Solomon W. Golomb

Arithmetica topologica

In: (ed.): General Topology and its Relations to Modern Analysis and Algebra, Proceedings of the symposium held in Prague in September 1961. Academia Publishing House of the Czechoslovak Academy of Sciences, Prague, 1962. pp. 179--186.

Persistent URL: http://dml.cz/dmlcz/700933

ARITHMETICA TOPOLOGICA

S. W. GOLOMB

Pasadena

The topology D for the positive integers is obtained when those arithmetic progressions $\{an + b\}$ with $(a, b) = 1$ are taken as a basis for the open sets. This topology is connected and Hausdorff, but is neither regular, compact, nor locally compact. In the topology D, there is a simple proof of the existence of infinitely many primes, and Dirichlet's theorem on primes in arithmetic progressions is equivalent to the assertion that the primes form a dense subset of the integers. The interior of the set of primes is empty.

Let $Q = \{q_i\}$ be an infinite subset of the primes, and let $g(x)$ denote the number of members of Q which do not exceed x. Call Q *rare* if $\sum 1/q_i$ converges and call Q *sparse* if $g(x) = o\,(x/\log x)$. The causal relation between rarity and sparsity is settled, and it is shown that a certain condition on the structure of Q implies both that Q is *nowhere dense* (as a topological subset of the integers) and that Q is *sparse* (in the purely metric-analytic sense), although neither nowhere-density alone nor sparsity alone implies the other. It thus becomes possible not merely to formulate prime density problems and other sieve method problems in purely topological terms, but to solve problems of the type treated in [8] without resort to computation. Thus many of the results in [8] have been improved and extended by viewing them in their topological setting.

The topological viewpoint is useful not merely to formulate and solve traditional problems of prime number theory, but also to suggest problems of an essentially different character. Thus, one can form "Cantor Sets" of integers and examine the density of the primes contained therein.

I. A CONNECTED TOPOLOGY FOR THE INTEGERS

A topology D for the positive integers is obtained when those arithmetic progressions $\{an + b\}$ with $(a, b) = 1$ are taken as a basis for the open sets. They form a basis because the intersection of two such progressions is of the same type, or empty, as is easily verified. Note that every nonempty open set, being a union of arithmetic progressions, must be infinite.

This topology furnishes an interesting proof of

Theorem 1. *The number of primes is infinite.*

Proof. If p is prime, the progression $\{np\}$ is closed, since its complement is

$$\{np + 1\} \cup \{np + 2\} \cup \ldots \cup \{np + (p - 1)\},$$

a union of open sets. Consider the union $X = U_p\{ap\}$ extended over all primes. If this is a finite union of closed sets, then X is closed. But the complement of X is $\{1\}$, which is neither empty nor infinite. Since the complement of X is not open, X cannot be closed, the union is not a finite one, and the number of primes is infinite.

(A similar proof, in a stronger and very disconnected topology, was given by H. FURSTENBERG [2].)

Theorem 2. *The topology D is Hausdorff.*

Proof. Given distinct positive integers s and t, choose a prime p (by Theorem 1) which exceeds max (s, t). Then $\{pn + s\}$ and $\{pn + t\}$ are disjoint open sets which separate s and t.

Theorem 3. *The topology D is connected.*

Proof. Suppose the integers could be represented as the union of two disjoint nonempty open sets O_1 and O_2. Let $\{a_1n + b_1\}$ be a basis set in O_1, and let $\{a_2n + b_2\}$ be a basis set in O_2. Let α be a multiple of a_1. If α were in O_2, we would have $\alpha = An_0 + B$, where $\{An + B\} \subset O_2$. Since $(A, B) = 1$, we would have $(\alpha, A) = 1$, and hence $(a_1, A) = 1$. But then $\{a_1n + b_1\}$ and $\{An + B\}$ would intersect infinitely often, contradicting disjointness of O_1 and O_2. Thus all multiples of a_1 must belong to O_1. Similarly the multiples of a_2 must belong to O_2. But then the *common* multiples of a_1 and a_2 must belong to both O_1 and O_2, which contradicts disjointness.

A proof of the connectedness of the topology D, without reference to number theory, was presented by MORTON BROWN at the April 1953 meeting of the American Mathematical Society in New York [1].

Theorem 4. *The topology D is not regular.*

Proof. Suppose that open coverings are given for the closed set $\{2n\}$ and for the point $\{1\}$ outside it. Any open covering of $\{1\}$ not intersecting $\{2n\}$ must include a progression $\{en + 1\}$, where e is even. That is, $e \in \{2n\}$. Let $\{an + b\}$ be the member of the open covering of $\{2n\}$ which contains e, so that $e = an_0 + b$. Since $(a, b) = 1$, we have $(a, e) = 1$, whereby $\{an + b\}$ intersects $\{en + 1\}$ infinitely often. Thus the closed set $\{2n\}$ and the point $\{1\}$ cannot have disjoint open neighborhoods.

Theorem 5. *The topology D is not compact.*

Proof. The union $\bigcup_p \{np - 1\}$ extended over all primes is an infinite open covering for the positive integers. Since the omission of any progression $\{nq - 1\}$ leaves the number $q - 1$ uncovered, the Heine-Borel property fails.

Actually, the topology D is not even locally compact, because every locally compact Hausdorff space is regular. For a proof of this, as well as for the more basic definitions of point-set topology, the reader is referred to [5].

Dirichlet's theorem, which asserts that every progression $\{an + b\}$ with $(a, b) = 1$ contains infinitely many primes, has an elegant formulation in terms of the topology D.

Theorem 6. *Dirichlet's theorem is equivalent to the assertion that the primes are a dense subset of the integers in the topology D.*

Proof. Assume first the validity of Dirichlet's theorem. Then every nonempty open set contains primes, so that the primes are a dense subset of the integers. Conversely, assume that the primes are a dense subset. Then every nonempty open set, and in particular all the progressions $\{an + b\}$ with $(a, b) = 1$, must contain primes. It is well known [4] that if every such progression contains at least one prime, then every such progression contains infinitely many primes. (In topological terminology: "If the closure of the primes is the integers, then the derived set of the primes is the integers.")

Although it is quite unlikely that a complete topological proof of Dirichlet's theorem could be given without the introduction of powerful new ideas and methods, the attempt should be well worth the effort. In particular, if the proof that works for the rational integers should also be valid in other rings of algebraic integers (where the corresponding topology, based on residue classes of ideals, is introduced), the enrichment of number theory would be enormous. Thus, the corresponding theorem for the Gaussian integers would imply infinitely many Gaussian primes in the progression $\{n + i\}$, and hence infinitely many rational primes of the form $n^2 + 1$, a classical unsolved problem.

Another familiar fact capable of topological formulation is

Theorem 7. *In the topology D, the interior of the set of primes is empty.*

Proof. If there were an open set consisting entirely of primes, there would be a progression $\{an + b\}$ with $1 \leqq b \leqq a$ consisting entirely of primes. But with

$$n_0 = a + b + 1\,, \quad an_0 + b = (a + b)(a + 1)\,,$$

which is composite.

It is interesting to consider also the topology D' for the positive integers, which has as a basis those progressions $\{an + b\}$ with $(a, b) = 1$ for all $n > N$. (Here N is allowed to assume all values.) This topology may appear *stronger* than D, although it is in fact equivalent to D. Moreover, certain theorems related to Eratosthenes' sieve are readily seen in terms of D'. In particular,

Theorem 8. *The set of positive integers m such that $6m - 1$ and $6m + 1$ are a pair of "prime twins" is closed in D', and hence in D.*

Proof. It is known [3], [10], that the numbers m in question are precisely those positive integers not expressible in the form $6ab \pm a \pm b$ for any $a \geqq 1$ and $b \geqq 1$. Thus the *complement* of our set is

$$\bigcup_{b \geqq 1} \{(6b \pm 1)\, a \pm b\}\,,$$

where each progression is restricted to $a \geqq 1$, and is open because $(6b \pm 1, b) = 1$. The union is open in D', because it is a union of open sets. Thus the integers m for which $6m - 1$ and $6m + 1$ are both prime form a closed set.

II. RARITY AND SPARSITY

Let $Q = \{q_i\}$ be an infinite subset of the primes, and let $g(x)$ denote the number of members of Q which do not exceed x. Call Q *rare* if $\sum 1/q_i$ converges, and call Q *sparse* if $g(x) = o\,(x/\log x)$ [7]. In this section the causal relation between rarity and sparsity is settled in the negative, i. e.:

Theorem 9. *Rarity is unnecessary and insufficient for sparsity.*

Proof. Four examples suffice to establish this Theorem. (The most surprising of these is the fourth.)

1. The set of all primes, $P = \{p_i\}$, is neither sparse nor rare. (Trivial.)
2. The subset $Q_{rs} = \{p_{n!}\}$ is both sparse and rare. (Trivial.)
3. Define $Q_s = \{q_n\}$ recursively by $q_1 = 2$, and

$$q_{n+1} = \max\left[p_{\pi(q_n)+1}\,,\; p_{\pi(n\log n\log\log n)}\right].$$

Since $p_{\pi(y)}$ is the largest prime not exceeding y, and since $p_n \sim n\log n$, we see that $q_n \sim n\log n\log\log n$, so that

$$\sum\frac{1}{q_n} > k\sum 1/n\log n\log\log n = \infty\,,\quad \text{while}\quad g(x)\sim x/\log x\log\log x = o(x/\log x)\,.$$

The use of the Prime Number Theorem here can easily be replaced by more elementary results. (See also the last paragraph of this paper.)

4. Let

$$Q_r = \bigcup_{i=1}^{\infty}\ \bigcup_{j=1}^{2^{2^i-i}} p_{\pi(2^{2^i})+j}\,.$$

That is, after each value $x = 2^{2^i}$, the set Q_r contains the next 2^{2^i-i} primes. Using the elementary (Chebycheff) results: $\pi(x) < ax/\log x$ and $p_n < bn\log n$, we have, letting $c = a\log_2 e + 1$,

$$p_{\pi(2^{2^i})+2^{2^i-i}} < b[\pi(2^{2^i}) + 2^{2^i-i}]\log[\pi(2^{2^i}) + 2^{2^i-i}] <$$
$$< bc\,\frac{2^{2^i}}{2^i}\,[\log c + (2^i - i)\log 2] < k\,.\,2^{2^i}\,,$$

where k is independent of i.

Hence,

$$\sum_{q\in Q_r}\frac{1}{q} \leqq \sum_{i=1}^{\infty}\frac{2^{2^i-i}}{2^{2^i}} = \sum_{i=1}^{\infty}\frac{1}{2^i} = 1 < \infty\,,$$

while for $x = k\,.\,2^{2^i}$, which includes arbitrarily large values of x,

$$g(x) = g(k\,.\,2^{2^i}) \geqq 2^{2^i-i} = \frac{x/k}{\log(x/k)}\log 2 > d\,\frac{x}{\log x}\,,$$

for some absolute constant $d > 0$. This precludes $g(x) = o(x/\log x)$.

Thus the supposition that rarity is a stronger condition than sparsity is false, at least in the case of sets of primes with sufficiently irregular distribution. However, rarity *does* imply

$$\liminf_{x\to\infty} \frac{g(x)}{x/\log x} = 0\,,$$

but this is a weaker condition than sparsity, which can be rephrased

$$\lim_{x\to\infty} \frac{g(x)}{x/\log x} = 0\,.$$

In some sense the *intersection* of the rarity and sparsity conditions is the requirement

$$\liminf_{x\to\infty} \frac{g(x)}{x/\log x} = 0\,.$$

A somewhat more general approach is to define a "moment-generating function" for Q by

$$g(x, s) = \sum_{q \leqq x} q^{-s}\,.$$

The rarity *vs.* sparsity problem then becomes one of the interrelationship between the asymptotic behaviors of the two "moments" $g(x, 0)$ and $g(x, 1)$.

III. SPARSITY AND TOPOLOGICAL DENSITY

Definition. Let $A = \{an + b\}$ be any arithmetic progression. Define the *π-measure* on A *by*

$$\pi(A) = \begin{cases} 0 & \text{if } (a, b) > 1\,, \\ 1/\varphi(a) & \text{if } (a, b) = 1\,, \end{cases}$$

where $\varphi(a)$ is Euler's function.

(For the empty set Φ, define $\pi(\Phi) = 0$.)

If A and B are two progressions, so too is $A \cap B$. The formula $\pi(A \cup B) = \pi(A) + \pi(B) - \pi(A \cap B)$ may be used to extend the definition of π-measure to all *finite* unions of arithmetic progressions.

Theorem 10. *As a measure, $\pi(A)$ is finitely additive but not absolutely additive.*

Proof. By the principle of cross-classification,

$$\pi\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum \pi(A_i) - \sum \pi(A_i \cap A_j) + - \ldots + (-1)^{n+1}\, A_1 \cap A_2 \cap \ldots \cap A_n\,.$$

To show that this measure is not absolutely additive, consider the set of progressions $A_i = \{p_i n\}$, where p_i is the *i-th* prime, $i = 1, 2, 3, \ldots$, and define $A_0 = \{qn + 1\}$ for any odd prime q. Although $\pi(A_i) = 0$ for $i \geqq 1$, and though $\pi(A_0) = 1/(q - 1)$ where q may be arbitrarily large, yet $\bigcup_{i=0}^{\infty} A_i = Z$, the set of *all* integers, and $\pi(Z) = 1$.

Note. By the asymptotic form of Dirichlet's Theorem, the number of primes in a progression $A = \{an + b\}$ which do not exceed x, denoted by $\pi(x; a, b)$, satisfies

$$\lim_{x \to \infty} \frac{\pi(x; a, b)}{x/\log x} = \pi(A) .$$

Hence an equivalent *definition* for π-measure would be

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

or even more suggestively,

$$\pi(A) = \lim_{x \to \infty} \frac{\pi(x, A)}{\pi(x)} ,$$

and this definition generalizes to *any* set A.

Definition. Any set A having π-measure 0 will be called *sparse*, and any subset of a sparse set is also called sparse, and defined to have π-measure 0. (When this definition is restricted to sets of *primes*, it clearly coincides with the definition of sparsity given in Section II of this paper.)

Theorem 11. *The set A will be called essentially sparse if* $\pi(Z - A) = 1$. *Essential sparsity is equivalent to sparsity.*

Proof.

$$\pi(Z - A) = \lim_{x \to \infty} \frac{\pi(Z - A, x)}{\pi(x)} = \lim_{x \to \infty} \frac{\pi(Z, x) - \pi(A, x)}{\pi(x)} = \lim_{x \to \infty} \left(1 - \frac{\pi(A, x)}{\pi(x)}\right),$$

which is 1 if and only if

$$\pi(A) = \lim \frac{\pi(A, x)}{\pi(x)} = 0 .$$

Definition. Relative to the topology D, a subset S of the integers Z is called *nowhere dense* if there is no non-empty open set in the closure of S.

Theorem 12. *If S is a nowhere dense subset of Z, then S may or may not be sparse, and conversely.*

Proof. The empty set is *both* nowhere dense and sparse; the set of all primes is *neither*.

Let $\{A_i\}$ be the denumerable collection of all the basis sets $\{a_i n + b_i\}$. From each A_i, pick a prime $s_i > 2^i$, and let $S = \{s_i\}$. Then S contains at least one, and hence infinitely many (*cf.* Theorem 6) members of every non-empty open set, and is therefore dense in D. However,

$$\pi(S) = \lim_{x \to \infty} \frac{\pi(x, S)}{\pi(x)} \leq \lim_{x \to \infty} \frac{\log_2 x}{x/\log x} = 0 ,$$

so that S is both dense and sparse.

Finally, from each progression A_i it is possible to remove a "small" subprogression B_i so that most of the primes of A_i are still in $A_i - B_i$, and so that $\mathcal{S} = Z - \bigcup B_i$

is a set which is nowhere dense, yet has $\pi(\mathcal{S})$ arbitrarily close to 1. Care must be exercised in the choice of B_i to prevent the infinite union $\bigcup B_i$ from containing too many (or even all) of the primes. This can be done by assuring that the smallest prime in B_i exceeds 2^i, along with

$$\pi(x, B_i) < \frac{1}{N \cdot 2^i} \frac{x}{\log x}$$

for all $x > 1$ and suitable large N. The remaining details are left as an exercise.

Note that the last set $\mathcal{S}$ is a kind of *Cantor set*. By removing smaller and smaller "intervals" B_i, one is left with a set which is nowhere dense, which is perfect (*i. e.* $\mathcal{S}$ is the set of its own limit points), but which has "measure" arbitrarily close to 1. This is an example of the way in which topological notions can be used to exhibit sets of primes which are more pathological than those previously studied.

Theorem 13. *If Q is a set of primes such that, with only finitely many exceptions, $Q \subset A_i = \{a_i n + b_i\}$ for arbitrarily large values of a_i, then Q is both sparse and nowhere dense.*

Proof. To show sparsity, let $\varepsilon > 0$ be given, and pick a_i so large that

$$\frac{1}{\varphi(a_i)} < \varepsilon/4 \quad \left(\text{using } \varphi(n) \to \infty \text{ as } n \to \infty\right).$$

There are only finitely many – say t – members of Q which do not belong to $A_i = \{a_i n + b_i\}$, by the hypothesis. Pick x_1 so large that

$$\frac{t}{x_1/\log x_1} < \varepsilon/2 .$$

Pick x_2 so large that for all $x \geqq x_2$,

$$\frac{\pi(x, A_i)}{x/\log x} \leqq \frac{2}{\varphi(a_i)} < \frac{\varepsilon}{2} .$$

Pick $x_0 = \max(x_1, x_2)$. Then

$$\frac{\pi(x, Q)}{x/\log x} \leqq \frac{t}{x/\log x} + \frac{\pi(x, A_i)}{x/\log x},$$

for all $x > x_0$, whence

$$\pi(Q) = \lim_{x \to \infty} \frac{\pi(x, Q)}{x/\log x} < \varepsilon$$

for all $\varepsilon > 0$, and $\pi(Q) = 0$.

To show that Q is nowhere dense, assume the contrary. Then there would be a progression B such that *every* subprogression of B satisfying the relative-prime condition has non-empty intersection with Q. Let $\pi(B) = \beta$. By hypothesis, with only finitely many exceptions (say t exceptions), $Q \subset A_i$ with $\pi(A_i) < \frac{1}{2}\beta$. Thus $\pi(B - A_i) \geqq \frac{1}{2}\beta$, where $B - A_i$ denotes the intersection of B with the complement of A_i. Any open progression in $B - A_i$ (and there must be at least one) can be decomposed into more than t disjoint open subprogressions, with only t members of Q available to be in them, contradicting the assumption that Q could be dense in B.

An immediate application of Theorem 12 is to the improvement of Theorem 4 in [8], which asserts that if P_n is any particular prime factor of $F_n = 2^{2^n} + 1$, then the set $\{P_n\}$ has "intermediate density", a weaker condition than sparsity.

Theorem 14. *The set of all prime factors of all the Fermat numbers* $F_n = 2^{2^n} + 1$ *form a sparse set of primes.*

Proof. As shown in [8], every prime factor of F_n belongs to $A_n = \{2^{n+1}K + 1\}$. Since $A_1 \supset A_2 \supset A_3 \supset \ldots$ with $\Pi(A_n) \to 0$, Theorem 13 applies, and asserts that the set of all prime factors of the F_n are a *sparse* set, and are also nowhere dense.

Using the concepts of the present paper, Theorem 3 of [8] can also be strengthened. Let Q be the set of odd primes defined inductively by starting with $3 \varepsilon Q$, and placing each subsequent prime into Q if and only if it fails to be congruent to 1 modulo any of the previously chosen members of Q. In [8] it was shown that Q has "intermediate density", defined as

$$\liminf_{x \to \infty} \frac{\|q \in Q, q \leqq x\|}{x/\log x} = 0 .$$

A review of the proof, however, shows that Q must be either *rare* or *sparse*, and as shown in Section II of this paper, each of these conditions is stronger than the intermediate density condition.

Using very powerful analytic methods, ERDÖS has recently shown [9] that $q_n \sim n \log n \log \log n$, where q_n is the n-th member of Q. It is unlikely that topological methods will ever replace analysis in obtaining results of this depth.

References

[1] *M. Brown*: A Countable Connected Hausdorff Space. Bull. Amer. Math. Soc., *59* (1953), 367.

[2] *H. Furstenberg*: On the Infinitude of Primes. Amer. Math. Monthly, *62* (1955), 353.

[3] *S. W. Golomb*: Problem E 969. Amer. Math. Monthly, *58* (1951), 338.

[4] *R. Spira*: Problem E 1218. Amer. Math. Monthly, *63* (1956), 342.

[5] *J. L. Kelley*: General Topoiogy. New York 1955.

[6] *S. W. Golomb*: A Connected Topology for the Integers, Amer. Math. Monthly, *66*, No. 8, October (1959), 663–666.

[7] *S. W. Golomb*: Advanced Problem 4970. Amer. Math. Monthly, *68*, No. 5, May (1961), 511.

[8] *S. W. Golomb*: Sets of Primes with Intermediate Density. Mathematica Scandinavica, *3* (1955), 264–274.

[9] *P. Erdös*: On a Problem of S. Golomb. Australian Math. Soc. Jour. vol. II, Part I, April (1961), 1–8.

[10] *W. J. Leveque*: Topics in Number Theory. Vol. I, p. 69, problem 4, Addison-Wesley (1956).