6. W. J. LeVeque, Topics in Number Theory, vol. 1, Reading, Mass., 1956.
7. I. Niven, Irrational Numbers, Carus Monograph No. 11, 1956.
8. G. Polya, Über die Verteilung der quadratischen Reste und Nichtreste, Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. IIa, 1918, pp. 1–9.
9. G. Polya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, vol. 1, Berlin, 1925.
10. I. J. Schoenberg, Über die asymptotische Verteilung reeller Zahlen mod 1, Math. Z., vol. 28, 1928, pp. 171–199.
11. I. J. Schoenberg, On asymptotic distributions of arithmetical functions, Trans. Amer. Math. Soc., vol. 39, 1936, pp. 315–330.

HEURISTIC REASONING IN THE THEORY OF NUMBERS

G. PÓLYA, Stanford University

A deep but easily understandable problem about prime numbers is used in the following to illustrate the parallelism between the heuristic reasoning of the mathematician and the inductive reasoning of the physicist. The experts may judge whether the parallelism is more serious than the tone of presentation which is adapted to a wider audience.

1. "Till now the mathematicians tried in vain to discover some order in the sequence of the prime numbers and we have every reason to believe that there is some mystery which the human mind shall never penetrate. To convince oneself, one has only to glance at the tables of primes which some people took the trouble to compute beyond a hundred thousand, and one perceives that there is no order and no rule. This is so much more surprising as the arithmetic gives us definite rules with the help of which we can continue the sequence of the primes as far as we please, without noticing, however, the least trace of order."*

So wrote Euler about two centuries ago, yet the prime numbers may inspire the contemporary mathematician with the same feeling of mystery that Euler so vividly expressed. The primes remain puzzling in spite of many important discoveries made in the meantime. Let us look at some of these discoveries.

The intervals between successive primes are irregular, but these intervals seem to become larger "on the whole" (the primes seem to become scarcer) as we proceed in the sequence of numbers. Since Euler's time a definite law of this phenomenon was discovered (conjectured by Legendre and Gauss, investigated by Chebyshev and Riemann, finally proved by Hadamard and de la Vallée Poussin, proved recently in an essentially different "elementary" manner by Atle Selberg and Paul Erdös). We may formulate this law, the "prime number theorem," intuitively although not quite precisely, as follows: The probability

* See L. Euler, Opera Omnia, ser. 1, vol. 2, p. 241 or G. Pólya, Mathematics and Plausible Reasoning, Princeton, vol. 1, p. 91.

that a large integer x should be a prime, is $1/\log x$ (where $\log x$ is the natural logarithm of x).*

The following short table exhibits the first primes (with two exceptions) classified according to their last digit.

	11		31	41		61	71			101		
3	13	23		43	53		73	83		103	113	
7	17		37	47		67			97	107		
	19	29			59		79	89		109		

If we set apart 2 and 5, the prime factors of 10, the last figure in the decimal symbol of a prime cannot be 0, 2, 4, 5, 6, or 8 (since neither 2 nor 5 should be a divisor) and must, therefore, be 1, 3, 7, or 9. Thus, with respect to ten (*modulo* 10) there are four kinds of primes which are listed in the four horizontal lines of the foregoing table, respectively. Since Euler's time, a general law has been discovered (most of the credit for its discovery is due to Dirichlet) which, applied to our particular case, asserts that there are infinitely many prime numbers of each kind and, what is more, that each kind is equally probable. Therefore, in an extensive table of prime numbers there must be roughly as many primes ending with 1 as primes ending with 3.

Euler mentions a table of primes that goes beyond 10^6. Since his time much more extensive tables have been computed, especially in the last decade with the help of machines. Data derived from these tables may suggest problems not yet considered by Euler.

2. The least possible distance between two consecutive primes is 2, if we set apart the unique case of the primes 2 and 3. Two primes having this minimum distance are called *twin primes*. Here is a list of the twin primes under 100:

3, 5 5, 7 11, 13 17, 19 29, 31 41, 43 59, 61 71, 73

We can generalize this situation and consider a prime p that is escorted at a given distance d by another prime $p'=p+d$. (This situation is uninteresting unless d is even; we do not care whether there are or are not other primes between p and p'.) Here is a list of all such pairs at the distance 6, in which the

* The irregular distribution of primes ("there is no order and no rule") strongly suggests the idea of probability and chance. Yet this is paradoxical: Whether any given integer is a prime or not, can be decided by the "definite rules" of arithmetic—where and how could chance enter the picture? The paradox can be somewhat explained (or deepened) by a physical analogy. The kinetic theory of matter considers the probability distribution of the velocities of the molecules in a gas. Yet this is paradoxical: The velocities resulting from the collision of two molecules can be exactly predicted from the data of the collision by the "definite rules" of classical deterministic mechanics —where and how could chance enter the picture? The determinateness of the simple single event and the probabilistic theory of the highly composite whole may seem to be equally compatible (or incompatible) in both cases.

first prime does not (but its escort may) exceed 100:

5, 11	7, 13	11, 17	13, 19	17, 23	23, 29	31, 37	37, 43
41, 47	47, 53	53, 59	61, 67	67, 73	73, 79	83, 89	97, 103

It is curious that the second kind of pairs is more numerous. We count 8 pairs of twin primes and exactly twice as many pairs of primes at the distance 6. Let us take now instead of 10^2 the considerably higher bound $3 \cdot 10^7$. Under thirty million there are 152892 primes followed by another prime at the distance 2, but nearly twice as many, namely 304867 primes followed by another at the distance 6.

The numbers of these prime pairs have been obtained by Professor and Mrs. D. H. Lehmer with the use of appropriate computing apparatus; they computed, up to the same limit $3 \cdot 10^7$, the number of primes escorted by another prime at the distance d for $d = 2, 4, 6, 8, \cdots, 70$. I wish to thank them here for their kind permission to use their interesting material. I wish to use some of their results to offer the unprejudiced reader a particularly suitable opportunity for an inductive investigation in pure mathematics.

It will be convenient to introduce here some notation. Let $\pi_d(x)$ stand for the number of those prime numbers p that satisfy two conditions:

$$p \leqq x, \qquad p + d \text{ is a prime number.}$$

For instance,

$$\pi_2(100) = 8\ , \qquad \pi_2(30\ 000\ 000) = 152892,$$
$$\pi_6(100) = 16, \qquad \pi_6(30\ 000\ 000) = 304867.$$

I set

$$\pi_d(3 \cdot 10^7)/\pi_2(3 \cdot 10^7) = R_d.$$

For instance, $R_6 = 304867/152892 = 1.9940$, approximately. A small part of the material computed by Professor and Mrs. Lehmer is collected in Table I.

d	R_d	12	1.9985	24	1.9976	36	1.9997	48	1.9965	60	2.6632
2	1.0000	14	1.1985	26	1.0910	38	1.0566	50	1.3308	62	1.0341
4	0.9979	16	1.0001	28	1.1974	40	1.3330	52	1.0892	64	0.9999
6	1.9940	18	1.9982	30	2.6632	42	2.3987	54	1.9981	66	2.2186
8	0.9996	20	1.3311	32	0.9970	44	1.1097	56	1.1957	68	1.0663
10	1.3317	22	1.1088	34	1.0645	46	1.0467	58	1.0349	70	1.5977

TABLE I. VALUES OF R_d

3. Now, let us start our inductive research. At any moment at which the reader feels inspired, he should interrupt the reading and try to guess the result by himself.

The four kinds of prime numbers that we have considered in Section 1 (ending with 1, 3, 7 or 9 in the decimal notation, respectively) are known to be equally frequent. Are the 35 kinds of prime numbers with which Table I is concerned also equally frequent? If it were so, all the ratios R_d contained in Table I should be approximately equal to one. In fact, remarkably enough, a few entries in Table I are pretty close to the value 1, but the majority seem to deviate significantly from 1. The analogy with the previous case does not seem to go far. Yet, perhaps, the analogy holds at least in one respect: the ratio $\pi_d(x)/\pi_2(x)$ may converge towards some limit (not necessarily 1) when x tends to infinity, and the ratio $R_d=\pi_d(3\cdot 10^7)/\pi_2(3\cdot 10^7)$ entered into Table I may be an approximation to that limit.

We face here a situation somewhat analogous to the situation that the chemists faced around 1800 when they were about to discover the Law of Multiple Proportions. They had to perceive behind their experimental data distorted by unavoidable errors of observation the ratios of simple multiples of the atomic weights, and we have to perceive behind the approximate ratios R_d collected in Table I the true limiting ratios. To guess these limiting ratios is a challenging task.

We have already observed that some values of R_d are very close to 1; they correspond to $d=2$, 4, 8, 16, 64. (For $d=2$ the value is exactly 1, but this is trivial.) We can scarcely fail to notice here the powers of 2. By the way, these values of R_d so close to 1 are also the smallest values in the table. Are there other entries in the table so nearly equal to each other?

In trying to answer this question we may notice that the entries corresponding to

$$d = 6,\ 12,\ 24,\ 48$$

are approximately equal to each other, and so are those corresponding to

$$d = 10,\ 20,\ 40$$

or those corresponding to

$$d = 14,\ 28,\ 56.$$

In general, multiplication of d by 2 seems to leave the value of R_d almost unchanged.

What about multiplication by 3? It approximately doubles the value of R_d in certain transitions, as from

2 to 6,	4 to 12,	8 to 24,	16 to 48,
10 to 30,	20 to 60,	14 to 42,	22 to 66.

Yet it is not so in other cases, as

$$6 \text{ to } 18, \qquad 12 \text{ to } 36, \qquad 18 \text{ to } 54;$$

in these latter cases the multiplication of d by 3 leaves the value of R_d almost unchanged. How can you account for this different behavior?

And so on, from question to question, by observation and tentative generalization, carefully checking each guess, the reader may discover that many of the values R_d contained in Table I come very close to simple fractions; see Table II.

d	2 16 4 32 8 64	6 36 12 48 18 54 24	10 20 40 50	14 28 56	22 44	30 60	42	66	70
R_d (approx.)	1	$\frac{2}{1}$	$\frac{4}{3}$	$\frac{6}{5}$	$\frac{10}{9}$	$\frac{8}{3}$	$\frac{12}{5}$	$\frac{20}{9}$	$\frac{8}{5}$

TABLE II. SIMPLE APPROXIMATIONS TO SOME R_d

Table II strongly suggests that *R_d depends only on the decomposition of d into prime factors*. More precisely, just the presence of a prime factor in, or its absence from, the decomposition seems to be relevant; for instance, to all values of d of the form $2^\alpha 3^\beta$ with $\alpha, \beta = 1, 2, 3, \cdots$ there corresponds the same value of R_d (approximately).

Moreover, to each prime factor of d there seems to correspond a factor of R_d; to the (unavoidable) factor 2 of d, the (trivial) factor 1 of R_d; to the prime factors

$$3, \qquad 5, \qquad 7, \qquad 11$$

of d, the following factors of R_d:

$$\frac{2}{1}, \qquad \frac{4}{3}, \qquad \frac{6}{5}, \qquad \frac{10}{9},$$

respectively. Then, when d is a product of different primes (or powers of different primes) R_d seems to be the product of the corresponding factors.

4. All such observations point to the (conjectural) formula

$$\pi_d(x) \sim \pi_2(x) \prod_{p \mid d} \frac{p-1}{p-2}, \tag{1}$$

where the product $\prod_{p \mid d}$ is extended over all different *odd* prime factors p of the even number d.* The sign $\sim$ can be interpreted either vaguely or strictly. In a

* The usual abbreviation $a \mid b$ means "a divides b" or "a is a divisor of b." We shall need later also the abbreviation $a \nmid b$ which means "a is not a divisor of b."

vague interpretation $\sim$ means "approximately equal;" in the strict sense it means "the ratio of the two sides tends to 1 when x tends to ∞." The formula is merely a conjecture which we can conceive quite naively by examining Table I. In Table III, the observed values of R_d, taken from Table I and styled now R_d (obs.), are compared with the corresponding conjectural limiting values, styled R_d (theor.). This comparison yields strong inductive evidence for the conjecture which could be further strengthened by use of other data computed by Professor and Mrs. Lehmer.

d	R_d (obs.)	R_d (theor.)	24	1.9976	2.0000	48	1.9965	2.0000
2	1.0000	1.0000	26	1.0910	1.0909	50	1.3308	1.3333
4	0.9979	1.0000	28	1.1974	1.2000	52	1.0892	1.0909
6	1.9940	2.0000	30	2.6632	2.6667	54	1.9981	2.0000
8	0.9996	1.0000	32	0.9970	1.0000	56	1.1957	1.2000
10	1.3317	1.3333	34	1.0645	1.0667	58	1.0349	1.0370
12	1.9985	2.0000	36	1.9997	2.0000	60	2.6632	2.6667
14	1.1985	1.2000	38	1.0566	1.0588	62	1.0341	1.0345
16	1.0001	1.0000	40	1.3330	1.3333	64	0.9999	1.0000
18	1.9982	2.0000	42	2.3987	2.4000	66	2.2186	2.2222
20	1.3311	1.3333	44	1.1097	1.1111	68	1.0663	1.0667
22	1.1088	1.1111	46	1.0467	1.0476	70	1.5977	1.6000

TABLE III. VALUES OF R_d, OBSERVED AND "THEORETICAL"

5. We have before us a precise, general, but enigmatic formula derived from, and quite well verified by, observations. Of course, we wish to understand it, we wish to explain it. When we are looking at it, our situation is similar to that of Newton looking at the laws of Kepler or to that of Niels Bohr looking at Balmer's formula. The word "similar" must be correctly understood. Similar figures may be very different in magnitude, but they show the same proportions, and so do in a sense the three situations we have just compared.

We wish to explain that conjectural formula about prime numbers. Both the irregular distribution of the primes and the structure of the conjectural formula strongly suggest an explanation by probability. I wish to present such an explanation. We shall arrive at it in two steps (of which the second is much more dangerous).

PROBLEM I. *Let p denote a given prime number, d a given integer, and x a large integer chosen at random. Find the probability that neither x nor $x+d$ is divisible by p.*

The reader may visualize the integers as successive intervals of equal length along an infinite straight line, some sort of super-roulette. The interval is red or green, according as the integer is, or is not, divisible by p; among any p consecutive intervals there is always just one that is red. A ball is rolled along the line and steps in the interval x.

We have to distinguish two cases.*

First case: $p \mid d$. In this case $x+d$ falls on a multiple of p (a red space) if, and only if, x itself falls on such a multiple. Therefore, out of any p consecutive numbers (spaces), $p-1$ are favorable (green) and so the required probability is $(p-1)/p$.

Second case: $p \nmid d$. Even if x does not fall on a multiple of p, $x+d$ may. Therefore, out of any p consecutive numbers just $p-2$ are favorable. The required probability is $(p-2)/p$.

PROBLEM II. *Let d denote a given even integer, and x a large integer chosen at random. Find the probability P_d that both x and $x+d$ are prime numbers.*

In order that both x and $x+d$ should be prime numbers, a sequence of conditions must be satisfied:

First, neither x nor $x+d$ is divisible by 2;

then, neither x or $x+d$ is divisible by 3;

then, neither x nor $x+d$ is divisible by 5;

and so on. The general form of this condition is: neither x nor $x+d$ is divisible by p where p is a prime number.

We have computed above the probability for the fulfillment of any single one of these conditions. Now we have to compute the probability that all these conditions are fulfilled at the same time, all these events are realized simultaneously.

Two difficulties arise here: Are these events independent? How far should we go with p? In fact, these two difficulties may be connected, but at this stage of the game it will be better not to examine them too thoroughly; let us now proceed quickly and see whether anything worthwhile turns up.

Are the events independent? We do not know, but let us assume it. Also the physicist is inclined to assume the independence of the probabilities he deals with—not because he knows that they are independent, but interdependent probabilities are so much more difficult to handle—and so let us assume independence in our case too, although we have no better reasons than the physicist.

Having made this assumption all we have to do is to multiply probabilities computed above. We distinguish three cases:

* For the symbols $\mid$ and $\nmid$, see footnote p. 379.

$p=2$ (which is a divisor of the even number d);
p is odd and is a divisor of d;
p is odd and is not a divisor of d.

Accordingly, the required probability P_d is a product of three kinds of factors:

$$P_d = \frac{1}{2} \prod_{p|d} \frac{p-1}{p} \prod_{p \nmid d} \frac{p-2}{p}. \tag{2}$$

In this formula (2) (and in the following formulas (3), (4)) the letter p stands for an *odd* prime number.

How far should we go with p? Of course, on the right hand side of formula (2) we extend the first product over all odd prime factors of the given number d. In the second product, we take all the odd primes not dividing d up to a certain large upper bound, depending on the considered large number x—but let us *postpone* the decision, how far to go, how large that upper bound should precisely be.

We can transform formula (2) as follows:

$$P_d = \prod_{p|d} \frac{p-1}{p-2} \cdot \frac{1}{2} \prod_{p} \frac{p-2}{p}; \tag{3}$$

the second product on the right hand side of (3) is extended over *all* odd primes p under a certain (large, but not yet definitely characterized) upper bound. The first product is extended over the odd prime divisors of d; if d happens to be 2 (or a power of 2) there are no odd prime divisors, that first product is empty, and has to be replaced by 1. Therefore

$$P_d = \prod_{p|d} \frac{p-1}{p-2} \cdot P_2. \tag{4}$$

Yet the ratio of the probabilities P_d/P_2 should be approximately the same as the ratio of the observed numbers $\pi_d(x)/\pi_2(x)$—and so the formula (4) just derived justifies the conjectural formula (1)—complete success!

6. Unfortunately, our reasoning is vulnerable and the success is illusory. We left a gap in our derivation (we did not decide how far to go with p) and if we try to fill this gap, we run into trouble. The trouble becomes manifest if we try to apply our reasoning to the simplest analogous problem, the result of which is well known.

PROBLEM III. *Find the probability that x, a large integer chosen at random, is a prime number.*

By reasoning as we did in solving Problem I and assuming the independence of the probabilities involved as we did in solving Problem II we obtain the answer $\prod(p-1)/p$; the product is extended to all primes p not surpassing a cer-

tain bound—but what should be the bound? The number x is certainly prime if it is not divisible by any prime $p<x$. This leads to the evaluation of the desired probability

$$\prod_{p<x} \frac{p-1}{p} \sim \frac{\mu}{\log x} \tag{5}$$

where $\mu=0.561459\cdots=e^{-c}$ and $c=0.577215\cdots$ is the familiar constant of Mascheroni and Euler; the asymptotic evaluation in (5) (on the right hand side of the sign $\sim$) which is valid for $x\to\infty$, is due to Mertens.*

Now, the value (5) is too small. The probability in question is known to be $1/\log x$; this is just the prime number theorem. And we can "explain" somehow why the result is wrong: If the integer x is not divisible by any prime p which does not exceed $x^{1/2}$, x itself must be a prime—and so divisibility by primes exceeding $x^{1/2}$ is, in fact, *not* independent of the smaller primes.

Let us try to modify (5) by considering only primes p not exceeding $x^{1/2}$. This leads us to

$$\prod_{p\leq x^{1/2}} \frac{p-1}{p} \sim \frac{\mu}{\log (x^{1/2})} = \frac{1.122\cdots}{\log x} \tag{6}$$

(we used Mertens' result (5)) and this value is too large.

Let us, however, imitate the physicists who, without hesitation, modify their theories to fit the observed facts. And so let us do a thing between (5) and (6) and extend the product to all *primes not exceeding* x^{μ}. We obtain so

$$\prod_{p<x^{\mu}} \frac{p-1}{p} \sim \frac{1}{\log x}, \tag{7}$$

the right result.

I do not pretend to understand why the introduction of the upper bound x^{μ} *should* yield the right result. For that matter, when the quanta were introduced, no physicist pretended to understand why energy should be obtainable (as salt or sugar is in the self-service store) only in uniform little packages, in multipla of a certain unit. Yet the criterion of a physical theory is its applicability. Let us apply the (unintelligible) trick that gave us the right expression for the prime number theorem to our formula (3). Extending the second product to *odd* primes p inferior to x^{μ}, we are led to

$$\begin{aligned} P_d &= \prod_{p\mid d} \frac{p-1}{p-2}\cdot\frac{1}{2} \prod_{p<x^{\mu}} \frac{p-2}{p} \\ &\sim \prod_{p\mid d} \frac{p-1}{p-2}\cdot 2 \prod_{p<x^{\mu}} \frac{(p-2)p}{(p-1)^2} \frac{1}{(\log x)^2}; \end{aligned} \tag{8}$$

* *Cf.* G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford, 1938, p. 349, Th. 430.

we have used Mertens' result (5). It is easily seen that (8) is equivalent to

$$P_d \sim 2C_2 \prod_{p|d} \frac{p-1}{p-2} \frac{1}{(\log x)^2}, \tag{9}$$

where C_2 stands for the convergent infinite product

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

extended to all odd primes $p = 3, 5, 7, 11, \cdots$. The asymptotic formula (9) is due to Hardy and Littlewood, yet even their argument, which is incomparably deeper and more difficult than the one presented here, does not prove (9); it just confers on (9) another kind of plausible evidence. Yet all available numerical data also seem to support (9).

Let us recall that we have attained (9) by combining two analogies, one of which was extremely "natural" and the other (the "trick of the magic μ") extremely "artificial." And let us try to draw the moral: mathematicians and physicists think alike; they are led, and sometimes misled, by the same patterns of plausible reasoning.*

A CHAIN OF CYCLIC GROUPS

ROY DUBISCH, Fresno State College

1. Introduction. Consider the chain of groups $\mathfrak{G}_0, \mathfrak{G}_1, \cdots, \mathfrak{G}_i, \cdots$ where $\mathfrak{G}_0$ is a cyclic group of order m, and $\mathfrak{G}_i$ is the automorphism group of $\mathfrak{G}_{i-1}$, $i = 1, 2, \cdots$. We ask when the chain consists entirely of cyclic groups. Obviously, when $m = 1$ this will be so, and we suppose henceforth that $m > 1$.

When $\mathfrak{G}_0$ is cyclic of order m with generator a it is well known that its automorphism group, $\mathfrak{G}_1$, is of order $t = \phi(m)$ and that $\mathfrak{G}_1$ is isomorphic to the multiplicative group modulo m of integers less than and relatively prime to m.

* See G. H. Hardy and J. E. Littlewood, Some problems of "Partitio numerorium": On the expression of a number as a sum of primes, Acta Math., vol. 44, 1922, pp. 1–70, especially Conjecture B on p. 42. The more general conjecture on p. 61 (Theorem X 1) is also obtainable by the foregoing reasoning. See also the literature quoted (and criticized) on pp. 32–34, especially the writings of Sylvester, concerning the use of probabilities in questions of similar nature. The crux of the matter may be so expressed: When we consider a fixed number of primes, the "probabilities" introduced can be regarded as "independent," but they cannot be so regarded when the number of primes considered increases in an arbitrary manner. (*Added in proof.* Profesor E. M. Wright drew my attention to a paper by the late Lord Cherwell in the Quart. J. Math., vol. 17, 1946, pp. 46–62, which has a certain contact with the present paper, and to a paper by Lord Cherwell and himself which is scheduled to appear in a coming volume of the Quarterly.)