Multiplicative structure of integers, shifted primes and arithmetic functions

Kevin Ford
http://www.math.uiuc.edu/~ford/erdos100.pdf

University of Illinois at Urbana-Champaign

July 2, 2013
For primes $p$, let $X_p$ be independent Bernoulli random variables with

$$\text{Prob}(X_p = 1) = \frac{1}{p}, \quad \text{Prob}(X_p = 0) = 1 - \frac{1}{p}.$$ 

Each models whether a random integer is divisible by $p$.

**Theorem (Kubilius, 1956. Universal transference principle)**

*For any $\varepsilon > 0$, the sequence $\{X_p : p \leq y^\varepsilon\}$ models “within $\varepsilon$” the prime factors $\leq y^\varepsilon$ of a random integer $\leq y$.***

Roughly speaking, for any theorem about the sequence $\{X_p : p \leq y^\varepsilon\}$, the corresponding theorem about prime factors of random integers will be true with a small error term.
Example: The Erdős-Kac theorem

Recall \( \text{Prob}(X_p = 1) = 1/p \) and \( \text{Prob}(X_p = 0) = 1 - 1/p \).

**Example.** From \( E X_p = 1/p \) and \( V X_p = 1/p - 1/p^2 \), get

\[
E \left( \sum_{p \leq y^\varepsilon} X_p \right) = \log \log y + O_\varepsilon(1), \quad V \left( \sum_{p \leq y^\varepsilon} X_p \right) = \log \log y + O_\varepsilon(1).
\]

From the Central Limit Theorem for \( \sum_{p \leq y^\varepsilon} X_p \), get

**Theorem (Erdős-Kac, 1939)**

Let \( \omega(n) \) be the number of distinct prime factors of \( n \). For each real \( z \),

\[
\lim_{y \to \infty} \frac{1}{y} \# \left\{ n \leq y : \frac{\omega(n) - \log \log y}{\sqrt{\log \log y}} \leq z \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} \, dt.
\]

**Hardy-Ramanujan:** \( \omega(n) \sim \log \log n \) for almost all \( n \)
Kubilius’ model and random walks

**Kubilius, Billingsly (1960s).** Connect \( \omega(n, t) = \#\{p | n : p \leq t\} \) to Brownian motion.

- (Erdős, 1930s–). Prime factors in any interval are Poisson.
  Provided \( I = [\exp \exp(t), \exp \exp(u)] \) isn’t too short,

\[
P(\text{random integer has } k \text{ prime factors in } I) \sim e^{t-u} \frac{(u-t)^k}{k!}.
\]

Normal number of prime factors is \( \sim u - t \).
- Prime divisors in disjoint intervals are independent.

**Probabilistic model 2 (Galambos, Maier, DeKoninck 1970s,80s).**
By a theorem of Rényi, these properties characterize the Poisson process: the sequence of (all but the smallest and the largest) prime factors of a random integer, taken on a log log –scale, behave like a random walk with exponentially distributed steps.

Recall: \( X \) has exponential distribution if \( P(X \geq y) = e^{-y} \) for \( y > 0 \).
Probabilistic model 2: The sequence of prime factors of a random integer, taken on a log log—scale, behave like a random walk with exponentially distributed steps.

Theorem (Maier, Tenenbaum (1984); was a 1948 conjecture of Erdős) Almost all integers have two divisors \(d_1, d_2\) satisfying \(d_1 < d_2 < 2d_1\).

Multiplication table problem (Erdős, 1955). Let

\[ A(N) = \#\{de : 1 \leq d \leq N, 1 \leq e \leq N\}. \]

Equivalently, count integers \(\leq N^2\) with a divisor near \(N\).

Easy (Erdős): \(A(N) = o(N^2)\). Proof: For most pairs \((d, e)\),

\[ \omega(de) \approx \omega(d) + \omega(e) \approx \log \log N + \log \log N = 2 \log \log (N^2) + O(1). \]
Improved bounds by Erdős (1960) and Tenenbaum (1984).

Theorem (KF, 2008)

\[ A(N) \asymp \frac{N^2}{(\log N)^c (\log \log N)^{3/2}}, \quad c = 1 - \frac{1 + \log \log 2}{\log 2} \approx 0.08607 \]

**Key**: Fine analysis of the prime factor random walk; small deviations of the prime factor random walk lead to large discrepancies in the distribution of divisors.

**Open problem.** Is there an asymptotic formula?

**Generalization.** Find the order of

\[ A_k(N_1, \ldots, N_k) = \#\{d_1 \cdots d_k : 1 \leq d_j \leq N_j \ (1 \leq j \leq k)\} \]

Order known for all \(N_1, \ldots, N_k\) for \(k = 2\) (KF, 2008), \(3 \leq k \leq 6\) (Koukoulopoulos 2010, 2013). Partial results for \(k > 6\).
Distribution of large prime factors

**Notation:** \( P_1(n) = \) largest prime factor of \( n \), \( P_2(n) = 2\)nd largest, etc.

**Distribution of \( P_1(n) \).** Early work of Ramanujan, Dickman, Erdős and others. \( \Psi(x, y) = \#\{n \leq x : P_1(n) \leq y\} \) is well understood now.

**Joint distribution of \( P_1(n), \ldots, P_k(n) \).** (Billingsly, 1972).
(Donnelly and Grimmett, 1993): It’s the Poisson-Dirichlet distribution

Simple description: Let \((x_1, x_2, \ldots)\) be a random partition of \([0, 1]\):

\[
\begin{array}{ccccccc}
  & & x_1 & & x_2 & & x_3 & \ldots \\
 0 & y_2 & & y_1 & & y_4 & y_3 & 1
\end{array}
\]

Let \( y_1 = \) largest \( x_i \), \( y_2 = \) the 2nd largest, etc.
Then \((y_1, y_2, \ldots)\) and \(\left(\frac{\log P_1(n)}{\log n}, \frac{\log P_2(n)}{\log n}, \ldots\right)\) have the same distribution.

Same distribution appears in the cycle lengths of random permutations, factor sizes of random polynomials in \( \mathbb{F}_q[t] \), certain physical processes, etc.
Sets $\mathcal{P}_a = \{ p + a : p \text{ prime} \}$, where $a \neq 0$ fixed.

Used to study arithmetic functions $\phi$, $\sigma$, orders in $\mathbb{Z}/p\mathbb{Z}$, primality testing, factorization algorithms, cyclotomic fields, Fermat’s Last Theorem, etc. Important cases $a = -1, 1$.

**Small and intermediate prime factors.** Essentially the same distribution as for a random integer via sieve methods, Bombieri-Vinogradov, Gallagher. Ideas originate from 1935 paper of Erdős.

- $\omega(p + a)$ has normal order $\log \log p$ (Erdős, 1935)
- $\omega(p + a)$ satisfies the same CLT as $\omega(n)$ (Halberstam, 1956).
- $\#\{d_1 d_2 \in \mathcal{P}_a : 1 \leq d_i \leq N\} \sim \frac{A(N)}{\log N}$ (Koukoulopoulos, 2011)

Large prime factors ($> p^{1/2}$) of shifted primes largely unknown due to lack of knowledge of primes in progressions to large moduli.
Anatomy of values of arithmetic functions

Let $\mathcal{V}_f = \{f(n) : n \in \mathbb{N}\}$, $\mathcal{V}_f(x) = \#\mathcal{V}_f(x) \cap [1, x]$.

**Pillai, 1929.** $V_\phi(x) = o(x)$. Idea: $\omega(n) \approx \log \log x$ for most $n \leq x$, and $2^{\omega(n) - 1} | \phi(n)$.

**Erdős, 1935.** $V_\phi(x) = x(\log x)^{-1+o(1)}$. Idea: $\omega(p - 1) \sim \log \log p$ for most $p | n$. Hence, for typical $n$, $\omega(\phi(n))$ is abnormally large.

Improvements by Erdős, Erdős-Hall, Pomerance, Maier-Pomerance.

**KF, 1998.** exact order of $V_\phi(x)$ found:

$$V_\phi(x) \asymp \frac{x}{\log x} \exp \left\{ C_1 (\log \log \log x - \log \log \log \log \log x)^2 
+ C_2 \log \log \log x + C_3 \log \log \log \log \log x \right\}.$$  

Same order for $V_\sigma(x)$ and for the counting function of the semigroup generated by $\mathcal{P}_a$, $a \neq 0$.

Open problem. Is there an asymptotic formula?
Euler’s function. More open problems

**Carmichael, 1907.** \( \forall m \in \mathcal{V}_\phi, \phi(x) = m \) has at least 2 solutions \( x \).

Known: such an \( m \), if it exists, exceeds \( 10^{10^{10}} \) (KF, 1998).

Known: \( \forall k \geq 2, \exists m \) so that \( \phi(x) = m \) has exactly \( k \) sol’s (KF, 1999).

**Erdős.** \( \forall C > 1, \) is there an \( m \in \mathcal{V}_\phi \) so that \( \phi(x) = m \implies x > Cm \) ?

**KF, 1998.** Is there an \( m \in \mathcal{V}_\phi \) so that \( \phi(x) = m \implies 6 \mid x \) ?

The corresponding question with 6 replaced by 2,3,4,5,7,8 or 9 is affirmative. I think for 6, the answer is no. Perhaps for 10 also.

**Erdős.** Are there infinitely many \( n \) with \( \phi(n) = \phi(n + 1) \)?

\( \forall \varepsilon, \) are there infinitely many \( n \) with \( |\phi(n) - \phi(n + 1)| < n^\varepsilon \)?

**Alkan-Ford-Zaharescu (2009).** True with \( \varepsilon = 0.84 \).
Prime Chains

Definition

Let \( a \prec b \) if \( b \equiv 1 \pmod{a} \); that is, \( a | (b - 1) \).

Prime chains: \( p_1 \prec p_2 \prec \cdots \prec p_k \)

Example: \( 2 \prec 5 \prec 11 \prec 23 \prec 47 \prec 283 \prec 2432669 \)

Prime chain problems arise in the study of iterates of \( \phi \) and applications thereof; value distribution of \( \phi, \sigma, \lambda \); primality certificates (complexity of the Pratt certificate).

Basic question. Are there arbitrarily long prime chains? Yes - Infinitely long (Dirichlet, 1837).
Prime chains with a given starting prime

**Prime chains:** $p_1 \prec p_2 \prec \cdots \prec p_k$, $p_{j+1} \equiv 1 \pmod{p_j}$ for each $j$.

---

**Theorem (Ford-Konyagin-Luca, 2010)**

Let $N(x; p)$ be the number of prime chains starting at a prime $p$ and ending at a prime $\leq xp$. Then for every $\varepsilon > 0$, $N(x; p) \leq C(\varepsilon)x^{1+\varepsilon}$.

Note $N(x; p) \geq \pi(xp; p, 1) \approx x/\log x$.

An (perhaps unexpected) application to a 1958 conjecture of Erdős.

---

**Theorem (Ford-Luca-Pomerance, 2010)**

$\phi(n) = \sigma(m)$ has infinitely many solutions (i.e., $\mathcal{V}_\phi \cap \mathcal{V}_\sigma$ is infinite)

---

**Theorem (Ford-Pollack, 2012)**

Almost all values of $\phi$ are not values of $\sigma$ and vice-versa. That is, the counting function of $\mathcal{V}_\phi \cap \mathcal{V}_\sigma$ is $o(V_\phi(x) + V_\sigma(x))$. 

---

K. Ford (Illinois)
The aggregate of all prime chains ending at a given prime $p$ has a *tree structure*, the **Pratt tree** of $p$ (related to the Pratt primality certificates).
Pratt tree height

**Height** $H(p)$, = length of longest prime chain ending at $p$.
Trivially, $H(p) \leq \frac{\log p}{\log 2} + 1$.

$H(p) = 2$ for Fermat primes.

**Conjecture (Erdős ?):** For each $k \geq 3$, there are infinitely many primes with $H(p) = k$.

**Katai, 1968.** $H(p) \gg \log \log p$ for almost all $p$.

**Ford-Konyagin-Luca, 2010.** $H(p) \ll (\log p)^{0.9503}$ for almost all $p$.

Assuming the large prime factors of the shifted primes in the Pratt tree obey the Poisson-Dirichlet distribution, and are all independent of one another, one can model $H(p)$ be a branching random walk. Fine analysis of this process leads to the following conjecture.

**Conjecture (Ford-Konyagin-Luca, 2010)**

For most primes $p$, $H(p) \approx e \log \log p - \frac{3}{2} \log \log \log p + "O(1)"$. 
Let $\mathcal{P}_q$ be the set of primes $p$ such that the Pratt tree for $p$ doesn’t contain the prime $q$. For example,

$$\mathcal{P}_3 = \{2, 5, 11, 17, 23, 41, 47, 83, 89, 101, 137, 167, 179, 251, \ldots \}$$

Sieve methods quickly imply the counting function is $O(x / \log^2 x)$. Numerical computations of $\mathcal{P}_3$ up to $10^{13}$ indicate that the counting function is $\approx x^{0.62}$.

**Theorem (KF, 2013)**

The counting function of $\mathcal{P}_q$ is $O(x^{1-c})$ for some positive $c = c(q)$.

**Open problem.** Show that $\mathcal{P}_q$ is infinite. Likely extremely hard. $\mathcal{P}_5$ infinite (almost) implies Carmichael’s conjecture.
Largest prime factors. Open problems.

**Expected.** $P_1(n)$ and $P_1(n + 1)$ are independent.

**Theorem (Erdős-Pomerance, 1978)**

*We have*

1. $P_1(n) < P_1(n + 1)$ for a positive proportion of $n$;
2. $P_1(n) > P_1(n + 1)$ for a positive proportion of $n$;
3. *certain orderings of* $P_1(n - 1), P_1(n), P_1(n + 1)$ *occur infinitely often.*

**Balog, 2001.** Showed $P(n - 1) > P(n) > P(n + 1)$ infinitely often.

**Open problem.** Does any particular ordering of $P_1(n - 1), P_1(n), P_1(n + 1)$ occur for a positive proportion of $n$?

**Open problem.** Do all patterns (orderings) of $P_1(n), \ldots, P_1(n + 3)$ occur infinitely often?
**Conjecture.** \((P_1(p + a), \ldots, P_k(p + a))\) has the same distribution as \((P_1(n), \ldots, P_k(n))\).

True assuming Elliott-Halberstam conjecture. Unconditionally, very little known due to lack of knowledge of primes in arithmetic progressions to large moduli.

**Smooth shifted primes.** Erdős (1935) showed that \(P_1(p + a) < p^c\) infinitely often for some \(c < 1\). **Baker-Harman, 1998:** \(c = 0.2931\). Applications to \(\phi\) and Carmichael numbers.

**Large prime factors.** \(P_1(p + a) > p^c\) infinitely often. **Baker-Harman, 1998:** \(c = 0.677\).

**Open problem (Buchstab).** (i) Are there infinitely many primes \(p\) such that all prime factors of \(p + a\) are \(3 \mod 4\)? (ii) Same with \(3 \mod 4\) replaced by an arbitrary \(a \mod q\).
Propinquity of divisors. Hooley’s \(\Delta\)-function.

**Theorem (Maier, Tenenbaum (1984); was a 1948 conjecture of Erdős)**

Almost all integers have two divisors \(d_1, d_2\) satisfying \(d_1 < d_2 < 2d_1\).

Let \(\Delta(n) = \max_y \#\{d|n : y < d \leq ey\}\) (a concentration function).

**Normal order (Maier-Tenenbaum, 1984; 2009).** For almost all \(n\),

\[
\left(\log n\right)^{c-\varepsilon} < \Delta(n) < \left(\log n\right)^{\log 2+\varepsilon}, \quad c \approx 0.33827
\]

They conjecture that the lower bound is closer to the truth.

**Average values (Hall-Tenenbaum (lower); Tenenbaum (upper)).**

\[
\log \log x \ll \frac{1}{x} \sum_{n \leq x} \Delta(n) \ll \exp \left\{ C \sqrt{\log \log x \log \log \log x} \right\}.
\]

**Twisted \(\Delta\)-functions (Daniel; de la Bretèche-Tenenbaum):**

\[
\Delta_f(n) = \max_{1 \leq y < z \leq ey} \left| \sum_{d|n, y < d \leq z} f(d) \right|, \quad f = \mu, \chi, \ldots
\]
Prime chains: $p_1 \prec p_2 \prec \cdots \prec p_k$, $p_{j+1} \equiv 1 \pmod{p_j}$ for each $j$.

Theorem (Ford-Konyagin-Luca, 2010)

Let $f(p)$ be the number of prime chains that end at a prime $p$. Then

$$\frac{1}{3} \log p \leq f(p) \leq 3 \log p$$

for almost all $p$.

$f(p)$ is also the number of nodes in the Pratt tree for $p$.

Open Problem. Are there infinitely many $p$ with $f(p) = o(\log p)$?

Observations: $f(p) = 2$ for Fermat primes. $f(p)$ is small if $p - 1$ is very smooth, e.g. $f(p) = 4$ if $p = 2^a3^b + 1$. 
D. H. Lehmer, 1930. Is there a composite $n$ with $\phi(n)|(n - 1)$?

Pomerance (1977): The counting function of such $n$ is $O(n^{1/2}(\log n)^{O(1)})$.

**Open Problem:** Prove there are infinitely many chains $p_1 \prec p_2 \prec p_3$ with $\frac{p_3-1}{p_2} = \frac{p_2-1}{p_1}$ (quasi-geometric progression of primes).