POINCARÉ AND ANALYTIC NUMBER THEORY

E. KOWALSKI

> Un domaine arithmétique où l'unité semble faire absolument défaut, c'est la théorie des nombres premiers ; on n'a trouvé que des lois asymptotiques et l'on n'en doit pas espérer d'autres; mais ces lois sont isolées et l'on n'y peut parvenir que par des chemins différents qui ne semblent pas pouvoir communiquer entre eux. Je crois entrevoir d'où sortira l'unité souhaitée, mais je ne l'entrevois que vaguement ; tout se ramènera sans doute à l'étude d'une famille de fonctions transcendantes qui permettront, par l'étude de leurs points singuliers et l'application de la méthode de Darboux, de calculer asymptotiquement certaines fonctions de très grands nombres.
>
> *A domain of arithmetic where unity seems completely missing, is the theory of prime numbers; only asymptotic laws have been found, and one can not hope for others; but these laws are isolated and one can only reach them by different paths which do not seem to be able to communicate. I believe I can glimpse where the desired unity will come from, but I see this only vaguely; all will probably be reduced to the study of a family of transcendental functions which will permit, through the study of their singular points and application of Darboux's method, the asymptotic computation of certain functions of very large numbers.*
>
> H. Poincaré, *L'avenir des mathématiques* ("The future of mathematics"), [P1])

1. Introduction

Analytic number theory is a relatively young branch of arithmetic, although the natural motivation from which its essential questions arise is as old as the theorem stating that there exist infinitely many prime numbers.

Poincaré did not really contribute directly to analytic number theory. Looking in his complete works, in particular in the fifth volume dedicated to arithmetic, only two papers can be considered as being related to it. One of them [P3] considers prime numbers of the form $4n+1$ or $4n+3$, and extends to these residue classes the methods used by Chebychev to give upper and lower bounds for the number $\pi(x)$ of primes up to x; this is not one of the great works of Poincaré! The second [P4], in a volume of Crelle's Journal in honor of Dirichlet, introduces various analytic functions and derives the analytic class-number formula for quadratic forms of a given discriminant, using a different approach than that of Dirichlet.

As shown by the quote at the beginning of this text, this rather modest contribution is not due to a lack of interest in (to simplify) the asymptotic properties of prime numbers. It is not entirely clear which analytic functions Poincaré had in mind when he wrote those lines. Probably they were what are now called Dirichlet series, and more precisely L-functions.

And yet, although this was only discovered fairly recently, one (not very well-known) achievement of Number Theory in the 20th Century was the introduction of the automorphic or Fuchsian functions, dear to Poincaré, in analytic number theory. Going through a path which is no less beautiful for being a bit off the beaten track, using many different intuitions along the way, a whole area of the theory of prime numbers was thus finally transformed (the main steps being the work of Hecke and Petersson, then Maass and Selberg, then Kuznetsov and finally Iwaniec and Deshouillers). Moreover, very recently, a link has appeared between these results and the old Twin Primes Problem i.e., the question, which remains open, of proving that there exist infinitely many primes p such that $p + 2$ is also prime, and if yes, the further problem of understanding their distribution.

In what follows, there will of course be no question of giving proofs. Readers who wish to know more may find more information, and often complete proofs, in books like [IK] or [S], which are accessible at the level of first or second-year graduate students.

2. Poincaré series, Kloosterman sums

The functions we will discuss are among those justly known as "Poincaré series". Let us first recall their definition. The general principle is the following: suppose that we have a group Γ which acts on a set X. We wish to construct (complex-valued) functions on X which are invariant under the action of Γ, i.e., which satisfy $f(\gamma \cdot x) = f(x)$ for all $\gamma \in \Gamma$ and $x \in X$. In full generality, this is a delicate question; but if we have at hand a function F on X which happens to already be invariant under a *subgroup* B of Γ, the beautiful idea of Poincaré series is that, formally at least, the average of F over the cosets $B\gamma$ of B in Γ gives us one such function f:

$$f(x) = \sum_{g \in B\backslash\Gamma} F(g \cdot x).$$

Indeed, it suffices to rearrange the corresponding expression for $f(\gamma \cdot x)$ by writing $g\gamma = g'$ (which permutes the cosets of B) to see that f is invariant... if it exists, which may not be the case, since the sum above might well be infinite.

In the case which is of interest to us, this existence question is indeed non-trivial, although it is fairly easy to handle. We recall the context of automorphic functions, in a relatively simple case which suffices for the applications to number theory to be discussed afterwards.

The set X is here what is now called the *Poincaré upper half-plane*

$$\mathbf{H} = \{z = x + iy \in \mathbf{C} \mid y > 0\}.$$

The group Γ is a "Fuchsian group of the first kind" in $SL(2, \mathbf{R})$; it will be enough to know that the Hecke congruence groups $\Gamma_0(q)$, defined by

$$\Gamma_0(q) = \Big\{g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,\ b,\ c,\ d \in \mathbf{Z},\ ad - bc = 1,\ c \equiv 0 \,(\mathrm{mod}\, q)\Big\},$$

are examples of such groups (though they are atypical, because of their arithmetic nature which is not general). Even the simplest case where $q = 1$, when $\Gamma = \Gamma_0(1)$ is the group $SL(2, \mathbf{Z})$, is enough to encounter most of the main points of the theory.

The action of such a group on $X = \mathbf{H}$ is given by

$$g \cdot z = \frac{az + b}{cz + d}.$$

In addition to *automorphic functions* which are complex-valued functions defined on $\mathbf{H}$ satisfying

$$f\Big(\frac{az+b}{cz+d}\Big) = f(z) \tag{1}$$

for all $\gamma \in \Gamma$ (and some additional regularity conditions which are very important, but which we will mostly avoid discussing; roughly, one asks that f be holomorphic – which means that it preserves angles – and that the function f grow only moderately when z approaches the "boundary" of $\mathbf{H}$ – for instance, $|f(z)|$ must be bounded by a power of $\mathrm{Im}(z)$ for $\mathrm{Im}(z) \geqslant 2$) it is classical to consider *automorphic forms* of weight k, where k is a positive integer (which is necessarily even in the situation below if f is not identically zero): these are functions f, still complex-valued and defined on $\mathbf{H}$, which have regularity properties similar to those we briefly mentioned, and which satisfy now

$$f\Big(\frac{az+b}{cz+d}\Big) = (cz+d)^k f(z) \tag{2}$$

for $\gamma \in \Gamma$. One of Poincaré's intuitions was precisely that such forms are easier to construct directly than automorphic functions, but that by defining $f = f_1/f_2$, where f_1 and f_2 are automorphic forms of the same weight (and where f_2 is non-zero), an automorphic function f is obtained.

The principle underlying Poincaré series applies here if one takes for B the subgroup of "unipotent" matrices in $\Gamma_0(q)$:

$$B = \Big\{ g = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbf{Z} \Big\}.$$

Because such a matrix acts by $z \mapsto z+n$, we see that in order to find a suitable function F, one can select an arbitrary 1-periodic function defined on $\mathbf{H}$. It is natural to consider

$$F(x+iy) = \exp(2i\pi mz) = \exp(-2\pi my)\exp(2i\pi mx),$$

for $m \in \mathbf{Z}$, since these functions satisfy this periodicity condition. The corresponding Poincaré series are, formally at least, given by the expressions

$$P_m(z) = \sum_{g \in B\backslash\Gamma_0(q)} \exp(2i\pi m\gamma\cdot z),$$

$$P_m^{(k)}(z) = \sum_{g \in B\backslash\Gamma_0(q)} (cz+d)^{-k}\exp(2i\pi m\gamma\cdot z),$$

where the second definition corresponds to the modification of the averaging principle above which is needed to construct an automorphic form instead of an automorphic function.

It is not very difficult here to establish that the first series *diverges*, for any value of m and any z; the second, however, converges and defines an automorphic form of weight k as soon as $k > 2$.

In the case of $SL(2,\mathbf{Z})$, these series are described explicitly by Poincaré in a paper published after his death [P2] where, in the last pages, he proceeds to the next step in their analysis: he computes the *Fourier expansion* (or "q-expansion") of $P_m^{(k)}(z)$. Indeed, the relation (2), applied to a matrix in B, implies that the Poincaré series are still 1-periodic. The theory of Fourier series implies that they may be expanded in a series of the type

$$P_m^{(k)}(z) = \sum_{h \in \mathbf{Z}} p_m^{(k)}(h)\exp(2i\pi hz)$$

for certain *Fourier coefficients* $p_m^{(k)}(h) \in \mathbf{C}$. The general formula, which is now classical, is of the form

$$(3) \qquad p_m^{(k)}(h) = \Big(\frac{h}{m}\Big)^{(k-1)/2} \Big\{\delta(m,h) - 2\pi i^{-k} \sum_{\substack{c\geqslant 1\\ q\mid c}} c^{-1} S(m,h;c) J_{k-1}\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\Big\}$$

(we recall that the notation $q \mid c$, for integers q and c, means that q divides c, i.e., that c/q is an integer). This formula, which may seem a bit intimidating, involves a Bessel function J_{k-1}, for which one can for instance give a power-series expansion (which converges over the whole complex plane), namely

$$J_{k-1}(z) = \Big(\frac{z}{2}\Big)^{k-1} \sum_{n\geqslant 0} \frac{(-1)^n}{2^{2n} n!(n+k-1)!} z^{2n}.$$

More importantly, it involves a *Kloosterman sum*

$$(4) \qquad S(m,h;c) = \sum_{\substack{1\leqslant x<c\\ (x,c)=1}} \exp\Big(\frac{mx + h\bar{x}}{c}\Big), \qquad \text{where } x\bar{x} \equiv 1 \pmod{c}.$$

In Figure 2, we reproduce that part of Poincaré's computation where he writes down such a sum, denoted ΣE; notice that he does not say anything more about it than that "it is not zero in general" (" n'est pas nulle en général. ").

These sums, despite looking rather innocuous, were promised to a brilliant future. They first reappeared (independently of the work of Poincaré) in an important paper of H. Kloosterman, published in 1926 [Kl] (see Figure 2). There, Kloosterman uses them to derive a remarkable arithmetic application, which amply justifies that the sums bear his name rather than Poincaré's: he manages to obtain an asymptotic formula for the number of representations of a large integer $n \geqslant 1$ by a positive-definite integral quadratic form in four variables, or in other words, for the number of solutions $(x_1, \ldots, x_4) \in \mathbf{Z}^4$ of an equation

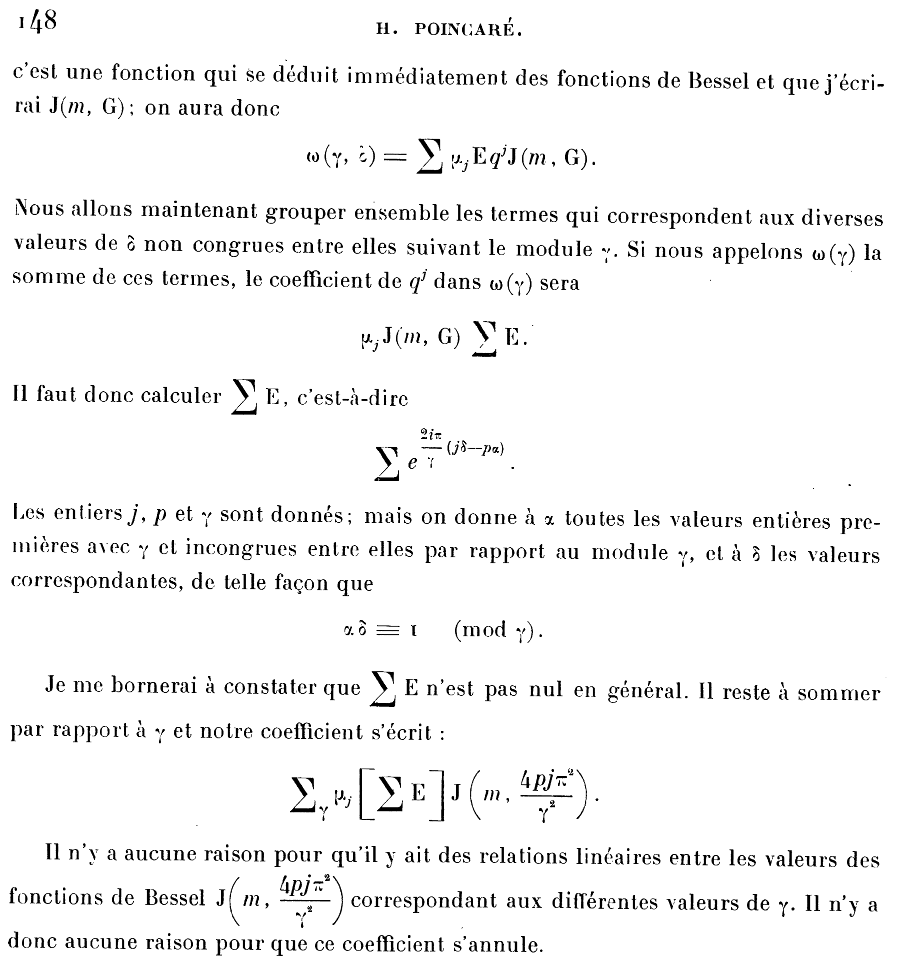
$$(5) \qquad n = a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_4^2,$$

where a_i, $1 \leqslant i \leqslant 4$, are fixed positive integers. The underlying context behind this result is the *circle method* of Hardy, Ramanujan and Littlewood. Indeed, Kloosterman describes a new, particularly refined, variant of this method. The original one had allowed Hardy and Littlewood to give a new solution of Waring's problem (the original one being due to Hilbert): for any fixed integer $k \geqslant 1$, there exists an integer $g = g_k \geqslant 1$ such that any integer $n \geqslant 1$ can be written as a sum of g integers which are k-th powers of non-negative integers:

$$n = x_1^k + \cdots + x_g^k, \text{ with } x_i \geqslant 0.$$

For $k = 2$, the circle method only led to the upper bound $g_2 \leqslant 5$ (at least for the problem of representing integers n which are "large enough"), whereas it had been known since the work of Lagrange that any positive integer is the sum of (at most) four squares, or in other words that $g_2 = 4$. Thus, Kloosterman's Theorem shows that the analytic approach of Hardy and Littlewood can be extended to reach a parity of results with the methods, more algebraic in nature, which were used to prove Lagrange's Theorem.

The crucial point in Kloosterman's argument was to prove a non-trivial estimate for the size of a sum $S(m,n;c)$ when m, n, c are pairwise coprime. The case when $c = p$ is a prime number is the most important, and one can reduce to it fairly easily. Then, one may notice that the definition (4) expresses $S(m,n;p)$ as the sum of $p-1$ complex numbers, each of which has modulus 1, but with (apparently) random arguments, because



148 H. POINCARÉ.

c'est une fonction qui se déduit immédiatement des fonctions de Bessel et que j'écrirai $J(m, G)$; on aura donc

$$\omega(\gamma, \delta) = \sum \mu_j E q^j J(m, G).$$

Nous allons maintenant grouper ensemble les termes qui correspondent aux diverses valeurs de δ non congrues entre elles suivant le module γ. Si nous appelons $\omega(\gamma)$ la somme de ces termes, le coefficient de q^j dans $\omega(\gamma)$ sera

$$\mu_j J(m, G) \sum E.$$

Il faut donc calculer $\sum E$, c'est-à-dire

$$\sum e^{\frac{2i\pi}{\gamma}(j\delta - p\alpha)}.$$

Les entiers j, p et γ sont donnés; mais on donne à α toutes les valeurs entières premières avec γ et incongrues entre elles par rapport au module γ, et à δ les valeurs correspondantes, de telle façon que

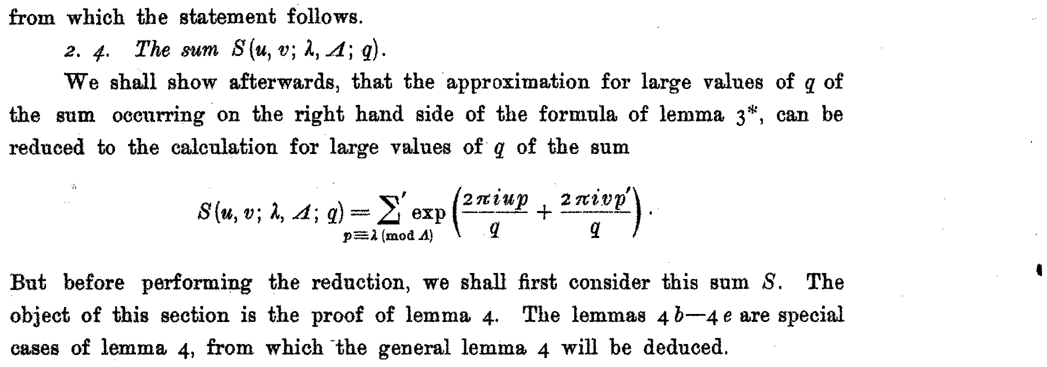
$$\alpha\delta \equiv 1 \pmod{\gamma}.$$

Je me bornerai à constater que $\sum E$ n'est pas nul en général. Il reste à sommer par rapport à γ et notre coefficient s'écrit :

$$\sum_\gamma \mu_j \left[\sum E\right] J\left(m, \frac{4pj\pi^2}{\gamma^2}\right).$$

Il n'y a aucune raison pour qu'il y ait des relations linéaires entre les valeurs des fonctions de Bessel $J\left(m, \frac{4pj\pi^2}{\gamma^2}\right)$ correspondant aux différentes valeurs de γ. Il n'y a donc aucune raison pour que ce coefficient s'annule.

FIGURE 1. Poincaré's article where a Kloosterman sum appears



from which the statement follows.

2. 4. The sum $S(u, v; \lambda, \Delta; q)$.

We shall show afterwards, that the approximation for large values of q of the sum occurring on the right hand side of the formula of lemma 3*, can be reduced to the calculation for large values of q of the sum

$$S(u, v; \lambda, \Delta; q) = \sum_{p \equiv \lambda \,(\mathrm{mod}\, \Delta)}' \exp\left(\frac{2\pi i u p}{q} + \frac{2\pi i v p'}{q}\right).$$

But before performing the reduction, we shall first consider this sum S. The object of this section is the proof of lemma 4. The lemmas 4 b—4 e are special cases of lemma 4, from which the general lemma 4 will be deduced.

FIGURE 2. Kloosterman sums in Kloosterman's paper

when x ranges over the integers from 1 to $p-1$, the inverse $\bar{x}$ of x modulo p varies rather chaotically. Except for some obvious symmetries, one may view the computation of the Kloosterman sum as realizing a kind of "random walk" in the plane. In Figure 3, we show

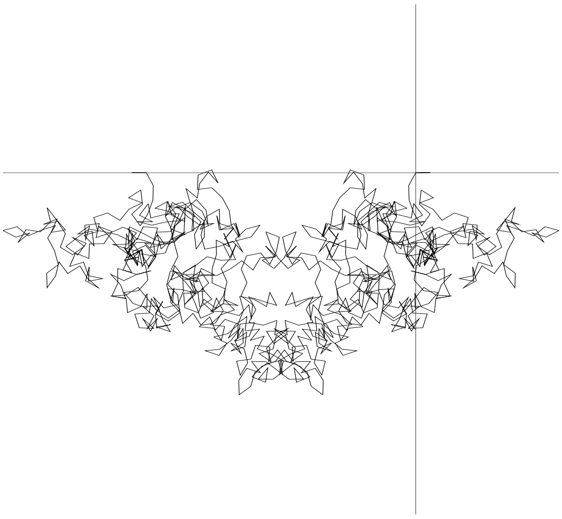


FIGURE 3. The Kloosterman sum $S(1, 1; 1021) = -18,608411\ldots$

such a path for the sum $S(1, 1; 1021)$: starting from the origin in the plane, line segments are drawn linking the points corresponding to the partial sums of the Kloosterman sum.

Because of this description, it is therefore natural to expect that $S(m, n; p)$ should have modulus quite a bit smaller than what the obvious bound – the triangle inequality – suggests, namely $|S(m, n; p)| \leqslant p - 1$. Indeed, in the present case, Kloosterman proves the following estimate:

$$|S(m, n; p)| < 2p^{3/4}.$$

Those readers who are keen on probability theory probably wonder now if the exponent 3/4 is the best possible, and may well have guessed that it should be possible to replace it by 1/2. This is indeed the case, but proving the estimate

$$|S(m, n; p)| \leqslant 2\sqrt{p}. \quad \text{(if } p \text{ does not divide } nm), \tag{6}$$

which is due to A. Weil, is much more difficult: it is a corollary of a special case of the Riemann Hypothesis for algebraic curves over finite fields, which was proved in general by Weil between 1940 and 1948.

On the other hand, it is fairly easy to check that the exponent 1/2 is indeed best possible, if one is allowed to vary m and n: precisely, one can always find some m such that $|S(1, m; p)| > \sqrt{2p - 2}$.

This being said, a very interesting feature can already be found in this first arithmetic application of Kloosterman sums: it is not really an individual sum which occurs in the application, but a *sum* of Kloosterman sums, possibly weighted with a "test function". One may then look at (3) slightly differently than before – remembering that equality is a symmetric relation –, and see this formula as expressing a weighted sum (we consider the case $q = 1$ to clarify the discussion) of Kloosterman sums like

$$\sum_{c \geqslant 1} c^{-1} S(m, n; c) f\Big(\frac{4\pi\sqrt{mn}}{c}\Big), \qquad f(x) = J_{j-1}(x), \tag{7}$$

in terms of Poincaré series, and therefore in terms of automorphic forms (of weight k). Note then that, by linearity, it is possible to obtain a "sum formula" for any function f which is a linear combination (or an infinite series) of Bessel functions J_{k-1}, assuming there is no convergence problem:

$$f(x) = \sum_{\substack{k\geqslant 4\\ k \text{ even}}} \alpha_k J_{k-1}(x).$$

However, the sums of Kloosterman sums which occur "in (arithmetic) nature" have no particular reason to be of this shape, even allowing infinite sums. This can be guessed quickly by remarking that (if the series converges) we have

$$\int_0^{\infty} f(x)(J_{it}(x) - J_{-it}(x))\frac{dx}{x} = 0$$

for *any* Bessel function J_{it} with pure imaginary index, $t \in \mathbf{R}$ being fixed. This means that the functions belonging to the space spanned by these Bessel functions J_{it} can *never* be expressed as combinations of the J_{k-1}...[1]

So, can one complete the inventory of sum formulas in order to incorporate the "missing" functions? This is indeed possible, but this discovery did not come without delays and difficulties.

Following our thread, the necessary ingredient appeared in work of A. Selberg, who was studying (for reasons which were mostly unrelated) some automorphic functions satisfying (1), but which are not holomorphic. Instead, they are solutions of the eigenfunction equation for the hyperbolic Laplace operator: for a constant λ (which is necessarily nonnegative in the cases we consider, because of some growth conditions imposed on the functions), we have

$$-y^2\Big(\frac{\partial^2 f}{\partial x^2}(z) + \frac{\partial^2 f}{\partial y^2}(z)\Big) = \lambda f(z)$$

for any $z \in \mathbf{H}$. (Such functions had been introduced originally by H. Maaß). Selberg introduced in his work a new type of Poincaré series, which avoids the divergence of those we have seen before when $k = 0$. The Poincaré series defined by Selberg depends on an "auxiliary" complex variable s,[2] and is given by

$$P_m(z;s) = \sum_{g\in B\backslash\Gamma_0(q)} \operatorname{Im}(\gamma\cdot z)^s \exp(2i\pi m\gamma\cdot z)$$

(here again, restricting to groups $\Gamma_0(q)$ is only a matter of simplification). Like the series considered earlier, it is possible to perform a Fourier expansion, since these functions are still 1-periodic. One obtains the relation

$$P_m(z;s) = \sum_{h\in\mathbf{Z}} p_m(h;y,s)\exp(2i\pi hz)$$

where

$$p_m(h;y,s) = \delta(m,h) + \sum_{\substack{c\geqslant 1\\ q\mid c}} c^{-2s} S(m,n;c) B(m,h,c,y,s)$$

and

$$B(m,h,c,y,s) = y^s \int_{-\infty}^{+\infty} (x^2+y^2)^{-s} \exp\Big(-2i\pi\Big(hx + \frac{m}{c^2(x+iy)}\Big)\Big)dx;$$

[1] There are authors who try to discourage the use of the (uncountably many) existing formulas for various integrals and transformations of Bessel functions; one should resist such temptation: who refrains from looking into tables of integrals and series risks missing quite a few discoveries...

[2] Using a factor of this type depending on s to "regularize" a diverging series was an idea introduced earlier by E. Hecke.

this is probably not so surprising anymore! (The dependency of the coefficient with respect to the variable y is more complicated than in the previous case, because the Poincaré series are not holomorphic anymore).

We now have an additional parameter (namely, s) and therefore many new sums of Kloosterman sums. Are there sufficiently many of them to represent the test functions arising in applications, for instance by a suitable (weighted) average over certain values of s? The answer is yes, because when $s = 1 + it$, it turns out that the (generally inextricable) function B is close to the functions J_{it}, and using an old theorem of Sears and Titchmarsh, one can see that any "reasonable" function defined on $[0, +\infty[$ can be expressed as the sum of a linear combination of the J_{k-1} and a weighted integral of the functions $J_{it} - J_{-it}$.

This step (which requires fairly deep arguments to be performed rigorously, and is quite delicate from the analytic point of view) was taken first by N. Kuznetsov (and independently by R. Bruggeman) in 1977, 1980, when $q = 1$, and then by Deshouillers and Iwaniec in 1982 for a general q. The *Kuznetsov formula* states that (7) can be expressed as a mixture of sums and integrals like

$$\sum_j \lambda_j(m)\overline{\lambda_j(n)}\hat{f}(t_j) + \int_{-\infty}^{\infty} \rho(m; \tfrac{1}{2} + it)\overline{\rho(n; \tfrac{1}{2} + it)}\hat{f}(t)dt,$$

involving a certain integral transform $\hat{f}$ of f (which is more complicated than the Fourier transform), and involving the Fourier coefficients $\lambda_j(n)$, $\rho(n; \frac{1}{2} + it)$ of automorphic functions and forms chosen to form an orthonormal basis of their respective spaces of automorphic functions and forms.

The arithmetic interest of such a formula brings us back to Kloosterman: as already described, it is more usual to have to consider, in applications, a *sum* of Kloosterman sums, rather than a single one. Even if individual bounds (especially Weil's bound (6)) can lead to very interesting results (for instance, to Kloosterman's first Theorem), having such general formulas available suggests the possibility of proving deep new results previously out of reach, by opening a door to an analysis of the oscillations of $S(m, n; c)$ as c varies... provided of course that one can understand the coefficients $\lambda_j(n)$ and $\rho(n; \frac{1}{2} + it)$, or in other words that one can understand well enough the space of automorphic functions and forms.

The great interest there would be to be able to understand these oscillation, for natural arithmetic applications, had been raised by Y. Linnik in his ground-breaking lecture at the International Congress of Mathematicians in Stockholm in 1962. He formulated a conjecture (see Figure 4) which is often stated in the following simplified form: for any fixed m and n, for any $\varepsilon > 0$, there exists a constant $C(\varepsilon)$, which may depend on m and n, such that

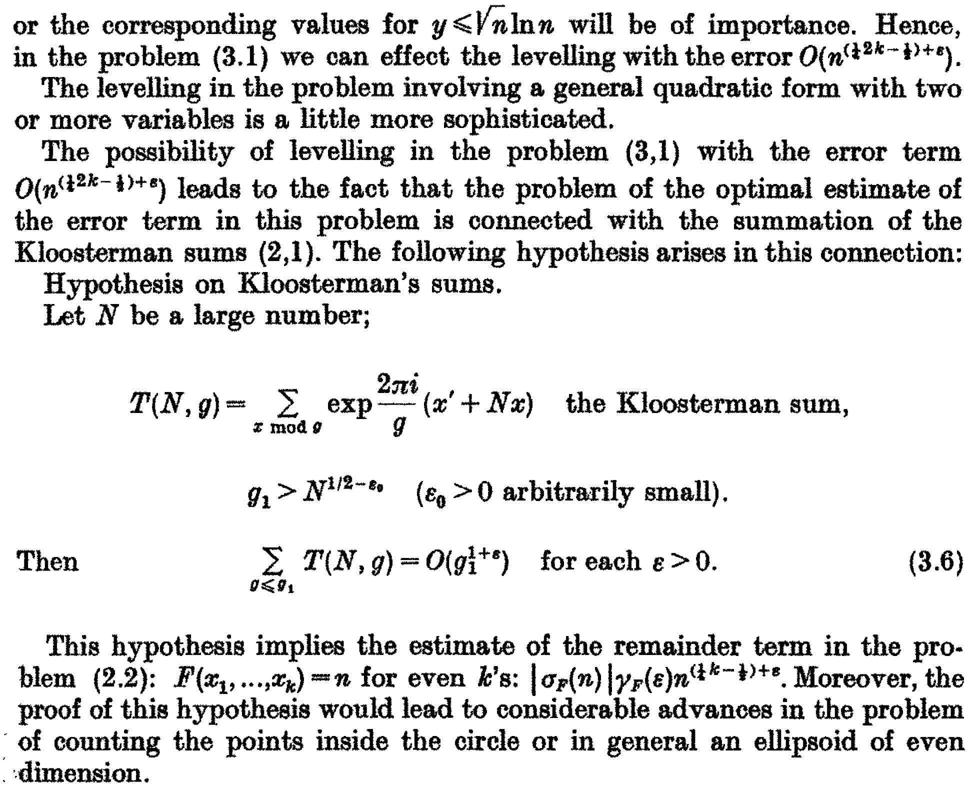
$$\Bigl|\sum_{c \leqslant X} c^{-1} S(m, n; c)\Bigr| \leqslant C(\varepsilon) X^{\varepsilon}, \qquad \text{for all } X \geqslant 1;$$

here again, if one thinks of $S(m, n; c)/\sqrt{c}$ as being a "random variable" which is essentially bounded but has changing sign, this prediction can be justified on probabilistic grounds. This conjecture, at least in a "smooth"[3] form, follows easily from the Kuznetsov formula.[4]

[3] I.e., after replacing the sum over $c \leqslant X$ by a sum of the type $\sum c^{-1} S(m, n; c) \exp(-c/X)$, which dampens the effect of the extreme terms and avoids purely analytic difficulties.

[4] It is interesting to note that it is only quite recently that É. Fouvry and P. Michel have succeeded in showing that this result is really a confirmation of the randomness of the signs of the Kloosterman sums, and not the effect of an hypothetical estimate of the type

$$|S(m, n; c)| \leqslant C(m, n) c^{1/2-\gamma}$$



or the corresponding values for $y \leqslant \sqrt{n}\ln n$ will be of importance. Hence, in the problem (3.1) we can effect the levelling with the error $O(n^{(\frac14 2k-\frac12)+\varepsilon})$.

The levelling in the problem involving a general quadratic form with two or more variables is a little more sophisticated.

The possibility of levelling in the problem (3,1) with the error term $O(n^{(\frac14 2k-\frac12)+\varepsilon})$ leads to the fact that the problem of the optimal estimate of the error term in this problem is connected with the summation of the Kloosterman sums (2,1). The following hypothesis arises in this connection:

Hypothesis on Kloosterman's sums.

Let N be a large number;

$$T(N,g) = \sum_{x \bmod g} \exp\frac{2\pi i}{g}(x' + Nx) \quad \text{the Kloosterman sum,}$$

$$g_1 > N^{1/2-\varepsilon_0} \quad (\varepsilon_0 > 0 \text{ arbitrarily small}).$$

Then
$$\sum_{g \leqslant g_1} T(N,g) = O(g_1^{\frac12+\varepsilon}) \quad \text{for each } \varepsilon > 0. \tag{3.6}$$

This hypothesis implies the estimate of the remainder term in the problem (2.2): $F(x_1,\dots,x_k) = n$ for even k's: $|\sigma_F(n)|\gamma_F(\varepsilon)n^{(\frac14 k-\frac12)+\varepsilon}$. Moreover, the proof of this hypothesis would lead to considerable advances in the problem of counting the points inside the circle or in general an ellipsoid of even dimension.

Figure 4. Linnik's conjecture

It would now be possible to describe the many applications of the summation formula for Kloosterman sums. We will only mention one however: it is an essential ingredient in the proof, by J.B. Conrey, of the fact that at least 60% of the zeros of the Riemann zeta function are located on the critical line.

But we will now go back to the original source of analytic number theory, the study of prime numbers...

3. Primes, arithmetic progressions, and a new hope

We recall that one denotes by $\pi(x)$ the number of prime numbers $p \leqslant x$. The most fundamental result in analytic number theory determines the *asymptotic* behavior of $\pi(x)$ as $x \to +\infty$:

Theorem 1 (Hadamard, de la Vallée-Poussin, 1895). *We have*

$$\lim_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} = 1,$$

which is also written

$$\pi(x) \sim \frac{x}{\log(x)} \quad \textit{as } x \to +\infty.$$

This result had been conjectured by Gauss, and Chebychev had given the first concrete evidence for it. The methods of Hadamard and de la Vallée-Poussin are fundamentally

for $c \geqslant 1$, with $\gamma > 0$, *for m and n fixed.* (Recall that if m and n can vary with c, the Weil bound is best possible.)

based on the revolutionary viewpoint of Riemann, namely, on the introduction of functions of one complex variable. There are now "elementary" proofs, but this adjective does not mean that they are simple.

It was quickly realized that, in this form, the result is not sufficient for many applications. Precisely, another basic question arises quickly: can one "count" similarly the primes satisfying additional conditions? For instance, can one count the primes $p \leqslant x$ for which the last (least significant) digit of p is equal to 1? Or more generally, count those p for which the remainder in the division of p by some integer $q \geqslant 1$ is equal to some fixed number a (the previous case corresponding to a remainder equal to 1 after division by 10); or even more, can one count the primes p for which $p-1$ is a square, i.e., those p such that $p = n^2 + 1$ for some integer $n \geqslant 1$; or can one count the primes $p \leqslant x$ such that $p+2$ is still prime (the "twin primes")?

These are many questions. Only the case where the constraint involves the remainder after some division is currently solved. Indeed, using ideas of Dirichlet (quite as revolutionary as those of Riemann, since they led to the theory of *representations of groups*), it was possible to adapt the methods used to prove the Prime Number Theorem to obtain:

Theorem 2. *Let $q \geqslant 1$ be a fixed integer, let $a \in \mathbf{Z}$ be a non-zero integer such that a is coprime with q. Then the number $\pi(x; q, a)$ of prime numbers $p \leqslant x$ such that $p - a$ is divisible by q, a property which is denoted $p \equiv a \pmod{q}$, and expressed as "p is congruent to a modulo q", satisfies*

$$\pi(x; q, a) \sim \frac{1}{\varphi(q)} \frac{x}{\log(x)} \sim \frac{1}{\varphi(q)} \pi(x) \text{ as } x \to +\infty,$$

where $\varphi(q)$, the Euler function, is the number of integers a such that $0 \leqslant a < q$ and a is coprime with q.

For instance, for $q = 10$, we have $\varphi(10) = 4$ (the possible values of a are $1, 3, 7, 9$), and this result, compared with the Prime Number Theorem, states roughly that the last (unit) digit of a prime number is equal to 1 for about 25% of primes, to 3 for another 25%, and similarly to 7 and 9 each for 25% (of course, if the last digit of p is even, the only possibility is that p be the unique even prime, namely $p = 2$).

But once more, when one wishes to apply this result in practice, in situations where other problems are reduced – one way or another – to considerations involving $\pi(x; q, a)$, it is almost the case that this result is grossly inefficient. In fact, the crucial issue is usually the uniformity of the dependency on q (most often) and a (less frequently... but see below...) of the "error term" $E(x; q, a)$ which appears after writing

$$\pi(x; q, a) = \frac{1}{\varphi(q)} \operatorname{li}(x) + E(x; q, a), \text{ where } \operatorname{li}(x) = \int_2^x \frac{dt}{\log(t)}.$$

The last function just introduced, which is often called the "logarithmic integral", satisfies the inequality

$$\left| \operatorname{li}(x) - \frac{x}{\log(x)} \right| \leqslant \frac{12x}{\log(x)^2}$$

for $x \geqslant 2$, and the Prime Number Theorem may just as well be expressed in the form

$$\pi(x) \sim \operatorname{li}(x) \text{ as } x \to +\infty.$$

This was indeed the way it had been predicted by Gauss, who interpreted $1/\log(x)$ as the "density" of primes around x, or as the "probability" that an integer of size x be prime.

The reason to use $\mathrm{li}(x)$ is that it makes the remainder term $E(x;q,a)$ much smaller (in other words, it gives a much better approximation of $\pi(x;q,a)$ than the elementary function $x/\log(x)$ does). One can, in particular, show fairly easily that the famous Riemann Hypothesis, generalized to the so-called Dirichlet L-functions, is *equivalent* to the following assertion: there exists a constant $C \geqslant 0$ such that

$$|E(x;q,a)| \leqslant Cx^{1/2}\log(x) \tag{8}$$

for any $q \geqslant 1$, any a coprime with q and any $x \geqslant 2$. What one should remember from this estimate is that the remainder $E(x;q,a)$ is *negligible* compared to the main term $\mathrm{li}(x)/\varphi(q)$, as long as, essentially, q is a bit smaller than $\sqrt{x}/\log(x)$.

Here is a simple and natural problem which leads quite directly to questions involving the dependency of $E(x;q,a)$ on q: the question is to estimate asymptotically the sum

$$S(x) = \sum_{p \leqslant x} d(p-1)$$

where p runs over primes and $d(n)$ denotes the number of positive divisors $d \geqslant 1$ of an integer $n \geqslant 1$. These sums may themselves appear in other applications, but one can also see the question as a simple test of our ability to understand the interaction between multiplicative properties of integers (as defining prime numbers and divisors) and their additive properties (involved in passing from p to $p-1$); or in other words: what (if any) are the specific multiplicative properties of integers of the form $p-1$?

This particular problem goes back to Titchmarsh, and it can be easily reduced to the study of $E(x;q,1)$ by writing

$$d(p-1) = \sum_{\substack{q,q' \geqslant 1 \\ p-1=qq'}} 1$$

and exchanging the order of the two sums:

$$\begin{aligned} S(x) &= \sum_{q \leqslant x-1} \sum_{\substack{p \leqslant x \\ p \equiv 1 \,(\mathrm{mod}\, q)}} 1 = \sum_{q \leqslant x-1} \pi(x;q,1) \\ &= \mathrm{li}(x)\Bigl(\sum_{q \leqslant x-1} \frac{1}{\varphi(q)}\Bigr) + \sum_{q \leqslant x-1} E(x;q,1). \end{aligned}$$

The first term is easy to handle and the difficulty is to deal with the other sum involving the remainders $E(x;q,1)$, where q may be very large compared with x. Using a trick of Dirichlet (if $ab=n$, one of the two divisors a or b is of size at most $\sqrt{n}$), one easily reduces to dealing only with $q \leqslant \sqrt{x}$ instead of $q \leqslant x$. But even then, *and even assuming the Generalized Riemann Hypothesis*, i.e., even by using (8), this sum is not well-controlled enough to conclude that the first term dominates.[5]

At this point, Kloosterman sums reappear. Using *sieve methods*, it is possible to analyze very precisely prime numbers by reducing their study to that of the sequences dd' of multiples of an integer d, *provided* this study is done uniformly with respect to d, where d is as large as possible compared with x, ideally $d \asymp \sqrt{x}$. But the equation $p \equiv a \,(\mathrm{mod}\, q)$, which can be expressed as $p-1=qq'$, becomes the simple equation

$$dd'-1=qq', \text{ or equivalently } dd'-qq'=1.$$

for these sequences of multiples.

[5] There is a certain analogy with the questions involving sums of Kloosterman sums: the Riemann Hypothesis is optimal if we allow q and a to vary, but for an average of remainder terms, one may expect compensation from changes of size and (especially) of sign.

Now the goal is to count the number of solutions of such an equation with high precision and uniformly when $dd' \leqslant x$. Why should Kloosterman sums occur? Note that $dd' - qq' = 1$ says that d' is the inverse of d modulo q, which recalls the definition (4). It is therefore not entirely surprising that, after appealing to various methods of harmonic analysis to "detect" the solutions of these equations, one ends up with sums very much like Kloosterman sums...

Well, this explanation is, obviously, only suggestive. But it is, indeed, by a reduction to sums of Kloosterman sums, and by an appeal to the results of Deshouillers and Iwaniec, obtained as consequences of the Kuznetsov formula generalized to $\Gamma_0(q)$, that Fouvry and Iwaniec first, and then Bombieri, Friedlander and Iwaniec, managed to *go beyond* the Riemann Hypothesis. The best result currently known takes the following shape: for any *fixed* integer $a \neq 0$, for any $A \geqslant 1$ and any $\varepsilon > 0$, there exists a constant $C(a, A, \varepsilon)$ such that, for any "well-factorable coefficients"[6] γ_q, and any $x \geqslant 2$, we have

$$\Big| \sum_{q \leqslant x^{4/7-\varepsilon}} \gamma_q E(x; q, a) \Big| \leqslant C(a, A, \varepsilon) \frac{x}{\log(x)^A}. \tag{9}$$

It is important to repeat: *even* under the Generalized Riemann Hypothesis, this result can not be derived by a direct individual estimation of $E(x; q, a)$. In final analysis, it depends of what is known about automorphic functions of (high) level q, and it particular on the properties (and the existence!) of Poincaré series.

We conclude by a very recent discovery due to Goldston, Pintz, and Yıldırım. By a new analysis of small gaps between primes, they have managed to prove the following remarkable result: *if* there exists $\delta > 0$ such

$$\sum_{q \leqslant x^{1/2+\delta}} \max_{y \leqslant x} \max_{(a,q)=1} |E(y; q, a)| \leqslant D(A) \frac{x}{\log(x)^A}, \tag{10}$$

holds for all $A \geqslant 1$ and $x \geqslant 2$, for a certain constant $D(A) \geqslant 0$,[7] then there exists at least one integer $k \geqslant 2$ for which infinitely many consecutive prime numbers satisfy $p' - p = k$. One expects, of course, that this last condition holds for any even integer k, with $k = 2$ corresponding to the Twin Prime Conjecture, but until this work, there were no clear links between these problems and well-established results of analytic number theory (with the exception of sieve methods).

Alas, (9) is not the same as (10) for $\delta = 4/7 - 1/2$! Indeed, in (10), one must be able to work with any a (instead of fixed a), and any interval $p \leqslant y \leqslant x$, not only $p \leqslant x$; and finally, the implicit coefficient $\gamma_q = \text{sign}(E(y; q, a))$ is *not* well-factorable...

However, there is no doubt that these differences are under close scrutiny, from the mathematicians already mentioned and from many others. Maybe new surprises will arise out of this devout attention...

References

[DI] J.M. Deshouillers and H. Iwaniec: *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. math. 70 (1982), 219–288.

[IK] H. Iwaniec and E. Kowalski: *Analytic Number Theory*, A.M.S Colloquium Publ. 53 (2004).

[6] This notion, which has been introduced by Iwaniec in sieve methods, is somewhat technical, and we omit the definition; the essential point is that the averages of $E(x; a, q)$ which occur in applications come most often with such weights.

[7] When $\delta = -\varepsilon$, with $\varepsilon > 0$ as small as desired, this statement still makes sense, and it is known: it is the famous Bombieri-Vinogradov Theorem, which follows also from (and is not stronger than) the Riemann Hypothesis.

[Kl] H.D. Kloosterman: *On the representation of numbers in the form $ax^2 + by^2 + cz^3 + dt^2$*, Acta Math. 49 (1926), 407–464.

[P1] H. Poincaré : *L'avenir des mathématiques*, Atti IV Congr. Internaz. Matematici, Roma, 1908 (dans Oeuvres, t. 5, p. 22).

[P2] H. Poincaré : *Fonctions modulaires et fonctions fuchsiennes*, Ann. Fac. Sci. Toulouse 3ème série, t. 3 (1911), 125–149 ; accessible sur `http://numdam.org` or in Oeuvres, t. 2, p. 592–618.

[P3] H. Poincaré : *Extension aux nombres premiers complexes des théorèmes de M. Tchebicheff*, Journal de Math. 4ème série, t. 8 (1891), 25–28 ; or in Oeuvres, t. 5, p. 442–479.

[P4] H. Poincaré : *Sur les invariants arithmétiques*, J. für die reine und angew. Math. Bd. 1829, Ht. 2 (1905), 89–150 ; or in Oeuvres, t. 5, p. 203–266.

[S] P. Sarnak: *Some applications of modular forms*, Cambridge Tract Math. 99, Cambridge Univ. Press (1990).

ETH Zürich – DMATH, Rämistrasse 101, 8092 Zürich, Switzerland
E-mail address: kowalski@math.ethz.ch