Some Number Theoretic Matching Problems
Carl Pomerance

§1. INTRODUCTION.

Let \( R \) be a relation on the integers. Then we may be able to define two number theoretic functions \( f_R, g_R \) as follows.

**Definition 1.** If \( n \) is a natural number, let \( f_R(n) \) denote the least natural number such that for every integer \( m \), there is a one-to-one mapping

\[
F: \text{dom} R \cap \{1, \ldots, n\} \to \{m+1, \ldots, m+f_R(n)\}
\]

such that for each \( i \in \text{dom} F \), we have \((i, F(i)) \in R\).

**Definition 2.** If \( n \) is a natural number, let \( g_R(n) \) denote the maximal integer such that there is a one-to-one mapping

\[
G: \text{rng} R \cap \{n+1, \ldots, n+g_R(n)\} \to \mathbb{Z}
\]

such that for each \( i \in \text{dom} G \), we have \((G(i), i) \in R\).

Note that it is possible for a given relation \( R \) for either \( f_R(n) \) or \( g_R(n) \) not to exist. However if \( R \) is a natural relation and if \( f_R(n) \) or \( g_R(n) \) happen to exist, we then have the interesting problem of finding exact or approximate formulae for them. In this short survey article we present some attractive problems attached to three natural relations \( R \).
Let \( C \) be the coprime relation where \((a,b) \in C\) if and only if \((a,b) = 1\). The problem of computing \( f_C(n) \) is due to D. J. Newman. It is clear \( g_C(n) \) does not exist.

Let \( \varphi \) be the prime factor relation where \((a,b) \in \varphi\) if and only if \(a\) is a positive prime and \(a|b\). The problem of computing \( f_\varphi(n) \) is due to P. Erdős and J. L. Selfridge, while the problem of computing \( g_\varphi(n) \) is due to C. A. Grimm.

Let \( \mathcal{D} \) be the proper divisor relation where \((a,b) \in \mathcal{D}\) if and only if \(a|b\) and \(1 < a < b\). The problems of computing \( f_\mathcal{D}(n) \) and \( g_\mathcal{D}(n) \) are due to P. Erdős and C. Pomerance.

§2. D. J. NEWMAN'S COPRIME MAPPING CONJECTURE.

About 20 years ago, D. J. Newman conjectured that \( f_C(n) = n \) for all \(n\). That is, for every pair of natural numbers \(m, n\), there is a one-to-one mapping

\[
F: \{1, \ldots, n\} \rightarrow \{m + 1, \ldots, m + n\}
\]

such that \((i, F(i)) = 1\) for each \(i, 1 \leq i \leq n\). Daykin and Baines [3] showed that a "coprime mapping" \(F\) always exists in the special case \(m = n\). Chvátal [2] proved Newman's conjecture for all \(n \leq 1002\). Recently, Pomerance and Selfridge [10] proved Newman's conjecture in complete generality. This task was accomplished by producing an algorithm for the construction of a coprime mapping \(F\). We now describe this algorithm.
Clearly it is a trivial matter to produce coprime mappings in the cases \( n = 1, 2 \). So assume \( n \geq 3 \) and assume we have given algorithms for the construction of coprime mappings in every prior case. Let \( N = n - \lfloor n/2 \rfloor \) denote the number of odd numbers in \( \{1, \ldots, n\} \). Then \( N \geq 2 \). Label these odd numbers \( k_1, \ldots, k_N \) where \( \phi(k_i)/k_i \leq \phi(k_{i+1})/k_{i+1} \) for \( 1 \leq i < N \), \( \phi \) denoting Euler's function. If \( 1 \leq i \leq N-2 \), let \( F(k_i) \) be the least even number in \( \{m+1, \ldots, m+n\} \) that is coprime to \( k_i \) and not in \( \{F(k_1), \ldots, F(k_{i-1})\} \). Thus \( F(k_i) \) will exist if there are at least \( i \) even numbers in \( \{m+1, \ldots, m+n\} \) that are coprime to \( k_i \). This condition is in fact always satisfied and follows as any easy corollary of the following result.

**Theorem A.** Let \( D(u,N) \) denote the number of positive odd \( a \leq 2N-1 \) with \( \phi(a)/a \leq u \). For each integer \( k \), let \( E(k,N) \) denote the maximal number of integers coprime to \( k \) that can be found among any set of \( N \) consecutive integers. Then if \( k \) is odd, \( 1 < k \leq 2N-1 \), and \( k \) is not the largest prime not exceeding \( 2N-1 \), then
\[
D(\phi(k)/k,N) < E(k,N).
\]

If \( \{m+1, \ldots, m+n\} \) has \( N \) even numbers, then \( k_{N-1} \) and \( k_N \) can be easily seen to be mapped in a coprime fashion to the two remaining even numbers. If \( \{m+1, \ldots, m+n\} \) has only \( N-1 \) even numbers, then \( k_N \) is sent to the last remaining even number and \( k_{N-1} \) is sent to one of \( m+1, m+n \), which are both odd.
In either case, we have exactly \([n/2]\) remaining numbers in \(\{m+1, \ldots, m+n\}\) which are all odd and consecutive. We must map the even numbers \(\{2, \ldots, 2[n/2]\}\) in a one-to-one, coprime fashion to this string of consecutive odd numbers. Here we use our induction hypothesis for \([n/2]\) to show we can complete this last step.

Thus a proof of Theorem A is all that remains to prove Newman's conjecture. This task is not exactly easy. It should be recognized that the function \(D(u,N)\) is related to the distribution function \(D_\varphi(u)\) of \(\varphi(n)/n\):

\[
D_\varphi(u) = \lim_{n \to \infty} \frac{1}{n} \cdot \text{card}(a \leq n : \varphi(a)/a \leq u).
\]

In the proof of Theorem A, an idea of Erdős [4] on estimating \(D_\varphi(u)\) is used to estimate \(D(u,N)\). But since an asymptotic estimate is not of very much use in Theorem A, explicit estimates on the prime counting function \(\pi(x)\) due to Rosser and Schoenfeld [12] must be used.

§3. GRIMM'S PROBLEM.

Grimm [9] conjectured that if \(p < p'\) are consecutive primes, then \(g_\varphi(p) \geq p' - p - 1\). That is, if \(n+1, \ldots, n+k\) are all composite, then there are distinct primes \(p_1, \ldots, p_k\) with \(p_i \mid n+i\). Grimm's problem is still unsolved, but there has been some progress on it and the more general problem of estimating \(g_\varphi(n)\).

Concerning the latter problem, Ramachandra, Shorey, and Tijdeman [11] showed

\[
(1) \quad g_\varphi(n) \gg (\log n/\log\log n)^3.
\]
Erdős and Selfridge [8] have showed that

\[ g_\varphi(n) \ll (n/\log n)^{1/2}. \]

Thus if Grimm's conjecture is true and \( p, p' \) are consecutive primes, then

\[ p' - p \ll (p/\log p)^{1/2}. \]

This extraordinary corollary, which does not even appear to follow from the Riemann Hypothesis, shows that Grimm's conjecture, if true, must lie very deep.

There is a wide gap between (1) and (2), so short of improving on these estimates, one might ask what can be done for infinitely many \( n \). Erdős and Selfridge [8] have shown there is a positive constant \( c_1 \) with

\[ g_\varphi(n) \geq \exp(c_1 (\log n \log \log n)^{1/2}) \]

for infinitely many \( n \). They also stated without proof that there is a positive constant \( c_2 \) and infinitely many \( n \) such that

\[ g_\varphi(n) \leq \exp(c_2 \log n \log \log \log n/\log \log n). \]

We now prove that there is a positive constant \( c_3 \) with

\[ g_\varphi(n) \leq \exp(c_3 (\log n \log \log n)^{1/2}) \]

for infinitely many \( n \).

Indeed, let \( \psi(x,y) \) denote the number of integers \( n \leq x \) not divisible by any prime exceeding \( y \). Let \( L = L(x) = \exp((\log x \log \log x)^{1/2}) \). By de Bruijn [1]
there is a positive constant $\alpha$ such that
\[
\psi(x, L^\alpha) > \frac{x}{L^{\alpha/2}}
\]
for all large $x$. Let $w_k = x/L^\alpha + kL^{2\alpha}$ for $k = 0, 1, 2, \ldots$. Then by an averaging argument, for some $k = x/L^{2\alpha}$, there are at least $L^\alpha$ integers in the interval $(w_k, w_{k+1})$ divisible only by primes not exceeding $L^\alpha$. But $\pi(L^\alpha) < L^\alpha$, so there is no way we can pick distinct prime factors for these $L^\alpha$ integers. Thus
\[
g_\phi(w_k) < w_{k+1} - w_k = L^{2\alpha}.
\]
This shows we may take $c_3 = 3\alpha$ in (4).

Let $\rho(u)$ denote Dickman's function, so that if $u \geq 1$, $\psi(x, x^{1/u}) \sim \rho(u)x$ (see de Bruijn [1]). Another averaging argument shows that for each $\varepsilon > 0$, the lower density of the set of $n$ with $g_\phi(n) < n^\varepsilon$ is at least $\rho(1/\varepsilon) > 0$.

In light of these results, we conjecture that there are positive constants $c_4$, $c_5$ with
\[
\exp(c_4(\log n \log\log n)^{1/2}) < g_\phi(n) < \exp(c_5(\log n \log\log n)^{1/2})
\]
for all large $n$.

§4. A PROBLEM OF ERDŐS AND SELFRIDGE.

Recently Erdős and Selfridge [5] considered the problem of estimating $f_\phi(n)$, the least number so that for each $m$ there are distinct integers $a_1, \ldots, a_\pi(n)$ in $(m, m+f_\phi(n)]$ with $p_i | a_i$, $i \leq \pi(n)$. Here $p_i$ denotes
the $i$-th prime. Surprisingly little can be proved about $f_\varphi(n)$. It is obvious that $f_\varphi(n) \geq p_\pi(n)$ (just take $m = 0$). In fact, from looking at

$$m = p_\pi(n)^2 - p_\pi(n) - 1,$$

we see that $f_\varphi(n) \geq 2p_\pi(n) - 1$. Thus for every $\varepsilon > 0$, $f_\varphi(n) \geq (2-\varepsilon)n$ for all large $n$. Erdős and Selfridge [5] have shown that $f_\varphi(n) \geq (3-\varepsilon)n$ for all large $n$. They accomplish this by showing that for each $k \geq k_0(\varepsilon)$, there is a set of $k^2$ primes $q_1 < \ldots < q_{k^2}$ and an interval of length at least $(3-\varepsilon)q_1^k$ which contains only $2k$ distinct multiples of the $q_i$. Their proof uses Brun's method.

On upper bounds for $f_\varphi(n)$, the best that is known comes from Erdős and Pomerance [6] who show $f_\varphi(n) < n^{3/2}(\log n)^{-1/2}$. To this observer, there certainly seems to be ample room for improvement on both the upper and lower bounds for $f_\varphi(n)$.

The proof of the upper bound just mentioned for $f_\varphi(n)$ uses the famous "Marriage Theorem" of Hall. This theorem states that if $R \subseteq S \times T$ and if for each $A \subseteq S$, the cardinality of $R(A)$ is at least as big as the cardinality of $A$, then $R$ contains a one-to-one function mapping $S$ into $T$.

§5. THE RELATION $\mathcal{R}$.

In [6], Erdős and Pomerance consider the function $f_\varphi(n)$, the least number so that for every $m$ there are distinct integers $a_1, \ldots, a_n$ in $(m, m+f_\varphi(n)]$ with
\( i \mid a_i \) for each \( i \leq n \). (Note that it is unimportant to insist \( 1 < i < a_i \) in this problem.) They show

\[(5) \quad n(\log n/\log\log n)^{1/2} \ll f(n) \ll n^{3/2}.\]

The upper bound in (5) is similar to the upper bound for \( f_G(n) \) mentioned in §4 and, in fact, the proofs are essentially the same. The lower bound in (5) comes from examining the special case \( m = n \). Thus if \( f(n) \) is the least integer for which there are distinct integers \( a_1, \ldots, a_n \) in \( (n, f(n)) \) with \( i \mid a_i \) for each \( i \leq n \), then in [6] it is shown that

\[(6) \quad f(n) \gg n(\log n/\log\log n)^{1/2}.\]

The proof of (6) uses an asymptotic formula for \( \log \psi(x, y) \) (see §3) when \( y \) is in the vicinity of \( \log x/\log\log x \) due to de Bruijn [1] plus an intricate geometric averaging argument that makes use of the convex hull of the graph of the function \( \log \psi(x, y) \) for fixed \( y \).

The connection between the function \( \psi(x, y) \) and the function \( f(n) \) comes from the following easy lemma:

\[\psi(n, y) - \psi(f(n)/y, y) \leq \psi(f(n), y) - \psi(n, y)\]

for every \( y \). Indeed, if \( f(n)/y < j \leq n \) and no prime factor of \( j \) exceeds \( y \), then \( jk \in (n, f(n)) \) implies no prime factor of \( jk \) exceeds \( y \).

Using the Marriage Theorem (see §4) and the Prime Number Theorem, Erdős and Pomerance show

\[(7) \quad f(n) \ll n(\log n)^{1/2}.\]
The gap between (6) and (7) is not large and perhaps an asymptotic formula for \( f(n) \) is attainable.

Recently Erdős and Pomerance [7] have considered the function \( g_\beta(n) \), an obvious analogue to Grimm's function \( g_\phi(n) \). Thus \( g_\beta(n) \) is the largest number so that corresponding to the composite \( n + i, 1 \leq i \leq g_\beta(n) \), there are distinct integers \( a_i \) with \( a_i | n+i \) and \( 1 < a_i < n+i \). By letting \( a_i \) be the largest proper divisor of \( n + i \), we immediately get \( g_\beta(n) \gg n^{1/2} \). Using results of Huxley and Warlimont on the frequency of large gaps between consecutive primes, we can prove

\[
(8) \quad g_\beta(n) \leq n^{7/12 + o(1)},
\]

while if the Riemann Hypothesis holds, we have
\[
g_\beta(n) \leq n^{1/2 + o(1)}. \quad \text{In fact, we believe} \quad g_\beta(n) \ll n^{1/2}. \]

This result follows from the following very strong generalization of the Goldbach and twin prime conjectures: For each sufficiently large integer \( y \equiv 2 \pmod{3} \), there is a \( t < \sqrt{y} \) with \( y+t, y+t+2, y+t+6, y-t, y-t+2 \) all prime. To see the connection with \( g_\beta(n) \), let \( y \geq \sqrt{n+1} \) be minimal with \( y \equiv 0 \pmod{3} \). The six integers

\[
(y+t)(y-t), \quad (y+t+2)(y-t), \quad (y+t+6)(y-t),
\]

\[
(y+t)(y-t+2), \quad (y+t+2)(y-t+2), \quad (y+t+6)(y-t+2)
\]

are all between \( n \) and \( n + 15\sqrt{n} \) and collectively have only 5 proper divisors larger than 1.

If \( S \) is a set of integers, let
\[ D(S) = \{d : \text{for some } s \in S, \ d|s, \ 1 < d < s\} \]

Then by the Marriage Theorem there is a set of composite integers \( S \subset \{n+1, \ldots, n+g_f(n)+1\} \) such that
\[ |D(S)| < |S| \]
and for every \( T \subset S \) with \( T \neq S \), we have
\[ |D(T)| \geq |T| \].
It is not hard to show that the set \( S \) is unique. We call \( S \) the "blocking configuration" for \( n \). As a corollary of (8), we can show that for all sufficiently large \( n \), each member of \( n \)'s blocking configuration must be either the square of a prime or the product of two distinct primes. We conjecture that this fact holds for every natural number \( n \).

REFERENCES


Mathematics Department
University of Georgia
Athens, Georgia 30602