Reflections on Paul Erdős on His Birth Centenary, Part II

Krishnaswami Alladi and Steven G. Krantz, Coordinating Editors

This is Part II of the two-part feature on Paul Erdős following his centennial. Part I had contributions from Krishnaswami Alladi and Steven G. Krantz, Vera T. Sós and László Lovász, Ronald Graham and Joel Spencer, Jean-Pierre Kahane, and Mel Nathanson.

Here in Part II we have six articles from contributors Noga Alon, Dan Goldston, András Sárközy, József Szabados, Gérald Tenenbaum, and Stephan Ramon Garcia and Amy L. Shoemaker.

Paul Erdős in a relaxed mood in Hungary.

Krishnaswami Alladi is professor of mathematics at the University of Florida. His email address is alladi@ufl.edu.

Steven G. Krantz is professor of mathematics at Washington University in St. Louis.

Noga Alon

Paul Erdős and the Probabilistic Method
Probabilistic Beginnings

The probabilistic method is one of the most significant contributions of Paul Erdős. Indeed, Paul himself said, during his eightieth birthday conference in Keszthely, Hungary, that he believes the method will live long after him. This was the only time I heard him making any comment about the significance and impact of his work. He was always more interested in discussing new problems and results than in trying to assess their long-time expected merits.

The method is a powerful technique with numerous applications in combinatorics, graph theory, additive number theory and geometry.

Noga Alon is professor of mathematics at the Sackler School of Mathematics and Blavatnik School of Computer Science, Tel Aviv University, and Institute for Advanced Study, Princeton. His email address is nogaa@tau.ac.il.

The basic idea is very simple: Trying to prove that a structure with certain desired properties exists, one defines an appropriate probability space of structures and then shows that the desired properties hold in this space with positive probability. The amazing fact is that this simple reasoning can lead to highly nontrivial results. The results and tools are far too numerous to cover in a few pages, and my aim here is to give only a glimpse of the topic by describing a few examples of questions and results that illustrate the method. All of these have been initiated by Erdős, motivated by his questions and results. The fact that there is still intensive ongoing work illustrates the influence and long-term impact of his work. More material on the subject can be found in the books [6], [8], [23], [26].

Ramsey Numbers

Ramsey theory is the study of the general phenomenon that every large structure, even if it looks totally chaotic, must contain a rather large well-organized substructure. This holds for many types of structures (though there are exceptions) and yields interesting applications in several mathematical areas. A detailed treatment of the subject can be found in [21].

Although several Ramsey-type theorems appeared earlier, the origin of Ramsey theory is usually credited to Frank Plumpton Ramsey, who proved in 1930 the fundamental theorem that edge colorings of finite or infinite graphs or hypergraphs satisfy such a theorem. The statement for finite graphs is as follows.

Let $H_1, H_2, \ldots, H_k$ be k finite, undirected, simple graphs. Then there is a finite number r such that in every edge coloring of the complete graph on r vertices by k colors, there is a monochromatic copy of H_i in color i for some $1 \le i \le k$. The smallest integer r that satisfies this property is called the (multicolor) *Ramsey number* of $H_1, \ldots, H_k$ and is denoted by $r(H_1, H_2, \ldots, H_k)$.

The determination or estimation of these numbers is usually a very difficult problem, one which fascinated Erdős since the '30s. When each graph H_i is a complete graph K_t with $t > 2$ vertices, the only values that are known precisely are those of $r(K_3, K_m)$ for $m \le 9$, $r(K_4, K_4)$, $r(K_4, K_5)$, and $r(K_3, K_3, K_3)$. The determination of the asymptotic behavior of Ramsey numbers up to a constant factor is also a very hard problem, and despite a lot of effort by various researchers there are only a few infinite families of graphs for which this behavior is known.

In one of the first applications of the probabilistic method in combinatorics, Erdős [12] proved that, if $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, then $R(K_k, K_k) > n$; that is, there

exists a 2-coloring of the edges of the complete graph on n vertices containing no monochromatic clique of size k. This implies that $R(K_k, K_k) > 2^{k/2}$ for all $k \ge 3$. The proof is extremely short: the probability that a random two-edge coloring of K_n contains a monochromatic copy of K_k is at most $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, and hence there is a coloring with the required property.

It is worth noting that, although this argument seems trivial today, it was far from being obvious when published in 1947. In fact, several prominent researchers believed, before the publication of this short paper, that $R(K_k, K_k)$ may well be bounded by a polynomial in k. In particular, Paul Turán writes in [28] that he had conjectured for a while that $R(K_k, K_k)$ is roughly k^2 and that Erdős's result showing that this quantity behaves very differently than expected came to him as a big surprise.

My own first meeting with Paul Erdős took place when I was finishing high school in the early '70s in Haifa, Israel. Paul had a special visiting position at the Technion, and I met him during one of his visits. A few months before that I had read his probabilistic lower bound for the Ramsey numbers $R(K_k, K_k)$, formulated as a counting argument without any mention of probability, and noticed that the argument could be used to provide several similar results. I (proudly) told Erdős about my observations, and he encouraged me to keep thinking about these problems and gave me a book *The Art of Counting* [15], which had just been published at that time. This book contains selected publications of Erdős and is the first serious mathematical book I ever read. Reading it, and taking notes of much of its content, I quickly realized that it contained far more sophisticated extensions of the basic probabilistic lower bound proof of Erdős than the ones I observed. Paul, who surely knew well this fact, chose to suggest that I read the book and keep thinking about these problems, realizing that this is more stimulating than quickly pointing out the relevant references. Indeed, he always felt that young people interested in mathematics should be encouraged, and I am convinced that this approach was fruitful in many cases as it was in mine.

Returning to the asymptotics of Ramsey numbers, a particularly interesting example of an infinite family for which the behavior of the Ramsey number is known is the following result of Kim and of Ajtai, Komlós and Szemerédi.

Theorem 1 ([25], [1]). *There are two absolute positive constants c_1, c_2 such that*

$$c_1 m^2 / \log m \le r(K_3, K_m) \le c_2 m^2 / \log m$$

for all $m > 1$.

The upper bound, proved in [1], is probabilistic and applies a certain random greedy algorithm. There are several subsequent proofs, all based on probabilistic arguments. The lower bound is proved by a "semi-random" construction whose detailed analysis is subtle, relying on several large deviation inequalities. An alternative way of establishing the lower bound, which provides a better constant, appears in two recent papers, [20] and [7], that analyze the so-called "triangle-free process" suggested by Bollobás and Erdős. In this process one starts with a graph on n vertices with no edges and keeps adding uniformly chosen random edges among those that do not create a triangle. At the end, all these chosen edges are colored red and the nonchosen edges are colored blue. Clearly the resulting coloring contains no red triangle, and a careful analysis shows that with high probability there is no blue clique K_m for an appropriate choice of the initial size n.

It is worth noting that the question of obtaining a superlinear lower bound for $r(K_3, K_m)$ was mentioned already in [12], and Erdős established in [13], by an elegant probabilistic construction, an $\Omega(m^2/\log^2 m)$ lower bound.

Even less is known about the asymptotic behavior of multicolor Ramsey numbers, that is, Ramsey numbers with at least three colors. The asymptotic behavior of $r(K_3, K_3, K_m)$, for example, has been very poorly understood for quite some time, and Erdős and Sós conjectured in 1979 (cf., e.g., [10]) that

$$\lim_{m \to \infty} \frac{r(K_3, K_3, K_m)}{r(K_3, K_m)} = \infty.$$

This has been proved in [5], where it is shown that in fact $r(K_3, K_3, K_m)$ is equal, up to logarithmic factors, to m^3. A more complicated, related result proved in [5] that supplies the asymptotic behavior of infinitely many families of Ramsey numbers up to a constant factor is the following.

Theorem 2. *For every $t > 1$ and $s \geq (t-1)! + 1$ there are two positive constants c_1, c_2 such that, for every $m > 1$,*

$$c_1 \frac{m^t}{\log^t m} \leq r(K_{t,s}, K_{t,s}, K_{t,s}, K_m) \leq c_2 \frac{m^t}{\log^t m},$$

where $K_{t,s}$ is the complete bipartite graph with t vertices in one color class and s vertices in the other.

The proof of the lower bound is probabilistic: each of the first three color classes is a randomly shifted copy of an appropriate $K_{t,s}$-free graph that contains a relatively small number of large independent sets, as shown by combining spectral techniques with character sum estimates.

Sum-Free Subsets

A set A of integers is called *sum-free* if there is no solution to the equation $a + b = c$ with $a, b, c \in A$. Erdős [14] showed that any set A of n positive integers contains a sum-free subset of size at least $n/3$. The proof is a short and simple, yet intriguing, application of the probabilistic method. It proceeds by choosing a uniform random x in $(0, 1)$ and by observing that the set of all elements $a \in A$ satisfying $ax \mod 1 \in (1/3, 2/3)$ is sum-free and its expected size is $n/3$.

In [3] the authors showed that a similar proof gives a lower bound of $(n + 1)/3$. Bourgain [9] has further improved this estimate to $(n + 2)/3$. For quite some time it was not clear whether or not the constant $1/3$ could be replaced by a larger constant, until Eberhard, Green, and Manners proved in [11] that the constant $1/3$ is tight. Their proof is a sophisticated argument that contains a crucial probabilistic ingredient. The problem of deciding whether or not every set of n nonzero integers contains a sum-free subset of cardinality at least $n/3 + w(n)$, where $w(n)$ tends to infinity with n, remains open. It will be extremely surprising if there is no such $w(n)$.

List Coloring and Euclidean Ramsey Theory

The *list chromatic number* (or *choice number*) $\chi_\ell(G)$ of a graph $G = (V, E)$ is the minimum integer s such that for every assignment of a list of s colors to each vertex v of G, there is a proper vertex coloring of G in which the color of each vertex is in its list. This notion was introduced independently by Vizing in [29] and by Erdős, Rubin, and Taylor in [19]. In both papers the authors realized that this is a variant of usual coloring that exhibits several new interesting properties and that in general $\chi_\ell(G)$, which is always at least as large as the chromatic number of G, may be arbitrarily large even for graphs G of chromatic number 2.

For about ten years after the initial papers of Vizing and of Erdős, Rubin, and Taylor there was essentially no work on list coloring. Starting in the late '80s, the topic, motivated to a great extent by the many problems raised by Erdős and his collaborators in [19], received a considerable amount of attention. Paul Erdős himself told me in the early '90s that, when they wrote their paper, he thought that the topic was not very exciting and was pleasantly surprised to see that it eventually stimulated so much activity. I view this as a sign showing that Paul was essentially unable to ask any noninteresting questions. When he asked a question, even if at first sight it seemed artificial or nonappealing (even to Paul himself!), almost always it eventually turned out to be interesting.

It is natural to extend the notion of list coloring to hypergraphs. A hypergraph H is an ordered

pair (V, E), where V is a set of vertices and E is a collection of subsets of V, called edges. It is r-uniform if every edge contains exactly r vertices. Thus graphs are 2-uniform hypergraphs. The list chromatic number $\chi_\ell(H)$ of a hypergraph H is the minimum integer s such that, for every assignment of a list of s colors to each vertex of H, there is a vertex coloring of H assigning to each vertex a color from its list, with no monochromatic edges.

An intriguing property of list coloring of graphs which is not shared by ordinary vertex coloring is the fact that the list chromatic number of any (simple) graph with a large average degree is large. Indeed, it is shown in [2] that the list chromatic number of any graph with average degree d is at least $\Omega(\log d)$. For $r \geq 3$, simple examples show that there is no nontrivial lower bound on the list chromatic number of an r-uniform hypergraph in terms of its average degree. However, such a result does hold for simple hypergraphs. Recall that a hypergraph is *simple* if every two of its distinct edges share at most one vertex. The following result is proved in [4].

Theorem 3. *For every fixed $r \geq 2$ and $s \geq 2$, there is a $d = d(r, s)$ such that the list chromatic number of any simple r-uniform hypergraph with n vertices and at least nd edges is greater than s.*

A similar result for the special case of d-regular 3-uniform simple hypergraphs has been obtained independently in [22]. A subsequent proof with a much better upper estimate for $d(r, s)$ appears in a recent paper of Saxton and Thomason [27].

The proof of the theorem is probabilistic. For the simpler case of graphs it shows that if $G = (V, E)$ is a graph with average degree $d > 10^s$, then when we assign to each vertex of G a randomly chosen list consisting of s colors among the colors $\{1, 2, \dots, 2s - 1\}$, then with high probability there is no proper coloring of G assigning to each vertex a color from its list. The precise argument requires some work, and the result suggests an interesting algorithmic question: given a graph $G = (V, E)$ with minimum degree $d > 10^s$, can we find, deterministically and efficiently, lists of size s for each $v \in V$ so that there is no proper coloring of G assigning to each vertex a color from its list? This problem is open, as is the simpler NP version of it, that is, that of exhibiting lists and providing a certificate that there is no proper coloring using them. Here the lists do not have to be found efficiently, and we require only that one will be able to check the certificate efficiently.

The last theorem has an interesting application in Euclidean Ramsey theory, yet another subject initiated by Erdős and his collaborators. A well-known problem of Hadwiger and Nelson is that of determining the minimum number of colors required to color the points of the Euclidean plane so that no two points at distance 1 have the same color. Hadwiger showed in 1945 that seven colors suffice, and Moser and Moser noted in 1961 that three colors do not suffice. These bounds have not been improved despite a considerable amount of effort by various researchers; see [24, pp. 150–152] and the references therein for more on the history of the problem.

A more general problem is considered in [16], [17], [18], where the main question is the investigation of finite point sets K in the Euclidean space for which any coloring of a Euclidean space of dimension d by r colors must contain a monochromatic isometric copy of K. There are lots of intriguing conjectures that appear in these papers. One of them asserts that, for any set K of three points which do not form an equilateral triangle, the minimum number of colors required for coloring the plane with no monochromatic isometric copy of K is three. The situation is very different for list coloring. A simple corollary of the theorem above is the following.

Theorem 4 ([4]). *For any finite set X in the Euclidean plane and for any positive integer s, there is an assignment of a list of size s to every point of the plane such that, whenever we color the points of the plane from their lists, there is a monochromatic isometric copy of X.*

The examples described in this brief survey include applications of the probabilistic method of Paul Erdős in graph theory, Ramsey theory, additive number theory, and combinatorial geometry. There have been recent results in the study of each of these examples, while the roots of all of them lie in the work and questions of Paul. There is no doubt that the study and application of probabilistic arguments will keep playing a crucial role in the development of many mathematical areas in the future, providing further evidence for the profound influence of Erdős. The comment he made at his eightieth birthday conference proved to be accurate: the probabilistic method does live and will stay alive long after him.

References

[1] M. AJTAI, J. KOMLÓS, and E. SZEMERÉDI, A note on Ramsey numbers, *J. Combinatorial Theory Ser. A* **29** (1980), 354–360.

[2] N. ALON, Degrees and choice numbers, *Random Structures and Algorithms* **16** (2000), 364–368.

[3] N. ALON and D. J. KLEITMAN, Sum-free subsets, in *A Tribute to Paul Erdős* (A. Baker, B. Bollobás, and A. Hajnal, eds.), Cambridge University Press, Cambridge, 1990, pp. 13–26.

[4] N. ALON and A. V. KOSTOCHKA, Hypergraph list coloring and Euclidean Ramsey Theory, *Random Structures and Algorithms* **39** (2011), 377–390.

[5] N. Alon and V. Rödl, Asymptotically tight bounds for some multicolor Ramsey numbers, *Combinatorica* **25** (2005), 125–141.

[6] N. Alon and J. H. Spencer, *The Probabilistic Method*, Third Edition, Wiley, New York, 2008.

[7] T. Bohman and P. Keevash, *Dynamic Concentration of the Triangle-Free Process*, CRM Series, 16, Ed. Norm., Pisa, 2013.

[8] B. Bollobás, *Random Graphs*, Second Edition, Academic Press, London, 2001.

[9] J. Bourgain, Estimates related to sumfree subsets of sets of integers, *Israel J. Math.* **97** (1997), 71–92.

[10] F. Chung and R. L. Graham, *Erdős on Graphs: His Legacy of Unsolved Problems*, A K Peters, Wellesley, MA, 1998.

[11] S. Eberhard, B. Green, and F. Manners, Sets of integers with no large sum-free subset, *Ann. of Math. (2)* **180** (2014), no. 2, 621–652.

[12] P. Erdős, Some remarks on the theory of graphs, *Bulletin of the Amer. Math. Soc.* **53** (1947), 292–294.

[13] _____, Graph theory and probability. II, *Canad. J. Math.* **13** (1961), 346–352.

[14] _____, Extremal problems in number theory, Proc. Sympos. Pure Math., Vol. VIII Amer. Math. Soc., Providence, RI, 1965, pp. 181–189.

[15] _____, *The Art of Counting: Selected Writings*, edited by Joel Spencer and with a dedication by Richard Rado, Mathematicians of Our Time, Vol. 5 The MIT Press, Cambridge, Mass.-London, 1973.

[16] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus, Euclidean Ramsey theorems, I. *J. Combinatorial Theory Ser. A* **14** (1973), 341–363.

[17] _____, Euclidean Ramsey theorems, II, Infinite and finite sets (Colloq., Keszthely, 1973), Vol. I, pp. 529–557. Colloq. Math. Soc. Janos Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.

[18] _____, Euclidean Ramsey theorems, III. Infinite and finite sets (Colloq., Keszthely, 1973), Vol. I, pp. 559–583. Colloq. Math. Soc. Janos Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.

[19] P. Erdős, A. L. Rubin, and H. Taylor, Choosability in graphs, *Proc. West Coast Conf. on Combinatorics, Graph Theory and Computing*, Congressus Numerantium, XXVI, Utilitas Math., Winnipeg, Man., 1980, 125–157.

[20] G. Fiz Pontiveros, S. Griffiths, and R. Morris, *The triangle-free process and $R(3, k)$*, arXiv:1302.6279, 2013.

[21] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, Second Edition, Wiley, New York, 1990.

[22] P. E. Haxell and J. Verstraete, List coloring hypergraphs, *Electron. J. Combin.* **17** (2010), no. 1, Research Paper 129, 12 pp.

[23] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[24] T. Jensen and B. Toft, *Graph Coloring Problems*, John Wiley and Sons Inc., New York, 1995.

[25] J. H. Kim, The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$, *Random Structures and Algorithms* **7** (1995), 173–207.

[26] M. Molloy and B. Reed, *Graph Coloring and the Probabilistic Method*, Springer-Verlag, Berlin, 2001.

[27] D. Saxton and A. Thomason, *Hypergraph containers*, arXiv:1204.6595v2, 2012.

[28] P. Turán, *On the theory of graphs*, Colloquium Math. **3** (1954), 19–30.

[29] V. G. Vizing, Coloring the vertices of a graph in prescribed colors (in Russian), *Diskret. Analiz.* No. 29, Metody Diskret. Anal. v. Teorii Kodov i Shem (1976), 3–10, 101.

D. A. Goldston

Erdős's Work on Primes

Prime numbers were an early and abiding interest of Erdős throughout his life. The second paper in his collected works [2] from 1932 gave a new proof of Tschebyschef's theorem on bounding the number of primes up to x. Many more papers on primes followed over more than sixty years. Of course Erdős published on topics in a large subset of mathematics, but from his talks later in life it was clear that the primes held a special attraction for him. In addition to his theorems, Erdős was constantly asking questions about primes and frequently offering money for solutions to his problems. These questions have led to surprising developments and new fields of study. Many remain unsolved, but a surprising number have been solved. In this short survey of some of Erdős's work on primes, I can mention only a selection of his results. I will also include a few of his lesser-known results that have interested me over many years.

The Elementary Proof of the Prime Number Theorem

By far the best known of Erdős's contributions to prime numbers is the elementary proof of the Prime Number Theorem (PNT) in 1949. This states, on letting $\pi(x)$ denote the number of primes $\leq x$ that

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

The PNT was first proved independently by Hadamard and de la Vallée-Poussin in 1896, but this and later proofs all depended on information on the zeros of the Riemann zeta-function $\zeta(s)$. It was eventually shown that the PNT follows from the fact that $\zeta(1 + it) \neq 0$ for all real t and conversely that the PNT implies this result. Thus it was argued that no elementary proof could be obtained and if such a proof was found, then possibly this would be a step towards proving the Riemann Hypothesis. It was an electrifying moment in 1949 when an elementary proof was discovered. Much has been written about the circumstances and controversy surrounding how Atle Selberg and Erdős obtained their proofs, but putting aside

D. A. Goldston is professor of mathematics at San José State University. His email address is daniel.goldston@sjsu.edu.

the human story, both Erdős's and Selberg's papers, [5] and [27], describe well the mathematical contributions that each made to the proof. Both papers being elementary require as a prerequisite only a beginning knowledge of number theory and analysis and are well worth reading. In 1951 Erdős won the Cole Prize for his work in number theory with his paper [5] specifically mentioned, and in 1952 Selberg won the Fields Medal with his work on the elementary proof of PNT included as one of his major accomplishments. Some commentators have expressed disappointment over how little the elementary proof has influenced current work, but one can equally well argue that it is often our rarely used but prized possessions that save the day.

Large Gaps between Consecutive Primes

Erdős was very interested in the sequence of differences (or gaps) between consecutive primes and returned often to questions on this sequence. By the PNT there are $\sim \frac{x}{\log x}$ primes in the interval $[x, 2x]$, and therefore the average distance between primes in this interval is $\sim \log x$. If p_n denotes the nth prime, then the nth difference between consecutive primes $d_n = p_{n+1} - p_n$ is on average $\log p_n$. Much work has been done on irregularity in the distribution of these differences. For large gaps the best result known is that

$$p_{n+1} - p_n > c \log p_n \frac{\log \log p_n (\log \log \log \log p_n)}{(\log \log \log p_n)^2}$$

for infinitely many n. Erdős in 1935 [3] obtained the above result except for the quadruple log factor, which was first obtained by Rankin in 1938 [23]. The constant c has been improved many times over the last fifty years; the best result currently known is $c = 2e^{\gamma}$, where γ is Euler's constant, which is due to Pintz [22] in 1997. Erdős offered US$10,000 for anyone who could prove that this result is true for arbitrarily large c, which compared to other Erdős offers is a surprisingly large amount for what appears to be a very specialized result.

Erdős was always eager to examine patterns within patterns. In 1949 [6] he proposed showing that

$$D_k(n) = \min(d_n, d_{n+1}, \ldots, d_{n+k-1})$$

is sometimes large; i.e., there exist k consecutive extremely large gaps between consecutive primes. He succeeded in proving that $\lim \sup_{n \to \infty} D_2(n) = \infty$. The general problem was finally answered in 1981 by Helmut Maier[16], who proved the surprisingly strong result that

$$D_k(n) > c(k) \log p_n \frac{\log \log p_n (\log \log \log \log p_n)}{(\log \log \log p_n)^2}$$

for infinitely many n. This paper of Maier introduced the "Maier matrix method," which Maier in 1985 [17] used in a startling fashion to prove that

Paul Erdős lecturing at the Institute of Mathematical Sciences, Madras, India, in January 1984. On the board is the inequality on large prime gaps alluded to in the Alladi-Krantz article and by Goldston. Also on the board is a US$500 problem of Erdős.

the number of primes in intervals $[x, x + (\log x)^M]$ for any number M is distributed too irregularly to have an asymptotic formula. There have been many further applications of Maier's method, and recently it has been developed by Granville and Soundararajan [12] into a general uncertainty concept for sequences. This snowballing effect set off by a seemingly off-hand question of Erdős has occurred over and over again.

Small Gaps between Consecutive Primes

One expects that there will be infinitely many primes differing by any given even number and in particular that there will be infinitely many twin primes with difference 2. Up until 2013 it was widely believed to be an extremely hard problem to prove that any such even number exists, i.e., that the primes do not become isolated further and further apart when one examines larger and larger numbers. The year 2013 will go down in history as a spectacular year for prime numbers, because for the first time it was proved by Yitang Zhang that there actually exist even numbers which are the difference of infinitely many primes. Thus there always exist bounded gaps between primes. Later in 2013 James Maynard, by a different method, also obtained this result, finding infinitely often there are primes differing by a number less than 600 and also proving the even more spectacular result that there are bounded gaps between k primes for any given number k.

Progress on this problem prior to 2013 was painfully slow, and the results obtained concern

the sequence of normalized gaps between primes

$$\left\{ \frac{p_{n+1} - p_n}{\log p_n} \right\}_{n=1}^{\infty}.$$

Since the average of this sequence is 1, the first goal is to prove that this sequence has a limit point less than 1 and thus establish the existence of infinitely many smaller-than-average gaps between primes. Letting

$$\Xi = \liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n},$$

our goal is to prove $\Xi < 1$. The first contribution to this problem was due to Hardy and Littlewood in an unpublished paper from around 1922. They proved, using the circle method, that assuming the Generalized Riemann Hypothesis (GRH) one can prove that $\Xi \le \frac{2}{3}$; Rankin [25] improved this, still assuming GRH, to $\Xi \le \frac{3}{5}$. The first unconditional result was due to Erdős [4], who proved $\Xi \le 1 - A$ for an absolute positive constant A. Erdős first observed from the Brun sieve that the number of primes differing by k is bounded accurately, and then by adding this bound over k in an interval, one finds that it is not possible for all consecutive prime differences to be exactly located at the average spacing. This method of Erdős is frequently available as an add-on to other methods. For example, in the influential paper of Bombieri and Davenport [1] in 1965, they were able to make Hardy and Littlewood's method unconditional and prove $\Xi \le \frac{1}{2}$ and then use Erdős's method to improve this to $\Xi \le 0.46550\ldots$. When Maier in 1988 found a third method for finding small gaps between primes he combined it with both of the earlier methods to obtain $\Xi \le 0.2484\ldots$. This paper also answered for the first time the Erdős question of showing for

$$E_k(n) = \max\left(\frac{d_n}{\log p_n}, \frac{d_{n+1}}{\log p_{n+1}}, \ldots, \frac{d_{n+k-1}}{\log p_{n+k-1}} \right)$$

that $\liminf_{n \to \infty} E_2(n) < 1$. Because of the recent work of Maynard we now know $\liminf_{n \to \infty} E_k(n) = 0$ and much more. In 2005 Goldston, Pintz, and Yıldırım (see [11]) finally proved $\Xi = 0$, and their method was used as a starting point by both Zhang and Maynard.

With the recent advances on small gaps, it is reasonable to ask whether the earlier work should now be consigned to the dustbin of history. The answer is "no." The new methods are exquisitely adapted to examine primes and almost primes in various configurations, but often they do not say much about the number of primes in longer intervals with length around the average spacing between primes. The Erdős and Bombieri-Davenport methods are adapted and best applied to questions in this regime rather than very short gaps between primes.

At the end of his 1940 paper on small gaps between primes, Erdős made the conjecture that

$$\sum_{p_n \le x} (p_{n+1} - p_n)^2 = O(x \log x),$$

a result not likely to be proved anytime soon, since the Riemann Hypothesis only implies the bound $O(x \log^3 x)$ [26]. Erdős continued:

> This result if true must be very deep. I could not even prove the following very much more elementary conjecture : Let n be any integer and let $0 < a_1 < a_2 < \cdots < a_x < n$ be the $\phi(n)$ integers relatively prime to n; then
>
> $$\sum_{i=1}^{x-1} (a_{i+1} - a_i)^2 < c \frac{n^2}{\phi(n)}.$$

Progress was slow on this problem. Consider the generalized conjecture that

$$\sum_{i=1}^{\phi(n)} (a_{i+1} - a_i)^\gamma = O\left(\frac{n^\gamma}{\phi(n)^{\gamma-1}} \right).$$

Hooley [15] proved in 1963 that this is true for $1 \le \gamma < 2$. The breakthrough was in 1986 when Montgomery and Vaughan [27] proved not just Erdős's conjecture but that this bound holds for all fixed $\gamma \ge 1$. The techniques developed for this problem have subsequently been applied conditionally to higher moments for the primes in short intervals by Montgomery and Soundararajan [19].

Arithmetic Progressions of Primes

In 2004 Ben Green and Terry Tao proved a long-standing conjecture that there are arbitrarily long arithmetic progressions of primes [13]. Previously it was known only that there were 3 primes in arithmetic progression, and current computation has found only 26 primes in arithmetic progression. The Green-Tao theorem makes use of many difficult techniques, including Szemerédi's Theorem of 1975, which was conjectured by Erdős and Turàn in 1936. In 1973 Erdős made the following conjecture [8], which implies both Szemerédi's theorem and the Green-Tao theorem: For any set of positive integers where the sum of the reciprocal of the integers diverges, the set contains arbitrarily long arithmetic progressions. Erdős offered $3,000 for a proof, but no proof is yet in sight.

Limit Points for Normalized Prime Gaps and Jumping Champions of Primes

I will conclude by mentioning two amusing problems for which much can be conjectured but nothing of consequence proved. The first concerns the normalized consecutive prime gaps mentioned earlier. Ricci [24] and Erdős [7] independently

proved that the set of limit points for this sequence has positive Lebesgue measure. Erdős's proof is a model of clarity and cleverness. Erdős conjectured that the sequence is actually everywhere dense in $(0, \infty)$, although he could not find a single number which is a limit point. At the time only ∞ was a known limit point. Since 1955 there have been a few new results obtained. We now know from [11] that 0 is also a limit point, but no other limit point is known. In 1988 Hildebrand and Maier [14] proved, using Maier's method, that there are so many large limit points that the set of limit points has infinite measure, which answered a question of Erdős to show there is a finite limit point larger than 1. Finally, in 2013 Pintz used Zhang's result to show that there is an interval $[0, c]$ all of whose points are limit points. However, c is ineffective. Thus neither of these results helps answer Erdős's question of finding specified numbers which are limit points.

A second problem is concerned with the question of finding the most frequent difference between consecutive primes $\leq x$. This is referred to as the jumping champion up to x. For example, the prime differences up to $x = 12$ are $3 - 2 = 1$, $5 - 3 = 2$, $7 - 5 = 2$, and $11 - 7 = 4$, so the jumping champion up to 12 is 2. In 1980 Erdős and Straus [9] proved, assuming the Hardy-Littlewood conjectured asymptotic formula for the number of prime pairs, that the jumping champions must go to infinity. Numerically 6 is the jumping champion for $x \geq 947$ and continues to be so as far as has been computed, which is currently around 10^{15}. Odlyzko, Rubinstein, and Wolf [21] provided evidence for the conjecture that jumping champions will eventually transition from 6 to 30 and continue to increase through the sequence of primorials $2, 6, 30, 210, 2310, \ldots$, where the kth primorial is formed from the product of the first k primes. In agreement with this conjecture, Ledoan and I [10] proved, assuming the same Hardy-Littlewood conjecture used by Erdős and Straus, that any given primorial will divide every sufficiently large jumping champion. Assuming stronger Hardy-Littlewood conjectures, we proved the Primorial Conjecture for sufficiently large x. Despite this, nothing unconditional beyond what computation reveals is known about jumping champions. We cannot prove or show through computation that there is any jumping champion larger than 6. We cannot even prove that 2 is not the jumping champion for all large x, although we would actually conjecture that 2 and powers of 2 are the biggest jumping losers.

Ernst Straus and Erdős first met in Princeton in 1944 and were close friends and coauthors of many papers together. One of Straus's sons, Dan Straus, is a chemistry professor and colleague of mine at San José State University. I asked him for one Erdős story for this article, and naturally he had many, many Erdős stories. I include here one he sent me which represents how I also remember Erdős.

> Paul Erdős visited UCSB sometime around 1975 when I was an undergraduate chemistry student, and my father arranged a meeting. Erdős took a college friend and me to lunch in Isla Vista and then we walked to the lagoon on campus near Storke Tower. Erdős was interested in chemistry and particularly in recent discoveries of complexes of the noble gasses, and it was fun talking with him, as always. We were standing at the top of a grassy slope scattered with ducks. Spontaneously, Erdős threw out his arms like a scarecrow and careened off after the ducks, which he chased tirelessly in wide arcs. He was a sight in his loose fitting pinstripe suit and sandals with dress socks, and he was very agile for a man of about sixty-three. It was a time of studied nonconformity, and students were everywhere in beads, sandals and such just watching Erdős. It was refreshing to watch one person we can all be sure never tried to be different or worried what people thought.

Note Added in Proof

2014 was another spectacular year for primes. The method introduced by Maynard, and independently obtained by Terry Tao, has been further developed by the Polymath 8b project http://www.resmathsci.com/content/1/1/12. It is now known (as of January 2015) that there are infinitely often two primes that differ by less than or equal to 246.

The US$10,000 Erdős problem on large gaps between primes was solved independently both by Maynard and by Ford, Green, Konyagin, and Tao. These five authors have now proven jointly that for the size of the large gaps in the Erdős problem the power of the triple logarithm in the denominator may be reduced from 2 to 1.

On the problem of limit points of normalized consecutive prime gaps, Banks, Freiberg, and Maynard have adapted the Maynard–Tao method for small gaps between primes with the method for large gaps to prove very strong results. In particular they have proved that at least 12.5 percent of all positive real numbers are limit points of this sequence. However we still do not know any specific positive real number that is a limit point.

References

[1] E. Bombieri and H. Davenport, Small differences between prime numbers, *Proc. Roy. Soc. Ser. A* **293** (1966), 1–18.

[2] P. Erdős, Beweis eines Satzes von Tschebyschef (in German), *Acta Litt. Sci. Szeged* **5** (1932), 194–198; Erdős Collected Papers 1932-01.

[3] _____, On the difference of consecutive primes, *Quart. J. Math., Oxford Ser.* **6** (1935), 124–128; Erdős Collected Papers 1935-07.

[4] _____, The difference of consecutive primes, *Duke Math. J.* **6** (1940), 438–441; Erdős Collected Papers 1940-10.

[5] _____, On a new method in elementary number theory which leads to an elementary proof of the prime number theorem, *Proc. Nat. Acad. Sci. USA* **35** (1949), 374–384; Erdős Collected Papers 1949-02.

[6] _____, Problems and results on the differences of consecutive primes, *Publ. Math. Debrecen* **1** (1949), 33–37; Erdős Collected Papers 1949-02.

[7] _____, Some problems on the distribution of prime numbers, *Teoria dei numeri, Math. Congr. Varenna, 1954*, 8 pp., 1955; Erdős Collected Papers 1955-12.

[8] _____, Résultats et problémes en théorie des nombres (in French), Séminaire Delange-Pisot-Poitou (14e année: 1972/73), *Théorie des nombres*, Fasc. 2, Exp. No. 24, 7 pp., Secretariat Mathematique, Paris, 1973; Erdős Collected Papers 1973-24.

[9] P. Erdős and E. G. Straus, Remarks on the differences between consecutive primes, *Elem. Math.* **35** (1980), no. 5, 115–118; Erdős Collected Papers 1980-36.

[10] D. A. Goldston and A. H. Ledoan, Jumping champions and gaps between consecutive primes, *Int. J. Number Theory* **7** (2011), no. 6, 1413–1421.

[11] D. A. Goldston, J. Pintz, and C. Y. Yildirim, Primes in tuples. I, *Ann. of Math. (2)* **170** (2009), no. 2, 819–862.

[12] Andrew Granville and K. Soundararajan, An uncertainty principle for arithmetic sequences, *Ann. of Math. (2)* **165** (2007), no. 2, 593–635.

[13] Ben Green and Terence Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math. (2)* **167** (2008), no. 2, 481–547.

[14] Adolf Hildebrand and Helmut Maier, Gaps between prime numbers, *Proc. Amer. Math. Soc.* **104** (1988), no. 1, 1–9.

[15] C. Hooley, On the difference of consecutive numbers prime to n, *Acta Arith.* **8** (1963), 343–347.

[16] Helmut Maier, Chains of large gaps between consecutive primes, *Adv. in Math.* **39** (1981), no. 3, 257–269.

[17] _____, Primes in short intervals, *Michigan Math. J.* **32** (1985), no. 2, 221–225.

[18] _____, Small differences between prime numbers, *Michigan Math. J.* **35** (1988), no. 3, 323–344.

[19] Hugh L. Montgomery and K. Soundararajan, Primes in short intervals, *Comm. Math. Phys.* **252** (2004), no. 1–3, 589–617.

[20] H. L. Montgomery and R. C. Vaughan, On the distribution of reduced residues, *Ann. of Math. (2)* **123** (1986), no. 2, 311–333.

[21] Andrew Odlyzko, Michael Rubinstein, and Marek Wolf, Jumping champions, *Experiment. Math.* **8** (1999), no. 2, 107–118.

[22] J. Pintz, Very large gaps between consecutive primes, *J. Number Theory* **63** (2) (1997), 286–301.

[23] R. A. Rankin, The difference between consecutive primes, *Journal London Math. Soc.* **13** (1938), 242–247.

[24] Giovanni Ricci, Recherches sur l'allure de la suite $pn + 1 - pn/logpn$ (French), *Colloque sur la Théorie des Nombres, Bruxelles, 1955*, pp. 93–106, Georges Thone, Liège; Masson and Cie, Paris, 1956.

[25] R. A. Rankin, The difference between consecutive prime numbers. II, *Proc. Cambridge Philos. Soc.* **36** (1940), 255–266.

[26] A. Selberg, On the normal density of primes in short intervals, and the difference between consecutive primes, *Arch. Math. Naturvid.* **47** (1943), no. 6, 87–105; reprinted in *Atle Selberg Collected Papers*, Vol. 1, pp. 160–178, Springer-Verlag, Berlin-Heidelberg-New York, 1989.

[27] _____, An elementary proof of the prime-number theorem, *Ann. of Math. (2)* **50** (1949), 305–313; reprinted in *Atle Selberg Collected Papers*, Vol. 1, pp. 379–387, Springer-Verlag, Berlin-Heidelberg-New York, 1989.

András Sárközy

Erdős and Sequences

Paul Erdős was one of the most prolific mathematicians in the history of mathematics. He was a leading personality in modern combinatorial mathematics. The large number of his disciples ("epsilons" as he called them) and coauthors also contributed to his great impact on twentieth-century mathematics.

I was nineteen years old, a second-year university student, when I received the following letter: "Dear Mr. Sárközy, I have heard about your nice results …from Paul Turán. Please, come and see me at the Mathematical Institute" (of the Hungarian Academy of Sciences). The letter was signed by Paul Erdős. I visited him soon after at the Mathematical Institute. I told him my results and I sketched the proofs. This was followed by a very fruitful and inspiring discussion, and at the end he asked a related question. As an answer to this question, I soon published my first paper based on an Erdős problem, and I was on the way to becoming one of his disciples. This first meeting of ours was followed by many others, and during one of them Erdős asked a further related problem. This I settled jointly with Endre Szemerédi, whom I introduced to Erdős, and from that point on the three of us worked jointly for several years. We usually met at Erdős's apartment, where he lived then with his mother. If one of our meetings was successful enough and we ended up with a nice result, Erdős asked his mother to give us a "few cubic centimeters of poison," which meant a few drops of *tokaji* (a rather sweet and heavy Hungarian dessert wine of good quality). Between 1966 and 1970 we published ten triple papers, all but one of

András Sárközy is professor of mathematics at Eötvös Loránd University. His email address is sarkozy@cs.elte.hu.

them written on divisibility properties of sequences of integers. Then Szemerédi switched to the Erdős-Turán problem on arithmetic progressions and later turned towards combinatorics, while I continued to work with Erdős in number theory (mostly on sequences of integers). I have written sixty-two joint papers with him, which puts me on the top in the list of his coauthors (András Hajnal is a close second with fifty-seven joint papers with him).

In classical number theory, one studies *special* sequences like primes, squares, etc. If in number theory we say just "sequence" (without any adjective), then we mean a general sequence; typically, we are looking for the connection between certain arithmetic properties and the density properties of sequences (or sets) of integers. The study of (general) sequences was started around 1930 by Schnirelmann's papers, and Erdős played a dominant role in the advance of this field. The first monograph written on sequences was the excellent book of Halberstam and Roth [19]. They write:

> Anyone who turns the pages of this book, will immediately notice the predominance of results due to Paul Erdős. In so far as the substance of this book may be said to define a distinct branch of number theory—and its wide range of topics in classical number theory appears to justify this claim—Erdős is certainly its founder. He was the first to recognize its true potential and has been the central figure in many of its developments.

In some of his papers written on *additive properties of sequences* Erdős studied Sidon sets. A set $\mathcal{A}$ of positive integers is said to be a Sidon set if all the sums $a + a'$ with $a \in \mathcal{A}$, $a' \in \mathcal{A}$, $a \leq a'$ are distinct. Let $F(n)$ denote the maximal cardinality of a Sidon set $\mathcal{A}$ with $\mathcal{A} \subset \{1, 2, \ldots, n\}$. Erdős and Turán [16] proved that $F(n) \leq (1 + o(1))n^{1/2}$, and later Chowla and Erdős [3] independently also showed that $F(n) \geq (1 + o(1))n^{1/2}$. Erdős [22] also proved that, if $\mathcal{A}$ is an infinite Sidon set, then its counting function $A(n) = \left| \{a : a \in \mathcal{A}, \ a \leq n\} \right|$ must satisfy

$$\liminf_{n \to \infty} A(n) \left(\frac{\log n}{n} \right)^{1/2} < \infty.$$

As an answer to a question of Sidon (who was a Hungarian mathematician working in harmonic analysis), Erdős [4] proved that there is an infinite sequence $\mathcal{A}$ of positive integers such that, denoting the number of solutions of the equation

$$(1) \qquad a + a' = n, \ a \in \mathcal{A}, \ a' \in \mathcal{A}$$

by $r(n)$, we have

$$c_1 \log n < r(n) < c_2 \log n$$

with some positive absolute constants c_1, c_2. He used a probabilistic method, and in [10] Erdős and

(From l to r, front row) **Vera Sós, Andras Sárközy, and Paul Erdős at the Institute of Mathematical Sciences, Madras, India, in January 1984 just before the start of Erdős's lecture.**

Rényi gave further applications of this method. Later Erdős and Sárközy [11] generalized and extended this result of Erdős significantly.

Write $R(x) = \sum_{n \leq x} r(n)$, where $r(n)$ is the function defined above. Erdős and Fuchs [6] proved that, if $\mathcal{A}$ is an infinite sequence of positive integers, then

$$(2) \qquad R(x) = cx + o\left(\frac{x^{1/4}}{(\log x)^{1/2}} \right)$$

cannot hold with a positive constant c. The significance of this result is based on the fact that it is closely related to the circle problem (which consists of the estimate of the number of lattice points in the circle $x^2 + y^2 \leq r^2$). A further interesting feature of this result is that this is one of the first occasions when analytical tools were used to prove a theorem on general sequences. Jurkat (unpublished) and later Montgomery and Vaughan proved that the $o(\ldots)$ term in (2) can be improved to $o(x^{1/4})$.

Erdős also studied *multiplicative properties* of sequences. A set $\mathcal{A}$ of positive integers is said to be primitive if there is no divisibility relation in it: if $a \in \mathcal{A}$, $a' \in \mathcal{A}$ and $a \neq a'$, then $a_i \nmid a_j$. Behrend proved that, if $\mathcal{A} \subset \{1, 2, \ldots, n\}$ and $\mathcal{A}$ is primitive, then

$$(3) \qquad \sum_{a \in \mathcal{A}} \frac{1}{a} < c \frac{\log n}{(\log \log n)^{1/2}}$$

for some absolute constant c. Erdős, Sárközy, and Szemerédi [12], [13] determined the smallest c with this property, and they also showed that, for infinite sets $\mathcal{A}$, the inequality (3) can be sharpened: if $\mathcal{A}$ is an infinite primitive set, then we have

$$\sum_{\substack{a \in \mathcal{A} \\ a \leq x}} \frac{1}{a} = o\left(\frac{\log x}{(\log \log x)^{1/2}} \right).$$

Erdős [2] also proved another theorem on positive sequences: he proved that, if $\mathcal{A}$ is a primitive set with $\mathcal{A} \neq \{1\}$, then

$$(4) \qquad \sum_{a \in \mathcal{A}} \frac{1}{a \log a} \leq c$$

with some absolute constant c. He conjectured that, for primitive sequences $\mathcal{A}$, the left-hand side of (4) is maximal if $\mathcal{A}$ is the set of the primes, and Erdős and Zhang [17] proved partial results in this direction.

Erdős also studied multiplicative Sidon sets, i.e., sets $\mathcal{A}$ of positive integers such that all the products aa' with $a \in \mathcal{A}$, $a' \in \mathcal{A}$, $a \leq a'$ are distinct. Let $G(n)$ denote the maximal cardinality of a multiplicative Sidon set $\mathcal{A}$ with $\mathcal{A} \subset \{1, 2, \ldots, n\}$. He proved [5] that there are positive absolute constants c_1, c_2 such that, for $n > n_0$ we have

$$c_1 n^{3/4} (\log n)^{-3/2} < G(n) - \pi(n) < c_2 n^{3/4} (\log n)^{-3/2},$$

where $\pi(n)$ is the number of primes not exceeding n. This result is a beautiful application of combinatorics: the proof is based on estimates for the number of edges in bipartite graphs containing no cycles of length 4.

Many further results and also recent developments related to Erdős's works are surveyed in [21] (in the additive case) and [20] (in the multiplicative case).

Erdős and Szemerédi [14] studied a problem involving both sums and products. Let $a_1 < a_2 < \cdots < a_n$ be a sequence of positive integers. Consider the products and sums, i.e., all the integers of the form

$$(5) \qquad a_i + a_j \quad \text{or} \quad a_i a_j \quad (\text{with } 1 \leq i \leq j \leq n).$$

Denote by $f(n)$ the largest integer so that, for every $\{a_1, a_2, \ldots, a_n\}$, there are at least $f(n)$ distinct integers of form (4). They proved that

$$n^{1+c_1} < f(n) < n^2 e^{-c_2 \log n / \log\log n} \quad \text{for} \quad n > n_0(\varepsilon)$$

with some positive absolute constants c_1 and c_2. This result has been sharpened and extended in various directions. A survey of these results has been given by Bourgain [1].

Erdős also studied "hybrid" problems in which both general sequences and special sequences (e.g., the sequence of the primes) occur; in particular, he studied the number of prime factors of sums $a + b$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$ (where $\mathcal{A}$, $\mathcal{B}$ are "large" or "dense" sets) jointly with Turán [15]. They proved in 1934 that, if $\mathcal{A}$ is a finite set of positive integers, then

$$\omega\left(\prod_{a, a' \in \mathcal{A}} (a + a') \right) > c \log |\mathcal{A}|,$$

where $\omega(n)$ denotes the number of distinct prime factors of n and c is a positive absolute constant. They also conjectured that this result can be extended to sums $a + b$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$, where $\mathcal{A}$, $\mathcal{B}$ are finite sets of positive integers with $|\mathcal{A}| = |\mathcal{B}|$. This conjecture (in a slightly sharper form) was proved more than fifty years later by Győry, Stewart, and Tijdeman. Erdős, Maier, and Sárközy [8] studied the distribution of the numbers $\omega(a + b)$, and Erdős, Pomerance, Sárközy, and Stewart [9] the maximum of $\omega(a + b)$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$ for "large" subsets $\mathcal{A}$, $\mathcal{B}$ of $\{1, 2, \ldots, n\}$.

This survey would not be complete without mentioning some of Erdős's most famous problems. In 1936 Erdős and Turán conjectured that, if for any positive integers k and n we denote the size of the largest subset of $\{1, 2, \ldots, n\}$ which does not contain an arithmetic progression of length k by $r_k(n)$, then for any fixed k we have $r_k(n) = o(n)$. First Roth in 1953 and then Szemerédi in 1969 settled the $k = 3$ and $k = 4$ special cases, respectively, and finally, in 1975, Szemerédi proved the conjecture in its general form. Szemerédi's theorem has been extended and sharpened in various directions (see [18] for a survey of these results). Erdős also conjectured that the following stronger statement is also true: if $\mathcal{A}$ is an infinite sequence of positive integers such that $\sum_{a \in \mathcal{A}} \frac{1}{a} = \infty$, then $\mathcal{A}$ contains arbitrarily long arithmetic progressions. A further related conjecture of his: if $e_1, e_2, \ldots$ is any infinite sequence with $e_i \in \{-1, +1\}$ for $i = 1, 2, \ldots$, then for an arbitrarily large real number K there exist positive integers n and d such that $|e_d + e_{2d} + \cdots + e_{nd}| > K$. These last two problems are still wide open.

Erdős and Turán conjectured that, if $\mathcal{A}$ is a set of positive integers such that $r(n) > 0$ for $n > n_0$ (where again $r(n)$ denotes the number of solutions of (1)), then $r(n)$ cannot be bounded: $\limsup_{n \to \infty} r(n) = \infty$. This conjecture is also still open.

There are many more problems in the Erdős–Graham book [7] and in his numerous problem papers; the majority of these problems remain unsolved.

References

[1] J. Bourgain, Around the sum-product phenomenon, in: *Erdős Centennial*, eds. L. Lovász et al., Bolyai Soc. Math. Studies, 25, Springer, 2013, pp. 111–128.

[2] P. Erdős, Note on sequences of integers no one of which is divisible by any other, *J. London Math. Soc.* **10** (1935), 126–128.

[3] ———, Addendum, On a problem of Sidon in additive number theory and on some related problems, *J. London Math. Soc.* **19** (1944), 208.

[4] ———, Problems and results in additive number theory, *Colloque sur la Théorie des Nombres (CBRM) (Bruxelles, 1955)*, Georges Thone, Liège; Masson and Cie, Paris, 1956, pp. 127–137.

[5] ———, Some applications of graph theory to number theory, in: *The Many Facets of Graph Theory*, Lecture Notes in Math., 110, Springer, 1969, pp. 77–82.

[6] P. Erdős and W. H. J. Fuchs, On a problem of additive number theory, *J. London Math. Soc.* **31** (1956), 67–73.

[7] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monographies de l'Enseignement Math., no. 28, Geneva, 1980.

[8] P. Erdős, H. Maier, and A. Sárközy, On the distribution of the number of prime factors of sums $a + b$, *Transactions Amer. Math. Soc.* **302** (1987), 269–280.

[9] P. Erdős, C. Pomerance, A. Sárközy, and C. L. Stewart, On elements of sumsets with many prime factors, *J. Number Theory* **44** (1993), 93–104.

[10] P. Erdős and A. Rényi, Additive properties of random sequences of positive integers, *Acta Arith.* **6** (1960), 83–110.

[11] P. Erdős and A. Sárközy, Problems and results on additive properties of general sequences. I, *Pacific J.* **118** (1985), 347–357; and II, *Acta Math. Hung.* **48** (1986), 201–211.

[12] P. Erdős, A. Sárközy, and E. Szemerédi, On an extremal problem concerning primitive sequences, *J. London Math. Soc.* **42** (1967), 484–488.

[13] ――――, On a theorem of Behrend, *J. Australian Math. Soc.* **7** (1967), 9–16.

[14] P. Erdős and E. Szemerédi, On sums and products of integers, in: *Studies in Pure Math.* (eds. L. Alpár et al.), Birkhäuser, Basel, 1983, pp. 213–218.

[15] P. Erdős and P. Turán, On a problem in the elementary theory of numbers, *Amer. Math. Monthly* **41** (1934), 608–611.

[16] ――――, On a problem of Sidon in additive number theory and some related problems, *J. London Math. Soc.* **16** (1941), 212–215.

[17] P. Erdős and Z. Zhang, Upper bound of $\sum 1/(a_i \log a_i)$ for primitive sequences, *Proc. Amer. Math. Soc.* **117** (1993), 891–895.

[18] W. T. Gowers, Erdős and arithmetic progressions, in: *Erdős Centennial* (L. Lovász et al. (eds.)), Bolyai Soc. Math. Studies, 25, Springer, 2013, pp. 265–287.

[19] H. Halberstam and K. F. Roth, *Sequences*, 2nd ed., Springer, 1983.

[20] A. Sárközy, On divisibility properties of sequences of integers, in: *The Mathematics of Paul Erdős. I*, 2nd extended edition (R. L. Graham et al. (eds.)), Springer, to appear.

[21] A. Sárközy and V. T. Sós, On additive representation functions, in: *The Mathematics of Paul Erdős. I*, 2nd extended edition (R. L. Graham et al. (eds.)), Springer, to appear.

[22] A. Stöhr, Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe. II, *J. Reine Angew. Math.* **194** (1955), 111–140.

József Szabados

Erdős and Polynomial Interpolation

It is not easy to write about Paul Erdős. The number of stories on his way of life, on his method of work, on his interest in politics, is all well known.

József Szabados is research professor emeritus of mathematics at the Alfréd Rényi Institute of Mathematics. His email address is szabados.jozsef@renyi.mta.hu.

Nevertheless, it is important to remember him, to tell how we admire his devotion to mathematics.

My connection with him was sporadic and occasional. Although approximation theory was not his main field of interest, he contributed numerous significant results to this area of mathematics. In the early years of his career, mainly together with Paul Turán and Géza Grünwald, he was interested in interpolation and properties of polynomials. In these papers, sometimes his formulations and proofs were vague, and later he kept returning to his original problems by urging us to give more precise and simpler proofs or even to correct faulty arguments. In what follows we collect some examples of his outstanding achievements and loose statements which required further discussions. We restrict ourselves to Lagrange interpolation, although properties of polynomials would be a related area where he made significant progress as well.

Let $-1 \leq x_{1n} < x_{2n} < \cdots < x_{nn} \leq 1$ be an arbitrary system of nodes, and let

$$L_n(f, x) =: \sum_{k=1}^n f(x_{kn}) \ell_{kn}(x)$$

be the uniquely determined *Lagrange interpolation polynomial* of degree at most $n - 1$ of a continuous function $f(x)$ in the interval $[-1, 1]$, where

$$\ell_{kn}(x) := \prod_{\substack{j=1 \\ j \neq k}}^n \frac{x - x_{jn}}{x_{kn} - x_{jn}}, \qquad k = 1, 2, \ldots, n,$$

are the *fundamental polynomials* of interpolation. The crucial quantity for the convergence of interpolation is the *Lebesgue constant* λ_n, which is the maximum over $[-1, 1]$ of the *Lebesgue function*

$$\lambda_n(x) := \sum_{k=1}^n |\ell_{kn}(x)|.$$

G. Faber [11] proved that, no matter what the system of nodes of interpolation is, the Lebesgue constant λ_n is always at least $c_1 \log n - c_2$, with some constants $c_1, c_2 > 0$. Erdős [3] proved that the best constant here is $c_1 = 2/\pi$ (which can be attained for example for the roots of the Chebyshev polynomials $T_n(\cos(n \arccos x))$).

A direct consequence of this behavior of the Lebesgue constant is that Lagrange interpolation can never be uniformly convergent for all continuous functions. Far more is true: the celebrated theorem of Erdős and Vértesi [10] states that, for any system of nodes, one can always construct a continuous function whose interpolating polynomials diverge *almost everywhere.*

The phenomenon of divergence is in analogy with the behavior of the trigonometric Fourier series. However there, as is well known, the first arithmetic means (Fejér sums) converge uniformly

to continuous functions. In a joint paper [4] with Géza Grünwald, Erdős claimed that this analogy does not hold for Lagrange interpolation: considering the arithmetic means

$$(6) \qquad \frac{1}{n} \sum_{k=1}^{n} L_k(f, x)$$

of Lagrange interpolation based on the nodes of the Chebyshev polynomial $T_n(x) := \cos(n \arccos x)$, they diverge for some continuous function f. It turned out later that the proof was faulty: instead of (6), the proof works only for the arithmetic means of the absolute values of the interpolating polynomials. This error was corrected much later in the joint work [5] with G. Halász: in fact, they proved that (6) is at least $o(\log \log n)$ for some f_0, and this lower estimate is sharp.

Most of the time Erdős was correct in making conjectures for difficult problems, but not always. In [2] he conjectured that the minimum of the integral of the sum of squares of the fundamental polynomials of Lagrange interpolation, i.e.,

$$\int_{-1}^{1} \sum_{k=1}^{n} \ell_{kn}(x)^2 \, dx,$$

is attained when the nodes are the roots of the integral of the Legendre polynomials. It was a reasonable conjecture, since these are the only nodes where the sum of the squares of the fundamental polynomials is bounded by 1 (in contrast to the Lebesgue constants mentioned above). However, I succeeded in proving that these nodes do not serve the minimum [12], and this was the occasion when Erdős started telling me problems. The solution of the mentioned problem of minimum is still unsolved, and it seems as hopeless as the exact construction of nodes of interpolation with minimal Lebesgue constant.

Our first joint work [6] is a kind of strengthening of the celebrated result of G. Faber mentioned above. We proved that the *integral* of the Lebesgue function over a fixed subinterval of interpolation

$$\int_{a}^{b} \lambda_n(x) \, dx, \qquad [a, b] \subset [-1, 1]$$

is always at least $c(a, b) \log n$, where $c(a, b) > 0$ depends only on the length of the interval in question. This, of course, implies Faber's result. The proof is based on the simple observation that the sum of a positive number and its reciprocal is always at least 2. Then, in [9] we generalized this result for the case when the interval in question depends on n and proved also the sharpness of the result. The problem was further generalized in another direction by considering the *weighted* integral (with Jacobi weights) of the sum of *even* powers of the fundamental functions of interpolation (cf. [8]).

Of course, he was extremely fast in thinking and was rather annoyed when it turned out that somebody could not follow his train of thought. I have some sheets with his characteristic handwriting, but not much. Talking was more important for him than writing down occasionally long formulas. A typical situation was when he said, "Something is disturbing me," when we were stuck with a problem or something seemed to be too good to hold. Not for long, since he was always able to circumvent the difficulty.

In [1], Erdős (among others) considered the problem of loosening the strict condition of Lagrange interpolation when for n nodes we look for an interpolating polynomial of degree at most $n - 1$. What if we allow polynomials of degree at most $n(1 + \varepsilon)$, where ε is a fixed positive number? He claimed that, under some regularity conditions on the distribution of nodes (asymptotic equidistant distribution of the arccos of the nodes and a lower estimate on the distance of adjacent nodes), the interpolation polynomials of degree at most $n(1 + \varepsilon)$ converge uniformly for all continuous functions. He did not prove this statement, but gave an indication how it can be deduced from another theorem of the same paper by a simple modification. We were unable to recover this "simple modification" and started working on the problem. Eventually, we came up with a rather long and sophisticated proof, gave an error estimate for the convergence, and proved the necessity of the mentioned regularity conditions (cf. [7]).

References

[1] P. Erdős, On some convergence properties of the interpolation polynomials, *Ann. of Math. (2)* **44** (1943), 330–337.

[2] ——, Problems and results on the theory of interpolation, II, *Acta Math. Acad. Sci. Hungar.* **12** (1961), 235–244.

[3] ——, Problems and results on the convergence and divergence properties of the Lagrange interpolation polynomials and some extremal problems, *Mathematica Cluj*, **10** (1968), 65–73.

[4] P. Erdős and G. Grünwald, Über die arithmetischen Mittelwerte der Lagrangeschen Interpolationspolynome, *Studia Math.* **7** (1938), 82–95.

[5] P. Erdős and G. Halász, On the arithmetic means of Lagrange interpolation, in: *Approximation Theory*, Coll. Math. Soc. J. Bolyai, Vol. 58, (1990) pp. 263–274.

[6] P. Erdős and J. Szabados, On the integral of the Lebesgue function of interpolation, *Acta Math. Acad. Sci. Hungar.* **32** (1978), 191–195.

[7] P. Erdős, A. Kroó, and J. Szabados, On convergent interpolatory polynomials, *J. Approx. Theory* **58** (1989), 232–241.

[8] P. Erdős, J. Szabados, A. K. Varma, and P. Vértesi, On an interpolation theoretical extremal problem, *Studia Sci. Math. Hungar.* **29** (1994), 55–60.

[9] P. ERDŐS, J. SZABADOS, and P. VÉRTESI, On the integral of the Lebesgue function of interpolation. II, *Acta Math. Hungar.* **68** (1995), 1–6.

[10] P. ERDŐS and P. VÉRTESI, On the almost everywhere divergence of Lagrange interpolatory polynomials, *Acta Math. Acad. Sci. Hungar.* **36** (1980), 71–89, and **38** (1981), 263.

[11] G. FABER, Über die interpolatorische Darstellung stetiger Funktionen, *Jahresber. der deutschen Math. Verein.* **23** (1914), 190–210.

[12] J. SZABADOS, On a problem of P. Erdős, *Acta Math. Acad. Sci. Hungar.* **17** (1966), 1155–1157.

Gérald Tenenbaum

Paul Erdős and the Divisors

Paul Erdős and I collaborated from March 1977 until his death (he called this a leave) in October 1996. During these nineteen years, most of our mathematical discussions were on the distribution of divisors of integers.

Analytic and probabilistic number theory is mainly concerned with understanding how the multiplicative and the additive structure of integers combine together or ignore each other. The twin primes conjecture and the Goldbach conjecture are enlightening examples: since nothing trivially forbids two primes from having difference two, this should happen with statistical frequency; similarly, if $2n$ is an arbitrary integer, the sequence $2n - p$, where p runs through all primes up to n, should contain a random quota of primes.

The famous *abc* conjecture of Masser and Oesterlé, dating from the early 1970s, is even more representative of this class of problems: addition should destroy multiplicative structure. Let us discuss this in more detail. The squarefree kernel, say $k(n)$, of an integer n is the product of all primes appearing in the canonical decomposition of n, ignoring exponents. A normal integer (in other words an integer belonging to a set that will almost surely show up when one picks an integer at random) has a "large" kernel (see [31] for recent progress on this question), whereas small kernels occur only for integers with a very special structure such as perfect powers or integers divisible by a large perfect power. In qualitative form, the *abc* conjecture states that when a and b are coprime and $c = a + b$, the three integers cannot be simultaneously abnormal. A thorough quantitative discussion of this question may be found in [30].

The distribution of divisors is a problem of similar type to those described above: a divisor has a very special multiplicative structure, since its prime factors, and even their exponents, are

restricted in a drastic manner. Thus it is a basic number theoretic challenge to try to understand how the sequence of divisors is distributed by size, that is to say, with respect to additive structure.

Erdős was interested in all aspects of this question: usual behavior of the sequence of divisors of a random, or normal, integer; extremal properties involving divisors, that is, small and large values of arithmetic functions defined in terms of divisors; structure of sequences defined by constraints on the divisors of their elements; stochastic variations of divisor functions; etc.

It is clear that the divisors are made up from the prime factors. So the first step in the problem is to describe the growth of the sequence of prime factors. The step Erdős made was that of a giant. Before anyone else, he understood that basic results in probabilistic number theory, such as the Turán-Kubilius inequality,[1] yield a very strong and very surprising fact: in first approximation, the size of the jth prime factor of a normal integer n does not depend on n but only on j; more precisely, if we let $\{p_j(n)\}_{j=1}^{\omega(n)}$ denote the increasing sequence of the distinct prime factors of n, then we have

$$(1) \qquad \log_2 p_j(n) \sim j \qquad (j \to \infty)$$

for almost all integers n.[2] The short proof may be retrieved in Erdős's paper [7]. It will not be reproduced here: we content ourselves with saying that the Turán-Kubilius inequality is the arithmetical analogue of the theorem from probability theory stating that the variance of a sum of independent random variables is the sum of the variances of its terms and that Erdős merely uses a Bienaymé-Chebyshev-type inequality to deduce his result. A much sharper result, actually exhibiting a Gaussian behavior of the prime factors around their means, is obtained in [5]. The reader may consult [18] (Chapter 1) for a detailed proof and [39] (Theorem III.3.10) for a simpler proof of a slightly weaker result.

From the above statement on prime factors, one could guess that the jth divisor of a normal integer is somewhat close to $\exp j^c$ with $c = 1/\log 2$. It was one of Erdős's outstanding qualities to devise a question that would precisely test how adequate the model is to the arithmetic nature of things. Since $c > 1$, we may deduce from the previous heuristics that the minimal ratio $E(n)$ between consecutive divisors, say d'/d, with $d < d'$, does not tend to 1 for almost all integers. However, as early as 1948 and probably much before, Erdős

Gérald Tenenbaum is professor of mathematics at the Université de Lorraine. His email address is gerald.tenenbaum@univ-lorraine.fr.

---

[1] *See, e.g., Elliott* [4] *or Tenenbaum* [39].

[2] *Here and in the sequel we denote by* $\log_k$ *the k-fold iterated logarithm.*
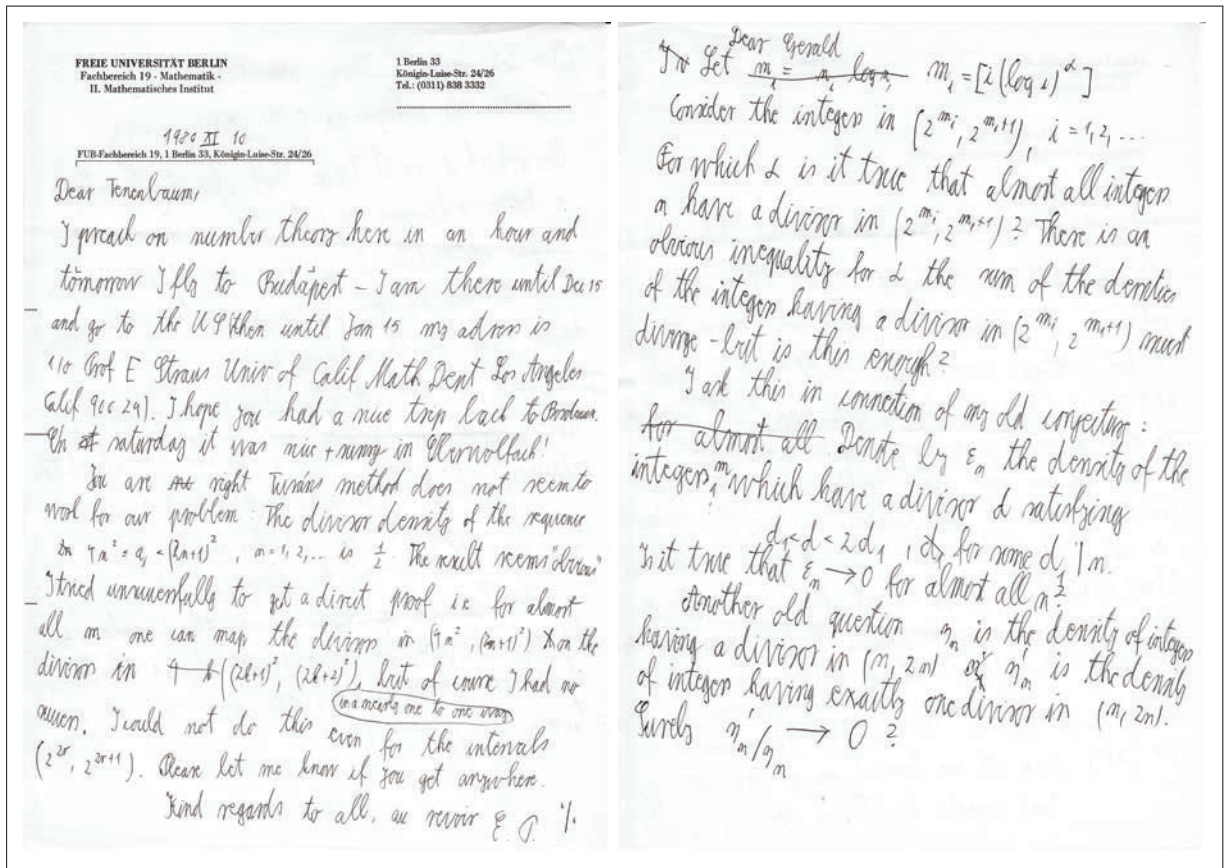
Letter from Erdős to Tenenbaum stating a problem on divisor density.

conjectured the exact opposite: for almost all n, one should have

(2) $\quad E(n) := \min_{dd'|n,\, d<d'} d'/d = 1 + (\log n)^{1-\log 3 + o(1)}.$

The reason for this is precisely that the distribution of the prime factors fluctuates significantly around the mean and hence the quantities $\log d'/d$ should be fairly evenly distributed in the interval $[-\log n, \log n]$. Since these quantities are at least $3^{\omega(n)}$ in number and, by the Turán-Kubilius inequality or indeed from (1), we have $\omega(n) \sim \log_2 n$ for almost all n, we naturally arrive at (2). This conjecture was proved by Erdős-Hall [8] for the lower bound and by Maier-Tenenbaum [24] for the upper bound. A further refinement [29] comes even closer to the heuristics: for almost all n, we have

(3) $\qquad E(n) = 1 + \dfrac{\log n}{3^{\omega(n)}} (\log_2 n)^{\vartheta_n}$

where $-5 \leqslant \vartheta_n \leqslant 10$.

One of the most challenging problems remaining on the so-called subject of propinquity of divisors concerns the functions

$E_r(n) := \min_{1 \leqslant j \leqslant \tau(n)-r} \log\{d_{j+r}(n)/d_j(n)\} \qquad (r \geqslant 2),$

where $\{d_j(n)\}_{j=1}^{\tau(n)}$ denotes the increasing sequence of the divisors of n. The precise normal behavior remains still unknown for all $r \geqslant 2$. Using techniques similar to that of the proof of Theorem 3 of [12], it can be shown that, on a sequence of natural density 1, we have

$$E_2(n) > (\log n)^{-\gamma_2 + o(1)}$$

for some $\gamma_2 < \log 3 - 1$. Moreover, the methods and results of [25] yield, still normally,

$$E_r(n) \leqslant (\log n)^{-\beta_r + o(1)},$$

with

$$\beta_r := \frac{(\log 3 - 1)^m}{(3\log 3 - 1)^{m-1}}, \quad 2^{m-1} < r + 1 \leqslant 2^m.$$

Thus, we have

$\beta_1 = \log 3 - 1 \approx 0.09861, \quad \beta_2 = \beta_3 \approx 0.00423,$

$\beta_r \approx 0.00018 \quad (4 \leqslant r \leqslant 7).$

Also, it is proved in [25] (Thm. 1.1) that $E_r(n) > \tau(n)^{-1/r+o(1)}$ holds for almost all integers, uniformly in $r \geqslant 1$, and thus, on a sequence of density 1,

$$E_r(n) = 1/(\log n)^{o(1)} \qquad (r = r(n) \to \infty),$$

a result which might look surprising at first sight.

Maier and I conjecture the existence of a strictly decreasing sequence $\{\alpha_r\}_{r=1}^{\infty}$ such that we have

$$E_r(n) = (\log n)^{-\alpha_r + o(1)}$$

on a sequence of density 1. It is particularly irritating, for instance, to be unable to find a better normal upper bound for $E_2(n)$ than for $E_3(n)$.

I refer the reader to the recent survey [40] for a further account of Erdős's motivations for the conjecture (2), in particular related to the concept of set of multiples. The link is particularly apparent in Erdős's letter dated November 10, 1980, reproduced herein.

On page 1, Erdős mentions *divisor density*. The definition, due to R. R. Hall [15], is as follows. Let $\tau(n, \mathcal{A})$ designate the number of divisors of an integer n belonging to a sequence $\mathcal{A}$, and write $\tau(n) = \tau(n, \mathbb{Z}^+)$. We say that the integer sequence $\mathcal{A}$ has divisor density z, and we write $D\mathcal{A} = z$ if we have $\tau(n, \mathcal{A}) = \{z + o(1)\}\tau(n)$ as n tends to infinity on a sequence of natural density 1.[3]

Divisor density is a fruitful and surprising notion. For instance, Hall proved in [15] that, for any pair $(z, w) \in [0,1]^2$, there is an integer sequence $\mathcal{A}$ with divisor density z and logarithmic density w.[4] A criterion for $D\mathcal{A} = z$ is given in [33].

With the above definition, Erdős's question on page 1 may be stated as follows: define $\mathcal{A} := \bigcup_{m \geqslant 1} [4m^2, (2m+1)^2[\cap \mathbb{Z}^+$; is it true that $D\mathcal{A} = \frac{1}{2}$? It needed a lot of work and the appeal to many deep results from analytic number theory, such as estimates of Karatsuba on exponential sums [22], to answer, positively, Erdős's question; see Hall-Tenenbaum [17] and Tenenbaum [38] (Theorem 11). To be slightly more precise, we observe that the sequence $\mathcal{A}$ may alternatively be defined by the condition $\langle \frac{1}{2}\sqrt{n} \rangle \leqslant \frac{1}{2}$ where $\langle x \rangle$ denotes the fractional part of the real number x. The two theorems above actually imply that, for any real number c, any $z \in [0,1]$, and any nonintegral positive number α, the sequence $\{n \geqslant 1 : \langle cn^\alpha \rangle \leqslant z\}$ has divisor density z.

On page 2 of the reproduced letter, Erdős asks a slightly different question: given a sequence $m_j := \lfloor j(\log j)^\alpha \rfloor$, and setting $\mathcal{A} := \bigcup_{j \geqslant 1} [2^{m_j}, 2^{m_j+1}[$, for which α do we have $\tau(n, \mathcal{A}) \geqslant 1$ on a sequence of asymptotic density 1? Here it is clear, and Erdős explicitly notes the fact that the problem is linked with conjecture (2): it deals with the distribution of divisors in dyadic intervals.

Once again, it took years of struggle to solve the problem: as proved in [37], the answer is positive for all α. As it turns out, one needs to take faster growing sequences to see a threshold: if we now define $m_j = j^\beta$, then Hall-Tenenbaum proved in 1992 [19] that the answer to Erdős's question is affirmative if, and only if, $\beta \leqslant 1/(1 - \log 2)$.

Aside from his interest in the normal behavior of the set of divisors, Erdős was also intrigued by extremal properties. Out of many of his problems, I extract two.

To describe the first, let us put, for $\alpha > 0$,

$$F_\alpha(n) := \sum_{1 \leqslant j < \tau(n)} \left(\frac{d_{j+1}(n)}{d_j(n)} - 1 \right)^\alpha.$$

The conjecture asserted that, for all $\alpha > 1$,

$$(4) \qquad \liminf_{n \to \infty} F_\alpha(n) < \infty.$$

Since $F_1(n) \geqslant \log n$, it is clear that the condition $\alpha > 1$ cannot be weakened. This conjecture may be seen as dual to another conjecture of Erdős, related to the sequence $\{a_j\}_{j=1}^{\varphi(n)}$ of integers in $[1, n]$ and coprime to n. The problem here was to show that, for all $\gamma > 0$, we have

$$\limsup_{n \to \infty} \frac{\varphi(n)^{\gamma-1}}{n^\gamma} \sum_{1 \leqslant j < \varphi(n)} (a_{j+1}(n) - a_j(n))^\gamma < \infty.$$

Improving a result of Hooley [20] disposing of the case $\gamma < 2$, Montgomery and Vaughan confirmed this conjecture in 1986 [27].

These two problems are specific to Erdős's particular way of thinking. He manufactured innocent-looking questions whose solution actually requires a deep understanding of the structure of integers defined by multiplicative constraints.

Conjecture (4) was solved by Vose in 1984. Since, from Holder's inequality, we have

$$(\log n)^\alpha \leqslant F_\alpha(n)\tau(n)^{\alpha-1},$$

candidates to bounded values of $F_\alpha(n)$ must have a large number of divisors.[5] This led Erdős to the further conjecture that F_α should be bounded by natural sequences with many divisors, such as $n!$, l.c.m.$\{1, 2, \ldots, n\}$, or $\prod_{p \leqslant n} p$. I could establish this in [35] as a consequence of a more general result proved by the saddle-point method.

The second extremal problem was asked by Erdős on numerous occasions and is referred to as Problem 23 in the appendix of Montgomery's book [28]. It states that, for suitable constant C, the inequality

$$(5) \qquad \sum_{d|n, t|n, d<t} \frac{1}{t-d} \leqslant C\tau(n)$$

---

[3] *The natural density of an integer sequence is, when it exists, the limit, as $N \to \infty$, of the frequency of $\mathcal{A}$ among the N first integers.*

[4] *The logarithmic density of an integer sequence $\mathcal{A}$ is, when it exists, the value of the limit*

$$\delta(\mathcal{A}) := \lim_{N \to \infty} \frac{1}{\log N} \sum_{n \leqslant N, n \in \mathcal{A}} \frac{1}{n}.$$

[5] *Recall that a normal integer n has about $(\log n)^{\log 2 + o(1)}$ divisors.*

should hold for all positive integers *n*. Here again, the aim is to test the lacunarity of the sequence of the divisors of an integer: despite the fact that *d* and *t* in the above sum may get fairly close, we expect that this happens sufficiently rarely so that (5) remains true. Thus we are really facing a sieve problem of a delicate nature. I could establish (5) in a strong form, improved by La Bretèche [1].

Erdős's interest in divisors was so constant and so intense that a whole book would be necessary to describe his problems, attempts at solutions, and original methods on this topic—and the two already written books, [18] and [16], largely dominated by the work of Erdős, would only be a small part of the story. The references of this short survey constitute an incomplete and partial list of articles related to the subject, either by Erdős himself and his collaborators or inspired by his appealing way of thinking of mathematical problems.

Acknowledgments

References

[1] R. DE LA BRETÈCHE, Sur une classe de fonctions arithmétiques liées aux diviseurs d'un entier, *Indag. Math. (N.S.)* **11** (2000), no. 3, 437–452.

[2] R. DE LA BRETÈCHE and G. TENENBAUM, Sur les lois locales de la répartition du *k*-ième diviseur d'un entier, *Proc. London Math. Soc. (3)* **84** (2002), 289–323.

[3] J.-M. DE KONINCK and G. TENENBAUM, Sur la loi de répartition du *k*-ième facteur premier d'un entier, *Math. Proc. Camb. Phil. Soc.* **133** (2002), 191–204.

[4] P. D. T. A. ELLIOTT, *Probabilistic Number Theory: Mean Value Theorems*, Grundlehren der Math. Wiss., 239, Springer-Verlag, New York-Berlin-Heidelberg, 1979.

[5] P. ERDŐS, On the distribution function of additive functions, *Ann. of Math. (2)* **47** (1946), 1–20.

[6] ———, On some applications of probability to analysis and number theory, *J. London Math. Soc.* **39** (1964), 692–696.

[7] ———, Some unconventional problems in number theory, *Astérisque* **61** (1979), 73–82.

[8] P. ERDŐS and R. R. HALL, The propinquity of divisors, *Bull. London Math. Soc.* **11** (1979), 304–307.

[9] P. ERDŐS and M. KAC, The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738–742.

[10] P. ERDŐS and G. TENENBAUM, Sur la structure de la suite des diviseurs d'un entier, *Ann. Inst. Fourier (Grenoble)* **31**, no. 1 (1981), 17–37.

[11] ———, Sur les diviseurs consécutifs d'un entier, *Bull. Soc. Math. France* **111** (1983), 125–145.

[12] ———, Sur les fonctions arithmétiques liées aux diviseurs consécutifs, *J. Number Theory* **31** (1989), 285–311.

[13] ———, Ensembles de multiples de suites finies, *Discrete Math.* **200**, nos. 1–3 (1999), 181–203.

[14] K. FORD, The distribution of integers with a divisor in a given interval, *Ann. of Math. (2)* **168**, no. 2 (2008), 367–433.

[15] R. R. HALL, A new definition of the density of an integer sequence, *J. Austral. Math. Soc. Ser. A* **26** (1978), 487–500.

[16] ———, *Sets of Multiples*, Cambridge Tracts in Mathematics, 118, Cambridge University Press, Cambridge, 1996.

[17] R. R. HALL and G. TENENBAUM, Les ensembles de multiples et la densité divisorielle, *J. Number Theory* **22** (1986), 308–333.

[18] ———, *Divisors*, Cambridge Tracts in Mathematics, 90, Cambridge University Press, 1988 (paperback ed., 2008).

[19] ———, On Behrend sequences, *Math. Proc. Camb. Phil. Soc.* **112** (1992), 467–482.

[20] C. HOOLEY, On the difference of consecutive numbers prime to *n*, *Acta Arith.* **8** (1962/1963), 343–347.

[21] ———, A new technique and its applications to the theory of numbers, *Proc. London Math. Soc. (3)* **38** (1979), 115–151.

[22] A. A. KARATSUBA, Estimates for trigonometric sums by Vinogradov's method and some applications, *Proc. Steklov Inst. Math.* **112** (1971), 251–265.

[23] H. MAIER, On the Mobius function, *Trans. Amer. Math. Soc.* **301**, no. 2 (1987), 649–664.

[24] H. MAIER and G. TENENBAUM, On the set of divisors of an integer, *Invent. Math.* **76** (1984), 121–128.

[25] ———, On the normal concentration of divisors, 2, *Math. Proc. Camb. Phil. Soc.* **147**, no. 3 (2009), 593–614.

[26] M. MENDÈS FRANCE and G. TENENBAUM, Systèmes de points, diviseurs, et structure fractale, *Bull. Soc. Math. France* **121** (1993), 197–225.

[27] H. L. MONTGOMERY and R. C. VAUGHAN, On the distribution of reduced residues, *Ann. of Math. (2)* **123**, no. 2 (1986), 311–333.

[28] H. L. MONTGOMERY, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, CBMS Regional Conference Series in Mathematics, 84, published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994, xiv+220 pp.

[29] A. RAOUJ, A. STEF, and G. TENENBAUM, Mesures quadratiques de la proximité des diviseurs, *Math. Proc. Camb. Phil. Soc.* **150** (2011), 73–96.

[30] O. ROBERT, C. L. STEWART, and G. TENENBAUM, A refinement of the *abc* conjecture, *Bull. London Math. Soc.* **46** (2014), 1156–1166.

[31] O. ROBERT and G. TENENBAUM, Sur la répartition du noyau d'un entier, *Indag. Math. (N.S.)* **24**, no. 4 (2013), 802–914.

[32] A. STEF, *L'ensemble exceptionnel dans la conjecture d'Erdős concernant la proximite des diviseurs*, Thèse de doctorat de l'Université Nancy 1, UFR STMIA, June 1992.

[33] G. TENENBAUM, Sur la densité divisorielle d'une suite d'entiers, *J. Number Theory* **15**, no. 3 (1982), 331–346.

[34] ———, Sur la concentration moyenne des diviseurs, *Comment. Math. Helvetici* **60** (1985), 411–428.

[35] ———, Sur un problème extrémal en arithmétique, *Ann. Inst. Fourier (Grenoble)* **37**, no. 2 (1987), 1–18; Corrigendum, *ibid.* **50**, 1 (2000), 311–312.

[36] ———, Une inégalité de Hilbert pour les diviseurs, *Indag. Math. N.S.*, **2**(1) (1991), 105–114.

[37] ———, On block Behrend sequences, *Math. Proc. Camb. Phil. Soc.* **120** (1996), 355–367.

Figure 1. The first few paragraphs of Erdős's paper [6].

[38] _____, Uniform distribution on divisors and Behrend sequences, *L'Enseignement Mathématique* **42** (1996), 153–197.

[39] _____, *Introduction à la théorie analytique et probabiliste des nombres*, third ed., coll. Échelles, Berlin, 2008, 592 pp.

[40] _____, Some of Erdős's unconventional problems in number theory, thirty-four years later, in: *Erdős Centennial Volume* (L. Lovász I. Z. Ruzsa, V. T. Sós (eds.)), Bolyai Society Mathematical Studies, 25, 2013, pp. 651–681.

Stephan Ramon Garcia and Amy L. Shoemaker

Wetzel's Problem, Paul Erdős, and the Continuum Hypothesis: A Mathematical Mystery

We are concerned here with the curious history of *Wetzel's problem*: If $\{f_\alpha\}$ is a family of distinct analytic functions (on some fixed domain) such that for each z the set of values $\{f_\alpha(z)\}$ is countable, is the family itself countable?

In September 1963, Paul Erdős submitted to the *Michigan Mathematical Journal* a stunning solution to Wetzel's problem (Figure 1). He proved that an affirmative answer is equivalent to the negation of the continuum hypothesis. Erdős ends in an understated manner: "Paul Cohen's recent proof of the independence of the continuum hypothesis gives this problem some added interest." Together these results render Wetzel's problem undecidable in ZFC.

Erdős had a knack for solving "innocent-looking problems whose solutions shed light on the shape of the mathematical landscape" [7, p. 2]. In this case,

Stephan Ramon Garcia is associate professor of mathematics at Pomona College. His email address is `Stephan.Garcia@pomona.edu`.

Amy L. Shoemaker is a graduate of Pomona College. Her email address is `amyshoemaker2@gmail.com`.

the landscape he revealed was one of underground tunnels, surprising links, and glittering mysteries. However, our interest lies not with the solution itself but rather with the story of how Erdős encountered Wetzel's problem in the first place.

Our first exposure to Wetzel's problem was in *Proofs from The Book* by Aigner and Ziegler. "Paul Erdős liked to talk about The Book," they write, "in which God maintains the perfect proofs for mathematical theorems, following the dictum of G. H. Hardy that there is no permanent place for ugly mathematics. Erdős also said that you need not believe in God but, as a mathematician, you should believe in The Book" [1]. Erdős asked Aigner and Ziegler to assemble a moderate approximation of The Book; included in it was Erdős's answer to Wetzel's problem [1, pp. 102–6].

Regarding the origin of the problem, Erdős simply asserted that Wetzel posed the question in the Ann Arbor Problem Book in December 1962. Ziegler suggested several mathematicians who might be the Wetzel in question, eventually putting us in contact with John E. Wetzel (professor emeritus at the University of Illinois Urbana-Champaign), who confirmed that the problem was indeed his.

John Wetzel, born on March 6, 1932, in Hammond, Indiana, earned a B.S. in mathematics and physics from Purdue University in 1954 and went on to study mathematics at Stanford University (see Figure 2). While studying spaces of harmonic functions on Riemann surfaces under Halsey Royden, he posed the following question in his dissertation:

> Let V be a collection of harmonic functions on a Riemann surface R such that for each point p of R the set $V_p = \{v(p) : v \in V\}$ is countable. Must V then be countable? [9, p. 98]

Figure 2. John E. Wetzel shortly after he left Stanford (L) and in recent years (R).

Figure 3. Page 11 of the Boneyard Book.

Regarding the origin of the question, Wetzel explained that Royden had asked him to investigate a specific conjecture:
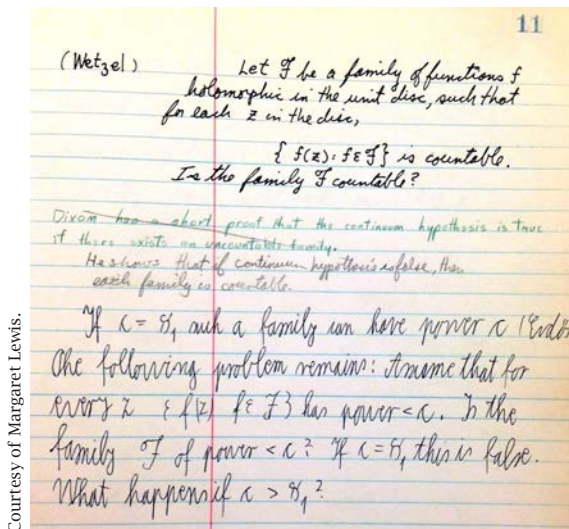
I thought about it for a while and eventually showed that what he conjectured was, in fact, not true. I remember reporting to him, thinking that all I had to do was write my work up in good form and I'd be finished with my dissertation; and I remember clearly Royden's reaction that my result would make up perhaps a third of an acceptable dissertation. The question might have had its genesis in the subsequent confusion. [8]

Wetzel left Stanford in 1961 to become an instructor at UIUC, having not yet finished his dissertation. He married Rebecca Sprunger in September 1962 and completed the writing of his dissertation in 1963. During this time, Paul Erdős visited the University of Illinois with his mother, who often accompanied him as he traveled from campus to campus [7, p. 7].

During one of his casual contacts with Erdős, Wetzel recounts,

I mentioned the question to him, rather timidly, if memory serves. He thought about it briefly and said it was interesting—and that was the extent of my mathematical contact with him. I don't think he ever told me that he had settled the matter, but every time he visited the campus in the next few years he always asked me if I had any new interesting questions. [8]

Upon learning (probably not before 1966, Wetzel said) about Erdős's proof, Wetzel wrote to Royden that "Erdős has showed that the answer to a question I asked in my dissertation is closely tied to the continuum hypothesis! So once again a natural analysis question has grown horns" [8].

The Ann Arbor Problem Book that Erdős mentions seemed most likely to be the Math Club book, which has its own fiery history. In the words of Peter Duren (professor emeritus at the University of Michigan):

When I came to Ann Arbor in 1962, I learned of a local tradition called "Math Club," an informal gathering of faculty and graduate students which met in the evening every month. A speaker was announced in advance, but the main attraction (in my view) was the series of 3-minute talks, unannounced and often spontaneous, that preceded the announced lecture. There people would tell their colleagues about neat mathematics they had come across, or raise questions, whatever they thought would be of interest. Afterwards each speaker was invited to write a short summary of his presentation in a book maintained for that purpose. The Secretary of the Math Club acted as guardian of the book, and both locals and visitors were invited to look through it. Unfortunately, the book was lost during the Christmas break of 1962–63, on the streets of Chicago. The man then serving as Secretary of the Math Club had carried the book (or books) with him when he drove to Chicago and had left it in his car overnight. Someone broke into the car and set it on fire, and the Math Club book was lost (among other items, including the car). The Math Club continued to meet, with a new book, but attendance gradually declined and the meetings were discontinued around 1990, as I recall. What Paul Erdős called the Ann Arbor Problem

Book must have been the Math Club book. But his reference can't be checked, since the original entries for December 1962 no longer exist. [5]

However, Wetzel declared, "I have never visited the University of Michigan; I've never even been to Ann Arbor" [8]. This initially led us to conclude that Erdős erred in his citation. Given his unique manner of doing mathematics and his myriad colleagues, such a slipup is understandable. Duren wrote:

I vividly recall him asking people (including me) at math conferences, "Where are you located?", which was the polite way of asking, "Who are you? I know I've met you somewhere." This was only natural, since he traveled so much and met so many mathematicians. It's easy to imagine that he didn't remember correctly where he had seen the problem. [5]

After much investigation, we are now able to trace the sequence of events from the original question sparked by Wetzel's dissertation to the problem treated by Erdős in his paper.

Within the UIUC mathematics library is a volume of particular importance for us. Wetzel explained:

The library maintains a bound volume of blank pages called the Boneyard Book, in which faculty and visiting faculty are encouraged to write problems, including whatever supporting information or commentary they care to include, and faculty looking for interesting problems can browse in it for inspiration.[8]

Senior Library Specialist Margaret Lewis recovered the relevant page of the Boneyard Book (see Figure 3), despite the fact that the volume had been collecting dust in the UIUC archives for decades. Duren recounted that, during the early 1960s, Erdős traveled frequently between the University of Michigan and the University of Illinois Urbana-Champaign, so conflating the two schools' problem books would have been a natural mistake.

However, the Boneyard Book led to many more questions. There are clearly four scribes who contributed to the page, so our next task was to identify the mathematicians involved. While the final entry of the page certainly boasts Erdős's distinctive penmanship, the remaining three entries required further scrutiny.

Wetzel explained that the first entry, though attributed to him, was not written by him: "I haven't a clue who wrote the problem in our Boneyard Book and wrote my name next to it. I thought for a few moments that it might have been Ranga Rao, with whom I shared an office during my (and his) first year at Illinois, 1961–62. But his handwriting

was recognizably 'Indian English penmanship,' and I think him unlikely" [8]. We therefore decided to investigate the second and third entries before tackling the first.

Jane Bergman, secretary of the chair of the UIUC math department, suggested that the Dixon mentioned in the Boneyard Book and Erdős's article (Figure 1) was Robert ("Bob") Dan Dixon:

I was able to find out that Mr. Robert Dan Dixon received his PhD from Ohio State University in 1962 in the same year he was hired here as an instructor in Mathematics. In 1964 he was recommended for promotion to assistant professor, but there is no result in his file of that recommendation. I would guess that he moved on at that point, but there is nothing in his file to support my guess. He was born in 1936. I hope you find this helpful. [8]

A colleague of Wetzel's, George Robert ("Bob") Blakley, had more information about Dixon. Bob Blakley and Bob Dixon both arrived at UIUC in September 1962 as new PhD's and both left in 1964. Blakley, in personal correspondence with the authors and Wetzel, wrote:

Bob [Dixon] went to the nascent Wright State University in the Dayton area as a founding father. First he founded the math department and headed it. Then he founded the computer science department and headed it…Late in the last century he retired from WSU, covered with glory…He still manages to fleece me regularly and disreputably in the most varied sorts of bets. But I think he has given up one hundred mile bike rides. [2]

Blakley provided us with three possible ways to contact the elusive Dixon: two email addresses which may or may not have been current, along with his home address. Wetzel sent a letter out to all three addresses, and luckily one route was successful. Dixon responded:

I was there [UIUC] from the fall of 1962 until the spring of 1964. I remember Erdős visiting and may even have had some time with him but I don't remember discussing this problem. The handwriting in the book is puzzling. The entry describing your problem is not familiar to me. The first entry that mentioned me could be by Bob [Blakley] as he generally printed. The second entry refers to me in third person but I could have written it…Although I can't remember any details there is a bit of familiarity. I worked in complex analysis at the time and I had a very interesting course in graduate school that covered the relevant set theory.

Lots of problems were thrown around in that group of young faculty. [4]

Dixon elaborated on "that group of young faculty" known as the SixtyTwo Illini Hall Group. In 1962 UIUC hired twenty new mathematics faculty to add to their faculty of one hundred; Illini Hall was located across from the mathematics building. According to Dixon:

> There was not a sense of privacy about the problems we were investigating. My own work was very specialized and detailed so I had no problems to share. Many others, like Jack Wetzel, did have problems that they proposed, or pointed us to, that they were curious about or needed to solve to get on to the problem they really wanted to work on. We fell then into three overlapping groups: proposers, solvers, and brokers. I was in the solvers group but not particularly successful, Jack may have been in all three. Bob Blakley was in all three but was effective as a broker…There were many others who participated in this interchange but my memory of names is bad. That said it was an experience that had more to do with my career than my own doctoral research. I suspect that was true of several of the other SixtyTwo Illini Hall Group. [4]

Upon receiving a copy of the Boneyard Book page, Blakley confirmed that he authored the entry in the Boneyard Book that reads, "Dixon has a short proof that the continuum hypothesis is true if there exists an uncountable family." Blakley also remarked that he feels "rather strongly…that Dixon is the third scribe." Given Erdős's parenthetical remark in his paper that "I have been informed that R. D. Dixon proved the first part of the theorem last year" [6], it is likely that Robert Dan Dixon was indeed the scribe of the third entry. Since Dixon expressed that he never spoke to Erdős about this problem during their overlapping time at UIUC, the information Erdős claimed was relayed to him almost certainly came from the Boneyard Book.

After immersion in the memories and details surrounding the problem, Wetzel recalled:

> I just remembered that I had given a faculty seminar and a departmental colloquium on the substance of my dissertation shortly after arriving at Illinois (even though the dissertation was not yet completely written), and that widened significantly the list of people who might have written the first entry in the Boneyard Book. [8]

A chance meeting between the first author and John P. D'Angelo (a professor at UIUC) finally revealed the most likely candidate for the first scribe.

D'Angelo was convinced that the handwriting was that of Lee A. Rubel:

> Rubel, who died in 1995, often contributed to the Boneyard Book. Furthermore, his many interests included the interplay between logic and function theory…Rubel would have been quite interested in this problem, and the handwriting is remarkably similar to that of notes he wrote to me around 1979–80. [3]

Of the possible candidates, Wetzel remarks:

> Lee certainly strikes me as the more likely…I truly don't doubt that Rubel was the author, but I confess that I still find it a little surprising that he never mentioned it to me—admitting always the possibility that he did and I have forgotten. [8]

The final piece of evidence was a sample of Rubel's handwriting, obtained by D'Angelo from the UIUC archives, which appears to validate this conclusion. In fact, D'Angelo tells us, Rubel was the creator of the Boneyard Book.

Let us now return to Erdős's paper [6]. It was easy to jump to the conclusion that Erdős had erred in his citation, since the problem appears in the Boneyard Book at the University of Illinois, and Erdős saw it written there. However, there is another scenario that seems to fit the facts more closely. In the first few lines of his paper (see Figure 1), Erdős cites the "Ann Arbor Problem Book" as the source of Wetzel's problem and mentions "an unsigned comment" and a remark by Roger Lyndon (then a professor at Michigan). Neither of those comments appears on the relevant page of the Boneyard Book (Figure 3).

The only explanation is that someone had transported the problem to Michigan and had recorded it either in the Math Club book or in a separate book of open problems. Peter Duren conjectures that the problem was transferred by Lee Rubel himself, who was making frequent trips to Ann Arbor in those days to work with Allen Shields [5]. Erdős first saw the problem written there and came up with the beautiful result presented in [6]. If Erdős saw the problem in December 1962, it could well have been in the Math Club book which perished in Chicago. However, Duren reports having examined the Math Club book for 1962–91, which now resides in the Bentley Historical Library at UM. There he found a record that Paul Erdős gave a lecture (entitled "Some Unsolved Problems") at the Math Club on September 10, 1963, just a week before his paper [6] was received by the *Michigan Mathematical Journal* (on September 18). Thus it seems far more likely that Erdős saw the problem in Ann Arbor during his short visit of 1963, in which case the

Ann Arbor Problem Book was not lost in Chicago and may yet be found. However, Duren has no memory of an independent problem book, nor do his fellow retirees at Michigan. In any case, it seems clear that, after having submitted his paper [6], Erdős saw the problem in the Boneyard Book and learned that Dixon had obtained part of the result independently.

No story can ever have the entirety of its details pinned down. As Wetzel said, "It may require transfinite induction to bring this matter to a close." However, we have identified with a high degree of certainty the trajectory of Wetzel's question as it made its way to Erdős. It began in 1961, when Wetzel posed the original question (for harmonic functions on Riemann surfaces) in his evolving dissertation. When he arrived at UIUC in 1962, he gave a talk on his graduate research. Lee Rubel was almost certainly one of the attendees and likely transmitted the problem to Ann Arbor. Rubel wrote Wetzel's question in the Boneyard Book in 1962, and Bob Blakley responded with an entry claiming that Bob Dixon had a proof, assuming the truth of the continuum hypothesis. Dixon crossed out Blakley's entry and wrote (in third person), "He showed that if the continuum hypothesis is false, then each family is countable." Dixon's short proof was rediscovered and published by Erdős, who went on to prove that an affirmative answer to Wetzel's problem is equivalent to the negation of the Continuum Hypothesis. Erdős's Boneyard Book entry is likely from the fall of 1963, after he had submitted his proof to the *Michigan Mathematical Journal*.

References

[1] MARTIN AIGNER and GÜNTER M. ZIEGLER, *Proofs from The Book*, Springer-Verlag, Berlin, third edition, 2004, including illustrations by Karl H. Hofmann.
[2] GEORGE ROBERT BLAKLEY, personal correspondence.
[3] JOHN P. D'ANGELO, personal correspondence.
[4] ROBERT DAN DIXON, personal correspondence.
[5] PETER DUREN, personal correspondence.
[6] PAUL ERDŐS, An interpolation problem associated with the Continuum Hypothesis, *Michigan Math. J.* **11**:9–10, 1964.
[7] RONALD L. GRAHAM and JAROSLAV NEŠETŘIL, editors, *The Mathematics of Paul Erdős. I*, volume 13 of Algorithms and Combinatorics, Springer-Verlag, Berlin, 1997.
[8] JOHN EDWARD WETZEL, personal correspondence.
[9] JOHN EDWARD WETZEL, *A Compactification Theory with Potential-Theoretic Applications*, Thesis (PhD)–Stanford University, 1964, ProQuest LLC, Ann Arbor, MI.