that \((S, X) ≤ (T, Y) ⇔ S \subseteq T \land Y \subseteq X\). Thus, we get dual Mathias
forcing from Mathias forcing by replacing subsets of \( ω \) with partitions of \( ω \).
However, as we shall see below, dual Mathias forcing is much stronger than
Mathias forcing (see also Related Result 151), but first, let us show that
dual Mathias forcing is at least as strong as Mathias forcing:

**FACT 26.1.** Dual Mathias forcing adds Mathias reals and consequently it also
adds dominating reals.

**Proof.** Firstly, let \( M_0 \) be the set of all \( \mathcal{M} \)-conditions \((s, x)\) for which we have
0 ∈ s, or, in case s = ∅, 0 ∈ x, and let \( M_0 = (M_0, \leq) \). Obviously, the forcing
notions \( M_0 \) and \( \mathcal{M} \) are equivalent. Secondly, define the function \( h : M^* → M_0 \)
by stipulating
\[h : M^* → M_0\]
\[(S, X) → (\text{Min}(S), \text{Min}(X) \setminus \text{Min}(S)) .\]
Then, the function \( h \) satisfies the following conditions:
- for all \( q_0, q_1 ∈ M^* \), if \( q_0 ≤_{M^*} q_1 \) then \( h(q_0) ≤_M h(q_1) \),
- for all \( q ∈ M^* \) and each \( p ∈ M_0 \) with \( h(q) ≤_M p \), there is a \( q' ∈ M^* \) with
\( q ≤_{M^*} q' \) such that \( p ≤_M h(q') \).
We leave it as an exercise to the reader to verify that this implies that whenever
\( G^* \) is \( M^* \)-generic, then \( \{ (\text{Min}(S), \text{Min}(X) \setminus \text{Min}(S)) : (S, X) ∈ G^* \} \)
is \( M_0 \)-generic. Thus, dual Mathias forcing \( M^* \) adds Mathias reals, and since
Mathias reals are dominating, it also adds dominating reals.

One of the main features of Mathias forcing is that it has pure decision.
This is also the case for dual Mathias forcing and the proof is essentially
the same as the proof for the corresponding result for Mathias forcing.
However, at a crucial point we have to use the Partition Ramsey Theorem 11.4 —
a dual form of Ramsey’s Theorem 2.1 — which will serve as a kind of Pigeon-
Hole Principle.
Theorem 26.2. Let \( (S_0, X_0) \) be an \( M^* \)-condition and let \( \varphi \) be a sentence of the forcing language. Then there exists an \( M^* \)-condition \( (S_0, Y_0) \geq (S_0, X_0) \) such that either \( (S_0, Y_0) \vdash M^* \varphi \) or \( (S_0, Y_0) \vdash M^* \neg \varphi \)(i.e., \( (S_0, Y_0) \) decides \( \varphi \)).

Proof. We follow the proof of Theorem 24.3: For any set \( \mathcal{O} \subseteq M^* \) which is open with respect to the dual Ellentuck topology, let

\[
\bar{\mathcal{O}} := \bigcup \left\{ (S, X)^\omega : (S, X) \in \mathcal{O} \right\}.
\]

With respect to a fixed open set \( \mathcal{O} \subseteq M^* \), we call the condition \( (S, X) \) good if there is a \( Y \in (S, Y)^\omega \) such that \( (S, X)^\omega \subseteq \bar{\mathcal{O}} \); otherwise, we call it bad. Furthermore, we call \( (S, X) \) ugly if \( (T^*, X) \) is bad for all \( S \preceq T^* \subseteq X \) with \( |T| = |S| \), where \( T^* := T \cup \{ \text{Dom}(T) \} \).

Claim 1. If the condition \( (S, X) \) is bad, then there is a \( Y \in (S, X)^\omega \) such that \( (S, Y) \) is ugly.

Proof of Claim 1. We follow the proof of Lemma 24.4: Let \( Z_0 := X \) and let \( T_0 := S \). Assume we have already defined \( Z_{n-1} \in (\omega)^\omega \) and \( T_{n-1} \in (\mathbb{N}) \) for some positive integer \( n \). Let \( T_n \) be such that \( S \preceq T_n \), \( |T_n| = |S| + n \), and \( T^*_n \leq Z_{n-1} \). Let \( \{U_i : i \leq m\} \) be an enumeration of all \( T \) such that \( S \preceq T \subseteq T_n \), \( |T| = |S| \) and \( \text{Dom}(T) = \text{Dom}(T_n) \). Further, let \( Z \vdash T_n := Z_{n-1} \).

Now, choose for each \( i \leq m \) a partition \( Z_i \in (\omega)^\omega \) such that \( Z_i \leq Z_{n-1} \), \( T^*_n \leq Z_i \) and either \( (U^*_i, Z_i)^\omega \) is bad or \( (U^*_i, Z_i)^\omega \) is bad, and let \( Z_{n+1} := Z^m \).

Finally, let \( Z \in (\omega)^\omega \) be the only partition such that for all \( n \in \omega \), \( T_n \vartriangleleft Z \).

By construction of \( Z \), for all \( T \in (S, Z)^{(\omega|\mathbb{N})} \), we have either \( (T^*, Z)^\omega \subseteq \bar{\mathcal{O}} \) or \( (T^*, Z)^\omega \) is bad. Now, for \( n = |S| \), define the sets \( C_0 := \{ T \in (S, Z)^{\omega|\mathbb{N}} : (T^*, Z)^\omega \) is bad \} and \( C_1 := \{ T \in (S, Z)^{(\omega|\mathbb{N})} : (T^*, Z)^\omega \subseteq \bar{\mathcal{O}} \} \). Then, by the properties of \( Z \), \( C_0 \cup C_1 = (S, Z)^{(\omega|\mathbb{N})} \). Hence, by the Partition Ramsey Theorem 11.4, there exists a \( Y \in (S, Z) \) such that either \( (S, Y)^{(\omega|\mathbb{N})} \subseteq C_0 \) or \( (S, Y)^{(\omega|\mathbb{N})} \subseteq C_1 \). Thus, since \( (S, X) \) is bad, \( (S, Y) \) is ugly.

Moreover, by a similar construction as in the proof of Lemma 24.5 we can prove the following

Claim 2. If the condition \( (S, X) \) is bad, then there is a \( Y \in (S, X)^\omega \) such that \( (S, Y)^\omega \cap \bar{\mathcal{O}} = \emptyset \).

Proof of Claim 2. By Claim 1, there is a \( Z_0 \in (S, X)^\omega \) such that \( (S, Z_0) \) is ugly; i.e., for all \( T \in (\mathbb{N}) \) with \( S \preceq T^* \subseteq Z_0 \) and \( |T| = |S| \), \( (T^*, Z_0) \) is bad. Let \( T_0 \in (\mathbb{N}) \) be such that \( T_0 \not\leq Z_0 \) and \( |T_0| = |S| \). Then, since \( (S, Z_0) \) is ugly, \( (T_0, Z_0) \) is bad. Assume that for some \( n \in \omega \) we have already constructed \( (T_n, Z_n) \preceq (T_0, Z_0) \) with \( T^*_n \leq Z_n \) and \( |T_n| = |S| + n \); such that for all \( T \in (\mathbb{N}) \) with \( T \preceq T^*_n \) and \( \text{Dom}(T) = \text{Dom}(T_n) \) and \( (T^*, T \cap Z_n) \) is bad or \( (T, Z_n)^\omega \subseteq \bar{\mathcal{O}} \). Let \( T_{n+1} \) be such that \( T_n \preceq T^*_{n+1} \) and \( T_{n+1} \preceq Z_{n+1} \).
and \(|T_{n+1}| = |T_n| + 1\). By applying Claim 1 to every \(T \in \langle N \rangle\) with \(T_0 \equiv T \subseteq T_{n+1}\) and \(\text{Dom}(T) = \text{Dom}(T_{n+1})\), we find a \(Z_{n+1} \in (T^*_n, Z_n)^\circ\) such that for all \(T \in \langle N \rangle\) with \(T_0 \equiv T \subseteq T_{n+1}\) and \(\text{Dom}(T) = \text{Dom}(T_{n+1})\), we have either \((T^*_n, T \cap Z_{n+1})^\circ \subseteq \mathcal{O}\). Let \(Y = \bigcup_{n \in \omega} T_n\), i.e., \(Y\) is the only (infinite) partition such that for all \(n \in \omega\), \(T_n \leq Y\).

Assume towards a contradiction that \((S, Y)^\circ \cap \mathcal{O} \neq \emptyset\). Then there are \(T \in \langle N \rangle\) with \(S \equiv T \subseteq Y\) such that \((T, Y)^\circ \subseteq \mathcal{O}\), i.e., \((T, T \cap Y)^\circ\) is good. Choose \(T_0\) (with \(S \equiv T_0 \subseteq Y\)) of least cardinality such that \((T_0, T_0 \cap Y)^\circ\) is good. Since \((S, Y)^\circ\) is ugly, \(|T_0| > |S|\). Hence, we find a \(T_1 \subseteq Y\) with \(S \equiv T_1^* \equiv T_0\) and \(|T_1| = |T_0| - 1\). By construction of \(Y\), \((T_1, T_1 \cap Y)^\circ\) is either ugly or good. In the former case, \((T_0, T_0 \cap Y)^\circ\) would be bad (a contradiction to the choice of \(T_0\)), and in the latter case, \(T_0\) would not be of least cardinality (again a contradiction to the choice of \(T_0\)). Thus, \((S, Y)^\circ \cap \mathcal{O} = \emptyset\), which completes the proof.

Now, let \(\varphi\) be a sentence of the forcing language. With respect to \(\varphi\) we define \(O_1 := \{q \in M^*: q \forces_{M^*} \varphi\}\) and \(O_2 := \{q \in M^*: q \forces_{M^*} \neg \varphi\}\). Notice that \(O_1 \cup O_2\) is an open dense subset of \(M^*\). If the \(M^*\)-condition \((S_0, X_0)^\circ\) is good with respect to \(O_1\), there is a \(Y_0 \in (S_0, X_0)^\circ\) such that \((S_0, Y_0)^\circ \subseteq O_1\). Otherwise, if \((S_0, X_0)^\circ\) is bad with respect to \(O_1\), by Claim 2 there is a \(Y_0 \in (S_0, X_0)^\circ\) such that \((S_0, Y_0)^\circ \cap O_1 = \emptyset\). In the former case, we have \((S_0, Y_0)^\circ \cap O_1 = \emptyset\), and we are done. In the latter case, we proceed as follows: Since \((S_0, Y_0)^\circ \cap O_1 = \emptyset\) and \(O_1 \cap O_2\) is dense, for every \((S_0, Z_0) \geq (S_0, Y_0)\), there exists a \((T, Z) \geq (S_0, Z_0)\) such that \((T, Z) \in O_2\). This implies that \((S_0, Y_0)^\circ\) cannot be bad with respect to \(O_2\) since otherwise, by Claim 2, we would find an \((S_0, Z_0) \geq (S_0, Y_0)\) such that \((S_0, Z_0)^\circ \cap (O_1 \cup O_2) = \emptyset\). Thus, \((S_0, Y_0)^\circ\) is good with respect to \(O_2\) and we find \((S_0, Y_0)^\circ\) \(\equiv (S_0, Y_0)^\circ\) \(\subseteq O_2\), i.e., \((S_0, Y_0)^\circ \forces_{M^*} \neg \varphi\).

Now, having Theorem 25.2 at hand, it is not hard to show that dual Mathias forcing is proper and has the Laver property: Firstly, notice that to each \(G \subseteq M^*\) which is \(M^*\)-generic over some model \(V\) there exists a unique infinite partition \(X_G \in (\omega)^\omega\) with the property that for all \(S \in \langle N \rangle\),

\[ S \prec X_G \iff \exists Y \in (\omega)^\omega ((S, Y) \in G). \]

Thus, every \(M^*\)-generic set \(G \subseteq M^*\) corresponds to a unique \(M^*\)-generic partition \(X_G \in (\omega)^\omega\), which we call Mathias partition. Following the proof of Corollary 24.6 we can show that if \(X_G\) is a Mathias partition over \(V\) and \(Y \subseteq X_G\) is an infinite partition, then \(Y\) is a Mathias partition over \(V\), too. Furthermore, by similar arguments as in the proofs of Corollaries 24.7 & 24.8, one can show that dual Mathias forcing is proper and has the Laver property, in particular, dual Mathias forcing does not add Cohen reals (the details are left as an exercise to the reader).

A feature of Mathias forcing is that it can be written as a two-step iteration. More precisely, \(\mathbb{M} \simeq \mathbb{U} * M\), where \(\mathbb{U}\) is the canonical \(\mathbb{U}\)-name for
the $U$-generic ultrafilter (see Lemma~24.10). Before we can prove the corresponding result with respect to dual Mathias forcing, we have to introduce a dual form of $U$ and have to define restricted dual Mathias forcing: Firstly, for $X, Y \in (\omega)^\omega$ let $Y \subseteq^* X \iff \exists F \in \text{fin}(\omega) (Y \cap \{F\} \subseteq X)$; notice that $\{F\}$ is a one-block partition with domain $F$. Now, let $U^* = ((\omega)^\omega, \leq)$, where

$$X \leq Y \iff Y \subseteq^* X.$$ 

Strictly speaking, $((\omega)^\omega, \leq)$ is not a partially ordered set since “$\leq$” is not anti-symmetric (i.e., $X \leq Y$ and $Y \leq X$ does not imply $X = Y$). However, it is slightly easier to drop anti-symmetry than to work with equivalence classes.

Furthermore, for any family of infinite partitions $\mathcal{F} \subseteq (\omega)^\omega$, let $M_{\mathcal{F}}^* = (M_{\mathcal{F}}, \leq)$, where $M_{\mathcal{F}}$ is the set of all $M^*$-conditions $(S, X)$ such that $X \in \mathcal{F}$. Now, the dual form of Lemma~24.10 reads as follows.

**Lemma 26.3.** $M_{\mathcal{F}}^* \approx U^* \ast M_{\mathcal{F}}^*$, where $\mathcal{F}^*$ is the canonical $U^*$-name for the $U^*$-generic filter.

Before we prove Lemma 26.3, we first show that the forcing notion $U^*$ is $\sigma$-closed and that it adds Ramsey ultrafilters.

**Lemma 26.4.** The forcing notion $U^*$ is $\sigma$-closed, and whenever $\mathcal{F}^*$ is $U^*$-generic over $V$, then there is a Ramsey ultrafilter in $V[\mathcal{F}^*]$.

**Proof.** $U^*$ is $\sigma$-closed: Let $X_0 \leq X_1 \leq \cdots$ be an increasing sequence of infinite partitions (i.e., for all $i \in \omega$, $X_{i+1} \subseteq^* X_i$). Choose a sequence $\langle F_i : i \in \omega \rangle$ of finite sets of natural numbers such that for all $i \in \omega$, $X_{i+1} \cap \{F_i\} \subseteq X_i$. For every $X \in (\omega)^\omega$, order the blocks of $X$ by their least element, and for $k \in \omega$, let $X(k)$ denote the $k$th block with respect to this ordering. Define $y_0 := X_0(0)$, and for positive integers $n$, let $y_n := X_n(k)$, where $k := n + \bigcup_{i \in \omega} (\bigcup F_i)$. Now, let $Y := \{y_i : i \in \omega\} \cup (\omega \backslash \bigcup_{i \in \omega} y_i)$. Then, for each $i \in \omega$ we have $Y \subseteq^* X_i$, which shows that $U^*$ is $\sigma$-closed.

$U^*$ adds Ramsey ultrafilters: We show that the set $\{ \text{Min}(X) \setminus \{0\} : X \in \mathcal{F}^* \}$ is a Ramsey ultrafilter over $\omega \setminus \{0\}$: Firstly, recall that a forcing notion which is $\sigma$-closed does not add new reals to the ground model (see Lemma~14.17). Let $\pi : [\omega]^2 \rightarrow 2$ be an arbitrary colouring and let $Y \in (\omega)^\omega$. Then, by Ramsey’s Theorem 2.1, there exists an infinite set $x \subseteq \text{Min}(Y)$ with $0 \notin x$ such that $\pi$ is constant on $[x]^2$. Now, let

$$\ X = \{b : b \in Y \land \min(b) \in x\} \cup \bigcup\ \{b : b \in Y \land \min(b) \notin x\}. \$$

Then $X \subseteq Y$, $X \in (\omega)^\omega$, and $\text{Min}(X) \setminus \{0\} = x$. Consequently we get that

$$D_\pi := \{X \in (\omega)^\omega : \pi|_{\text{Min}(X) \setminus \{0\}} \text{ is constant}\}$$

is open dense, which implies that $D_\pi \cap \mathcal{F}^* \neq \emptyset$. Finally, since the colouring $\pi$ was arbitrary, this shows that $\{ \text{Min}(X) \setminus \{0\} : X \in \mathcal{F}^* \}$ is a Ramsey ultrafilter over $\omega \setminus \{0\}$. 


As a consequence we get the following

**Fact 26.5.** Forcing with \( U^* \) does not add new partitions to the ground model.

**Proof.** First, notice that partitions \( X \) can be encoded by real numbers \( r_X \subseteq \omega \), for example let

\[
  r_X = \{ k \in \omega : \exists n, m \in \omega (k = n \land \exists \exists ((n, m) \subseteq X(l))) \},
\]

where \( \eta \) is a bijection between \( \omega \times \omega \) and \( \omega \), and \( X(l) \) is as above.

Now, by **Lemma 14.17** we know that \( \sigma \)-closed forcing notions do not add new reals — and therefore no new partitions — to the ground model. \( \dashv \)

Now we are ready to give the

**Proof of Lemma 26.3.** Since \( U^* \) does not add new partitions, for every \( U^* \)-name \( (T, Y) \) for an \( M_{\text{gr}}^* \)-condition, and for every partition \( Z \in (\omega)^\omega \), there is an \( M^* \)-condition \( (S, X) \) in the ground model as well as a partition \( Z' \subseteq^* Z \) such that

\[
  Z' \Vdash_{U^*} (S, X) = (T, Y).
\]

We leave it as an exercise to the reader to show that

\[
  h : M^* \longrightarrow (\omega)^\omega \times M_{\text{gr}}^*,
\]

\[
  (S, X) \longmapsto (X, (S, X))
\]

is a dense embedding. Hence, by **Fact 14.3**, dual Mathias forcing \( M^* \) is equivalent to the two-step iteration \( U^* \ast M_{\text{gr}}^* \).

At this point, we would like to say a few words about the two-step iterations \( U \ast M_{\text{gr}} \) and \( U^* \ast M_{\text{gr}}^* \), respectively: At first glance, the iterations look very similar and in both cases we start with a forcing notion which is \( \sigma \)-closed. However, \( M_{\text{gr}} \) satisfies \( \text{ccc} \), which is not the case for \( M_{\text{gr}}^* \). The reason for this is that partitions of \( \omega \) — in contrast to subsets of \( \omega \) — do not have “complements”, which changes the situation drastically, especially when we work with partition ultrafilters (see below).

In order to investigate dual Mathias forcing in greater details, we have to define first a dual form of the shattering cardinal \( \mathcal{H} \): Two partitions \( X, Y \in (\omega)^\omega \) are called **almost orthogonal**, denoted \( X \perp Y \), if \( X \cap Y \notin (\omega)^\omega \), otherwise they are called **compatible**. A family \( \mathcal{A} \subseteq (\omega)^\omega \) is called **maximal almost orthogonal** (MAO) if \( \mathcal{A} \) is a maximal family of pairwise almost orthogonal partitions. Furthermore, a family \( \mathcal{H} \) of MAO families of partitions **shatters** a partition \( X \in (\omega)^\omega \), if there are \( \mathcal{A} \in \mathcal{H} \) and two distinct partitions \( Y, Y' \in \mathcal{A} \) such that \( X \) is compatible with both \( Y \) and \( Y' \). Finally, a family of MAO families of partitions is **shattering**, if it shatters each member of \( (\omega)^\omega \). Now, the **dual shattering number** \( \mathcal{H} \) is the smallest cardinality of a shattering family; more formally
\[ \mathcal{H} = \min \{ |\mathcal{H}^*| : \mathcal{H}^* \text{ is shattering} \} . \]

What can we say about the size of \( \mathcal{H} \)? Now, like for \( \mathfrak{h} \) we can show that the cardinal \( \mathcal{H} \) is uncountable and less than or equal to \( \mathfrak{c} \).

**Fact 26.6.** \( \omega_1 \leq \mathcal{H} \leq \mathfrak{c} \).

**Proof.** \( \omega_1 \leq \mathcal{H} \): Let \( \mathcal{H}_n^* = \{ \mathcal{A}_n^* : n \in \omega \} \) be a countable set of mao families. We construct a partition \( X \in (\omega)^\omega \) which is not shattered by \( \mathcal{H}_n^* \): Let \( X_0 \in \mathcal{A}_n^* \), and for \( n \in \omega \), let \( X_{n+1} = X_n \cap Y_n \), where \( Y_n \in \mathcal{A}_{n+1}^* \) is such that \( X_n \cap Y_{n+1} \in (\omega)^\omega \). Then, by the first part of Lemma 26.4, there exists an \( X \) such that for all \( n \in \omega \), \( X \subseteq X_n \).

\( \mathcal{H} \leq \mathfrak{c} \): Recall that each partition \( X \in (\omega)^\omega \) can be encoded by a real \( r_X \). Now, for each \( X \in (\omega)^\omega \) choose a mao family \( \mathcal{A}_X \) which contains two distinct partitions \( Y_0, Y_1 \in (\omega)^\omega \) such that both, \( Y_0 \) and \( Y_1 \), are compatible with \( X \). Then \( \{ \mathcal{A}_X : X \in (\omega)^\omega \} \) is a shattering family of cardinality less than or equal to \( \mathfrak{c} \). ⊤

Compared to other cardinal characteristics of the continuum, \( \mathcal{H} \) is quite small, in fact we get

**Proposition 26.7.** \( \mathcal{H} \leq \mathfrak{h} \).

**Proof.** Notice first that for every mad family \( \mathcal{A} \subseteq [\omega]^{\omega} \) there is a mao family \( \mathcal{A}^* \subseteq (\omega)^\omega \) consisting of partitions \( X \in (\omega)^\omega \) such that \( \text{Min}(X) \setminus \{ 0 \} \) is contained in some element of \( \mathcal{A} \). Let \( \mathcal{H} = \{ \mathcal{A}_\xi : \xi \in \mathfrak{h} \} \) be a shattering family of mad families and let \( \mathcal{H}^* = \{ \mathcal{A}^*_\xi : \xi \in \mathfrak{h} \} \) be the corresponding family of mao families. By contraposition we show that if \( \mathcal{H}^* \) is not shattering, then also \( \mathcal{H} \) is not shattering. So, suppose that \( \mathcal{H}^* \) is not shattering. Then there is a partition \( X \in (\omega)^\omega \) which is not shattered by \( \mathcal{A}^*_\xi \) (for any \( \xi \in \mathfrak{h} \)). Thus, for every \( \xi \in \mathfrak{h} \), we find an \( X_\xi \in \mathcal{A}^*_\xi \) such that \( X \nsubseteq X_\xi \), and therefore, \( \text{Min}(X) \subseteq \text{Min}(X_\xi) \). Hence, \( \text{Min}(X) \) is not shattered by any \( \mathcal{A}^*_\xi \), which shows that \( \mathcal{H} \) is not a shattering family. ⊥

Another small cardinal characteristic which is less than or equal to \( \mathfrak{h} \) is \( p \). So, it is natural to compare \( \mathcal{H} \) with \( p \). On the one hand, one can show that \( p = \mathcal{H} < \mathfrak{h} \) is consistent with ZFC (see Related Result 151). On the other hand, one can show that also \( \mathcal{H} < \mathfrak{h} = p \) is consistent with ZFC (see Related Result 152). Hence, \( \mathcal{H} \) can be small even in the case when \( p \) or \( \mathfrak{h} \) is large. However, by a countable support iteration of dual Mathias forcing we can enlarge \( \mathcal{H} \) without changing the size of \( p \) and show that also \( p < \mathcal{H} = \mathfrak{h} \) is consistent with ZFC.

**Proposition 26.8.** \( p = \text{cov}(\mathcal{M}) < \mathcal{H} = \mathfrak{h} \) is consistent with ZFC.

**Proof (Sketch).** Since \( p \leq \text{cov}(\mathcal{M}) \) (by Theorem 21.5), and since \( \omega_1 \leq p \), it is enough to show that \( \omega_1 = \text{cov}(\mathcal{M}) < \mathcal{H} = \omega_2 \) is consistent with ZFC.
We can just follow Proposition 24.12 (replacing Mathias forcing with dual Mathias forcing). Thus, let $P_{\omega_2} = \langle Q_\alpha : \alpha \in \omega_2 \rangle$ be a countable support iteration of dual Mathias forcing and let $G$ be $P_{\omega_2}$-generic over some model $V$ of $\text{ZFC + CH}$.

Firstly, show that $V[G] \models \mathfrak{b} = \mathfrak{h} = \omega_2$: For this, use the fact that dual Mathias forcing, like Mathias forcing, is proper, that $M^* \approx \mathbb{U} \ast M_{\mathbb{U}^*}'$, and that $\mathfrak{b} \leq \mathfrak{h}$.

Secondly, show that $V[G] \models \omega_1 = \text{cov}(M)$: For this, use the fact that dual Mathias forcing, like Mathias forcing, has the Laver property and therefore does not add Cohen reals. Furthermore, recall that the Laver property is preserved under countable support iteration of proper forcing notions and that $\text{cov}(M)$ remains unchanged if no Cohen reals are added. Thus, since $V \vDash \text{CH}$, we get $V[G] \models \omega_1 = \text{cov}(M)$.

A Dual Form of Ramsey Ultrafilters

In Chapter 10 we have seen several equivalent definitions of Ramsey ultrafilters. For example, a filter $\mathcal{U} \subseteq [\omega]^\omega$ is a Ramsey ultrafilter if for every colouring $\pi : [\omega]^2 \to 2$ there is an $x \in \mathcal{U}$ such that $\pi|_{\mathcal{U}^2}$ is constant, which is equivalent to saying that the Maiden does not have a winning strategy in the game $G_\mathcal{U}$, defined by

$$G_\mathcal{U} : \quad \text{Maiden} \quad x_0 \quad x_1 \quad x_2 \quad \ldots$$

$$\quad a_0 \quad a_1 \quad a_2 \quad \text{Death}$$

in which Death wins the game $G_\mathcal{U}$ if and only if $\{ a_i : i \in \omega \}$ belongs to $\mathcal{U}$.

Moreover, by Chapter 10 (Related Result 71, $\mathcal{U} \subseteq [\omega]^\omega$ is a Ramsey ultrafilter if the Maiden does not have a winning strategy in the game $G_\mathcal{U}'$, defined by

$$G_\mathcal{U}' : \quad \text{Maiden} \quad (a_0, x_0) \quad (a_1, x_1) \quad (a_2, x_2) \quad \ldots$$

$$\quad y_0 \quad y_1 \quad y_2 \quad \text{Death}$$

in which the Maiden wins the game $G_\mathcal{U}'$ if and only if $\{ a_i : i \in \omega \}$ does not belong to $\mathcal{U}$. The dual form of the latter game is in fact just the game $G_\mathcal{U}^*$, which we introduced in Chapter 11:

$$G_\mathcal{U}^* : \quad \text{Maiden} \quad (S_0, X_0) \quad (S_1, X_1) \quad (S_2, X_2) \quad \ldots$$

$$\quad Y_0 \quad Y_1 \quad Y_2 \quad \text{Death}$$
In that game, we require that the first move \((S_0, X_0)\) of the Maiden is such that \(X_0 \in \mathcal{U}^*\) and that \((S_0', X_0')\) is a dual Ellentuck neighbourhood. Furthermore, we require that for each \(n \in \omega\), the \(n\)th move of Death \(Y_n\) is such that \(Y_n \in (S_n, X_n)^\omega\) and \(Y_n \in \mathcal{U}^*\), and that the Maiden plays \((S_{n+1}, X_{n+1})\) such that

- \(S_n \equiv S_{n+1}, |S_{n+1}| = |S_n| + 1, S^*_{n+1} \subseteq Y_n\), and
- \(X_{n+1} \in (S^*_{n+1}, Y_n)^\omega \cap \mathcal{U}^*\).

Finally, the Maiden wins the game \(\mathcal{G}_{\mathcal{U}^*}\) if and only if the (unique) infinite partition \(X \in (\omega)^\omega\) such that \(S_n \equiv X\) (for all \(n \in \omega\)) does not belong to the family \(\mathcal{U}^*\).

With respect to the game \(\mathcal{G}_{\mathcal{U}^*}\), we define dual Ramsey ultrafilters as follows (for another dual form of Ramsey ultrafilters see Related Result 158): A family \(\mathcal{F}^* \subseteq (\omega)^\omega\) is a partition-filter if \(\mathcal{F}^*\) is closed under refinement and finite coarsening, and if for all \(X, Y \in \mathcal{F}^*\) we have \(X \cap Y \in \mathcal{F}^*\). Furthermore, a partition-filter \(\mathcal{U}^* \subseteq (\omega)^\omega\) is a partition-ultrafilter if \(\mathcal{U}^*\) is not properly contained in any partition-filter. Finally, a partition-ultrafilter \(\mathcal{U}^* \subseteq (\omega)^\omega\) is a Ramsey partition-ultrafilter if the Maiden does not have a winning strategy in the game \(\mathcal{G}_{\mathcal{U}^*}\).

It is easy to show that every Ramsey partition-ultrafilter \(\mathcal{U}^* \subseteq (\omega)^\omega\) generates a Ramsey ultrafilter \(\mathcal{U} \subseteq [\omega]^\omega\). In fact, if \(\mathcal{U}^*\) is a Ramsey partition-ultrafilter, then \(\{ \text{Min}(X) \setminus \{0\} : X \in \mathcal{U}^*\} \subseteq [\omega]^\omega\) is a Ramsey ultrafilter over \(\omega \setminus \{0\}\). On the other hand, it is not at all clear whether Ramsey ultrafilters also generate Ramsey partition-ultrafilters — in fact it seems that Ramsey partition-ultrafilters are much stronger than Ramsey ultrafilters. However, the following result shows that the existence of Ramsey partition-ultrafilters is consistent with ZFC.

**Theorem 26.9.** If \(\mathcal{U}^*\) is \(U^*\)-generic over \(V\), then \(\mathcal{U}^*\) is a Ramsey partition-ultrafilter in \(V[\mathcal{U}^*]\).

**Proof.** Because \(\mathcal{U}^*\) is \(U^*\)-generic over \(V\), \(\mathcal{U}^* \subseteq (\omega)^\omega\) is a partition-filter in \(V[\mathcal{U}^*]\). Furthermore, since \(U^*\) is \(\sigma\)-closed (by Lemma 26.4), \(U^*\) does not add new partitions which implies that \(\mathcal{U}^*\) is a partition-ultrafilter in \(V[\mathcal{U}^*]\).

It remains to show that in \(V[\mathcal{U}^*]\), the Maiden does not have a winning strategy in the game \(\mathcal{G}_{\mathcal{U}^*}\). For this, let \(\sigma\) be a \(U^*\)-name for a strategy for the Maiden in the game \(\mathcal{G}_{\mathcal{U}^*}\), i.e.,

\[\mathcal{U} \models \sigma \text{ is a strategy for the Maiden in the game } \mathcal{G}_{\mathcal{U}^*}\,\]

where \(\mathcal{U}^*\) is the canonical \(U^*\)-name for the \(U^*\)-generic filter. Let us assume that the Maiden follows the strategy \(\sigma[\mathcal{U}^*]\) in the model \(V[\mathcal{U}^*]\). Furthermore, let \(Z_0 \in (\omega)^\omega\) be such that

\[Z_0 \models \sigma(\emptyset) = (S_0, X_0)\]
In particular, since $\sigma$ is the $\name{U}$-name for a strategy,

$$Z_0 \forces \langle X_0 \rangle \in \name{U}^*.$$ 

Assume that for some $n \in \omega$ we have already constructed an $M^*$-condition $Z_n \geq Z_0$ such that

$$Z_n \forces \langle (S_0, X_0), Y_0, \ldots, (S_{n-1}, X_{n-1}), Y_{n-1} \rangle = (S_n, X_n).$$

Then, since does not add new partitions, we find a $\name{U}^*$-condition $Z'_n \geq Z_n$ (i.e., $Z'_n \subseteq^* Z_n$) and a dual Ellentuck neighbourhood $(S_n, X_n)$ in $V$ such that

$$Z'_n \forces (S_n, X_n) = (S_n, X_n).$$

Because $Z'_n \geq Z_n$, we have

$$Z'_n \forces \langle (S_0, X_0), Y_0, \ldots, (S_{n-1}, X_{n-1}), Y_{n-1} \rangle = (S_n, X_n).$$

In particular, $Z'_n \forces X_n \in \name{U}^*$, which implies that $Z'_n$ and $X_n$ are compatible. Finally, Death plays a partition $Y_n$ such that $Y_n \subseteq^* (Z'_n \cap X_n)$ and $Y_n \in (S'_n, X_n)^c$. Proceeding this way, we get an increasing sequence $S_0 \leq S_1 \leq \cdots$ of partitions of $(\omega)^\omega$.

Now, let $W \in (\omega)^\omega$ be the unique partition such that for all $n \in \omega$, $S_n \leq W$. Notice that $W$ belongs to $V$. Then $W$ is an infinite partition (i.e., an $\name{U}^*$-condition), $W \forces \name{U}^*$, $W \in \name{U}^*$, and for each $n \in \omega$, $W \subseteq^* (Z'_n \cap X_n)$. Thus, by construction we get

$$W \forces \text{"}\sigma\text{" is not a winning strategy for the Maiden in the game $\varnothing_{\name{U}^*}$},$$

and since $\sigma$ was an arbitrary strategy, the Maiden does not have a winning strategy at all.

As a consequence we get that the existence of Ramsey partition-ultrafilters is consistent with $\text{ZFC} + \text{CH}$ (just force with $\name{U}^*$ over a model in which $\text{CH}$ holds). Unlike for Ramsey ultrafilters, it is not known whether $\text{CH}$ implies the existence of Ramsey partition-ultrafilters. On the other hand, replacing $\name{U}$ with $\name{U}^*$ in the proof that ultrafilter forcing $\name{U}$ collapses $\omega$ to $\emptyset$ (see Chapter 25 | Result 144), one can show that the forcing notion $\name{U}^*$ collapses $\omega$ to $\emptyset$, and since $\emptyset > \omega_1$ is consistent with $\text{ZFC}$ (by Proposition 26.8), we get that the existence of Ramsey partition-ultrafilters is also consistent with $\text{ZFC} + \neg\text{CH}$.

**Notes**

Dual Mathias forcing was introduced and investigated by Carlson and Simpson in [4] (e.g., they showed that dual Mathias forcing has pure decision). The dual shattering number was introduced and investigated by Cichoń, Krawczyk, Majcher-Iwanow, and Węglorz in [5] (e.g., they showed that $\emptyset^* \leq \emptyset$). However, most of the results presented in this chapter are from Halbeisen [6, 7].
Related Results

150. Dualising cardinal characteristics of the continuum. The first who studied systematically the dual forms of cardinal characteristics of the continuum were Cichoń, Krawczyk, Majcher-Iwanow, and Weglorz. For example they showed that \( \mathfrak{b} \) is regular, that \( \mathfrak{b} \leq \mathfrak{h} \), and that \( \mathfrak{b} \leq \mathfrak{c} \). Before their work [5] was published in 2000, the paper was already available as a preprint in 1994 and motivated for example the work of Brendle [2], Spinas [15] and Halbeisen [6].

151. On the consistency of \( p = \mathfrak{b} < \mathfrak{h} \). Spinas [15, Theorem 4.2] showed that in Mathias’ model, which is the model we get if we have a cardinals of the continuum \( \omega_2 \) of Mathias forcing starting in a model of ZFC + CH, we have \( p = \mathfrak{b} < \mathfrak{h} \). In particular, this shows that Mathias forcing does not add Mathias partitions; otherwise, by the proof of Proposition 26.8 (originally proved in Halbeisen [6]), we would have \( \mathfrak{b} = \mathfrak{h} \) in Mathias’ model.

152. On the consistency of \( \mathfrak{b} < p \). Brendle [2] showed that \( \mathfrak{b} < \mathfrak{h} \) is consistent with \( \text{ZFC} + \text{MA} \). In particular, also \( \mathfrak{b} < p = \mathfrak{h} \) is consistent with \( \text{ZFC} \). To some extent this shows that dual Mathias forcing is far away from being a ccc forcing notion, even in the case when we restrict dual Mathias forcing to a partition-ultrafilter.

153. Dualisations of \( a \) and \( t \). We have seen above how one could dualise the shattering cardinal \( h \), and we have seen that both statements, \( \mathfrak{b} = \omega_1 = \mathfrak{b} \) and \( \mathfrak{b} = \omega_2 \), are consistent with \( \text{ZFC} \). Now, it is somewhat surprising that the dual forms of \( a \) and \( t \) are absolute (i.e., they cannot be moved). In particular, Krawczyk proved in [5] that the size of a maximal almost orthogonal family (i.e., the dualisation of a mad family) is always equal to \( c \), and Carlson proved that the dual tower is always equal to \( \omega_1 \) (see Matet [13, Proposition 43]).

154. Converse dual cardinal characteristics. If we replace the ordering “\( \leq \)” on \( (\omega)^{\omega} \) with “\( \subseteq \)”, we obviously get other kinds of dual cardinal characteristics: The so-called converse dual cardinal characteristics were first introduced and investigated by Majcher-Iwanow [12], whose work was continued by Brendle and Zhang in [3], where it is shown for example that the converse dual tower number is equal to \( p \).

155. The dual Ramsey property. In Chapter 9 we have seen that the shattering cardinal \( h \) is closely related to the Ramsey property. Now, one can show in a similar way that the dual shattering cardinal \( \mathfrak{b} \) is closely related to the so-called dual Ramsey property, which was introduced and investigated by Carlson and Simpson in [4], and further investigated by Halbeisen in [6, 7] and by Halbeisen and Löwe in [9].

156. Ultrafilter spaces on the semilattice of partitions. There is essentially just one way to define a topology on the set of ultrafilters over \( \omega \). This topological space is usually denoted by \( \beta \omega \) (cf. Chapter 9 Related Result 63). On the other hand, there are four natural ways to define a topology on the set of partition-ultrafilters. Moreover, one can show that the corresponding four spaces of partition-ultrafilters are pairwise non-homeomorphic, but still have some of the nice properties of \( \beta \omega \) (see Halbeisen and Löwe [16]).

157. Partition-filters. In [14], Matet introduced partition-filters associated with Hindman’s Theorem and the Milliken-Taylor Theorem respectively (see
Chapter 2 \textit{(Related Result 3)} and investigated the existence as well as combinatorial properties of these partition-filters. For a slightly different approach to filters associated to Hindman’s \textit{Theorem} see Blass [1].

158. \textit{Ramsey partition-ultrafilters versus Ramseyan ultrafilters}. Above, we have introduced Ramsey partition-ultrafilters in terms of the game $G_{\mathcal{W}^*}$, which is, by Chapter 10 \textit{(Related Result 71)}, related to Ramsey ultrafilters $\mathcal{W} \subseteq [\omega]^\omega$. Furthermore, we have seen that the existence of these Ramsey partition-ultrafilters is consistent with ZFC (see also Halbeisen [7, Theorem 5.1]). Ramsey partition-ultrafilters have very strong combinatorial properties (see for example Halbeisen and Matet [11]), and it seems that they are significantly stronger than Ramsey ultrafilters. For example it is not known whether CH implies the existence of Ramsey partition-ultrafilters, whereas CH implies the existence of $2^\omega$ mutually non-isomorphic Ramsey ultrafilters (see Chapter 10 \textit{(Related Result 64)}. Now, instead of defining Ramsey partition-ultrafilters in terms of the game $G_{\mathcal{W}^*}$, we could equally well take another approach: In Chapter 10 we defined Ramsey ultrafilters in terms of colourings of $[\omega]^2$, i.e., $\mathcal{W} \subseteq [\omega]^2$ is a Ramsey ultrafilter if for every colouring $\pi : [\omega]^2 \to 2$ there is an $x \in \mathcal{W}$ such that $\pi(x)2$ is constant. Dualising — and slightly strengthening — this property, we get what is called a Ramseyan ultrafilter. A partition-ultrafilter $\mathcal{W}^* \subseteq (\omega)^\omega$ is a Ramseyan ultrafilter if for every finite colouring of $(\omega)^n$, there is an $X \in \mathcal{W}^*$ such that $(X)^{(n)}$ is monochromatic. Unlike for Ramsey partition-ultrafilters, it is known that CH implies that there are $2^\omega$ mutually non-isomorphic Ramseyan ultrafilters (see Halbeisen [8, Theorem 2.2.1]). Thus, it seems that Ramseyan ultrafilters are somewhat weaker than Ramsey partition-ultrafilters — but it is also possible that they are equivalent.

References


In this chapter we shall demonstrate how the tools we developed in the previous chapters can be used to shed new light on a classical problem in Measure Theory.

Assuming the Continuum Hypothesis, Banach and Kuratowski proved a combinatorial theorem which implies that a finite measure defined for each subset of \( R \) vanishes identically if it is zero for points (for the notion of measure we refer the reader to Oxtoby [3, p. 14]). We shall consider this result — which will be called Banach-Kuratowski Theorem — from a set-theoretical point of view, and among others it will be shown that the Banach-Kuratowski Theorem is equivalent to the existence of a \( K \)-Lusin set of size \( c \) and that the existence of such a set is independent of ZFC + \( \neg \text{CH} \).

The original proof of the Banach-Kuratowski Theorem is due to Banach and Kuratowski [1]. Theorem 27.1 is due to Halbeisen, and the non-classical results of this chapter are all due to Bartoszyński. References and some more results related to the Banach-Kuratowski Theorem can be found in Bartoszyński and Halbeisen [2].

Prelude

**Historical background.** In a paper of 1929, Banach and Kuratowski investigated the following problem in Measure Theory: Does there exist a non-vanishing finite measure defined for each subset of \( R \) which is zero for points? They showed that such a measure does not exist if one assumes \( \text{CH} \). In fact, assuming \( \text{CH} \), they proved the following combinatorial theorem and showed that it implies the non-existence of such a measure (notice that it is sufficient to consider just measures on subsets of the unit interval \([0, 1]\)).

**Theorem of Banach and Kuratowski.** If \( \text{CH} \) holds, then there is an infinite matrix \( A_k^i \subseteq [0, 1] \), where \( i, k \in \omega \), such that:

(a) For each \( i \in \omega, [0, 1] = \bigcup_{k \in \omega} A_k^i \).
(b) For each \( i \in \omega \), if \( k \neq k' \) then \( A_k^i \cap A_{k'}^i = \emptyset \).

(c) For every infinite sequence \((k_0, k_1, \ldots, k_i, \ldots)\) of natural numbers,

\[
\bigcap_{i \in \omega} \left( A_0^i \cup A_1^i \cup \ldots \cup A_k^i \right) \text{ is countable.}
\]

Below, we call an infinite matrix \( A_k^i \subseteq [0, 1] \) (where \( i, k \in \omega \)) for which (a), (b), and (c) hold a **BK-Matrix**.

Concerning the measure-theoretical problem we would like to mention that Ulam [4] proved the following generalisation of the Banach-Kuratowski Theorem: If no cardinal less than or equal to \( \varepsilon \) is weakly inaccessible, then every finite measure defined for all subset of \( \mathbb{R} \) which is zero for points vanishes identically. For further results in this context we refer the reader to Oxtoby [3, Chapter 5].

**Allemande**

**A cardinal characteristic called 1.** Before we give a slightly modified version of the original proof of the Banach-Kuratowski Theorem we introduce the following notion.

Recall that for functions \( f, g \in \omega^\omega \), \( f \leq g \) \( \iff \) \( f(n) \leq g(n) \) for all \( n \in \omega \). Now, for \( \mathcal{F} \subseteq \omega^\omega \), let \( \lambda(\mathcal{F}) \) denote the least cardinality such that for each \( g \in \omega^\omega \), the cardinality of the set \( \{ f \in \mathcal{F} : f \leq g \} \) is strictly less than \( \lambda(\mathcal{F}) \). For any family \( \mathcal{F} \subseteq \omega^\omega \) we obviously have \( \lambda(\mathcal{F}) \leq \varepsilon^+ \). Furthermore, for families \( \mathcal{F} \subseteq \omega^\omega \) of size \( \varepsilon \) one can easily show that \( \omega_1 \leq \lambda(\mathcal{F}) \). Thus, for families \( \mathcal{F} \subseteq \omega^\omega \) of size \( \varepsilon \) we have \( \omega_1 \leq \lambda(\mathcal{F}) \leq \varepsilon^+ \), which leads to the following definition:

\[
l = \min \{ \lambda(\mathcal{F}) : \mathcal{F} \subseteq \omega^\omega \wedge |\mathcal{F}| = \varepsilon \}
\]

If one assumes CH, then one can easily construct a family \( \mathcal{F} \subseteq \omega^\omega \) of cardinality \( \varepsilon \) such that \( \lambda(\mathcal{F}) = \omega_1 \), hence, CH implies \( l = \omega_1 \).

In our notation, the crucial point in the original proof of Banach and Kuratowski reads as follows.

**Theorem 27.1.** The existence of a BK-Matrix is equivalent to \( l = \omega_1 \).

**Proof.** \((\Leftarrow)\) Let \( \mathcal{F} \subseteq \omega^\omega \) be a family of cardinality \( \varepsilon \) with \( \lambda(\mathcal{F}) = \omega_1 \). In particular, for each \( g \in \omega^\omega \), the set \( \{ f \in \mathcal{F} : f \leq g \} \) is at most countable. Since the interval \([0, 1] \) has cardinality \( \varepsilon \), there is a one-to-one function \( h \) from \([0, 1] \) onto \( \mathcal{F} \). For \( x \in [0, 1] \), let \( n^*_x := h(x)(i) \). Now, for \( i, k \in \omega \), define the sets \( A_k^i \subseteq [0, 1] \) as follows:

\[
x \in A_k^i \iff k = n^*_x.
\]
We leave it as an exercise to the reader to check that these sets satisfy the conditions (a) and (b) of a BK-Matrix. For (c), take any sequence \((k_0, k_1, \ldots, k_n, \ldots)\) of \(\omega\) and pick an arbitrary \(x \in \bigcap_{i \in \omega} (A_0 \cup A_1 \cup \ldots \cup A_{k_i})\). By definition, for each \(i \in \omega\), \(x\) is in \(A_0 \cup A_1 \cup \ldots \cup A_{k_i}\). Hence, for each \(i \in \omega\) we get \(x_i \leq k_i\), which implies that for \(g \in \omega\) with \(g(i) := k_i\) we have \(h(x) \leq g\). Now, since \(\lambda(\mathcal{F}) = \omega_1\), \(h(x) \in \mathcal{F}\) and \(x\) was arbitrary, the set \(\{x \in [0, 1] : h(x) \leq g\} = \bigcap_{i \in \omega} (A_0 \cup A_1 \cup \ldots \cup A_{k_i})\) is at most countable.

\((\Rightarrow)\) Let \(A_i \subseteq [0, 1]\), where \(i, k \in \omega\), be a BK-Matrix and let \(\mathcal{F} \subseteq \omega\) be the family of all functions \(f \in \omega\) such that \(\bigcap_{i \in \omega} A_i^{f(i)}\) is non-empty. It is easy to see that \(\mathcal{F}\) has cardinality \(\kappa\). Now, for any sequence \((k_0, k_1, \ldots, k_n, \ldots)\) of natural numbers, the set \(\bigcap_{i \in \omega} (A_0 \cup A_1 \cup \ldots \cup A_{k_i})\) is at most countable, which implies that for \(g \in \omega\) with \(g(i) := k_i\), the set \(\{f \in \mathcal{F} : f \leq g\}\) is at most countable. Hence, \(\lambda(\mathcal{F}) = \omega_1\). \(-\)

**Courante**

**Lusin and K-Lusin sets.** Before we can define the notions of *Lusin* and *K-Lusin* sets respectively, we have to introduce the notion of a *compact set* (for the notions open, closed, dense, and meagre we refer the reader to Chapter 21). A set \(X \subseteq \omega\) is **compact** if for every set \(\mathcal{G} \subseteq \text{seq}(\omega)\) of finite sequences in \(\omega\) such that \(X \subseteq \bigcup_{\mathcal{G}} O_\mathcal{G}\), there exists a finite subset \(\{s_0, \ldots, s_{m-1}\} \subseteq \mathcal{G}\) such that \(X \subseteq \bigcup_{i \in [m]} O_{s_i}\). In other words, \(X \subseteq \omega\) is compact if every open cover of \(X\) has a finite subcover.

The following lemma gives a combinatorial characterisation of compact subsets of \(\omega\).

**Lemma 27.2.** The closure of a set \(A \subseteq \omega\) is compact if and only if there is a function \(f_0 \in \omega\) such that \(A \subseteq \{f \in \omega : f \leq f_0\}\).

**Proof.** For \(A \subseteq \omega\) let \(T_A = \{g|_n : g \in A \cap n \in \omega\}\). Then \((T_A, \subseteq)\) is obviously a tree. Notice that if \(A\) denotes the closure of \(A\), then \(T_A = T_A\). Now, \((T_A, \subseteq)\) is finitely branching if and only if for each \(n \in \omega\), \(\{g|_n : g \in A\}\) is finite; in which case we can define \(f_0 \in \omega\) by stipulating \(f_0(n) := \max \{g|_n : g \in A\}\) (and get that for all \(g \in A\), \(f_0 \leq f_0\)). Thus, it is enough to prove that a closed set \(A\) is compact if and only if \((T_A, \subseteq)\) is finitely branching.

\((\Rightarrow)\) If \((T_A, \subseteq)\) is not finitely branching, then there is an \(n_0 \in \omega\) such that \(\mathcal{N}_{n_0} = \{g|_n : g \in A\}\) is infinite. On the one hand, \(A \subseteq \bigcup \{O_s : s \in \mathcal{N}_{n_0}\}\), but on the other hand, for any finite subset \(\{s_0, \ldots, s_{m-1}\} \subseteq \mathcal{N}_{n_0}\) we have \(A \not\subseteq \bigcup_{s \in m} O_s\), hence, \(A\) is not compact.

\((\Leftarrow)\) Assume that \((T_A, \subseteq)\) is finitely branching. Let \(\mathcal{G} \subseteq \text{seq}(\omega)\) be such that \(A \subseteq \bigcup_{\mathcal{G}} O_\mathcal{G}\) and let \(T_{\mathcal{G}} = \{g|_n : g \in A \cap n \in \omega \land \forall k \leq n(g|_k \notin \mathcal{G})\}\). First we show that \(T_{\mathcal{G}}\) is finite: Assume towards a contradiction that \(T_{\mathcal{G}}\) is infinite. Then, by Kö nig’s Lemma, \((T_{\mathcal{G}}, \subseteq)\) contains an infinite branch,
say $g_0 \in \omega$. Now, $g_0$ belongs to $A$ (since $A$ is closed), but by construction $g_0 \notin \bigcup_{s \in \mathcal{S}} O_s$, a contradiction. We say that $t \in \hat{T}_A$ is a leaf of $\hat{T}_A$ if for all $n \in \omega$, $\hat{t}n \notin \hat{T}_A$. Let $L(\hat{T}_A)$ denote the finite set of leaves of $\hat{T}_A$. Now, let $\mathcal{S}_0 = \{ \hat{t}n : t \in \hat{T}_A \land n \in \omega \land \hat{t}n \in T_A \}$. Notice that $\mathcal{S}_0 \cap T_A = \emptyset$. Then, since $(T_A, \subseteq)$ is finitely branching, $\mathcal{S}_0$ is finite, and by definition we get $\mathcal{S}_0 \subseteq \{ \hat{t}n : t \in \hat{T}_A \land n \in \omega \land \hat{t}n \in \mathcal{S} \}$. Moreover, $A \subseteq \bigcup\{ O_s : s \in \mathcal{S}_0 \}$, which shows that $A$ is compact.

An uncountable set $X \subseteq \omega$ is a Lusin set if for each meagre set $M \subseteq \omega$, $X \cap M$ is countable, and an uncountable set $X \subseteq \omega$ is a $K$-Lusin set if for each compact set $K \subseteq \omega$, $X \cap K$ is countable.

**Fact 27.3.** Every Lusin set is a $K$-Lusin set.

**Proof.** By Lemma 27.2, every compact set $K \subseteq \omega$ is meagre (even nowhere dense), and therefore, every Lusin set is a $K$-Lusin set.

Let $Q$ be a countable dense subset of $\omega$. Then $X \subseteq \omega$ is concentrated on $Q$ if every open subset of $\omega$ containing $Q$, contains all but countably many elements of $X$. Finally, a subset of $\omega$ is called concentrated if it is concentrated on some countable dense subset of $\omega$.

**Proposition 27.4.** The following statements are equivalent:

(a) There exists a $K$-Lusin set of cardinality $\mathfrak{c}$.

(b) There exists a concentrated set of cardinality $\mathfrak{c}$.

**Proof.** (b)⇒(a) Let $X \subseteq \omega$ be concentrated on some countable dense set $Q \subseteq \omega$. One can show that there exists a homeomorphism between $\omega \setminus Q$ and $\omega$, i.e., there exists a bijection $h : \omega \setminus Q \rightarrow \omega$ which maps open sets to open sets and closed sets to closed sets (the details are left to the reader). Let $K$ be an arbitrary compact subset of $\omega$, then $h^{-1}[K]$ is also compact, and therefore $\omega \setminus h^{-1}[K]$ is an open set containing $Q$. Thus, since $X$ is concentrated on $Q$, $\omega \setminus h^{-1}[K]$ contains all but countably many elements of $X$ and consequently $h[X] \cap K$ is countable; and since $K$ was arbitrary, this implies that the image under $h$ of a set concentrated on $Q$ of cardinality $\mathfrak{c}$ is a $K$-Lusin set of the same cardinality.

(a)⇒(b) Similarly, if $Q \subseteq \omega$ is a countable dense set and $h : \omega \setminus Q \rightarrow \omega$ is a homeomorphism, then the pre-image under $h$ of a $K$-Lusin set of cardinality $\mathfrak{c}$ is a concentrated set of the same cardinality.

---

**Sarabande**

**The cardinal $1$ and the existence of large $K$-Lusin sets.** The following result—even though it follows quite easily from the definitions—is in fact the heart of our set-theoretical investigation of the Banach-Kuratowski Theorem.
Theorem 27.5. $l = \omega_1$ if and only if there is a $K$-Lusin set of cardinality $\mathfrak{c}$.

Proof. ($\Rightarrow$) Assume $l = \omega_1$ and let $\mathcal{F} \subseteq ^\omega \omega$ be a set of cardinality $\mathfrak{c}$ such that for each $g \in ^\omega \omega$, $\{ f \in \mathcal{F} : f \leq g \}$ is countable. By Lemma 27.2, for each closed and compact set $K \subseteq ^\omega \omega$ there is a function $g_K \in ^\omega \omega$ such that $K \subseteq \{ g \in ^\omega \omega : g \leq g_K \}$. Thus, for every closed and compact set $K$ we have $\mathcal{F} \cap K \subseteq \{ f \in \mathcal{F} : f \leq g_K \}$ is countable, hence, $\mathcal{F}$ is a $K$-Lusin set of cardinality $\mathfrak{c}$.

($\Leftarrow$) Let $X \subseteq ^\omega \omega$ be a $K$-Lusin set of cardinality $\mathfrak{c}$. By Lemma 27.2, for each $g \in ^\omega \omega$ the set $K_g = \{ f \in ^\omega \omega : f \leq g \}$ is closed and compact. Thus, $X \cap K_g = \{ f \in X : f \leq g \}$ is countable. Hence, $\lambda(X) = \omega_1$ and since $|X| = \mathfrak{c}$ we have $l = \omega_1$.

Gavotte I & II

$K$-Lusin sets and the cardinals $\mathfrak{b}$ and $\mathfrak{d}$.

Proposition 27.6. The existence of a $K$-Lusin set of cardinality $\mathfrak{c}$ implies $\mathfrak{b} = \omega_1$ and $\mathfrak{d} = \mathfrak{c}$.

Proof. Let $X \subseteq ^\omega \omega$ be a $K$-Lusin set of cardinality $\mathfrak{c}$. On the one hand, every uncountable subset of $X$ is unbounded, so $\mathfrak{b} = \omega_1$. On the other hand, every function $g \in ^\omega \omega$ dominates only countably many elements of $X$. Hence, no family $\mathcal{F} \subseteq ^\omega \omega$ of cardinality strictly less than $\mathfrak{c}$ can dominate all elements of $X$, and thus, $\mathfrak{d} = \mathfrak{c}$.

By the definition of $K$-Lusin sets we get that $K$-Lusin sets are exactly those (uncountable) subsets of $^\omega \omega$ all whose uncountable subsets are unbounded, which explains that $K$-Lusin sets are also called strongly unbounded; $K$-Lusin sets play an important role in preserving unbounded families in iterations of proper forcing notions.

The existence of $K$-Lusin sets of cardinality $\mathfrak{c}$.

Lemma 27.7. If $G$ is $\mathfrak{c}_1$-generic over $V$, then

$$V[G] \models \text{there is a Lusin set of cardinality } \mathfrak{c}.$$ 

Proof. With $G$ we can construct a set $C = \{ c_\alpha : \alpha \in \mathfrak{c} \}$ of Cohen reals of cardinality $\mathfrak{c}$. Further, let $\tau$ be a $\mathfrak{c}_1$-name for the code of a meager $\mathcal{F}$ set $A_\tau \in V[G]$ and let $I = \text{supp}(\tau)$ (cf. Chapter 21). Clearly, $I \subseteq \mathfrak{c}$ is countable, and by Proposition 21.7, for each $\alpha \in \mathfrak{c} \setminus I$ we have $V[G] \models c_\alpha \notin A_\tau$. Hence, $C \cap A_\tau$ is countable in $V[G]$, and since $\mathcal{C}_\mathfrak{c}$ preserves cardinalities and $A_\tau$ was arbitrary, $V[G] \models \text{“$C$ is a Lusin set of cardinality } \mathfrak{c} \text{”}.$

Theorem 27.8. The existence of a $K$-Lusin set of cardinality $\mathfrak{c}$ is independent of ZFC $+ \neg \text{CH}$. Equivalently, the existence of a BK-Matrix is independent of ZFC $+ \neg \text{CH}$. 


Proof. Firstly, notice that by Theorem 27.1 and Theorem 27.5 the existence of a BK-Matrix is equivalent to the existence of a $K$-Lusin set of cardinality $\kappa$. Now, by Lemma 27.7 and Fact 27.3 it is consistent with $\text{ZFC}$ that there is a $K$-Lusin set (even a Lusin set) of cardinality $\kappa$. On the other hand, it is consistent with $\text{ZFC}$ that $b > \omega_1$ or that $\supseteq \kappa$ (cf. Chapter 18). Therefore, by Proposition 27.6, it is consistent with $\text{ZFC}$ that there are no $K$-Lusin sets of cardinality $\kappa$.

$K$-Lusin sets and the cardinals $b$ and $\supseteq$. As an immediate consequence of Proposition 27.6 and Theorem 27.8 we get that $\omega_1 = b < \supseteq = \kappa$ is consistent with $\text{ZFC}$. Since Cohen reals are unbounded and since Cohen forcing does not add dominating reals (see Chapter 21), Proposition 27.6 is in fact just a consequence of Lemma 27.7.

In the next section, a very similar construction will be used to show that the converse of Proposition 27.6 is not provable in $\text{ZFC}$.

Gigue

A model without $K$-Lusin sets in which $b = \omega_1$ and $\supseteq = \kappa$.

Proposition 27.9. It is consistent with $\text{ZFC}$ that $b = \omega_1$ and $\supseteq = \kappa$, but there is no $K$-Lusin set of cardinality $\kappa$.

Proof. Let $V$ be a model of $\text{ZFC}$ in which $p = c = \omega_2$. Let $G = \langle c_\alpha : \alpha \in \omega_1 \rangle$ be $\subset^*\omega_1$-generic over $V$. In the resulting model $V[G]$ we have $b = \omega_1$ and $\supseteq = \omega_2$ (see Proposition 21.13). On the other hand, there is no $K$-Lusin set of cardinality $\kappa$ in $V[G]$. Why? Suppose $X \subseteq \omega_\omega$ has cardinality $\omega_2$. Take a countable ordinal $\alpha$ and a subset $X' \subseteq X$ of cardinality $\omega_2$ such that $X' \subseteq V[G|\alpha]$, where $G|\alpha = \langle c_\beta : \beta \in \alpha \rangle$. Now, $V[G|\alpha] = V[c]$ for some Cohen real $c$ (by Fact 18.4), and $V[c] \models p = c$ (by Theorem 19.4), and since $p \leq b$ we have $V[c] \models b = \omega_2$. Thus, there is a function which dominates uncountably many elements of $X'$. Hence, by the remark after Proposition 27.6, $X$ cannot be a $K$-Lusin set. \hfill \dashv

One after another, the bells jangled into silence,
lowered their shouting mouths and were at peace.

Dorothy L. Sayers
The Nine Tailors, 1934
References

Index

Symbols

logic

\( M \models \Phi, 309 \)
\( \text{Con}(T), 38 \)
\( T + \varphi, 42 \)
\( T \not\models \psi, 37 \)
\( T \vdash \psi, 36 \)
\( \exists \) (exists), 32
\( \exists!, 45 \)
\( \forall \) (for all), 32
\( \text{free}(\varphi), 33 \)
\( \in, 44 \)
\( \leftrightarrow \) (iff), 32
\( I, 39 \)
\( I \models \varphi, 40 \)
\( M \models \varphi, 309 \)
\( M \not\models \varphi, 40 \)
\( N \prec M, 308 \)
\( \models, 40 \)
\( \rightarrow \) (not), 32
\( \rightarrow \text{Con}(T), 38 \)
\( \varphi(x/t), 34 \)
\( \varphi \equiv \psi, 38 \)
\( \varphi \equiv M, 309 \)
\( \rightarrow \) (implies), 32
\( \forall, 39 \)
\( \lor \) (or), 32
\( \land \) (and), 32

axioms

\( \text{AC}, 2, 111 \)
\( \text{AD}, 146 \)
\( C(N_\omega, < N_0), 134 \)
\( C(N_\omega, N_\omega), 134 \)
\( C(N_\omega, \infty), 134 \)
\( C(\mathbb{R}, n), 134 \)
\( C(\infty, < N_\omega), 134 \)
\( C(\infty, n), 135 \)
\( \text{CH}, 4, 85, 190 \)
\( \text{C}_n, 135 \)
\( D_C, 145 \)
\( \text{GCH}, 190 \)
\( \text{KL}, 135 \)
\( \text{MA}, 278, 365 \)
\( \text{MA}(\text{countable}), 279 \)
\( \text{MA}(\omega), 277 \)
\( \text{MA}(\sigma\text{-centered}), 279 \)
\( \text{PIT}, 131 \)
\( \text{RPP}, 135 \)
\( \text{ZF}, 58 \)
\( \text{ZFA}, 168 \)
\( \text{ZFC}, 111 \)
\( \text{ZFC^*}, 272, 308 \)

forcing

\( G(\alpha), 356 \)
\( G_{\infty}, 356 \)
\( \mathbb{B}, 396 \)
\( \mathbb{C}_\lambda, 346 \)
\( \mathbb{C}_\alpha, \mathbb{C}, 286 \)
\( \mathbb{C}_\lambda, 346 \)

\( \Delta_\psi, 293 \)
\( L, 414 \)
\( L_{\omega}, 415 \)
\( M^*, 442 \)
\( M, M_{\infty}, 417 \)
\( M, 406 \)
\( G, G, 288 \)
\( P \approx Q, 290 \)
\( P \ast Q, 351 \)
\( P_n, 355 \)
\( P_{\text{a,p}}\text{-condition}, 432 \)
\( U, 286 \)
\( S, 402 \)
\( S_{\infty}, 399 \)
\( x, 288 \)
\( U^*, 445 \)
\( U^*, 292-294 \)
\( x, 287 \)
\( \|\|, 293 \)
\( \text{rk}(x), 287 \)
\( \text{supp}(p), 345, 355 \)
\( p \mid q, 275 \)
\( p \perp q, 113, 275 \)
\( \text{ccc}, 276, 301 \)

classes and models

\( \Omega, 46 \)
\( L, 112 \)
\( V, 58 \)
\( V[G], 288 \)
\(\triangle, 20\)
\(a \mid b, 13\)
\(f <^* g, 191\)
\(s\bar{x}, 173\)

\(x \subseteq^* y, 190\)
\(|A| = |B|, 53\)
\(|A| \leq |B|, 53\)
\(|A| < |B|, 53\)

\(|A| \leq^* |B|, 84\)
Names

A page number is given in italics when that page contains a biographical note about the person being indexed.

Adermann, Wilhelm, 65
Aniszczyk, Bohdan, 222
Argyros, Spiros A., 25, 262
Aristotle, 31, 35, 64

Bachmann, Heinz, 68, 69, 105, 107, 140, 142, 143
Bakar, Bohuslav, 201, 206, 222, 438
Banach, Stefan, 154, 164, 165, 455, 456
Bamakh, Taras O., 24
Bar-Hillel, Yehoshua, 67, 138, 139
Bartończyński, Tomek, 241, 242, 370, 383, 384, 395, 396, 414, 415, 424, 439, 455
Baumgartner, James E., 207, 360, 382, 384, 403
Bell, John L., 141
Bell, Murray G., 283, 370
Benson, David J., 6
Bernays, Paul, 65, 66, 139, 142, 145
Bernstein, Felix, 69, 107
Blas, Andreas, XI, 140, 141, 204, 205, 206, 241, 264, 283, 339, 340, 383, 396, 402, 403, 414, 415, 452
Bocheński, Joseph M., 64
Bohano, Bernard, 66
Boole, George, 64, 140–141
Booth, David, 241, 212
Borel, Émile, 69, 70, 415
Bourbakis, Nicolas, 139
Brendel, Jörg, 22, 205, 206, 360, 403, 414, 415, 424, 425, 451
Brown, Jack B., 222
Brown, Tom, 262
Brunel, Antoine, 25

Campbell, Paul J., 139
Canjar, R. Michael, 283, 424
Carlson, Timothy J., 222, 261, 262, 263, 450, 451
Chang, Chen Chung, 314
Church, Alonzo, 105
Ciekoń, Jacek, 450, 451
Cohen, Paul J., 271, 305, 339
Corazza, Paul, 222

De la Vallée Poussin, Charles J., 6
Dedekind, Richard, 65, 66, 69, 104
Detlovs, Vilnis, 65
Devlin, Dennis, 24
Dimitriou, Joanna, XI, 340
Descartes, René, XI
Dordal, Peter Lars, 206
Dow, Alan, 206
Doxiadis, Apostolos, 65
Díaz-Monja, Mirna, 264

Easton, William B., 322, 361
Ebbinghaus, Heinz-Dieter, 65, 68, 139
Elieentuch, Erik, 221, 425
Engelking, Ryszard, 221
Erdős, Paul, 21, 22, 23, 204, 262
Euclid, 66
Euler, Leonhard, 69
Faber, Georg, 106
Farah, Ilijas, 262
Feferman, Anita Burdman, 64
Feferman, Solomon, 305, 339
Feligner, Ulrich, 140, 145, 186, 204
Fenouli, Vaggelis, 262
Fichtenholz, Grigorii, 204
Fischer, Vera, 206
Flum, Jörg, 65
Flummi, Dandolo, XI
Font, Josep Maria, 141
Forster, Thomas E., 108, 143, 144
Frankiewicz, Ryszard, 222
Frege, Gottlob, 64, 65, 70
Fremlin, David H., 283, 284, 370
Galilei, Galileo, 6, 66
Calvin, Fred, 221, 241
Names

Gauntt, Robert J., 144
Geschke, Stefan, 403
Gödel, Kurt, 65-66, 112, 139, 190, 271, 304
Goldstein, Rebecca, 66
Goldstern, Martin, 65, 206, 360, 382, 383, 384, 415
Goodstein, Reuben L., 106
Gowers, W. Timothy, 25
Graham, Ronald L., 21, 22, 23, 24, 261, 263
Grassmann, Hermann, 65
Grattan-Guinness, Ivor, 67
Gray, Charles W., 415
Grigorieff, Sergei, 402
Hajnal, Andras, 23
Halbeisen, Stephanie, XI
Hales, Alfred W., 246, 261
Hallett, Michael, 138
Halpern, James D., 142, 184, 185, 261, 262, 263, 264
Harrington, Leo, 22
Hartogs, Friedrich, 70, 70, 140
Hausdorff, Felix, 68, 140, 142, 164, 204, 206, 315, 340
Henkin, Leon, 65, 143
Hermes, Hans, 65
Hernández-Hernández, Fernando, 206
Herrlich, Horst, 145, 146
Hessenberg, Gerhard, 140
Hilbert, David, 65, 67
Hindman, Neil, 23, 24, 222, 264
Hodges, Wilfried, 314
Howard, Paul, 141, 142, 144
Hrushik, Michael, 206, 283
Hungerbühler, Norbert, 24, 108, 262
Jansana, Ramon, 141
Jech, Thomas, 22, 139, 140, 141, 143, 144, 145, 146, 184, 185, 186, 221, 283, 307, 314, 315, 322, 339, 340, 361, 370, 384, 396, 397, 402, 430
Jewett, Robert L., 246, 261
Jourdain, Philip E. B., 65, 140
Julma, Stasys, 261
Kakutani, Shizuo, 204
Kalemba, Piotr, 221
Kanamori, Akihiro, 67, 68, 130, 146, 204, 305
Kanellopoulos, Vassilis, 262
Kantorovitch, Leonid V., 204
Kaye, Richard, 65
Kečkris, Alexander S., 414
Keisler, H. Jerome, 241, 314
Kellner, Jakob, 403
Keränen, Veikko, 262
Keremedis, Kyriakos, 142
Ketonen, Jussi A., 204, 241
Khomskii, Yuri, 206
Kirby, Laure, 106
Kleene, Stephen Cole, 65, 67, 68, 70
Klemberg, Eugene M., 141
Kneser, Hellmuth, 139
Komjáth, Péter, 23
König, Julius, 140
König, Dénes, 6
Korselt, Alwin, 60
Krawczyk, Adam, 450, 451
Kuratowski, Casimir, 68, 139, 455, 456
Kurepa, Djuro, 105, 140
Kurilčik, Miks S., 396, 403
Laczkovich, Miklós, 165
Laflamme, Claude, 241, 243
Lagrange, Joseph-Louis, 30
Landman, Bruce M., 262
Larson, Jean A., 264
Läuchli, Hans, 105, 142, 143, 145, 185, 186, 261, 263, 264
Laver, Richard, 264, 403, 414, 415, 438
Leader, Imre, 264
Lebesgue, Henri, 105
Lesniewski, Stanislaw, 140
Lévy, Azriel, 66, 67, 105, 138, 139, 143, 144, 184, 185, 262, 314, 339
Lewin, Mordechai, 21
Lindenbaum, Adolf, 69, 104, 105, 140, 141, 184
Lothaire, 266
Louveau, Alain, 222, 414
Löwe, Benedikt, 22, 340, 403, 451
Löwenheim, Leopold, 68
Müller Aloys, 64
MacHale, Desmond, 141
Majcher-Iwanow, Barbara, 450, 451
Mancosi, Paolo, 66
Marczewski, Edward, 145
Martin, Donald A., 283, 370
Matét, Pierre, 221, 247, 261, 263, 451, 452
Mathias, Adrian Richard David, 145, 146, 214, 242, 402, 424, 425
Mendelson, Elliott, 68, 184
Mijares, José G., 222
Miller, Arnold W., 283, 396, 403, 414, 438, 439
Milliken, Keith R., 23, 264
Mirimanoff, Dimitry, 68
Mitchell, William J., 264
Montague, Richard, 314
Montenegro, Carlos H., 143
Moore, Gregory H., 138, 139, 140, 142, 305
Morris, Walter D., 21
Mostowski, Andrzej, 65, 144, 145, 184, 186, 314
Mycielski, Jan, 143, 146
Neumann, John von, 66, 67, 68, 69, 105, 140, 164
Nilli, Alon, 261
Noether, Emmy, 65
Odell, Edward W., 25
Oxtoby, John C., 455, 456
Papadimitriou, Christos H., 65
París, Jeff B., 22, 106
Pawlowski, Janusz, 396
Peano, Giuseppe, 64–65, 69, 137–138
Peirce, Charles S., 66
Pelant, Jan, 204, 222, 438
Perron, Oskar, 69
Pelczyński, Aleksander, 207
Pigozzi, Don, 141
Pín, Jean-Eric, 261
Pinch, David, 144, 145, 262
Piotrowski, Zbigniew, 306
Piper, Greg, 206
Plato, 29, 66
Pleasants, Peter A. B., 262
Plewik, Szymon, 221, 222
Podnieks, Karlis, 65
Prikry, Karel, 221
Protsasov, Igor V., 24
Prümel, Hans J., 261
Putnam, Hilary, 65
Quickert, Sandra, 403
Rado, Richard, 22
Radziszowski, Stanisław P., 23
Razoumier, Jean, 165
Ramović, Goran, 264
Ramsey, Arthur M., 21
Ramsey, Frank P., 12, 21, 141
Rang, Bernhard, 67
Rasiowa, Helena, 141, 143
Repicky, Miroslav, 415
Robertson, Aaron, 262
Robinson, Raphael M., 153, 154, 164
Roitman, Judy, 370, 439
Rosenthal, Haskell, 207
Roslanowski, Andrzej, 402
Rothschild, Bruce L., 21, 22, 261, 263
Rubin, Herman, 139, 142
Rubin, Jean E., 139, 141, 142, 144
Rudin, Mary Ellen, 283
Rudin, Walter, 241, 242
Russell, Bertrand, 65, 66, 67, 70, 138, 139, 140
Sacks, Gerald E., 402, 403
Sayers, Dorothy L., VII, 460
Schmidt, Erhard, 138
Schoenflies, Arthur M., 67
Schröder, Ernst, 66, 69, 105
Schur, Issai, 21
Scott, Dana, 305, 339
Shanin, Nikolai A., 283
Shapiro, Stewart, 138
Shelah, Saharon, 104, 105, 165, 184, 185, 186, 204, 205, 206, 241, 242,
Names

247, 261, 264, 283, 370, 382, 383, 384, 403, 414, 415, 424, 425, 430
Sierpiński, Wacław, 104, 105, 107, 138, 140, 142, 164, 165, 204, 207
Siikovski, Roman, 141, 143
Silver, Jack, 221, 425
Simon, Petr, 204, 222, 403, 438
Simpson, Steve G., 222, 261, 262, 263, 450, 451
Sjört, Jörg, XI
Skolem, Thoralf, 44, 67, 68
Skane, Neil J. A., 24
Skamson, Alan B., 141
Sobociński, Bolesław, 142
Sochor, Antonín, 339
Soifer, Alexander, 22
Solovay, Robert M., 165, 283, 370, 396
Soltan, Valeriu, 21
Specker, Ernst, 105, 184
Spencer, Joel H., 21, 22, 261
Spinias, Otmar, 415, 451
Spivak, Ladislav, 66, 105
Stephenson, Fabian, VIII, 30
Steinhaus, Hugo, 146
Stephūns, Juris, 402
Stern, Jacques, 396
Strauss, Donna, 23, 24, 222, 254
Sucheston, Louis, 25
Sukalov, Vladimir N., 207
Sudan, Gabriel, 142
Szemeréti, George, 21
Szpilrajn, Edward, 145
Szymański, Andrzej, 396

Tarski, Alfred, 69, 104, 105, 107, 140, 141, 142, 143, 144, 154, 164, 165, 207
Tarsy, Michael, 21
Taylor, Alan D., 23
Teichmüller, Oswald, 139
Thomas, Wolfgang, 65, 67

Todorčević, Stevo, 25, 222, 262, 370
Truss, John K., 105, 108, 142, 143, 144, 145, 186, 396
Tukey, John W., 139
Ulam, Stanisław, 456
Van der Waerden, Bartel L., 246, 261
Van Douwen, Eric K., 204, 205, 206
Van Mill, Jan, 222, 242
Van Heijenoort, Jean, 64
Vaughan, Jerry E., 204
Verbitski, Oleg V., 24
Voigt, Bernd, 261
Vojtěš, Peter, 66, 105, 205
von Plato, Jan, 68
Vorobets, Yaroslav B., 24
Vulkanović, Vojkan, 24

Węglorz, Bogdan, 450, 451
Wagon, Stan, 165
Wang, Hao, 65
Wagner, Leonard M., 165
Weiss, William, 283
Whitehead, Alfred North, 65
Wieferich, Arthur, 24
Wilson, Trevor M., 165
Wimmers, Edward L., 242
Wiśniewski, Kazimierz, 143
Wojciechowska, Anna, 221
Woodin, W. Hugh, 146, 424

Yatabe, Shunsuke, 403

Zapletal, Jindřich, 206
Zarliino, Gioseffo, 1, 6, 11, 29, 79, 111, 153, 167, 189, 211, 225, 245, 269, 375
Zermelo, Ernst, 31, 44, 63, 67-68, 69, 111, 114, 138-139, 140
Zhang, Shuguo, 451
Zorn, Max, 139
Subjects

Proper definitions as well as theorems with their proofs are indicated by boldface page numbers, theorems without proofs are indicated by page numbers without serifs, and historical notes are indicated by page numbers given in italics.

addition
  cardinal, 123
  ordinal, 56

algebra
  algebra of sets, 128
  Boolean algebra, 127, 140
  Lindenbaum algebra, 129, 141

atoms, 168

Axiom
  of Atoms, 168
  of Choice (AC), 3, 18, 20, 44, 58, 68, 69, 111, 133–122, 137–139, 157, 159, 177, 184, 185, 186, 298
  of Determinacy, 146
  of Empty Set, 45
  of Empty Set (for ZFA), 168
  of Extensionality, 45
  of Extensionality (for ZFA), 168
  of Foundation, 58, 68, 68, 88, 106, 119
  of Infinity, 22, 48, 66, 314
  of Pairing, 46, 298
  of Power Set, 51–52, 315
  of Regularity, 68
  of Union, 47
  Schema of Replacement, 55, 68, 91, 314
  Schema of Separation, 48

axiom, 34
  logical, 34–35
  non-logical, 35
  schema, 34

systems:
  finite fragments of ZFC, 272, 308
  Group Theory, 35–36
  Peano Arithmetic, 36, 106
  Zermelo-Fraenkel, 45–58

axiom-like statements:
  Axiom A, 382
  Continuum Hypothesis (CH), 85, 139, 190, 204, 455
  Generalised Continuum Hypothesis, 139, 143, 190
  Martin’s Axiom (MA), 4, 241, 242, 278, 283, 365
  Pigeon-Hole Principle, 2
  infinite version, 2, 13, 14
  Singular Cardinal Hypothesis, 143

Baire Category Theorem, 389
Baire property, 215
binary mess, 131
consistent with, 132
Borel’s conjecture, 415, 424
Canonical Ramsey Theorem, 20, 22
Cantor normal form, 88
Cantor products, 54, 69
Cantor’s diagonal argument, 66
Cantor’s Normal Form Theorem, 88, 105
Cantor’s Theorem, 61, 66, 81, 122
Cantor-Bernstein Theorem, 53–54, 55, 69, 70–71, 80, 81, 105
cardinal characteristics, 3
cardinality, 11, 53
Cardon’s Lemma, 255–258, 259, 261
Cartesian product, 52
Cayley graph, 155
choice principles:
  Axiom of Choice for Finite Sets, 134, 141
  Compactness Theorem for Propositional Logic, 132–133, 141
  Consistency Principle, 132–133
  Countable Axiom of Choice, 134, 141, 145
  Hausdorff’s Principle, 142
  König’s Lemma, 2–3, 6, 15, 16, 135, 141, 143, 173, 457
  Kuratowski-Zorn Lemma, 116–117, 139
  Kurpaa’s Principle, 116, 119, 140, 185, 292, 340
Linear-Ordering Principle, 140
Multiple Choice, 118–119, 140, 185, 340
Order-Extension Principle, 144, 186
Ordering Principle, 144, 186
Prime Ideal Theorem, 131, 133–134, 141, 177–180, 186, 331
Principle of Dependent Choices, 145
Ramsey Partition Principle, 135
Teichmüller’s Principle, 116–117, 131, 139
Trichotomy of Cardinals, 69, 121–122, 140
Tukey’s Lemma, 113
Ultrafilter Theorem, 131, 133, 141
class, 50
Cohen forcing $\mathbb{C}$, 346
adds splitting reals, 388
adds unbounded reals, 387
Cohen real, 345
does not add dominating reals, 388
equivalent forms, 346
is proper, 387
Compactness Theorem, 43, 143, 312, 318
comparable, 113
compatible, 113, 275
c-condition, 275
stronger, 286
weaker, 286
consistency
of ZF, 63
constructible universe, 112, 130
countable chain condition (ccc), 276
cumulative hierarchy, 58–59

De Morgan laws, 128
Deduction Theorem, 37
Delta-System Lemma, 276–277
doughnut, 21
property, 21, 22
dual Mathias forcing $\mathbb{M}^*$, 442
adds dominating reals, 442
adds Mathias reals, 442
has pure decision, 443–444
has the Laver property, 444
is proper, 444
Mathias partition, 444
Dual Ramsey Theorem, 261, 263
Ellentuck topology (on $|\omega|^\omega$), 216
Ellentuck’s Theorem, 217–218, 221, 222
equivalence class, 12
Euler number $e$, 55, 96
exponentiation
cardinal, 123
ordinal, 57
family
$P$-family, 237
$\sigma$-reaping, 205
almost disjoint, 194
refining, 202
dominating, 191
free, 226
happy, 226, 241
independent, 195
maximal almost disjoint, 194
strongly, 243
maximal independent, 196
Ramsey, 237
reaping (unsplittable), 3, 193
shattering, 200
refining, 202
splitting, 192
strong finite intersection property
(sfi), 190
tower, 205
ultrafilter base, 206
unbounded, 191
filter, 130
$\mathbb{P}$-generic, 276
dual, 130
free, 226
normal, 169, 326
prime, 130
principal, 130
rapid, 242, 437
trivial, 130
ultrafilter
$P$-point, 232
$Q$-point, 232, 437
Ramsey, 5, 229, 397
simple $P_\alpha$-point, 242
unbounded, 242
Finite Ramsey Theorem, 14–15, 17, 21, 102, 179, 183, 247, 263
forcing, 305
P-generic filter, 290
P-name, 287
collapsing of cardinals, 301
condition, 286
\( \mathfrak{p} \)-generic, 381
support, 345, 355
generic extension, 292
generic model, 292
iteration, 355
countable support, 355
finite support, 355
language, 289
name, 287
canonical, 288
hereditarily symmetric, 326
nice, 321
symmetric, 326
notion, 285
\("\omega\)-bounding, 378
\(\kappa\)-closed, 303
\(\sigma\)-closed, 299
dense embedding, 290
equivalent, 290
Laver property, 379
preserve \(P\)-points, 383
proper, 381
satisfies \(\kappa\)-cc, 322
satisfies \(\text{ccc}\), 301
preservation of cardinals, 301
preservation of cofinalities, 301
product
finite support, 345
real
dominating, 377
Mathias, 417
Miller, 406
minimal (degree of constructability), 384
Sacks, 402
splitting, 377
unbounded, 377
relationship (\(\Leftrightarrow\)), 292–294
symmetric submodel, 326
forcing notions:
\( K_\alpha, K_\delta \), 303
Cohen forcing \( \mathbb{C}, \mathbb{C}_\alpha \), 286

dual Mathias forcing \( \mathbb{M}^* \), 442
Grigorieff forcing, 399
Laver forcing \( L \), 414
Mathias forcing \( \mathbb{M}, \mathbb{M}_\alpha \), 417
Miller forcing \( \mathbb{M} \), 406
random forcing \( \mathbb{R} \), 396
rational perfect set forcing,
see Miller forcing, 406
restricted Laver forcing \( L_\alpha \), 415
Sacks forcing \( S \), 402
Silver forcing, 399
Silver-like forcing \( S_\alpha \), 399
ultrafilter forcing \( U \), 287
Forcing Theorem, 294–296
Fréchet
filter, 130, 225
ideal, 130
function, mapping, 52
automorphism of \( P \), 325
bijective, 52
choice, 111
domain, 52
finite-to-one, 233
image, 52
injective, 52
one-to-one, 52
onto, 52
range, 52
surjective, 52

Gödel’s Completeness Theorem, 41, 42, 65, 143, 317
Gödel’s Incompleteness Theorem, 42–44, 65–66
Gödel’s Second Incompleteness Theorem, 44, 63, 65–66, 315
game, 236
run, 236
strategy, 237
winning strategy, 237
Generic Model Theorem, 297, 318
Goodstein sequences, 106
Gower’s Dichotomy Theorem, 25
graph, 1
connected, 1
cycle-free, 1
edge, 1
infinite, 1
vertex, 1
Hales-Jewett function, 246
Hales-Jewett Theorem, 246–249, 256, 261, 263
Halpern-Läuchli Theorem, 184, 261–262, 263–264
Hartogs's Theorem, 62–63, 70, 84, 85, 120
Hausdorff’s Paradox, 154
Hindman's Theorem, 23
ideal, 129
  dual, 130
  normal, 170
  prime, 130
  principal, 130
  trivial, 130
incomparable, 113
incompatible, 113, 275
Induction Schema, 63
Inequality of König-Jourdain-Zermelo, 126–127, 140
Jech-Sochor Embedding Theorem, 334–338
König's Theorem, 140
kernel, 169
Läuchli’s Lemma, 100–104, 105, 171
Löwenheim-Skolem Theorem, 68, 308, 314, 315
Laver forcing \(\mathbb{L}\), 414
  has the Laver property, 415
  is minimal, 415
  is proper, 415
  satisfies Axiom A, 415
Laver forcing \(\mathbb{L}_\omega\)
  adds dominating reals, 415
  adds splitting reals, 415
  has pure decision, 415
  has the Laver property, 415
  satisfies ccc, 415
logic
  \(\varepsilon\)-automorphism, 169
  \(\varepsilon\)-isomorphism, 334
  \(\mathcal{L}\)-formulae, 32
  absolute, 311
  assignment, 39
  atomic formula, 33
  bound variable, 33
  complete theory, 42
  consistent, 38
  consistent relative to, 42
  consistent with, 42
  constant symbols, 32
  definable over, 112
  definition, 36
  domain of \(\mathfrak{A}\), 39
  elementary substructure, 308
  equality symbol, 32
  equivalent formulae, 38
  first-order, 32–42
  formal proof, 36–37
  formula, 33
  Polish notation, 33, 132
  free variable, 33
  function symbols, 32
  higher-order, 31
  incomplete theory, 42
  inconsistent, 38
  independent, 42
  inference rules, 36
  Generalisation, 36
  Modus Ponens, 36
  interpretation I, 39
  isomorphic structures, 307
  language \(\mathcal{L}\), 32
  logical operators, 32
  logical quantifiers, 32
  logical symbols, 32
  model, 40
  non-logical symbols, 32
  propositional, 132
  realisation, 132
  reflect, 309
  relation symbols, 32
  relativisation, 309
  satisfiable, 132
  satisfies, 132
  sentence, 34
  set model, 308
  structure \(\mathfrak{A}\), 39
  substitution, 34
  admissible, 34
tautology, 38
term, 33
theory, 42
variables, 32
propositional, 132

Mathias forcing $\mathcal{M}, \mathcal{M}_e$, 417
adds dominating reals, 417–418
adds splitting reals, 418
has pure decision, 418–419
has the Laver property, 420–421
is proper, 420
Mathias real, 417
stem of a condition, 417
Miller forcing $\mathcal{M}$, 406
adds unbounded reals, 407
does not add dominating reals, 410
does not add splitting reals, 407–410
has the Laver property, 414
is minimal, 414
is proper, 406–407
Miller real, 406
preserves $P$-points, 410–413
satisfies Axiom A, 414
Milliken-Taylor Theorem, 23, 25
Mostowski’s Collapsing Theorem, 312, 318
multiplication
  cardinal, 123
  ordinal, 56
natural numbers, 11
  non-negative integers, 11
number
  $\sigma$-reaping, 205
  additivity of $\mathcal{R}_0$, 213
  additivity of $\mathcal{M}$, 396
  algebraic, 80
  almost disjoint, 194
  bounding, 191
  cardinal, 60, 69–70, 122
  D-finite (Dedekind-finite), 79
  aleph, 60, 79
  cofinality, 125
  Dedekind-infinite, 79
  finite, 60, 79
  inaccessible, 315
  infinite, 60, 79
  limit, 123
  measurable, 340
  regular, 125
  singular, 125
  successor, 123
  transfinite, 79
  covering of $\mathcal{R}_0$, 213
  covering of $\mathcal{M}$, 396
  dominating, 191
dual shattering, 446
homogeneity, 199
homogeneous, 5
independence, 196
natural, 51
ordinal, 46, 68
addition, 56
exponentiation, 57
limit, 50
multiplication, 56
order type, 62
successor, 50
partition, 5, 199
pseudo-intersection, 191
reaping, 3, 194
shattering, 201
splitting, 192
tower, 205
transcendental, 81
ultrafilter, 206

operation
  associative, 35

partition, 250
  almost orthogonal, 446
  block, 250
coarser, finer, 250
compatible, 446
domain, 250
dual Ellentuck neighbourhood, 251
  dual Ellentuck topology, 251
family
  complete, 251
  free, 251
  Ramsey, 252
filter, 449
finite, 250
infinite, 250
maximal almost orthogonal, 446
Ramsey ultrafilter, 449
segmented, 254
shattering family, 446
ultrafilter, 449
Partition Ramsey Theorem, 255–259, 260, 261, 262–263, 443
Peano Arithmetic, 64–65
permutation model, 169, 184
second Fraenkel model, 173
in which \( \text{seq}(m) < \text{fin}(m) \), 180
in which \( m^2 < |m|^2 \), 182
Prime Ideal Theorem, 184

Ramsey numbers, 22
Ramsey property, 21, 22, 212
random forcing, \( \mathbb{B} \), 396
random real, 396
Reflection Principle, 309–311, 313, 314, 318
relation
\( \tau \)-ary, 53
almost contained, 190
almost disjoint, 194
antisymmetric, 113, 285
binary, 53
dominates, 191
equivalence relation, 11–12
extensional, 312
linear ordering, 113
membership, 44
partial ordering, 113
reflexive, 11, 113
splits, 192
symmetric, 12
transitive, 12, 113
well-founded, 312
well-ordering, 53, 114
representatives, 12
restricted Laver forcing, \( L_\omega \), 415
Russell's Paradox, 31, 67
Sacks forcing \( \mathbb{S} \), 402
does not add splitting reals, 403
has the Laver property, 402
is \( \omega \)-bounding, 402
is minimal, 403
is proper, 402
Schröder-Bernstein Theorem, 69
Schur's Theorem, 16, 21–22
set
\( K \)-Lusin, 458
D-finite (Dedekind-finite), 79, 104
\( F_\omega \) (in \( \omega \)), 389
\( G_\delta \) (in \( \omega \)), 389
\( \in \)-minimal, 46
almost homogeneous, 5, 19, 199
anti-chain, 113, 276, 290
chain, 113
closed
in \( \omega \), 389
compact, 457
concentrated, 458
on \( Q \), 458
congruent, 153
countable, 11
dense
above \( p \), 292
in \( \omega \), 389
in \( P \), 276, 290
difference, 48
directed, 276, 290
downwards closed, 276
equidecomposable, 153
extension (of \( x \)), 312
filter (on \( P \)), 276, 290
finite, 51
finite character, 116
hereditarily \( < \kappa \), 315
hereditarily symmetric, 169
homogeneous, 5, 12
inductive, 47
infinite, 51
intersection, 48
linearly ordered, 113
Lusin, 458
maximal anti-chain, 113
meagre
in \( \omega \), 389
monochromatic, 12
natural numbers, 51
open, 276, 290
in “ω, 388
ordered by ∈, 46
ordered pair, 47
partially ordered, 113
σ-centred, 279
σ-linked, 283
centred, 279
countable, 279
strict sense, 113
partition, 87
perfect, 402
power set, 52
proper subset, 45
pseudo-intersection, 190
real numbers, 54
sequence, 53
subset, 45
transfinite, 79
transitive, 46
transitive closure, 59
uncountable, 60
union, 47
well-orderable, 114
well-ordered by ∈, 46
Set Theory, 66–67
Silver-like forcing $S_\omega$, 399
adds splitting reals, 401
Grigorieff forcing: $\mathcal{E}$ a P-point, 399
has the Laver property, 402
is “ω-bounding, 400–401
is minimal, 402
is proper, 400
Silver forcing: $\mathcal{E} = [\omega]^\omega$, 399
Silver real, 399
Skolem Paradox, 68
Soundness Theorem, 41
support, 170
Susslin operation, 219
generalised, 219
symmetric, 169
symmetric difference, 20
symmetry group (of x), 169
topology
π-base, 222
base, 215
closed, 214
closure, 215
dense, 215
interior, 215
meagre, 215
nowhere dense, 215, 389
on X, 214
open, 214
basic, 215
points, 214
P-point, 241
space, 214
tree π-base, 222
Transfinite Induction Theorem, 51
Transfinite Recursion Theorem, 55–56,
69, 92, 93
tree, 1, 259, 405
branch, 2
finitely branching, 2
height, 222
node, 405
perfect, 260
root, 1
superperfect, 405
urelements, 168
Van der Waerden numbers, 262
Van der Waerden’s Theorem, 4, 246,
261, 262
Weak Halpern-Läuchli Theorem, 260,
361, 403
Wieferich primes, 13, 24