Ramifications of Probabilistic Number Theory

P. D. T. A. ELLIOTT*

University of Colorado Boulder, 395 UCB, Boulder, CO 80309-0395

1. Introduction

This paper is based upon a talk that I gave at the conference on Number Theory held in Kyoto, October 23–27, 2000. I thank the organizers, Murata, Leo and Wakabayashi, Isao, for their kind invitation. I thank the organisers and Motohashi, Yoichi for their financial support. Moreover, I thank Fujiwara, Masahiko, Motohashi, Yoichi and Murata, Leo and all the other participants of the meeting for their hospitality and for their kindness to me and my family.

Since the aim of the meeting was to look forward into the next century as well as back into the twentieth, I chose to illustrate the mathematical influence of ideas of probability in number theory as manifest in the study of additive and multiplicative functions. The theory of probability was not axiomatised until the nineteen thirties, so its influence could hardly have been formalised before the twentieth century. In this sharp sense the aesthetic that it brings is relatively new.

As a background I begin with a notion already to be found in the work of Euler.

2. The method of Dirichlet series

To each sequence of complex numbers a_n, $n = 1, 2, \ldots$, one may attach the Dirichlet series

$$g(s) = \sum_{n=1}^{\infty} a_n n^{-s}, \quad s = \sigma + i\tau, \quad \sigma = \mathrm{Re}\,(s).$$

If the series converges at some point, then it does so absolutely in a half-plane, and defines there an analytic function of s. If, further, the members of the sequence are the values of a multiplicative function, i.e. satisfy $a_{mn} = a_m a_n$ whenever $(m, n) = 1$, then the series possesses an Euler product. The paradigm is given by

*Partially supported by NSF contract DMS–0070496.

a_n identically 1, and the resulting Riemann zeta function has a representation

$$\zeta(s) = \prod (1 - p^{-s})^{-1}, \quad \sigma > 1,$$

the product taken over the primes.

Formally, the a_n may be recovered from a knowledge of the function $g(s)$ by means of Fourier analysis in the complex plane or, more accurately, on the multiplicative group of positive reals:

$$\sum_{n \le x} a_n = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} g(s) \frac{x^s}{s} ds, \quad c > 0,$$

where x is not an integer. A standard procedure is to analytically continue $g(s)$, say by an integral representation, choose c sufficiently large and move the line $\text{Re}\,(s) = c$ to the left, passing over or around the various singularities of $g(s)$.

This procedure, begun by Riemann in the 19th century, is still vital, the analytic continuation of $g(s)$ perhaps achieved using the resolvent operator of a Laplacian rather than with an integral representation. However, it is easy to give examples of Dirichlet series absolutely convergent in the half-plane $\sigma > 1$ and possessing an Euler product there, but for which the line $\text{Re}\,(s) = 1$ is a natural boundary. Indeed, a few irregularly large $|a_n|$ may spoil all hope even of the convergence of the attached Dirichlet series.

3. The discipline of Probabilistic Number Theory

Although in this paper I am concerned with the ramifications of Probabilistic Number Theory, it seems appropriate to give a simplified, three step description of the birth of the discipline.

In 1917, Hardy and Ramanujan proved that $\omega(n)$, the number of distinct prime divisors of the integer n, has *normal order* $\log\log n$, i.e. *if* $\nu_N(n; \ldots)$ *denotes the number of positive integers in the interval* $[1, N]$ *that possess the property* $\ldots$*, then for each* $\varepsilon > 0$

$$\lim_{N \to \infty} \nu_N(n; |\omega(n) - \log\log n| > \varepsilon \log\log n) = 0.$$

Their argument was elementary and largely combinatorial in nature, [22].

Additive arithmetic functions, such as $\omega(n)$, are real valued and satisfy $f(ab) = f(a) + f(b)$ whenever the integers a and b are mutually prime. They are *completely additive* if we may omit the condition $(a, b) = 1$.

The well-known theorem of Erdős and Wintner, [20] 1939, asserts that *for an additive function f there is a distribution function $F(z)$ towards which the frequencies $\nu_N(n; f(n) \le z)$, $N = 1, 2, \ldots$, converge (weakly) if and only if the three series*

$$\sum_{|f(p)|>1} \frac{1}{p}, \qquad \sum_{|f(p)|\le 1} \frac{f(p)^2}{p}, \qquad \sum_{|f(p)|\le 1} \frac{f(p)}{p},$$

taken over the prime number, converge.

That the convergence of the three series is sufficient Erdős had established with an elementary and largely combinatorial argument. For its necessity Erdős and Wintner appeal to the celebrated Erdős–Kac theorem, [19]: *If the additive function f satisfies $f(p^m) = f(p)$, $m = 1, 2, \ldots$,*

$$B(N) = \left(\sum_{p \le N} f(p)^2 p^{-1} \right)^{1/2} \to \infty, \quad N \to \infty,$$

and we define

$$A(N) = \sum_{p \le N} f(p) p^{-1},$$

then

$$\nu_N(n; f(n) - A(N) \le zB(N)) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du, \quad N \to \infty.$$

Here the weak convergence of measures asserts that the limiting relation holds for all z, since the *normal law* towards which the frequencies converge is everywhere continuous.

It is interesting to observe the development of the new aesthetic. In their 1917 paper Hardy and Ramanujan wished to draw attention to the relative infrequency of those integers n for which $\omega(n)$ is appreciably large or rather small, and such integers exist. At that time the theory of Probability lacked a sound foundation and Hardy, who had worked hard to introduce clarity and rigour into the teaching of mathematics in England, no doubt viewed the ideas of Probability theory with suspicion. Hardy and Ramanujan express their result in terms of *asymptotic density*, a notion familiar from Analytic Number Theory and the study of primes.

By the nineteen thirties, in the work of Erdős and others, the parameter z has been introduced and the weak convergence of the measures induced by the frequencies $\nu_N(n; f(n) \leq z)$ is being studied. It should be mentioned that in his work on additive functions Erdős arranged that the limit law would be continuous, and the weak convergence became convergence for each fixed value of z. Implicitly these results may still be viewed in terms of asymptotic density.

The form of the Erdős–Kac theorem is determined by Kac' appeal to a version of the Central Limit Theorem from the theory of probability proper. The additive function f is modelled by a sum of independent random variables X_p, one for each prime p, where X_p assumes the value $f(p)$ with probability $1/p$, zero with probability $1-1/p$. Erdős implements Kac' idea that divisibility by differing primes be viewed as independent events; he uses the sophisticated elementary sieve of Brun. It was possible to employ the once ill-defined notions of independence and random variables because in 1933 a satisfactory axiomatisation of the theory of probability had been given by A. N. Kolmogorov. The 'renormalising constants' $A(N), B(N)$ in the frequencies

$$\nu_N(n; f(n) - A(N) \leq zB(N)), \quad N = 1, 2, \ldots,$$

depend upon the parameter N; asymptotic density has been supplanted by the (weak) convergence of distribution functions.

In this new aesthetic very large but reasonably rare members of a real number sequence b_n, $n = 1, 2, \ldots$, hardly affect the limiting behaviour of the frequencies $\nu_N(n; b_n \leq z)$, $N = 1, 2, \ldots$. The same may be said concerning short sequences $h_n(N)$, $n = 1, \ldots, N$, and the weak convergence of the associated frequencies $\nu_N(n; h_n(N) \leq z)$, $N = 1, 2, \ldots$. The Erdős–Kac theorem corresponds to the case $h_n(N) = B(N)^{-1}(f(n) - A(N))$, $n = 1, \ldots, N$.

Before the axiomatisation of probability there were two well-developed methods for establishing the convergence of a distribution function to the normal or to some other law, and these methods continued to be vital following the acceptance of Kolmogorov's axioms. The second method, developed by Liapounov at the beginning of the twentieth century, required the asymptotic estimation of the Fourier transform of the relevant function, here the mean-value

$$N^{-1}\sum_{n=1}^{N}\exp(ith_n(N)), \qquad t \text{ real},$$

with $N = 1, 2, \ldots$. Note that each term in a typical sum lies in the complex unit circle. The first and earlier method was introduced by Chebyshev in the second half of the nineteenth century: estimate asymptotically the moments

$$N^{-1}\sum_{n=1}^{N}h_n(N)^k$$

for each positive integer k. There are further requirements in each method but these I omit.

Three examples of what might be viewed as an influence of this nexus of ideas now follow. That Erdős, Pal plays a rôle in each of them may not surprise.

4. A first example: character sums

The reduced residue class group to a prime modulus is cyclic. An estimate for the least positive representative of a generator for the group that is anywhere near the estimate guaranteed by the Riemann hypothesis for Dirichlet L-series seems very far away.

In fact an old conjecture of I. M. Vinogradov that $n_2(p)$, the least positive integer that is not a square $(\operatorname{mod} p)$ is $O(p^\varepsilon)$ for each fixed $\varepsilon > 0$ and all odd primes p seems hardly nearer. In 1919, Pólya and Vinogradov independently established the inequality

$$\left|\sum_{n\le y}\chi(n)\right| \le cq^{1/2}\log q$$

for some constant c, valid uniformly for all real y and all non-principal Dirichlet characters $(\operatorname{mod} q)$. The Pólya–Vinogradov inequality may be employed to establish Vinogradov's 1919 bound $n_2(p) \ll p^{1/(2\sqrt{e})}(\log p)^2$ and, indeed, slightly better.

With the uniformities given, apart from a replacement of $c\log q$ by a constant multiple of $\log\log q$, the bound of the Pólya–Vinogradov inequality cannot be improved. Nevertheless, it seems likely that there is much cancellation between the values of Dirichlet characters even over an interval short compared to $q^{1/2}$. As a consequence, $n_2(p) \ll (\log p)^{1+\varepsilon}$ for each fixed $\varepsilon > 0$ may be valid.

In a study of quadratic and power residues Davenport and Erdős, [5] 1952, considered the value distribution of the sums

$$S_h(x) = \sum_{n=x+1}^{x+h} \left(\frac{n}{p}\right),$$

where $\left(\dfrac{n}{p}\right)$ is the Legendre symbol and $p$ is a prime. They proved that *if* $h \to \infty$, $\log h/\log p \to 0$ *as* $p \to \infty$, *then for each fixed* λ

$$\frac{1}{p} \sum_{\substack{x=0 \\ S_h(x)\le \lambda h^{1/2}}}^{p-1} 1 \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-u^2/2} du.$$

Their argument employs the

Lemma. *If* $0 < h < p$, *and* $r$ *is a fixed positive integer, then*

$$\sum_{x=1}^{p} (S_n(x))^{2r} = 1.3 \cdots (2r-1)(p-\theta r)(h-\theta' r)^r + O(h^{2r}p^{\alpha_r}),$$

and

$$\sum_{x=1}^{p} (S_h(x))^{2r-1} = O(h^{2r}p^{\alpha_r}),$$

where α_r *depends only upon* $r$, $0 < \alpha_r < 1$, $0 \le \theta \le 1$, $0 \le \theta' \le 1$.

They mention that the work of Weil would allow a smaller α_r. It would be of no advantage to them.

Under their hypotheses on $h$ and the prime moduli $p$,

$$p^{-1} \sum_{x=1}^{p} (h^{-1/2} S_h(x))^k \to c_k, \quad k = 1, 2, \ldots,$$

where c_k is the kth moment of the normal distribution with mean zero and variance

By the time of this paper, Erdős had founded Probabilistic Number Theory with Kac and Wintner, had launched the invariance principle in probability proper with Kac, had refined the law of the iterated logarithm and had worked with Dvoretski and Kakutani in Brownian motion. Moreover, if s is real and $s > 3/4$, then in 1951, in a joint paper with Chowla, [4], Erdős had established the existence of a continuous limiting distribution for the frequencies amongst the integers d, not squares but satisfying $d \equiv 0$ or $1 (\bmod 4)$, of those for which the series $L_d(s) = \sum_{n=1}^{\infty} n^{-s}(d/n)$ formed with the Kronecker symbol (d/n) satisfies $L_d(s) < z$. It therefore seems likely that Erdős provided the aesthetic for this result with Davenport.

It should be mentioned that whilst establishing the convergence of arithmetically defined frequencies to the normal law by adopting the procedure of Chebyshev and estimating asymptotically their moments may seem natural, in the period when Davenport and Erdős were writing their paper such an approach perhaps did not automatically suggest itself. On page 355 of volume one of his *Collected Works* [26], A. Selberg states that in 1946 he did not know that asymptotic estimates that he had obtained for the moments of a quantity connected with the distribution of the zeros of the Riemann zeta function guaranteed convergence of corresponding frequencies to the normal distribution function.

I recall that at a tea before a number theory seminar in Cambridge in the early nineteen sixties, Davenport was discussing the above joint paper with Erdős. ... "We got", here Davenport paused, cocked his head on one side and with a twinkle in his eye continued: "precisely nowhere." So I think it would have been Davenport who added to their theorem on $S_h(x)$ the comment: 'It does not seem to throw any light on the problem of the magnitude of the least quadratic non-residue'.

Davenport gave to his student David Burgess the problem of improving these results and David did. By 1963, [3], David had established that

$$\sum_{A<n\leq A+H} \chi(n) \ll H^{1-1/r} p^{(r+1)/(4r^2)} \log p, \quad r = 1, 2, \ldots,$$

Uniform in A, H and non-principal characters $(\bmod p)$. In particular,

$$n_2(p) \ll (p^{1/4} p^{1/4r} (\log p)^r)^{1/\sqrt{e}}, \quad r = 1, 2, \ldots.$$

To establish his character sum estimate Burgess sets out from a version of the Davenport–Erdős lemma sharpened by the use of a complete character sum estimate of A. Weil. On the way Burgess adds ideas not appearing in the paper of Davenport and Erdős. Well might Davenport smile.

The consideration of sums

$$\sum_{x=1}^{p}\left|\sum_{n=x}^{x+h}\chi(n)\right|^{2r}$$

by Burgess is surely an outcome of the adoption of explicitly a method and implicitly an aesthetic from the theory of probability. In spite of Davenport's assessment, his work with Erdős partly turned interest away from the Fourier analytic methods employed by Pólya and Vinogradov, methods that seemed not to offer the opportunity of further advance.

The character sum estimate of Burgess has hardly been improved in forty years. Is this particular application of the ideas of probability to the estimation of an individual character sum in number theory isolated?

5. A second example: products of rationals

In the first example adoption of an approach from the theory of probability lead to progress in an existing problem. In the second example a result in probabilistic number theory appears to catalyze a new discipline: the arithmetical study of denumerably infinite abelian groups.

If $r_1, r_2, \ldots$ is a sequence of positive rationals, not necessarily distinct, we define Q^* to be the multiplicative group of positive rationals, Γ, the subgroup of Q^* generated by the r_j, G the quotient group Q^*/Γ. Since Q^* is free on the prime numbers, we may regard G as presented by the primes as generators, with relations $r_1 = 1$, $r_2 = 1, \ldots$, and so on. We may realise every denumerable abelian group in this manner.

A serious difficulty appears at once: there can be no recursive algorithm to decide whether an arbitrary positive rational is generated in Γ. The structure of G may not be generally discerned by computation; particular properties of the sequence r_j must come into play. What might these properties be?

From a number theoretical point of view it would be most interesting if the properties involved the arithmetic nature of the r_j. For example, let $h(x)$ be a rational function with rational number coefficients. For a positive integer $k$, define the r_i to be the values $|h(n)|$, $n = k+1, k+2, \ldots$. For $k$ large enough they will be defined. What is the form of $G$? If, further, the r_i are taken to be the sequence of values $|h(p)|$, where $p$ runs through the primes exceeding $k$, what then is $G$? In general, questions of this type appear difficult.

Consider the case $h(x) = x+1$. Every positive integer $m$ has a representation $(k+1)m(k+1)^{-1}$, and the group $G$ is trivial. If we specialize $x$ to prime values, then the group is almost certainly again trivial. Indeed, it would follow from an old conjecture of Dickson that every positive rational had infinitely many representations of the form $(p+1)(q+1)^{-1}$ with p, q prime. Currently the best that is known is that $G$ then has order at most 3, Elliott [13].

Consider next the case $h(x) = (ax+b)(Ax+B)^{-1}$ with integers $a > 0, b$, $A > 0, B$ for which $aB \neq Ab$, so that $h$ is not a constant. For $x$ confined to integer values it is known that the quotient group $G$ is finitely generated, Elliott [12]. It is not difficult to show that the torsion group of $G$ may have arbitrarily large order. When $x$ is confined to prime values nothing is known.

As a further example, set $h(x) = x^2+1$, $r_n = n^2+1$. An elementary argument by induction shows $G$ to be freely generated by the primes of the form $4m+3$. When $r_j = p_j^2+1$, with p_j the jth prime, nothing is known of $G$ beyond the obvious fact that if the (prime) generators $q \equiv 3 (\bmod 4)$ of Q^* remain free under the map $Q^* \to Q^*/\Gamma$.

What have such questions to do with the probabilistic theory of numbers? In a 1946 paper [18], Erdős proved that an additive function satisfies (what a probabilist would call the concentration function estimate)

$$\limsup_{N\to\infty} \sup_{k\in\mathbb{Z}} N^{-1} \sum_{\substack{n=1 \\ k<f(n)\leq k+1}}^{N} 1 > 0$$

if and only if for some constant c the series

$$\sum_{|f(p)-c\log p|>1} \frac{1}{p}, \qquad \sum_{|f(p)-c\log p|\le 1} \frac{(f(p)-c\log p)^2}{p},$$

taken over the prime numbers, converge. Such additive functions he called *finitely distributed.* As one of his corollaries he proved that the only (real-valued) additive functions f for which $f(n+1) - f(n) \to 0$ as $n \to \infty$ are those of the form $f(n) = c\log n$. In the same paper Erdős made several conjectures, one of which was that any additive function for which $f(n+1) - f(n)$ is uniformly bounded above must have the form $c\log n + O(1)$.

In a series of four papers at the end of the nineteen sixties, [23], [24], [25], [26], Kátai widened the scope of this and related problems raised by Erdős. In the third paper of this series Kátai asked for a characterization of those additive functions f which satisfy

$$f(an+b) - f(An+B) \to c \quad \text{as} \quad n \to \infty,$$

where the integers $a > 0, b, A > 0, B$ satisfy $aB \neq Ab$.

Eduard Wirsing settled Erdős' conjecture in the affirmative in 1968 [29]. Using a different method, I solved Kátai's problem in 1980 [10], [11]: on the integers prime to aA, f is a constant multiple of the logarithmic function.

In the first of his four papers Kátai had defined a sequence of positive integers to be a *set of uniqueness* if every completely additive functions which vanished at each member of the sequence vanished identically. In his second paper Kátai proved that adjoining finitely many integers to the set of shifted primes $p+1$ gave a set of uniqueness. His conjecture that the set of shifted primes alone is a set of uniqueness I established in 1974 [8]. I began by proving that integers of the form $(p+1)(q+1)^{-1}$, with p, q prime, have positive logarithmic density and showed that as a consequence any additive function vanishing on the shifted primes must be finitely distributed.

All these results may be viewed as an outcome of probabilistic number theory, induced consequences of Erdős' interest in the value distribution of additive functions. They were given large added value in 1978 when Dress and Volkmann [7], and Wolke [30], showed that a sequence a_n can be a set of uniqueness if and only if every positive integer m has a representation

$$m^h = \prod_{j=1}^{k} a_j^{d_j}$$

with h and the d_j integers. If Γ is the subgroup of Q^* generated by the a_n, then a completely additive function which vanishes on the a_n is a homomorphism from Q^*/Γ into the additive reals. In establishing Kátai's conjecture I had given a harmonic proof that every positive integer m had a representation of the form

$$m^h = \prod_{p \le P} (p+1)^{d_p},$$

with p prime. Probabilistic number theory began to bear upon the algebraic structure of denumerably infinite abelian groups.

My two volume book on probabilistic number theory [9], may be viewed as a treatment of additive and multiplicative functions informed by the theory of probability. My subsequent book on arithmetic functions and integer products [12], in the same series and in which I obtain results that settle the problem of Kátai and generalize and deepen the theorem of Wirsing, may be viewed as a treatment of additive and multiplicative functions informed by the algebraic study of denumerable abelian groups.

To give direction to this section I summarise with a central question. Let $F_i(x)$, $i = 1, \ldots, k$, be rational functions with rational coefficients. Let N be a positive integer and let Γ_i be the subgroup of Q^* generated by the $|F_i(n)|$, $n = N, N+1, \ldots$. It is assumed that N is large enough for the $F_i(n)$ to be defined and non-zero. *What is the direct product group*

$$\bigotimes_{i=1}^{k} Q^*/\Gamma_i?$$

A complex-valued multiplicative function satisfies $g(ab) = g(a)g(b)$ whenever a and b are mutually prime. If it is completely multiplicative, then we may omit the condition $(a, b) = 1$.

An equivalent question is then: *describe those completely multiplicative functions g_j, $j = 1, \dots, k$ with values in the group of roots of unity, that satisfy*

$$g_1(|F_1(n)|) \cdots g_k(|F_k(n)|) = 1$$

for all integers $n \geq N$.

This question is related to, but possibly easier than, that of deciding *when do the limits*

$$\lim_{x \to \infty} x^{-1} \sum_{n \leq x} g_1(|F_1(n)|) \cdots g_k(|F_k(n)|)$$

exist, particularly when are they zero? To follow tradition it would be more natural in this (harder) question to allow the g_j to assume values in the complex unit disc.

We may *ask similar questions with the integers n replaced by primes p.* In this case the weight x^{-1} in the limit would be replaced by $\pi(x)^{-1}$, where $\pi(x)$ counts the number of primes not exceeding x.

Answers to these questions may satisfy a group theorist, but as a number theorist I have further, particular interests. For simplicity of exposition, I revert to the case of a single group G, generated by the values of a single rational function $h(x)$. *Can one give a reasonably explicit upper bound for a set of coset representatives for G?*

This question bears upon a related question of independent interest: *How many terms does one need to effect a* (multiplicative) *representation of a given rational by values of $h(x)$ at the integers, and is there a number s so that no more than s terms are ever needed?*

It should be mentioned that there need not be a universal s. According to Theorem (19.2) of Elliott, [12], if the positive integers m_1 and m_2 satisfy $m_1 \equiv m_2 \equiv 1$, 4 or 7 $(\bmod 9)$, then there is a simultaneous representation

$$m_1 = \prod_{i=1}^{k} (3n_i - 17)^{\varepsilon_i}, \quad m_2 = \prod_{i=1}^{k} (3n_i + 19)^{\varepsilon_i},$$

with integers $n_i \geq 7$, $\varepsilon_i = \pm 1$. A positive integer $m \equiv 1 (\bmod 9)$ hence has a representation

$$m = \prod_{i=1}^{k} \left(\frac{3n_i - 17}{3n_i + 19} \right)^{\varepsilon_i}.$$

Since the absolute value of a typical term in the product does not exceed 10, the number of terms needed to so represent m satisfies $k \gg \log m$. Argument given in the earlier part of that same reference, particularly that in Chapter 4, shows that there is such a representation with $k \ll (\log m)^{\alpha}$, the implied constant depending upon α, for each (fixed) value of $\alpha > (\log 4/3)^{-1} \log 3$.

The properties of $h(x)$ needed to guarantee the existence of a universal s are quite unclear.

Let S_k denote the set of rationals that can be generated using a group product of at most k elements taken from a sequence $r_1, r_2, \ldots$. Pedro Berrizbeitia and I proved that in order for the corresponding group Q^*/Γ to be finite and for there to exist a universal s of the above type in this case, it is necessary and sufficient that the lower asymptotic density of the integers in the set $m^{-1}S_k$ be bounded below (and away from zero) for some k and all positive integers m, [2]. Moreover, if a value is known for the bound, then a value can be given for s. Here $m^{-1}S_k$ is formed from S_k by dividing each member of S_k by m.

When the f_j are given by the shifted primes $p+1$ I had established the existence of a universal s for which I could not give a value, [13]. Pedro and I could later prove that $s = 19$ is permissible, [2], a bound which I subsequently reduced to $s = 9$, [15]. Of course, s should be 2.

In these cases no method is given to provide a group product representation for a given rational. The same is true for the following result with which I end this section.

Let $w(x)$ be a polynomial with integer coefficients, leading coefficient positive. Suppose that $w(n)$ is positive for all $n \geq N$. Then the group $G_N = Q^*/\Gamma$, with Γ generated by the ratios $(p+1)w(q)^{-1}$ with p, q prime, $p, q \geq N$, has order at most 4.

If $M \geq N$, then Γ_M is a subgroup of Γ_N; G_M may be mapped homomorphically onto G_N and $G_M/(\Gamma_N/\Gamma_M) \simeq G_N$. In particular,

$$|G_M| = |G_N| \prod_{i=1}^{M-N} |\Gamma_{M-i}/\Gamma_{M-i+1}|$$

where each term in the product is an integer. It follows from the uniformity of the bound $|G_N| \leq 4$ with respect to N that for all N sufficiently large the groups G_N are isomorphic. I think that a similar phenomenon holds for many arithmetically defined groups. Let g be the order of the limiting group. Then every positive rational r has infinitely many representations

$$r^g = \prod_{j=1}^{d} \left(\frac{p_j + 1}{w(q_j)} \right)^{\varepsilon_j}, \quad \varepsilon_j = \pm 1,$$

with p_j, q_j prime and $d \leq 11$.

It seems to me that this area contains challenging questions and offers many interesting possibilities. General references may be found in my volume [12], mentioned above. I draw attention, also, to my paper in Acta Arithmetica [14], in which I discuss the possibility of obtaining (group) product representations using numbers of the form $q+1$ or $N-p$ with q, p prime, $p < N$, by employing integration with respect to the Haar measure on the group dual to Q^*.

6. A third example: the distribution of primes

In my Kyoto lecture I ran out of time and concerning this third topic I could only make a few remarks. In this paper I have much overrun my allotted space and must do likewise, postponing a fuller account to another occasion.

Following the approach of Liapounov, to establish the weak convergence of the frequencies $\nu_N(n; f(n) \leq z)$ for an additive function f is to establish the uniform convergence, on compact sets of real t values, of the characteristic functions

$$N^{-1} \sum_{n=1}^{N} \exp(itf(n)), \qquad N = 1, 2, \dots .$$

For each t the function $n \mapsto \exp(itf(n))$ is multiplicative. A treatment of the Erdős–Winter theorem along these lines was given by Delange [6], and had there been no other reason this would have sufficed to promote the investigation of the mean-values of multiplicative functions with values in the complex unit disc.

In the event, the breakthough work of Wirsing, [28], was clearly influenced by the elementary proof of the prime number theorem given by Erdős and Selberg in

1948. In turn, Wirsing's studies catalyzed the now well-known theorem of Halász that achieved a complete taxonomy of multiplicative functions g, with values in the complex unit disc, for which

$$\lim_{N\to\infty} N^{-1}\sum_{n=1}^{N} g(n)$$

exists, and in terms of their behaviour on the primes, [21]. Although the proof applies the theory of complex variables to study the Dirichlet series $\sum_{n=1}^{\infty} g(n)n^{-s}$, there is no appeal to an analytic continuation of the sum function of this series which may, indeed, not have one. The influence of Halász' argument on probabilistic number theory may be seen in my two volume work [9], and is touched upon in my plenary address to the conference in memory of Erdős, held in Budapest in July, 1999, [16].

Suffice it to say that combining ideas of Halász and Wirsing with other arguments, in particular using sieves, I was able in 2000 to give a new proof of Linnik's theorem that there is a universal c so that each reduced residue class (mod D) contains a prime not exceeding D^c. Moreover, the proof makes no use of the zeros of L-series in the so called 'critical strip' $0 < \mathrm{Re}(s) < 1$; there is no appeal to the Deuring–Heilbronn phenomenon, or to estimates for the density of the zeros of L-series, or even to zero-free regions of those series beyond their non-vanishing at the point $s = 1$, [17].

Perhaps a completely elementary proof of Linnik's theorem may be given.

Besides the explicit application of probability to number theory, there is the implicit influence of the field of probabilistic number theory itself. I think the ideas associated with this field are very far from exhausted.

References

[1] Berrizbeitia, P., Elliott, P.D.T.A. On products of shifted primes, *The Ramanujan Journal* **2** (1998), 219–223.

[2] Berrizbeitia, P., Elliott, P.D.T.A. Product bases for the rationals, *Bulletin Canadian Math. Soc.* **42** (1999), 441–444.

[3] Burgess, D.A. On character sums and $L$-series, II, *Proc. London Math. Soc.* (3) **13** (1963), 524–536.

[4] Chowla, S., Erdős P. A theorem on the distribution of the values of $L$-functions, *J. Indian Math. Soc.* **15** (1951), 11–18.

[5] Davenport, H., Erdős, P. The distribution of quadratic and higher residues, *Publ. Math. Debrecen* **2** (1952), 252–265.

[6] Delange, H. Sur les fonctions arithmétiques multiplicatives, *Ann. Scient. Éc. Norm. Sup. 3e série*, t. **78** (1961), 273–304.

[7] Dress, F., Volkmann, B. Ensembles d'unicité pour les fonctions arithmétiques additives on multiplivatives, *C. R. Acad. Sci. Paris, Sér. A* **287** (1978), 43–46.

[8] Elliott, P.D.T.A. A conjecture of Kátai, *Acta Arithmetica* **26** (1974), 11–20.

[9] ____________. *Probabilistic Number Theory, I: Mean Value Theorems, II: Central Limit Theorems*, Grund. der math. Wiss. **239**, **240**, Springer Verlag, Berlin, New York, 1979, 1980.

[10] ____________. On the distribution of the roots of certain congruences and a problem for additive functions, *J. Number Theory* **16** (1983), 267–282.

[11] ____________. On additive arithmetic functions $f(n)$ for which $f(an+b) - f(cn+d)$ is bounded, *J. Number Theory* **16** (1983), 285–310.

[12] ____________. *Arithmetic Functions and Integer Products*, Grund. der math. Wiss. **272**, Springer-Verlag, New York, Berlin, Tokyo, 1984.

[13] ____________. The multiplicative group of rationals generated by the shifted primes, I, *J. reine angew. Math.* **463** (1995), 169–216.

[14] ____________. Products of shifted primes: Multiplicative analogues of Goldbach's problem, *Acta Arithmetica* **88** (1999), 31–50.

[15] ____________. The multiplicative group of rationals generated by the shifted primes, II., *J. reine angew. Math.* **519** (2000), 59–71.

[16] ____________. 'Cast thy bread upon the waters...' A personal view of the mathematician Paul Erdős, Proceedings of the conference *Paul Erdős and his mathematics*, Budapest, July 1999, to be published by the J. Bolyai Society.

[17] ______________. The last prime primitive root and Linnik's theorem, to be published in the proceedings of *The Millenial Conference*, Urbana–Champaign, Illinois, 2000.

[18] Erdős, P. On the distribution function of additive functions, *Annals Math.* **47** (1946), 1–20.

[19] Erdős, P., Kac, M. On the Gaussian law of errors in the theory of additive functions, *Proc. Nat. Acad. Sci. U.S.A.* **25** (1939), 206–207.

[20] Erdős, P., Wintner, A. Additive arithmetical functions and statistical independence, *Amer. J. Math.* **61** (1939), 713–721.

[21] Halász, G. Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen, *Acta Math. Acad. Sci. Hung.* **19** (1968), 365–403.

[22] Hardy, G.H., Ramanujan, S. The normal number of prime factors of a number n, *Quart. J. Math.* (Oxford) **48** (1917), 76–92.

[23] Kátai, I. On sets characterizing number-theoretical functions, *Acta Arithmetica* **13** (1968), 315–320.

[24] ______________. On sets characterizing number-theoretical functions, II. (The set of "prime plus one"'s is a set of quasi-uniqueness), *Acta Arithmetica* **16** (1968),1–4.

[25] ______________. Some results and problems in the theory of additive functions, *Acta. Sc. Math. Szeged.* **30** (1969), 305–311.

[26] ______________, I. On number theoretical functions, *Colloquia Mathematica Societas Janos Bolyai*, Vol. 2, North Holland, Amsterdam 1970, 133–136.

[27] Selberg, A. *Collected Papers*, two volumes, Springer Verlag, Berlin, New York, Tokyo, 1989, 1991.

[28] Wirsing, E. Das asymptotische Verhalten von Summen über multiplikative Funktionen, II, *Acta Math. Acad. Sci. Hung.* **18** (1967), 411–467.

[29] ______________. A characterization of $\log n$ as an additive arithmetic function, *Symposia Mathematica IV*, 45–57, Academic Press, London and New York, 1970.

[30] Wolke, D. Bemerkungen über Eindeutigkeitsmengen additiver Funktionen, *Elem. der Math.* **33** (1978), 14-16.

<pdtae@euclid.colorado.edu>