Chapter 4
The Prime Distribution

> *"All this is amusing,*
> *though rather elementary*[1] *…, Watson."*
> *– Sir Arthur Conan Doyle*
> *(Sherlock Holmes)*

How are the primes distributed among the integers? Here "distribution" is a misleading term because a given positive integer either is a prime or is not a prime – there is nothing chancy about primality. Yet superficially, the occurrence of primes appears to be rather haphazard, and, indeed, many properties can be derived by playing "dumb" and assuming nothing more than that "every other integer is divisible by 2, every third is divisible by 3", etc., and letting complete randomness reign beyond the most obvious. The result of this loose thinking suggests that the average interval between two successive primes near n is about $\ln n$. This is not easy to prove rigorously, especially if one forgoes such foreign tools as complex analysis. Yet fairly simple probabilistic arguments come very close to the truth. In fact, probabilistic thinking as introduced here can reveal a lot about primality and divisibility [4.1], and we shall make ample use of the probabilistic approach throughout this book to gain an intuitive understanding of numerous number-theoretic relationships. For a formal treatment of probability in number theory see [4.2].

4.1 A Probabilistic Argument

Two facts about the distribution of the primes among the integers can be noticed right away:

1) They become rarer and rarer the larger they get.
2) Apart from this regularity in their mean density, their distribution seems rather irregular.

In fact, their occurrence seems so unpredictable that perhaps probability theory can tell us something about them – at least that is what the author thought in his second (or third) semester at the Georg-August University in Göttingen. He had just

[1] In number theory *elementary* methods are often the most difficult, see [4.5].

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_4, © Springer-Verlag Berlin Heidelberg 2009

taken a course in *Wahrscheinlichkeitsrechnung* at "Courant's" famous Mathematics Institute, and one afternoon in 1948, in the excruciatingly slow "express" train from Göttingen to his parents' home in the Ruhr, he started putting some random ideas to paper. His train of thought ran roughly as follows.

The probability that a given "arbitrarily" selected integer is divisible by p_i is $1/p_i$. In fact, starting with 1, precisely every p_ith number is divisible by p_i (every third is divisible by 3, every fifth by 5 and so forth). Thus, the "probability" that a given selected number is *not* divisible by p_i is $1 - 1/p_i$.

Assuming that divisibility by different primes is an *independent*[2] property, the probability that x is not divisible by any prime below it is given by the product

$$W(x) \approx \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\ldots \approx \prod_{p_1 < x}\left(1 - \frac{1}{p_i}\right). \tag{4.1}$$

If x is not divisible by any prime below it, it is, of course, not divisible by *any* smaller number, i. e., x is prime.

More strictly, we could limit the product to primes $p_1 < \sqrt{x}$ (see Sect. 3.2 on the sieve of Eratosthenes). In fact, in that 1948 train the author did limit the product to primes smaller than the square root of x. But since the end result is not much affected, we will not bother about this "refinement".

If one feels uncomfortable with a product, it can be quickly converted into a sum by taking (naturally) logarithms:

$$\ln W(x) \approx \sum_{p_i < x} \ln\left(1 - \frac{1}{p_i}\right). \tag{4.2}$$

If one does not like the natural logarithm on the right-hand side, expanding it and breaking off after the first term does not make much difference, especially for the larger primes:

$$\ln W(x) \approx -\sum_{p_i < x} \frac{1}{p_i}. \tag{4.3}$$

There is something about the sum that is still bothersome: it is not over consecutive integers, but only over the primes. How can one convert it into a sum over *all* integers below x? Again, one can use a probability argument: a given term $1/n$ in the sum occurs with probability $W(n)$. Thus, let us write (and this is the main trick here):

$$\ln W(x) \approx -\sum_{n=2}^{x} \frac{W(n)}{n}. \tag{4.4}$$

[2] Simultaneous independence for all primes is never exactly true, but there is near independence that suffices for our argument.

By now sums may have become boring, and one wishes the sum were an integral. Thus, we write with our now customary nonchalance:

$$\ln W(x) \approx -\int_2^x \frac{W(n)}{n}\,\mathrm{d}n. \tag{4.5}$$

The next thing that may strike one as offensive is the minus sign on the right-hand side. Introducing the *average distance* $A(x) = 1/W(x)$ between primes, we get a positive expression:

$$\ln A(x) \approx \int_2^x \frac{\mathrm{d}n}{nA(n)}. \tag{4.6}$$

Now, suddenly, the integral has served its purpose and can go; most people would rather solve differential equations than integral ones. Differentiating will of course be the appropriate integral vanishing trick:

$$\frac{A'(x)}{A(x)} \approx \frac{1}{xA(x)}, \quad \text{or} \tag{4.7}$$

$$A'(x) \approx \frac{1}{x}. \tag{4.8}$$

And the unexpected has happened: we have an answer (fortuitously correct)! The average distance between primes ought to be

$$A(x) \approx \ln x, \tag{4.9}$$

and the mean density becomes

$$W(x) \approx \frac{1}{\ln x}. \tag{4.10}$$

Example: $x = 20$, $\ln 20 \approx 3.00$, and, indeed, the average spacing of the 3 primes closest to 20, namely 17, 19 and 23, is exactly 3.

Around $x = 150$, the average spacing should be about 5, and in the neighbourhood of $x = 10^{50}$, every 115th number, on average, is a prime.

4.2 The Prime-Counting Function $\pi(x)$

If we accept the estimate (4.10) of the average prime density, the number of primes smaller than or equal to x, usually designated by $\pi(x)$, is approximated by the "integral logarithm":

$$\pi(x) \approx \int_2^x \frac{\mathrm{d}x'}{\ln x'} =: Li(x), \tag{4.11}$$

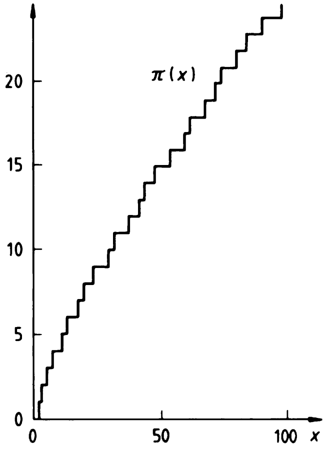


Fig. 4.1 The prime-counting function $\pi(x)$ for $2 \leq x \leq 100$

where the sign =: indicates that the notation $Li(x)$ is *defined* by the integral on the left.

The prime-counting function $\pi(x)$ is plotted in Fig. 4.1 for $x \leq 100$. Every time x equals a prime, $\pi(x)$ jumps up by 1. But apart from the "jumpiness" of $\pi(x)$, a smoother, slightly concave trend is also observable. This smoothness becomes more obvious when we plot $\pi(x)$ for x up to 55,000 as in Fig. 4.2. On this scale, the jumpiness has disappeared completely.

The inadequacy of Gauss's original estimate $\pi(x) \approx x/\ln x$ is illustrated by Fig. 4.3. By contrast, the integral logarithm, which we "derived" above (and which was also conjectured by Gauss) gives seemingly perfect agreement with $\pi(x)$ in the entire range plotted in Fig. 4.4.

However, even $Li(x)$, labelled "Gauss" in Fig. 4.5, shows noticeable deviations when we expand the ordinate by a factor 10^4 as was done in that figure (see [4.3]). In fact, for $x = 10^7$, the excess of $Li(x)$ over $\pi(x)$ is about 300 and remains positive for all $x < 10^9$. Nevertheless, $\pi(x) - Li(x)$ has infinitely many zeros, at least one of which occurs below $x = 10^{10^{10^{34}}}$; in fact, it may be near $x = 10^{370}$. (A number such as $10^{10^{10^{34}}}$, introduced by S. Skewes in 1933, was once considered a large number. But *much much much larger* numbers have now become important in connection with Gödel's famous "incompleteness" theorem [4.4].)

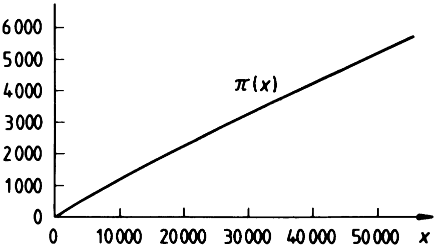


Fig. 4.2 Same as Fig. 4.1 but $x \leq 55{,}000$

Fig. 4.3 $\pi(x)$ and $x/\ln x$

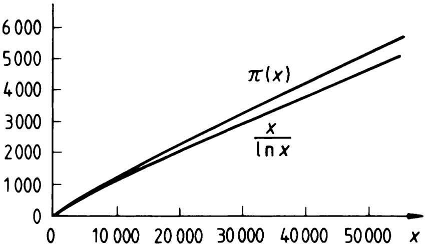
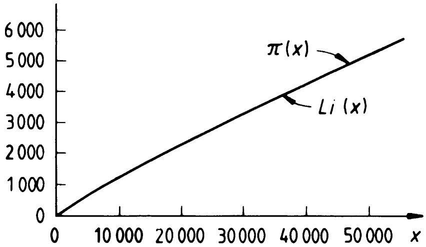


Fig. 4.4 $\pi(x)$ and the integral logarithm $Li(x)$



Legendre, independently of Gauss, gave the following formula in 1778:

$$\pi(x) \approx \frac{x}{\ln x - 1.08366}, \tag{4.12}$$

a closer approximation than (4.11) up to about $x = 4 \cdot 10^6$, as can be seen in Fig. 4.5. However, above $x = 5 \cdot 10^6$ Legendre's formula begins to go to pieces. (Expanding $Li(x)$ gives 1 as the constant in (4.12), but Legendre missed that.)

Either formula (4.11) or Legendre's (4.12) says that there are about $7.9 \cdot 10^{47}$ 50-digit primes – plenty to go around for the "trap-door" encryption schemes to be discussed later in Chap. 10.

In our "derivation" of $\pi(x)$, we considered primes up to x and pointed out that consideration of primes up to $\sqrt{x}$ would have sufficed. This idea was further pursued by Bernhard Riemann, who showed that

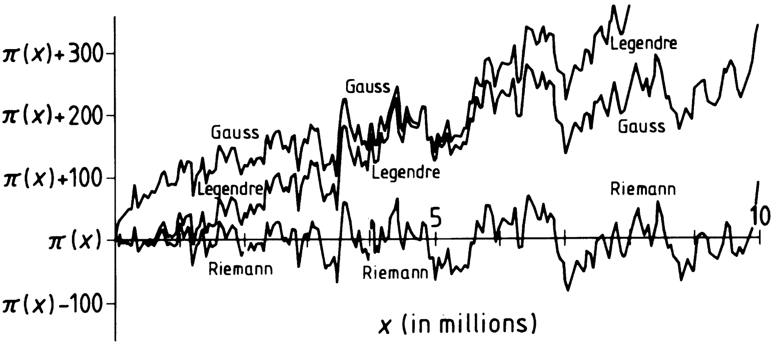


Fig. 4.5 Comparison of formulas by Gauss: $Li(x)$, Legendre: $x/(\ln x - 1.08366)$, and Riemann: $Li(x) - 1/2Li(x^{1/2}) - 1/3Li(x^{1/3}) - \ldots$ (Courtesy of D. Zagier)

$$\pi(x) \approx R(x) := Li(x) - \tfrac{1}{2}Li\left(\sqrt{x}\right) - \tfrac{1}{3}Li\left(\sqrt[3]{x}\right) - \ldots. \tag{4.13}$$

Figure 4.5 demonstrates how good an approximation $R(x)$ is; the curve labelled "Riemann" does not seem to have any deviant trend up to $x = 10^7$.

The closeness of $R(x)$ to $\pi(x)$ is further emphasized by Table 4.1, which shows that, even for $x = 10^9$, the error of $R(x)$ is only 79 (out of $5 \cdot 10^7$).

It is interesting to note that it was not until 1896, almost a hundred years after Gauss's and Legendre's conjectures, that Hadamard and de la Vallée Poussin proved the "Prime Number Theorem" in the form

$$\lim_{x \to \infty} \frac{\pi(x)\ln(x)}{x} = 1 \tag{4.14}$$

using "analytic" methods, i. e., mathematical tools from outside the domain of integers. The first "elementary" proof not using such tools did not come until 1948 and is due to *Erdös* [4.5] and Selberg. This illustrates the vast gap between obtaining an easy estimate, as we have done in the preceding pages, and a hard proof.

Perhaps one of the most surprising facts about $\pi(x)$ is that there "exists" an *exact* formula, given by a limiting process of analytic functions $R_k(x)$:

$$\pi(x) = \lim_{k \to \infty} R_k(x), \quad \text{where} \tag{4.15}$$

$$R_k(x) := R(x) - \sum_{l=-k}^{k} R(x^{\rho_l}). \tag{4.16}$$

Here ρ_l is the lth zero of the Riemann zetafunction [4.6]:

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{4.17}$$

Figure 4.6 shows $\pi(x)$ and the two approximations $R_{10}(x)$ and $R_{29}(x)$, the latter already showing a noticeable attempt to follow the jumps of $\pi(x)$.

Table 4.1 Comparison of prime-counting function $\pi(x)$ and Riemann's approximation $R(x)$

x	$\pi(x)$	$R(x)$
100000000	5761455	5761552
200000000	11078937	11079090
300000000	16252325	16252355
400000000	21336326	21336185
500000000	26355867	26355517
600000000	31324703	31324622
700000000	36252931	36252719
800000000	41146179	41146248
900000000	46009215	46009949
1000000000	50847534	50847455

Fig. 4.6 Riemann's approximation to $\pi(x)$. [After H. Riesel, G. Göhl: Math. Comp. *24*, 969–983 (1970)]

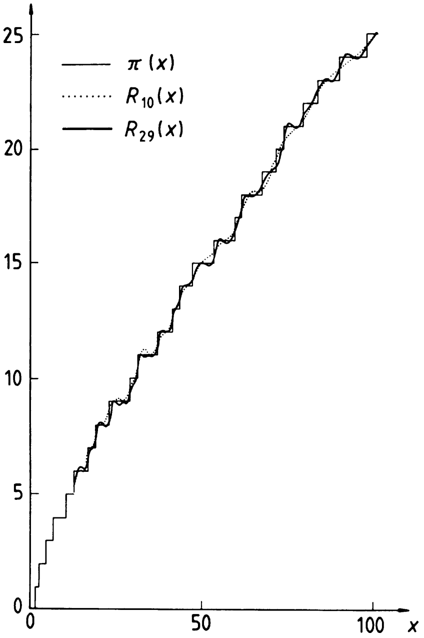


Table 4.2 The first five zeros of the zetafunction with real part equal to 1/2

$\rho_1 = \frac{1}{2} + 14.134725\,i$
$\rho_2 = \frac{1}{2} + 21.022040\,i$
$\rho_3 = \frac{1}{2} + 25.010856\,i$
$\rho_4 = \frac{1}{2} + 30.424878\,i$
$\rho_5 = \frac{1}{2} + 32.935057\,i$

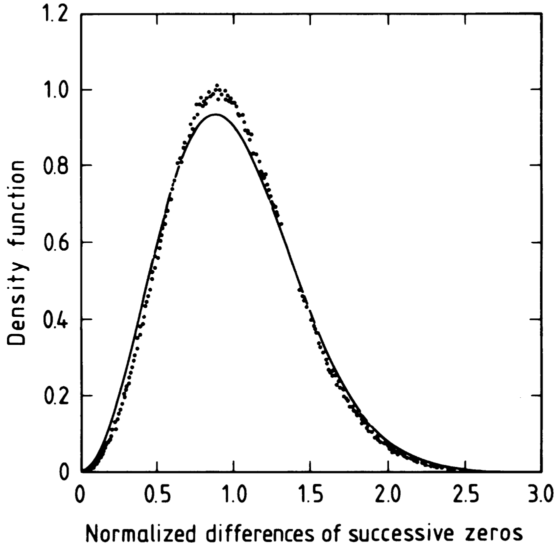
The zeros for $l = 1, 2, \ldots, 5$ of $\zeta(s)$ are shown in Table 4.2. The real parts are all equal to $1/2$. In fact, more than 100 years ago Riemann enunciated his famous hypothesis that *all* complex zeros of $\zeta(s)$ have real part $1/2$. Riemann thought at first that he had a proof, but the Riemann Hypothesis (and the so-called Extended Riemann Hypothesis, abbreviated ERH) has remained unproved to this day, although hundreds of millions of zeros have been calculated, all with real part $1/2$. In fact, the ERH is *so* widely believed today that a sizable edifice is based on it, and will collapse when the first $\mathrm{Re}(\rho_l) \neq 1/2$ makes its appearance.

(In late 1984, a *possible* proof was presented by Matsumoto in Paris. Mind boggling! If it can only be confirmed...)

4.3 David Hilbert and Large Nuclei

In conclusion, we mention that David Hilbert once conjectured that the zeros of the Riemann zetafunction were distributed like the eigenvalues of a certain kind of random Hermitian matrix. This same kind of matrix, incidentally, later gained

Fig. 4.7 Interval distribution between successive zeros $(x = 1/2)$ of zetafunction. (——) Conjecture by Hilbert (Courtesy of A. Odlyzko, Bell Laboratories)



prominence in the physics of large atomic nuclei, where its eigenvalues correspond to the energy levels of the nucleons (protons and neutrons) [4.7]. In physics the resulting distribution of energy level differences is called the *Wigner distribution* after Eugene Wigner, who derived it. It is shown in Fig. 4.7 as a solid line. The dots are the results of computer calculations by Andrew Odlyzko of Bell Laboratories (private communications) of the zeros of the Riemann zetafunction around $x = 10^8$. Since the density of zeros increases logarithmically with their distance from the real line, the spacing of zeros normalized by their average spacing is shown. The close agreement between the solid line (Hilbert) and the dots (Odlyzko) shows how close Hilbert's conjecture, made almost a century ago, is.

Even so, there are noticeable differences between Hilbert's conjecture and the numerical data. Was Hilbert off? Of course not. More recent calculations by Odlyzko of hundreds of millions of zeros around the 10^{20}th zero show no discernible differences with the conjecture. In other words, convergence to the asymptotic result is very very slow. But this is not unusual for number theory where not a few results go with the twice or thrice iterated logarithm, and $\ln\ln\ln 10^{20}$ is just a little more than 1 (1.34 to be more exact). While for some problems in physics 3 is already a large number ("almost infinity" in the physicist's book, reminiscent of the sayings "three's a crowd" or "period three means chaos"), even such a monster as 10^{20} is not all that large in some corners of number theory.

4.4 Coprime Probabilities

What is the probability that two arbitrarily and independently selected numbers from a large range do not have a common divisor, i. e., that they are coprime? The probability that one of them is divisible by the prime p_i is, as we have seen, $1/p_i$, and the probability that both of them are divisible by the same prime, assuming independence, is $1/p_i^2$. Thus, the probability that they are *not* both divisible by p_i equals $1-1/p_i^2$. If we assume divisibility by different primes to be independent, then the probability of coprimality becomes

$$W_2 \approx \prod_{p_i}\left(1-\frac{1}{p_i^2}\right), \quad \text{or} \tag{4.18}$$

$$\frac{1}{W_2} \approx \prod_{p_i}\frac{1}{1-1/p_i^2} = \prod_{p_i}\left(1+\frac{1}{p_i^2}+\frac{1}{p_i^4}+\dots\right), \tag{4.19}$$

where we have expanded the denominator into an infinite geometric series.

Now if, for simplicity, we extend the product over *all* primes, then – as Euler first noted – the result is quite simple: one obtains exactly every reciprocal square integer once (this follows from the *unique* decomposition of the integers into prime factors). Thus,

$$\frac{1}{W_2} \approx \sum_{n=1}^{\infty}\frac{1}{n^2} = \zeta(2) = \frac{\pi^2}{6}, \tag{4.20}$$

and the probability of coprimality W_2 should tend towards $6/\pi^2 \approx 0.608$ for large numbers.

The probability that a randomly selected integer n is "squarefree" (not divisible by a square) also tends to $6/\pi^2$. The reasoning leading to this result is similar to that applied above to the coprimality of two integers: for an integer to be squarefree it must not be divisible by the same prime p_i more than once. Either it is not divisible by p_i or, if it is, it is not divisible again. Thus,

$$\text{Prob}\left\{p_i^2 \nmid n\right\} = \left(1-\frac{1}{p_i}\right)+\frac{1}{p_i}\left(1-\frac{1}{p_i}\right) = 1-\frac{1}{p_i^2}.$$

Taking the product over all p_i (assuming again independence of the divisibility by different primes) gives the above expression for $W_2 \approx 6/\pi^2$.

How fast is this asymptotic value reached? The sum over the reciprocal squares in (4.20) converges quite rapidly and the value of $6/\pi^2$ might already hold for the coprimality and squarefreeness of small numbers. In fact, 61 of the first 100 integers above 1 are squarefree and of the 100 number pairs made up from the integers 2 to 11, exactly 60 are coprime. This is the closest possible result because the answer has to be an even number and 62 is further away from $600/\pi^2$ than 60.

Figure 4.8 shows a computer-generated plot of coprimality in the range from 2 to 256: a white dot is plotted if its two coordinates are coprime. As expected, the density of white dots is quite uniform. All kinds of interesting micropatterns can be observed, and a number of long-range structures at angles whose tangents are simple ratios: 0, 1/2, 1, 2, etc., are also visible. Does such a plot pose new questions or suggest new relationships for number theory?

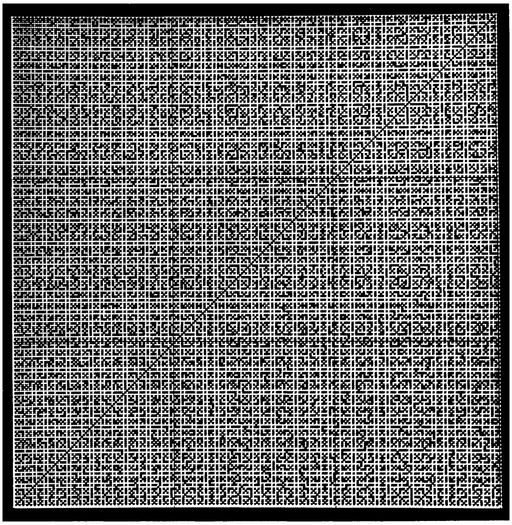


Fig. 4.8 The coprimality function, a simple number-theoretic function, in the range $2 \leq x \leq 256$ and $2 \leq y \leq 256$. A white dot is shown if $(x,y) = 1$. Whenever $(x,y) > 1$, there is no dot (*black*)

When the author first had this plot prepared (by Suzanne Hanauer at Bell Laboratories), he thought that a two-dimensional Fourier transform should make an interesting picture because the Fourier transformation brings out periodicities. And, of course, divisibility *is* a periodic property.

Figure 4.9 shows the result, which with its prominent starlike pattern would make a nice design for a Christmas card (and has, in fact, been so used). What is plotted here (as increasing brightness) is the *magnitude* of the two-dimensional discrete Fourier transform of the number-theoretic function $f(n,m)$, for $n,m = 1,2,\ldots,256$ with $f = 1$ if the GCD $(n,m) = 1$ and $f = -1$ otherwise.

Since the original function is symmetric around the 45° diagonal, so is the Fourier transform. Since only magnitude is plotted, there is another symmetry axis: the

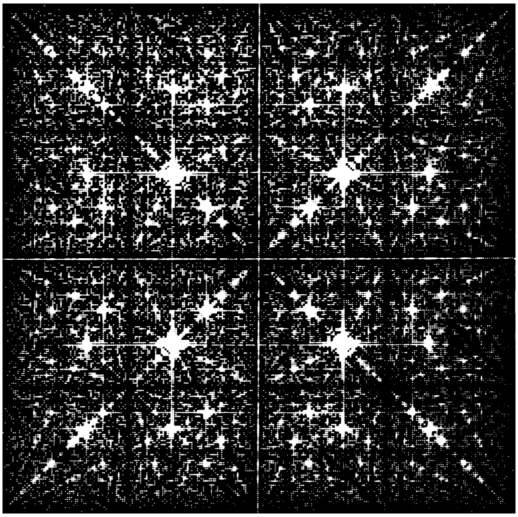


Fig. 4.9 The magnitude of the Fourier transform (simulated by increasing brightness) of the number-theoretic function shown in Fig. 4.8. The presence of a white dot, $(x,y) = 1$, is interpreted as $+1$, and the absence of a white dot as -1

$-45°$ diagonal. In addition, there are *near* symmetries about the horizontal and vertical axes which are not so easy to explain. We leave it as an exercise to the reader to explain both this near symmetry and each of the stars in Fig. 4.9. (See also [4.8].)

The coprimality probability for more than two randomly selected integers is obtained in the manner that led to (4.20). The general result that k integers are coprimes is

$$W_k := \text{Prob}\{(n_1, n_2 \ldots, n_k) = 1\} \approx [\zeta(k)]^{-1}, \tag{4.21}$$

where $\zeta(k)$ is Riemann's zetafunction as defined in (4.17). For $k = 3$ one obtains $W_3 \approx 0.832$, and for $k = 4$, $W_4 \approx 90/\pi^4 = 0.9239\ldots$. (The actual proportions in the range from 2 to 101 are 0.85 and 0.93, respectively.)

The probabilities that a randomly selected integer is not divisible by a cube, a fourth power, and in general by a kth power, also tend towards (4.21). Thus, roughly 84% of all integers are "cubefree".

A somewhat more difficult problem is posed by the probability of *pairwise* coprimality of three (or more) randomly selected integers. The probability that none of k integers has the prime factor p_i is

$$\left(1 - \frac{1}{p_i}\right)^k$$

and that exactly one has p_i as a factor is

$$\frac{k}{p_i}\left(1 - \frac{1}{p_i}\right)^{k-1}.$$

The sum of these two probabilities is the probability that at *most* one of the integers has p_i as a factor. The product over all primes p_i then approximates the probability that the k integers are pairwise coprime, i. e., that

$$(n_j, n_m) = 1 \quad \text{for all} \quad j \neq m.$$

The reader may want to show that for $k = 3$, this probability can be written

$$\frac{36}{\pi^4} \prod_{p_i} \left(1 - \frac{1}{(p_i + 1)^2}\right) = 0.28\ldots.$$

Thus, only about 28% of three randomly selected integers are *pairwise* coprime. (Compare this with the above result $W_3 \approx 0.832$.)

Jobst von Behr of Hamburg, who read the first edition of this book, generalized this problem by considering the probability $P_k(d)$ that the greatest common divisor (GCD) of k integers equal $d > 1$. By Monte Carlo computation on his home computer he obtained numerical results that looked suspiciously like

$$P_k(d) = d^{-k}\zeta^{-1}(k).$$

Can the reader of this edition prove this seductively simple scaling law? Summing over all d gives of course 1, as it should for a proper probability.

For the probability that the GCD of k random integers is *even* the above formula gives 2^{-k}; is this in conformity with elementary probability?

Here is another charming problem amenable to the probabilistic viewpoint: In the prime factor decomposition of a randomly selected integer $n > 1$, what is the probability $P(p)$ that the smallest prime divisor equals p? The probability (density) that n is divisible by p is of course $1/p$. For p to be the *smallest* prime factor, all prime smaller than p must *not* divide n. Hence

$$P(p) = \frac{1}{p}\prod_{q<p}\left(1-\frac{1}{q}\right)$$

where the product is extended over all primes q smaller than p. For example, for $p = 7$, $P(7) = 4/105$, i.e roughly 4% of all integers have 7 as their smallest prime factor.

Now let us *sum* the above expression over *all* primes, 2, 3, 5, 7, 11, . . ., so we get

$$P = \sum_p \frac{1}{p}\prod_{q<p}\left(1-\frac{1}{q}\right)$$

the probability that an integer $n > 1$ has some prime number as its smallest prime divisor which, given the sets $P(p)$ are pairwise disjoint, equals of course 1! Every integer > 1 has prime divisors one of which is necessarily the smallest.

To what extent does the above result for P depend on the exact values of the primes? For the (very rough) approximation $m \log m$ for the mth prime, the sum converges to 1.5 (instead of 1).

4.5 Primes in Progressions

A famous theorem by Dirichlet (1837), Gauss's successor in Göttingen, states: there are infinitely many primes in every *linear progression*

$$a \cdot n + b, \qquad n = 1, 2, 3, \ldots, \tag{4.22}$$

provided the constants a and b are coprime: $(a, b) = 1$. Thus, for example, with $a = 10$ and $b = 1$, 3, 7 or 9, we see that there are infinitely many primes whose last digit is 1, 3, 7 or 9. (In fact, as we shall see later, these four kinds of primes occur in equal proportion.)

The longest sequence known in early 1982 for which $a \cdot n + b$ gives primes for *consecutive* n is the progression

$$223092870 \cdot n + 2236133941,$$

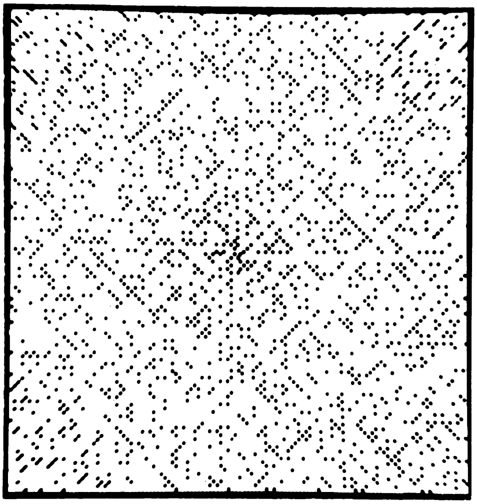


Fig. 4.10 Primes (*dots*) plotted on a spiral. Many primes fall on straight lines

which is prime for sixteen consecutive values: $n = 0, 1, 2, \ldots, 15$. Since then a 19-member progression has been discovered.

The record for a *quadratic* progression stands at 80 consecutive primes, namely

$$n^2 + n + 41 \quad \text{for } n = -40, -39, \ldots, 0, \ldots, 39. \tag{4.23}$$

This is remarkable because it would ordinarily take a polynomial in n of degree 80 to get 80 primes for consecutive values of n.

Many primes are of the form $4n^2 + an + b$, which makes them lie on straight lines if n is plotted along a square spiral. This fact is illustrated by so plotting the primes (see Figs. 4.10, 4.11).

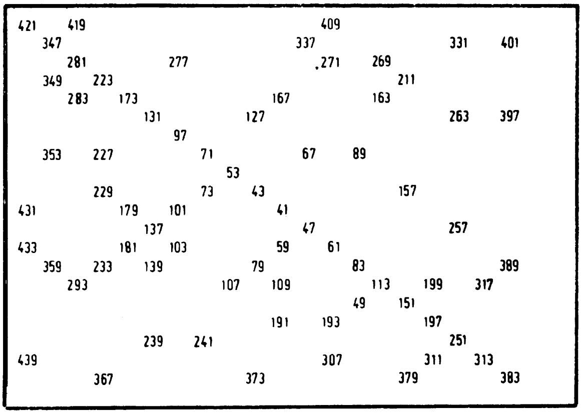


Fig. 4.11 The primes between 41 and 439 plotted on a square spiral beginning with 41 in the centre. Note the "solid-prime" diagonal

However, there is no polynomial, no matter how high its degree, which yields primes for all values of n. If there was a polynomial in n of finite degree r generating primes for all n, then $r+1$ primes would determine the $r+1$ coefficients of the polynomial and infinitely many other primes could then be calculated from these $r+1$ primes. The location of primes among the integers is simply too unpredictable to be "caught" by something as regular and finite as a polynomial.

4.6 Primeless Expanses

On the other hand, there is always a prime between n^3 and $(n+1)^3-1$ for large enough n. This fact was exploited by W. H. Mills (1947) to construct a constant A such that $\lfloor A^{3^n} \rfloor$ is prime for all n [4.9, p. 160]. But of course, the Mills expression is not a polynomial, and as we remarked before (Sect. 3.4), the primes thus generated have been "smuggled" into A first.

Somewhat paradoxically, there are also *arbitrarily* large intervals without a single prime! For example, the one *million consecutive* integers

$$(10^6+1)!+n, \qquad n=2,3,4,\ldots,1{,}000{,}001 \tag{4.24}$$

are all composite! In fact, there is even a set of one million somewhat smaller consecutive integers that are all composite, namely those in which the additive term n in the above expression is replaced by $-n$.

In *relative terms*, the primeless expanse of one million integers is, of course, rather small. A (weak) upper bound on the relative size of primeless ranges is an "octave" of integers; i. e., there is always a prime p in the range n and $2n$ (inclusive):

$$n < p \leq 2n, \tag{4.25}$$

or, equivalently, each prime is less than twice its predecessor [4.10]:

$$p_{k+1} < 2p_k. \tag{4.26}$$

Check: $3 < 2 \cdot 2$, $5 < 2 \cdot 3$, $7 < 2 \cdot 5$, $11 < 2 \cdot 7$, etc.

In fact, the number of primes in the interval from n to $2n$ is of the same order as those below n. This follows directly from the asymptotic expression for $\pi(x)$:

$$\pi(x) \approx \frac{x}{\ln x}, \tag{4.27}$$

so that

$$\begin{aligned}\pi(2x)-\pi(x) &\approx \frac{2x}{\ln x+\ln 2}-\frac{x}{\ln x}\\ &\approx \frac{x}{\ln x}-\frac{2x\ln 2}{\ln^2 x}.\end{aligned} \tag{4.28}$$

4.7 Squarefree and Coprime Integers

The probability that a given integer is squarefree approaches $6/\pi^2$ (see Sect. 4.4) and $6/\pi^2$ is also the asymptotic probability that two randomly chosen integers are coprime. Are these two properties independent? No! Among 500 random integers von Behr found 205 that were both squarefree *and* coprime (to another random integer), instead of only $500 \cdot 36/\pi^4 \approx 185$ if these two properties were independent. Thus, there seems to be a *positive* correlation between squarefreeness and coprimality. The reader may wish to show that the joint probability equals $36/\pi^4$ times a peculiar product, which is larger than 1:

$$\prod_i \left(1 + \frac{1}{p_i^3 + p_i^2 - p_i - 1} \approx 1.16\right).$$

4.8 Twin Primes

Primes not infrequently come in pairs called twin primes, like 11 and 13 or 29 and 31. How often does it happen? An *estimate* [4.2] shows their density to be proportional to $1/(\ln x)^2$, i.e. the square of the density of single primes, suggesting that they may occur independently. But one must be careful here because prime triplets of the form $(x, x+2, x+4)$ can never happen (other than the triplet 3, 5, 7), since one member of such a triplet is always divisible by 3. (Reader: try to show this – it is easy.)

On the other hand, triplets of the form $(x, x+2, x+6)$ or $(x, x+4, x+6)$ are not forbidden and do happen, for example, 11, 13, 17 or 13, 17, 19. Is their asymptotic density proportional to $1/(\ln x)^3$? In number theory, what is not explicitly forbidden often occurs, and often occurs randomly – resembling total chaos rather than neat order.

Let us try to estimate the density of twin primes. The probability that a natural number x is not divisible by a prime $p < x$ equals about $1 - 1/p$. The probability $W(x)$ that x is prime is therefore approximately

$$W(x) = \prod_p^{x'} \left(1 - \frac{1}{p}\right). \tag{4.29}$$

Here p ranges over all primes below some "cut-off" value x', where x' is about $x^{0.5}$. On the other hand the prime number theorem [4.10, Theorem 6] tells us that, asymptotically,

$$W(x) = \frac{1}{\ln x}. \tag{4.30}$$

How are (4.29) and (4.30) related? According to Mertens' Theorem [4.10, Theorem 429]

$$\prod_{p}^{x}\left(1-\frac{1}{p}\right) \to \frac{\mathrm{e}^{-\gamma}}{\ln x}, \tag{4.31}$$

where $\gamma = 0.5772$ is Euler's constant. To have (4.29) agree with (4.30) we set $x' = x^{\mathrm{e}^{-\gamma}} \approx x^{0.56}$.

While such heuristic estimates may appear rather rough, the numerical evidence is quite reassuring. Thus, for $x = 10^3$, (4.29) yields $W(x) = 0.139$, whereas (4.30) and an actual count of primes around $x = 10^3$ give 0.145 and 0.144, respectively. For $x = 10^6$, the corresponding results are 0.0723, 0.0724 and 0.0726. For $x = 10^9$, we get 0.0482, 0.0483 and 0.0484.

The probability $W_2(x)$ that both x *and* $x+2$ are primes is obtained from the following two "inequalities" or rather *incongruences*:

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -2 \mod p. \tag{4.32}$$

(For the definition and rules of congruences see Chap. 6.) For $p = 2$, both parts of (4.32) amount to just *one* condition (x must be odd) yielding the probability factor $\frac{1}{2}$. For $p > 2$, the two parts of (4.32) are two independent conditions, yielding the probability factor $(1-2/p)$. Thus,

$$W_2(x) = \frac{1}{2}\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right). \tag{4.33}$$

To connect (4.33) with (4.29) we rewrite the product in (4.33) as follows

$$\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right) = \frac{\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right)}{\prod_{p>2}^{x'}\left(1-\frac{1}{p}\right)^2}\prod_{p>2}^{x'}\left(1-\frac{1}{p}\right)^2. \tag{4.34}$$

Here the ratio of the two products converges for large x' to $0.66016\ldots$, called the twin-prime constant. The remaining product equals, according to (4.29) and (4.30), $4/\ln^2 x$. Thus,

$$W_2(x) = \frac{1.32032}{\ln^2 x}, \tag{4.35}$$

showing the expected trend with $1/\ln^2 x$. Note, that the sum over all x diverges, implying an (unproven) infinity of twin primes.

Although there is no mathematical *proof* for an infinity of twin primes, (4.35) is supported by excellent numerical evidence. (Curiously, Kummer, the great mathematician of early Fermat fame, obtained, for some reason, an erroneous answer for $W_2(x)$ that, in the pre-computer age, remained long undetected.)

Interestingly, while the sum of $1/p$ over all primes diverges (albeit very slowly – namely like $\ln(\ln x)$), the sum of $1/p_2$, where p_2 is the smaller member of a twin-prime pair, *converges*. Thus, although there probably are infinitely many twin primes (according to our heuristic estimate (4.35)), their density is not sufficient to make $\sum 1/p_2$ diverge. To what value does the $\sum 1/p_2$ converge? Numerical evidence obtained by summing $1/p_2$ up to $p_2 = 1299451$ gives 0.9652. Adding to this the integral of $1/p$ over the density (4.35) beyond 1299451 yields approximately 1.06 for $\sum 1/p_2$.

What can we deduce about twin primes with a spacing of 4 instead of 2? Instead of the incongruences (4.32) we now have

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -4 \quad \bmod p. \tag{4.36}$$

Again, for $p > 2$, two remainders, 0 and 4, are forbidden. Thus, the density $W_4(x)$ of $(x, x+4)$ twins equals the density of $W_2(x)$ of $(x, x+2)$ twins:

$$W_4(x) = W_2(x). \tag{4.37}$$

If both x and $x+4$ are primes, $x+2$ cannot be prime. In fact, $x+2$ must be divisible by 3. Thus, x and $x+4$ are true twins, i. e. primes with no intervening primes.

4.9 Prime Triplets

There are two kinds of prime triplets, those with a 2;4 spacing pattern and others with a 4;2 pattern. We first consider the 2;4 pattern, that is those cases for which x, $x+2$ and $x+6$ are primes. Thus the following three incongruences must be obeyed

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -6 \qquad \bmod p. \tag{4.38}$$

For $p = 2$, this amounts again to just one condition: x must be odd, which yields the probability factor $\frac{1}{2}$. For $p = 3$, the incongruences (4.38) impose two conditions, namely neither x nor $x+2$ must be divisible by 3, yielding the probability factor $\frac{1}{3}$. For $p > 3$, all three incongruences of (4.38) are "active", eliminating three out of p cases and yielding the probability factor $1 - 3/p$. The probability $W_{2;4}(x)$ for a prime triplet with differences between successive primes of 2 and 4, respectively, is therefore

$$W_{2;4}(x) = \frac{1}{6} \prod_{p>3}^{x'} \left(1 - \frac{3}{p}\right), \tag{4.39}$$

or

$$W_{2;4}(x) = \frac{1}{6} \frac{\prod_{p>3}^{x'} \left(1 - \frac{3}{p}\right)}{\prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^3} \prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^3. \tag{4.40}$$

Here the ratio of the two products converges for large x' to 0.63516... (which might be called the triple-prime constant). With (4.29) and (4.30), the remaining product is seen to approach $27 \ln^3 x$. Thus, approximately,

$$W_{2;4}(x) = \frac{2.858}{\ln^3 x}. \tag{4.41}$$

For prime triplets with x, $x+4$ and $x+6$ prime, the incongruences are

$$x \not\equiv 0, \quad x \not\equiv -4, \quad x \not\equiv -6 \qquad \bmod p, \tag{4.42}$$

leading to equinumerous sets of restrictions as (4.38). Thus, the corresponding probability $W_{4;2}(X)$ is given by

$$W_{4;2}(x) = W_{2;4}(x). \tag{4.43}$$

4.10 Prime Quadruplets and Quintuplets

As with prime triplets, there are two different spacing patterns for quadruplets of close primes: $2;4;2$ and $4;2;4$. Curiously, numerical evidence suggests that the latter pattern $(4;2;4)$ seems to be twice as numerous as the other pattern $(2;4;2)$. We would like to understand why.

For the $2;4;2$ pattern the four incongruences are

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -6, \quad x \not\equiv -8 \qquad \bmod p. \tag{4.44}$$

The *effective* number of incongruences for $p = 2$ equals one; for $p = 3$ the number equals two; for $p > 3$ all four incongruences are active. The probability of the $2;4;2$ pattern is therefore

$$W_{2;4;2}(x) = \frac{1}{6} \prod_{p>3}^{x'} \left(1 - \frac{4}{p}\right), \tag{4.45}$$

or

$$W_{2;4;2}(x) = \frac{1}{6} \frac{\prod_{p>3}^{x'} \left(1 - \frac{4}{p}\right)}{\prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^4} \prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^4. \tag{4.46}$$

Here the ratio of the two products converges to 0.307496... (the "quadruplet constant") and the remaining factor tends to $81/\ln^4 x$. Thus

$$W_{2;4;2}(x) = \frac{4.15}{\ln^4 x}. \tag{4.47}$$

For the spacing pattern $4;2;4$ the four congruences are

$$x \not\equiv 0, \quad x \not\equiv -4, \quad x \not\equiv -6, \quad x \not\equiv -10 \qquad \bmod p. \tag{4.48}$$

Here, for $p = 5$, the last incongruence is automatically fulfilled by the first ($x \not\equiv 0$). Thus there are only *three* (instead of four) active incongruences for $p = 5$. The probability factor for $p = 5$ therefore doubles from $(1 - 4/5)$ to $(1 - 3/5)$. As a result, we have

$$W_{4;2;4}(x) = 2W_{2;4;2}(x), \tag{4.49}$$

confirming the numerical evidence.

The next closely spaced cluster of primes is a quintuplet with the spacing pattern 2;4;2;4, such as 5, 7, 11, 13, 17 which repeats at 11, 13, 17, 19, 23 and 101, 103, 107, 109, 113, etc. Another possible quintuplet type has the spacing pattern 4;2;4;2, such as 7, 11, 13, 17, 19, which repeats at 97, 101, 103, 107, 109 but not again until 1867, 1871, ... The reader should have no difficulty deriving the asymptotic densities for these quintuplets following the recipe of Sects. 4.8–4.10.

However, not all patterns are possible, as we have already seen in Sect. 4.9 with prime triplets: the spacing pattern 2;2 is impossible after 3, 5, 7 because one member of such a triplet must be divisible by 3. Another way to demonstrate the impossibility of the spacing pattern 2;2 uses the appropriate incongruences: for x, $x+2$, $x+4$ to be prime. These are

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -4 \qquad \bmod p.$$

For $p = 3$, the last incongruence may be written as $x \not\equiv -1 \bmod 3$. Thus *all three* possible remainders modulo 3 $(0, -2, -1)$ are forbidden and 3, 5, 7 is the only triplet with the 2;2 pattern.

In a similar manner, we can show that the sextet with the spacing pattern 2;4;2;4;2 occurs only once (5, 7, 11, 13, 17, 19) and never again. The six incongruences for the sextet considered are

$$x \not\equiv 0,\ -2,\ -6,\ -8,\ -12,\ -14 \quad \bmod p.$$

For $p = 5$ we may write

$$x \not\equiv 0,\ -2,\ -1,\ -3,\ -2,\ -4 \quad \bmod 5,$$

which excludes *all five* possible classes of remainders modulo 5.

What if any spacing pattern of length 7, 8 etc. is unique in the sense that, like 2, 3 and 3, 5, 7 and 5, 7, 11, 13, 17, 19, it occurs only once and never again?

4.11 Primes at Any Distance

In Sect. 4.8 we derived the formula $W_2(x) = 1.32032/\ln^2 x$ for the asymptotic density of twin primes with a distance of 2 (such as 3 and 5 etc.). We also noted that twin primes with a distance of 4 (like 7 and 11 etc.) are equally probable: $W_4(x) = W_2(x)$.

What is the density of pairs of odd primes, not necessarily adjacent, with a distance of $\Delta = 6$ (like 5 and 11) or a distance of $\Delta = 8$ (like 3 and 11) or any other distance Δ irrespective of any intervening primes?

For $\Delta = 6$, the incongruences (4.32) have to be replaced by

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -6 \qquad \bmod p. \tag{4.50}$$

For $p = 2$ and $p = 3$, only one of these two incongruences is active, yielding a probability factor $(1 - 1/2)\ (1 - 1/3) = 1/3$. For $p > 3$ both incongruences are active yielding probability factors $(1 - 2/p)$. Hence, with (4.33),

$$W_6(x) = \frac{1}{3} \prod_{p>3}^{x'} \left(1 - \frac{2}{p}\right) = \prod_{p>2}^{x'} \left(1 - \frac{2}{p}\right) = 2W_2(x), \tag{4.51}$$

i.e. pairs of primes with a distance $\Delta = 6$ are twice as numerous as twin primes.

For $\Delta = 8$, the second of the two incongruences (4.50) becomes $x \not\equiv -8 \bmod p$. Since 8 does not contain any prime factors $p > 2$, we can proceed as in the derivation of the twin-prime density with $\Delta = 2$ and $\Delta = 4$. The result is

$$W_8(x) = W_2(x). \tag{4.52}$$

In fact, for any $\Delta = 2^k$, $k = 1, 2, 3 \ldots$, we obtain

$$W_{2^k}(x) = W_2(x). \tag{4.53}$$

More generally, for arbitrary Δ, only the prime factors of Δ are important. Of the two incongruences

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -\Delta \qquad \bmod p \tag{4.54}$$

only one is active for all odd prime factors q of Δ. Thus

$$W_\Delta(x) = \frac{1}{2} \prod_q \frac{1 - \frac{1}{q}}{1 - \frac{2}{q}} \prod_{p>2}^{x'} \left(1 - \frac{2}{p}\right), \tag{4.55}$$

where q are the different odd prime factors of Δ and p are *all* primes above 2 and below x'. With (4.33) we get

$$W_\Delta(x) = \Pi(\Delta) W_2(x), \quad \text{where} \quad \Pi(\Delta) = \prod_{\substack{q>2 \\ q|\Delta}} \frac{q-1}{q-2}, \tag{4.56}$$

a charming generalization of our earlier result (4.51) that $W_6(x)$ equals $2W_2(x)$. But W_6 is not the largest value. For $\Delta = 30$ we get $W_{30} = 2.66 W_2$. In fact, there is no upper bound for $\Pi(\Delta)$ because the product over q diverges if all possible primes are included as prime factors of Δ. More specifically, the product $\Pi(\Delta)$ diverges as $1.349 \ln q_{\max}$, where $q_{\max}$ is the largest prime factor of Δ. Are there two different coprime values of Δ having the same $\Pi(\Delta)$? Let the reader decide – or rather find out.

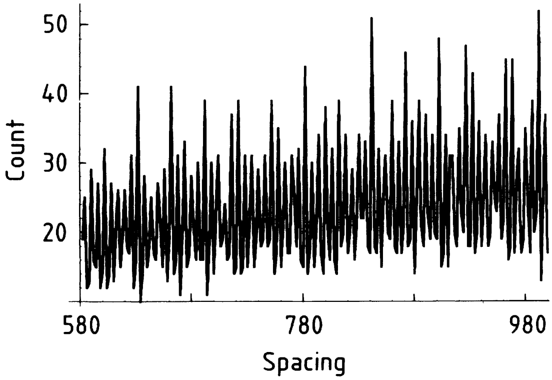


Fig. 4.12 The number of primes with a given spacing

Of the two factors in (4.56) one is slowly varying, representing a monotonic trend, $W_2(x)$, whereas the product $\Pi(\Delta)$ fluctuates appreciably with changing values of Δ. Figure 4.12 shows a plot of $\Pi(\Delta)$ as a function of Δ. For every value of Δ that is divisible by 3, i.e. every third value, the quotient $(q-1)/(q-2)$ contributes a factor of 2 to the product. Similarly, for every value of Δ divisible by 5, i.e. every 5th value, a factor of $4/3$ is contributed – and so on for 7, 11 etc.

These inherent periodicities are brought out nicely by the Fourier transform of $\Pi(\Delta)$: see Fig. 4.13, which shows a pronounced peak at one third the sampling frequencies, corresponding to $p = 3$, and smaller peaks corresponding to $p = 5$ and 7 and their "harmonics". As we shall see, $\Pi(\Delta)$ has a certain universal air about it because it also governs the sum of two primes (cf. Sect. 4.13).

$\Pi(\Delta)$ is also related to Euler's Φ function $\Phi(m)$ (see Sect. 8.3) and in fact resembles $m/\Phi(m)$:

$$\frac{m}{\Phi(m)} = \prod_{p|m} \frac{p}{p-1},$$

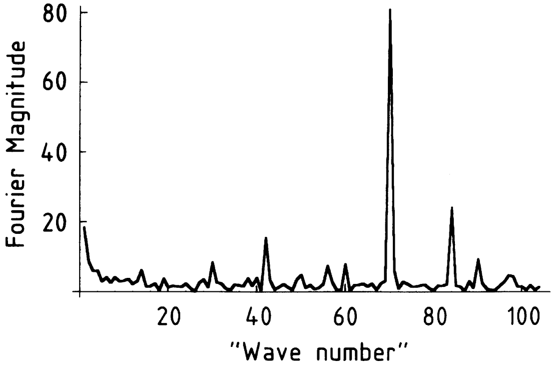


Fig. 4.13 The magnitude of the Fourier transform showing the preponderance of spacings at multiples of 6

whereas

$$\Pi(m) = \prod_{\substack{p>2 \\ p|m}} \frac{p-1}{p-2}. \tag{4.57}$$

4.12 Spacing Distribution Between Adjacent Primes

The prime number theorem prescribes the average spacing $\bar{s}$ between adjacent primes. For $n \gg 1$, this spacing is about $\ln n$. (For example, for $n = 20$, $\bar{s} \approx \ln 20 \approx 3$.)

All spacings between odd primes are of course even. If the distribution of spacings were otherwise unconstrained, then, given an average spacing, the maximum-entropy principle [4.11] would tell us that the distribution of $k = s/2$ is the geometric distribution:

$$d(k) = \frac{1}{\bar{k}}\left(\frac{\bar{k}}{1+\bar{k}}\right)^k, \quad \bar{k} = \frac{1}{2}\ln n, \qquad k = 1, 2, \ldots. \tag{4.58}$$

In reality, the distribution of spacings is anything but unconstrained. We already know (4.37) that the frequencies of the spacings $s = 2$ (i.e. $k = 1$) and $s = 4$ (i.e. $k = 2$) are equal, while (4.58) would predict a ratio of $d(1)/d(2) = 1 + 1/\bar{k}$. Figure 4.14 shows a plot of the logarithm of $d(k)$ in the range p_{20000} to p_{40000} for $k = 1$ to 36. There is the predicted overall linear trend: $\ln d(k) \sim r/k$ with $r = -\ln(1 + 1/\bar{k})$. But the "bumps" at $k = 3$ (i.e. $s = 6$), $k = 15$ (i.e. $s = 30$) and other places (especially multiples of 3) are also visible.

The average slope obtained by regression of the data in Fig. 4.14 is $r = -0.148$, in good agreement with the theoretical value of $r = -0.146$.

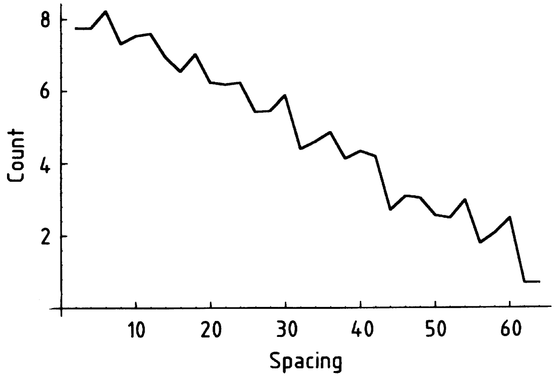


Fig. 4.14 The distribution of the spacings between *adjacent* primes

4.13 Goldbach's Conjecture

One of the most enduring (if not endearing) mathematical conjectures, made by the Russian mathematician Christian Goldbach (1690–1764) in a letter to Leonhard Euler, is the famous *Goldbach conjecture*. It asserts that every even number $n > 4$ is the sum of two odd primes. Some progress has been made on related weaker assertions, and the Goldbach conjecture itself has been *numerically* confirmed up to very large numbers. But alas, even if it had been shown to hold up to $10^{10^{10}}$, there would be no guarantee that it would not fail for $10^{10^{10}} + 2$.

Let us try our heuristic muscle at this recalcitrant conjecture. We want to get a feeling for the number of representations

$$x + y = n,$$

where x and y are odd primes and $n > 4$ is even. We want to count permuted representations, such as $3 + 5 = 5 + 3 = 8$, as only *one case*. Without loss of generality we assume $x \leq n/2$.

The number of odd primes below $n/2$ equals approximately $n/(2 \ln n)$. Each such prime is a candidate for $x + y = n$ and contributes to the count in question if $y = n - x$ is also prime. What is the probability of $n - x$ being prime, given that x is an odd prime?

To answer this question, we have to distinguish two cases. For a given potential prime divisor p of $n - x$: does p divide n or does it not? In the first case $(p|n)$ the two incongruences

$$x \not\equiv 0 \quad \text{and} \quad n - x \not\equiv 0 \qquad \bmod p \tag{4.59}$$

amount to only *one* condition because, for $p|n$, $n - x \equiv x \bmod p$. Thus the probability factor is $(1 - 1/p)$. In the other case $(p \nmid n)$ and $p > 2$, the incongruences (4.59) are independent of each other. The probability factor is therefore $(1 - 2/p)$. For $p = 2$ the factor is $1/2$. Multiplying these probability factors yields

$$\prod_{p|n} \left(1 - \frac{1}{p}\right) \prod_{\substack{p>2 \\ p|n}} \left(1 - \frac{2}{p}\right) \cdot \frac{1}{2}.$$

This factor has to be divided by the unrestricted probability factor

$$\prod_{p} \left(1 - \frac{1}{p}\right)^2$$

and multiplied by the *a priori* probability that $n - x$ is prime. Since $n - x$ is already odd (because n is even and x is odd), this probability equals

$$\frac{2}{\ln(n - x)} \approx \frac{2}{\ln n}$$

(because $x \leq n/2$).

Thus, the total estimated count $C_2(n)$ of cases for which both x and $n-x$ are prime is given by

$$C_2(n) = \prod_{p>2} \frac{\left(1-\frac{2}{p}\right)}{\left(1-\frac{1}{p}\right)^2} \cdot \prod_{\substack{p>2 \\ p|n}} \frac{p-1}{p-2} \cdot \frac{n}{\ln^2 n}. \tag{4.60}$$

Here the first factor approaches our old friend, the twin-prime constant 0.66016... (4.34). This factor combined with the last factor ($n/\ln^2 n$) is nothing but the twin-prime density $W_2(n)$, see (4.35), multiplied by 1/2. Thus

$$C_2(n) = \frac{n}{2} W_2(n) \prod_{\substack{p>2 \\ p|n}} \frac{p-1}{p-2}. \tag{4.61}$$

To simplify this formula even further, we recall the function $W_\Delta(n)$ [see (4.55)]. With $\Delta = n$

$$C_2(n) = \tfrac{n}{2} W_n(n). \tag{4.62}$$

Thus the number of Goldbach representations $C_2(n)$ equals, within a factor $n/2$, the density of prime pairs $W_n(n)$ with a spacing of n in the neighbourhood of n.

Another interesting aspect of this result is the product in (4.61). Suppose n is a multiple of 3, then this product, for $p = 3$, contributes a factor of 2 to the count. If n is divisible by 5, the product contributes a factor 4/3, which is still appreciably larger than 1. Thus, we see that the count $C_2(n)$ has a pronounced periodic component with a period of 3 (and weaker periodicities with periods of 5, 7, 11, etc.), just like $\Pi(n)$ (see Figs. 4.12 and 4.13).

4.14 Sum of Three Primes

Another problem from additive number theory that Goldbach posed in 1742 in correspondence with Euler, then at St. Petersburg, concerned the sum of *three* primes. Goldbach asked whether every sufficiently large odd n can be written as

$$n = p_1 + p_2 + p_3. \tag{4.63}$$

Positive proof that this is indeed possible had to wait almost 200 years – until 1937 when I. M. Vinogradov [4.12] furnished a proof based on Fourier-like trigonometric sums. Trigonometric sums have played an important role in number theory ever since, forging a strong link between additive and multiplicative number theory.

How did Vinogradov get from (4.63) to trigonometric sums? The number of cases including permutations $C_3(n)$ for which (4.63) holds can be written as follows

$$C_3(n) = \sum_{p_1<n}\sum_{p_2<n}\sum_{p_3<n}\int_0^1 \exp\left[2\pi\mathrm{i}(p_1+p_2+p_3-n)x\right]\mathrm{d}x, \tag{4.64}$$

because when (4.63) holds the integral equals 1 and the triple sum is augmented by 1; otherwise the integral is 0 and nothing is added.

As a next step, Vinogradov converted the triple sum into a single sum:

$$C_3(n) = \int_0^1 \sum_{p<n} [\exp(2\pi\mathrm{i}xp)]^3 \exp(-2\pi\mathrm{i}n)\,\mathrm{d}x, \tag{4.65}$$

which he was able to convert into the product of $(1-1/(p^2-3p+3))$ taken over all prime divisors of n.

The final result of Vinogradov's method is that the number of representations $C_3(n)$ of an odd integer as the sum of three primes equals, asymptotically,

$$C_3(n) = C\prod_{p|n}\left(1-\frac{1}{p^2-3p+3}\right)\frac{n^2}{\ln^3 n}. \tag{4.66}$$

This formula resembles (4.60) for the sum of *two* primes, except that the slowly varying factor $n/\ln^2 n$ (the "trend") has been replaced by $n^2/\ln^3 n$ and the product over $p|n$ is over a quadratic expression in p. However, like $\Pi(n)$ [see (4.57)] this product, too, depends sensitively on the small prime factors of n. Thus, if n is divisible by 3, the term for $p=3$ contributes a factor 2/3 to the product.