DISTRIBUTION OF SQUAREFREE INTEGERS IN NON-LINEAR SEQUENCES

IVAN ERNEST STUX
DISTRIBUTION OF SQUAREFREE INTEGERS
IN NON-LINEAR SEQUENCES

IVAN E. STUX

I. This paper investigates the occurrences of the squarefree integers in sequences $s_n = [f(n)]$, $n = 1, 2, 3, \cdots$ where $f(x)$ belongs to classes of functions described by 'smoothness' conditions. The result obtained is an extension of the well known fact that $Q(x) = \beta/\pi^2 x + O(x^{1/2})$, where $Q(x) =$ number of squarefree integers $\leq x$; it states that $Q_s(x) \sim 6x^2 g(x)$ where $Q_s(x) =$ number of squarefree integers $\leq x$ in the sequence $s_n$, and $g(x)$ is the inverse function of $f(x)$.

This result relates to the deep theorem of Piateskii-Shapiro which states that if $1 < c < 12/11$ then the sequence $[n^c]$ has the proper rate of primes occurring, namely, $\pi_e(x) \sim x^{1/c}/\log x$.

The classes of functions used is described by the following:

**Definition 1.** for given $1 < c < 2$, $0 < \delta < 1$

(1) $S(c, \delta) =$ set of functions $f(x)$ such that for some constant $a > 0$ depending on $f$, and for sufficiently large $x$'s, depending on $f$,

$$
(ax^c)^{(i)} \leq (f(x))^{(i)} < (ax^{c+\delta})^{(i)}
$$

holds for $i = 0, 1, 2$, the superscripts indicating the $i^{th}$ derivative.

Functions like $x^c$, $1 < c < 2$, or more generally $\sum_{i=1}^{k} a_i x^i (\log x)^{d_i}$, where the leading term has $a > 0$, $1 < c < 2$, belong to these classes of functions.

The following theorem will be proved:

**Theorem 1.** Let $1 < c < 4/3$, then there exists $\delta = \delta(c) > 0$, some small value depending on $c$ such that if $f(x) \in S(c, \delta)$, $0 \leq \delta < \delta(c)$, then

$$
Q_s(x) = 6/\pi^2 x^2 + O((g(x))^{1-\varepsilon})
$$

holds for some $\varepsilon > 0$ depending on $c$ and $\delta$, where $Q_s(x) =$ number of $\{s_n \leq x, s_n = [f(n)], s_n =$ squarefree, $n = 1, 2, 3, \cdots\}$, $g(x) =$ inverse function of $f$, $[z] =$ integer part of $z$.

II. Following are the lemmas that will be used in the proof:

**Lemma 1 (Piateskii-Shapiro, [2]).** Let $f(x)$ be a continuously
Let $g(x)$ be its inverse function. Then, for an integer $m$ such that $m = \lfloor f(n) \rfloor$, either $\{g(m)\} = 0$ or $1 - g'(m - 1) < \{g(m)\} < 1$. Conversely, if $\{g(m)\} = 0$ or $1 - g'(m + 1) < \{g(m)\} < 1$, then it follows that for some $n$, $m = \lfloor f(n) \rfloor$.

(The curly brackets indicate the fractional part of the real number, the straight brackets the integer part, as usual.)

**Lemma 2** (a theorem of Erdős-Turán, [1]). If $\mu_1, \mu_2, \ldots$ is a real sequence and if $D_N$ denotes its discrepancy modulo one, then for each integer $m \geq 1$ we have

$$ND_x \leq K \left( \frac{N}{m+1} + \sum_{t=1}^{m} \frac{1}{t} \left| \sum_{\tau=1}^{N} e(t \mu_\tau) \right| \right)$$

(where $K$ is a constant and $e(z) = e^{2\pi iz}$, as usual).

**Lemma 3** (Van der Corput, pg. 64, [3]). Let $g(x)$ be a real function with a continuous and steadily decreasing derivative $g'(x)$ in $(a, b)$, and let $g'(b) = \alpha, g'(a) = \beta$. Then

$$\sum_{\gamma < \nu < \beta + \gamma} e(\nu x) dx + O(\log(\beta - \alpha + 2))$$

where $\gamma$ is any positive constant less than one.

**Lemma 4** (Van der Corput, pg. 61, [3]). Let $F(x)$ be a real function, twice differentiable, and let $F''(x) \geq r > 0$, or $F''(x) \leq -r < 0$ throughout the interval $(a, b)$, then

$$\left| \int_{a}^{b} e^{tF(x)} dx \right| \leq \frac{8}{\sqrt{r}} .$$

III. The first part of the proof is aimed at establishing the uniform distribution modulo one and the discrepancy of that distribution for sequences $g(q)$ where $q$ are squarefree integers and $g(x)$ is the inverse function of a function in $S(c, \delta)$ (where $\delta$ is usually small, depending on $c$). The following is the result in this direction:

**Theorem 2.** For given $1 < c < 2$, and $\delta > 0$, small enough depending on $c$ alone, let $f(x) \in S(c, \delta)$ and let $g(x)$ be the inverse function of $f(x)$. Then the sequence $\{g(q): q \leq K, q$ the squarefree integers$\}$ is uniformly distributed modulo one and

$$N(K, \xi) = \xi Q(K) + Q(K)D_{q,K}(g) ,$$

and
differentiable function with \( f'(x) > 0, f''(x) \geq 0 \), for \( x \geq 1 \), and let
(7) \[ Q(K)D_Q(K)(g) \leq K^{\delta+\varepsilon\phi} + K^{1-\frac{1}{2\varepsilon\phi}} + K^{1-\frac{1}{2\varepsilon\phi}} \]
where \( Q(K) = \text{number of squarefree integers} \leq K, N(K, \xi) = \text{number of elements in the sequence} \ g(q), q \leq K, q \text{ squarefree}, \) which fall into a fixed interval of length \( \xi \) (< 1) modulo one, and \( D_{Q(K)}(g) \) is the discrepancy, modulo one, of the sequence \( g(q) \).

Clearly, uniform distribution holds whenever \( \delta > 0 \) is small enough to make the exponents in the estimate (7) less than one.

Proof. For \( h \geq 1 \), consider
(8) \[ T_h(K) = \sum_{q \leq K} e(hg(q)) \quad e(z) = e^{2\pi i z} \]
Suppose that \( K_o \) is the large value from where on the estimates of \( g, g', g'' \) induced by the definition 1 hold, and let \( K > K_o \), then
(8') \[ T_h(K) = \sum_{K_o \leq q \leq K} e(hg(q)) + O(K_o) \]
and
(9) \[ \sum_{K_o \leq q \leq K} e(hg(q)) = \sum_{K_o \leq q \leq K} e(hg(n)) \sum_{d^2 \mid (n, p^2)} \mu(d) \]
where \( p = \prod_{p \leq K^{1/2}} p, p = \text{primes, } (a, b) = \text{greatest common divisor}, \mu(d) = \text{Möbius function}. \) We can further write
\[
\sum_{d \leq K^{1/2}} \mu(d) \sum_{K_0 \leq d^2 \leq K/d^2} e(hg(d^2m))
\]
(10) \[ = \sum_{d \leq K^{1/2}} \mu(d) \sum_{K_0 \leq d^2 \leq K/d^2} e(hg(d^2m))
+ \sum_{A < d \leq K^{1/2}} \mu(d) \sum_{K_0 \leq d^2 \leq K/d^2} e(hg(d^2m)) \]
We will pick the value of \( A \) later. The second sum in (10) can be estimated trivially as
(11) \[ \sum_{A < d \leq K^{1/2}} \left| \sum_{K_0 \leq d^2 \leq K/d^2} e(hg(d^2m)) \right| \leq \sum_{A < d \leq K^{1/2}} K \leq K A. \]
The first sum, on the other hand, is estimated by
(12) \[ \sum_{d \leq A} \left| \sum_{K_0 \leq d^2 \leq K/d^2} e(hg(d^2m)) \right|. \]
To estimate the inner sum, divide the interval \( K_0/d^2 < m < K/d^2 \) up into pieces of type \( 1/2^r K/d^2 < m \leq 1/2^{r-1} K/d^2 \), to get
We will estimate the last inner sum by using Lemma 3 and then Lemma 4. The conditions in definition 1 give that
\[
\left( \frac{y}{a} \right)^{(e+\delta)-1} < g(y) \leq \left( \frac{y}{a} \right)^{(e+\delta)-1 - \frac{1}{e}}, \quad \frac{1}{c + \delta} \left( \frac{y}{a} \right)^{(e+\delta)-1} < g'(y) \leq \frac{1}{c} \left( \frac{y}{a} \right)^{(e+\delta)-1},
\]
the chain rule tells us that \((d/dx)g(d^2x) = [(d/dz)g(z)] \cdot d^2, z = d^2x, and so we have, by Lemma 3, for each \(r\)
\[
(13) \sum_{\nu \in \mathbb{N}} e(hg(d^2m)) = \sum_{\nu} I_{\nu} + E,
\]
where the \(\sum_{\nu}\) extends over \((1/\alpha \cdot (c + \delta))h((1/2^r - 1)K)^{(e+\delta)-1} d - 1/2 < \nu < (1/\alpha \cdot c)h((1/2^r)K)^{(e+\delta)-1} d^2 + 1/2, a_1 = a^{(e+\delta)-1}, a_2 = a^{1/e}\), and
\[
I_{\nu} = \int_{x^{-r+1}Kd^2}^{x^{-r}Kd^{-2}} e(hg(d^2x) - \nu x) dx,
\]
and
\[
E = O(\log (\max \nu - \min \nu + 2)).
\]
In (14), first we change variables to \(y = d^2x\), and then apply Lemma 4
\[
(14') I_{\nu} = \frac{1}{d^2} \int_{2^{-r+1}K}^{2^{-r}K} e(hg(y) - \nu y) dy
\]
but here \(d^2/dy^2(hg(y) - \nu y/d^2) \geq (c + \delta)^{-1}((c + \delta)^{-1} - 1)(1/\alpha)h^{(e+\delta)-1}.\) thus, we get, applying Lemma 4 that
\[
(15) I_{\nu} \ll \frac{1}{d^2} \left[ h \left( \frac{1}{2^r^{-1} K} \right)^{(e+\delta)-1} \right]^{-1/2}.
\]
We thus have for (12) the estimate:
\[
(16) \ll \sum_{d \in A} \sum_{r} \sum_{\nu} \frac{1}{d^2} h^{-1/2} 2^{2(r-1)(e+\delta)} c 2^{-r} K^{1-1/(2c+2d)} + \sum_{d \in A} \sum_{r} (E)
\]
(for largest \(r\) we might get a shorter range of integration in (14), but the upper bound estimates still clearly hold in (16)). where \(\sum_{\nu}\) is over
\[
\frac{1}{a(c + \delta)} h \left( \frac{1}{2^r^{-1} K} \right)^{(e+\delta)-1} d^2 - \frac{1}{2} < \nu < \frac{1}{ac} h \left( \frac{1}{2^r^{-1} K} \right)^{e-1} d^2 + \frac{1}{2}.
\]
From here, we have that the \(\nu\) summation is bounded by
\[
\ll hd^2 2^{r} \left( \frac{1}{2^r} \right)^{1/(e+\delta)} K^{r-1} + 1.
\]
and so we can further estimate (12) by

\[ (16') \quad \ll \sum_{d \neq 1} \sum_r \left( h d^2 \frac{2^r}{2^{r-1}} \frac{K^{\delta/(c+\delta)}}{r^{(c+\delta)-1}} + 1 \left( \frac{h^{-1/2}}{d^2} \frac{2^r}{2^{r-1}} \frac{K^{1-1/(2c+2\delta)}}{r^{(c+\delta)-1}} \right) \right) + \sum_{d \neq 1} \sum_r E \ll Ah^{1/2} K^{\delta/(c+\delta)-1} h^{-1/2} K^{1-1/(2c+2\delta)} + A \log K. \]

(The last step is because \( \sum_r \) in the first term was just a geometric sum and so it converges, while in the second part, the number of terms of the \( \sum_r \) is \( O(\log K) \).) The estimate \((16')\) together with \((11)\) now gives us that

\[ (17) \quad T_h(K) \ll Ah^{1/2} K^{\delta/(c+\delta)-1} + h^{-1/2} K^{1-1/(2c+2\delta)} + \frac{K}{A} + K_0. \]

Here, for \( K \) sufficiently large the last error term absorbs into the first one if \( A \geq 1 \) (which will anyway be the case). We now pick \( A \) so as to balance the 1st and 3rd terms of \((17)\), i.e. let \( A = \left[ K^{\delta/(c+\delta)-1} + (c+\delta)/(c+\delta) \right] \). With this choice we obtain

\[ (18) \quad T_h(K) \ll h^{1/4} K^{\delta/(c+\delta)-1} + h^{-1/2} K^{1-1/(2c+2\delta)}. \]

Finally we use Lemma 2 to write:

\[ (19) \quad Q(K)D_{Q(K)}(g) \ll \frac{K}{m+1} + \sum_{h=1}^m \frac{1}{h} \left( h^{1/4} K^{\delta/(c+\delta)-1} + h^{-1/2} K^{1-1/(2c+2\delta)} \right) \]

We pick the optimal \( m \), i.e. \( m = \left[ K^{\delta/(c+\delta)-1} + (c+\delta)/(c+\delta) \right] \), and thus we have

\[ (20) \quad Q(K)D_{Q(K)}(g) \ll K^{\delta/(c+\delta)-1} + K^{1-1/(2c+2\delta)}. \]

**COROLLARY 1.** If \( 1 < c < 4/3 \) then there exists \( \delta_* > 0 \) depending on \( c \) such that if \( f(x) \in S(c, \delta) \) for \( 0 < \delta < \delta_* \), and \( g(x) \) is its inverse function then:

\[ (21) \quad Q(K)D_{Q(K)}(g) \ll K^{(c+\delta)-1-\varepsilon} \ll (g(K))^{1-\varepsilon} \]

for some \( \varepsilon > 0 \), depending on \( c \) and \( \delta \).

**Proof.** All we need to show is that

\[ \frac{3}{5} + \frac{\delta}{5(c+\delta)} < \frac{1}{c+\delta} \quad \text{and} \quad 1 - \frac{1}{2(c+\delta)} < \frac{1}{c+\delta} \]

hold for some \( \delta > 0 \). By continuity it is enough to check that \( \frac{3}{5} + 1/5c < 1/c \) and \( 1 - 1/2c < 1/c \) hold. But the first of these holds if
c < 4/3, the second if c < 3/2.

IV. We can now prove Theorem 1. Let

\begin{equation}
T_s(x, y) = \text{number of } \{s_n = [f(n)], y < s_n \leq x, s_n = \text{squarefree}, n = 1, 2, 3, \ldots\}
\end{equation}

Clearly, \( T_s(1, y) = Q_s(y) \). Lemma 1 can now be used together with expressions (6) and (21). \( \xi \) in (6) will be taken \( g'(y + 1) \) or \( g'(x - 1) \) to give upper and lower bounds on \( T(x, y) \), where \( g(x) \) is as usual the inverse function of \( f(x) \). We obtain:

\begin{equation}
T_s(x, y) \begin{cases}
< g'(x - 1)(Q(y) - Q(x)) + O(y^{(\varepsilon + \delta)-1}) \\
> g'(y + 1)(Q(y) - Q(x)) + O(y^{(\varepsilon + \delta)-1})
\end{cases}
\end{equation}

where \( Q(x) = \# \text{ squarefree integers } \leq x \). Or

\begin{equation}
T_s(x, y) \begin{cases}
< g'(x)(Q(y) - Q(x)) + O(y^{(\varepsilon + \delta)-1}) + O(x^{(\varepsilon + \delta)-1}(Q(y) - Q(x))) \\
> g'(y)(Q(y) - Q(x)) + O(y^{(\varepsilon + \delta)-1}) + O(y^{(\varepsilon + \delta)-1}(Q(y) - Q(x)))
\end{cases}
\end{equation}

Thus, for \( 0 < \alpha < 1 \), using the well-known fact that \( Q(x) = 6x^{1/2} + O(x^{1/2}) \),

\begin{equation}
T_s(x, (1 + \alpha)x) = \begin{cases}
< \frac{6}{\pi^2} x \cdot \alpha \cdot g'(x) + O(x^{(\varepsilon + \delta)-1} + x^{1/2}) \\
> \frac{6}{\pi^2} x \cdot \alpha \cdot g'((1 + \alpha)x)
\end{cases}
\end{equation}

On the other hand, clearly

\begin{equation}
Q_s(x) = \sum_{k=1}^{L(x)} T_s\left(\frac{x}{(1 + \alpha)^k}, \frac{x}{(1 + \alpha)^{k-1}}\right) + O(1)
\end{equation}

holds for an appropriate function \( L(x) \) which tends to \( \infty \) for \( x \to \infty \), if \( \alpha = \alpha(x) > 0 \) is some given function of \( x \) which tends to zero as \( x \to \infty \) (the relation is \( (1 + \alpha(x))^{L(x)} \approx x \)).

Using (25) in the expression (26) we obtain

\begin{equation}
Q_s(x) = \begin{cases}
< \frac{6}{\pi^2} \sum_{k=1}^{L(x)} \frac{x \cdot \alpha}{(1 + \alpha)^k} g'\left(\frac{x}{(1 + \alpha)^k}\right) + O(L(x) \cdot x') \\
> \frac{6}{\pi^2} \sum_{k=1}^{L(x)} \frac{x \cdot \alpha}{(1 + \alpha)^k} g'\left(\frac{x}{(1 + \alpha)^{k-1}}\right) + O(L(x) \cdot x')
\end{cases}
\end{equation}

where \( \gamma = \max\{1/(c + \delta) - \varepsilon, 1/2\} \) and so it is actually \( 1/(c + \delta) - \varepsilon \). The main terms of the expressions on the right of (27) are exactly the upper and lower approximating sums of the Riemann integral.
To see how closely these sums approximate the integral, it suffices to find out how closely they are to each other, i.e. to estimate:

\[
\Delta(x) = \left| \sum_{k=1}^{L(x)} \frac{x \alpha}{(1 + \alpha)^k} g'\left(\frac{x}{(1 + \alpha)^k}\right) \right|
\]

\[
- \sum_{k=1}^{L(x)} \frac{x \cdot \alpha}{(1 + \alpha)^k} g'\left(\frac{x}{(1 + \alpha)^k}\right)
\]

\[
= \left| \sum_{k=1}^{L(x)-1} \frac{x \cdot \alpha}{(1 + \alpha)^{k+1}} g'\left(\frac{x}{(1 + \alpha)^k}\right) \right|
\]

\[
- \sum_{k=1}^{L(x)} \frac{x \cdot \alpha}{(1 + \alpha)^k} g'\left(\frac{x}{(1 + \alpha)^k}\right)
\]

\[
\leq \alpha \cdot \sum_{k=1}^{L(x)-1} \frac{x}{(1 + \alpha)^k} \left(\frac{1}{1 + \alpha} - 1\right) g'\left(\frac{x}{(1 + \alpha)^k}\right)
\]

\[
+ \left| x \cdot \alpha \cdot g'(x) \right| + O(\alpha)
\]

\[
\leq \left| \alpha \cdot \sum_{k=1}^{L(x)-1} \frac{x \cdot \alpha}{(1 + \alpha)^{k+1}} g'\left(\frac{x}{(1 + \alpha)^k}\right) \right|
\]

\[
+ O \left| x \cdot \alpha \cdot g'(x) \right| + O(\alpha)
\]

The last sum is now the lower estimating sum of the integral, so one can write for \( \alpha = \alpha(x) \)

\[
\Delta(x) \ll \alpha(x)(g(x) - g(1)) + \alpha(x) \cdot x \cdot g'(x)
\]

so

\[
\Delta(x) \ll \alpha(x)g(x) + O(1)
\]

Equation \((1 + \alpha(x))^{L(x)} \cong x\) gives us that \(\alpha(x)\) and \(L(x) = \lfloor \log x/\alpha(x) \rfloor\)

is a pair for which expression (26) holds; picking in particular \(\alpha(x) \approx (\log x(g(x))^{-\epsilon})^{1/2}\), gives

\[
\Delta(x) \ll \sqrt{\log x(g(x))^{1-\epsilon/2}}
\]

and

\[
L(x) \cdot x^\epsilon \ll L(x)(g(x))^{1-\epsilon} = \sqrt{\log x(g(x))^{1-\epsilon/2}}
\]

Calling \(\epsilon'\) some value \(0 < \epsilon' < \epsilon/2\) yields

\[
Q_\epsilon(x) = \frac{6}{\pi^2} g(x) + O((g(x)^{1-\epsilon'}))
\]

I take this opportunity to express my appreciation to Professor Harold N. Shapiro for his continued advice and valuable suggestions.
REFERENCES


Received June 30, 1975. This work was partially supported by the National Science Foundation, Grant NSF-GP-33019X.

Columbia University, New York City
Aharon Atzmon, A moment problem for positive measures on the unit disc .......... 317
Peter W. Bates and Grant Bernard Gustafson, Green’s function inequalities for two-point boundary value problems .................................................. 327
Howard Edwin Bell, Infinite subrings of infinite rings and near-rings .............. 345
Grahame Bennett, Victor Wayne Goodman and Charles Michael Newman, Norms of random matrices .......................................................... 359
Beverly L. Brechner, Almost periodic homeomorphisms of $E^2$ are periodic .......... 367
Beverly L. Brechner and R. Daniel Mauldin, Homeomorphisms of the plane ....... 375
Jia-Arng Chao, Lusin area functions on local fields ..................................... 383
Frank Rimi DeMeyer, The Brauer group of polynomial rings ........................ 391
M. V. Deshpande, Collectively compact sets and the ergodic theory of semi-groups ......................................................................................... 399
Raymond Frank Dickman and Jack Ray Porter, $\theta$-closed subsets of Hausdorff spaces ................................................................. 407
Charles P. Downey, Classification of singular integrals over a local field ............ 417
Daniel Reuven Farkas, Miscellany on Bieberbach group algebras ...................... 427
Peter A. Fowler, Infimum and domination principles in vector lattices ............. 437
Barry J. Gardner, Some aspects of $T$-nilpotence. II: Lifting properties over $T$-nilpotent ideals ................................................................. 445
Gary Fred Gruenhage and Phillip Lee Zenor, Metrization of spaces with countable large basis dimension ...................................................... 455
J. L. Hickman, Reducing series of ordinals ....................................................... 461
Hugh M. Hilden, Generators for two groups related to the braid group ............. 475
Tom (Roy Thomas Jr.) Jacob, Some matrix transformations on analytic sequence spaces ...................................................................................... 487
Elyahu Katz, Free products in the category of $k_w$-groups .................................... 493
Tsang Hai Kuo, On conjugate Banach spaces with the Radon-Nikodým property ... 497
Norman Eugene Liden, $K$-spaces, their antispaces and related mappings ............ 505
Clinton M. Petty, Radon partitions in real linear spaces .................................... 515
Alan Saleski, A conditional entropy for the space of pseudo-Menger maps .......... 525
Michael Singer, Elementary solutions of differential equations ......................... 535
Eugene Spiegel and Allan Trojan, On semi-simple group algebras. I ..................... 549
Charles Madison Stanton, Bounded analytic functions on a class of open Riemann surfaces ............................................................................. 557
Sherman K. Stein, Transversals of Latin squares and their generalizations .......... 567
Ivan Ernest Stux, Distribution of squarefree integers in non-linear sequences ........ 577
Lowell G. Sweet, On homogeneous algebras .................................................. 585
Lowell G. Sweet, On doubly homogeneous algebras ........................................ 595
Florian Vasilescu, The closed range modulus of operators ................................. 599
Arthur Anthony Yanushka, A characterization of the symplectic groups $\text{PSp}(2m,q)$ as rank 3 permutation groups ................................................. 611
James Juei-Chin Yeh, Inversion of conditional Wiener integrals ....................... 623