Primes and squares with preassigned digits

Cathy Swaenepoel

IMJ-PRG, Université Paris Cité.

Numeration 2023,
Liège, May 25.

Digits

Representation of an integer $k \geqslant 0$ in base $g \geqslant 2$:

$$k = \sum_{j \geqslant 0} \varepsilon_j(k) g^j$$

where $\varepsilon_j(k) \in \{0, \ldots, g-1\}$ is the **digit** of k at the position j.

independence?

base g expansion $\longleftrightarrow$ multiplicative representation (as a product of prime factors)

The study of the independence between the additive and the multiplicative structure of the integers is one of the most important topics in number theory.

Sum of digits of primes and polynomial values in AP

$s_g(k)$ = sum of digits of k in base g.

Gelfond's problems (1968): estimate, as $x \to +\infty$,

$$|\{p \leqslant x,\ s_g(p) \equiv a \bmod m\}| \quad \text{and} \quad |\{n \leqslant x,\ s_g(P(n)) \equiv a \bmod m\}| \quad (\deg P \geqslant 2).$$

- For primes: solved by Mauduit–Rivat (2010) in any base.
- For polynomials P of degree 2: solved by Mauduit–Rivat (2009) in any base.
- For polynomials P of degree $\geqslant 3$:
  - solved by Drmota–Mauduit–Rivat (2011) in all large enough prime bases,
  - lower bounds in any base by Dartyge–Tenenbaum (2006) and Stoll (2012),
  - still open in small bases.

Primes and polynomial values with restricted digits

Let $d \in \{0, \ldots, g-1\}$.

Problem: estimate the number of primes or polynomial values with no digit d.

- For almost primes: lower bounds by Dartyge–Mauduit (2000, 2001).
- For primes :
  - solved by Maynard (2021) in any large enough base,
  - solved by Maynard (2019) in base 10 (lower and upper bounds of the same order of magnitude).
- For polynomials P of degree $\geqslant 2$: solved by Maynard (2021) in any large enough base.

Integers with preassigned digits

Representation of an integer $k \in [0, g^n[$ in base $g \geqslant 2$:

$$k = \sum_{j=0}^{n-1} \varepsilon_j(k) g^j, \qquad 0 \leqslant \varepsilon_j \leqslant g-1.$$

- $A \subset \{0, \ldots, n-1\}$: set of positions,
- $\boldsymbol{d} = (d_j)_{j \in A}$: preassigned digits at these positions.

	d_{n-2}					d_6		d_4			d_1	
$n-1$	$n-2$					6		4			1	0

$$|\{k < g^n : \forall j \in A,\ \varepsilon_j(k) = d_j\}| = g^{n-|A|}$$

Integers with preassigned digits

Representation of an integer $k \in [0, g^n[$ in base $g \geqslant 2$:

$$k = \sum_{j=0}^{n-1} \varepsilon_j(k) g^j, \qquad 0 \leqslant \varepsilon_j \leqslant g-1.$$

- $A \subset \{0, \dots, n-1\}$: set of positions,
- $\boldsymbol{d} = (d_j)_{j \in A}$: preassigned digits at these positions.

	d_{n-2}					d_6		d_4			d_1	
$n-1$	$n-2$					6		4			1	0

$$\Big| \underbrace{\{k < g^n : \forall j \in A,\ \varepsilon_j(k) = d_j\}}_{\text{sparse set}} \Big| = g^{n-|A|}$$

if $|A| \to +\infty$ as $n \to +\infty$

General problem

Problem: For an interesting subset $E \subset \mathbb{N}$, estimate

$$|\{k < g^n : k \in E, \forall j \in A, \varepsilon_j(k) = d_j\}|$$

for A as large as possible.

Expected estimate (as $n \to +\infty$) ?
If the digits of the integers of E are expected to be "random" then this should be about

$$|\{k < g^n : k \in E\}| \cdot \frac{1}{g^{|A|}} \sim \begin{cases} \dfrac{g^{n-|A|}}{\log g^n} & \text{for } E = \text{ primes,} \\ g^{\frac{n}{2}-|A|} & \text{for } E = \text{ squares.} \end{cases}$$

(Recall that $\dfrac{|\{k < g^n : \forall j \in A, \varepsilon_j(k) = d_j\}|}{g^n} = \dfrac{1}{g^{|A|}}$.)

Prime numbers with preassigned digits

Historical background

Goal: estimate $|\{p < g^n : \forall j \in A,\ \varepsilon_j(p) = d_j\}|$ as $n \to +\infty$.

- **Kátai (1986)**.
- **Wolke (2005)**: asymptotic, $|A| \leqslant 2$
  ($|A| \leqslant (1-\varepsilon)\sqrt{n}$ under GRH).
- **Harman (2006)**: lower bound, $|A| \leqslant$ constant.
- **Harman-Kátai (2008)**: asymptotic, $|A| \ll \sqrt{n}(\log n)^{-1}$.
- **Bourgain (2013)**: asymptotic, $|A| \ll n^{4/7}(\log n)^{-4/7}$, in base 2.
- **Bourgain (2015)**: asymptotic, $|A| \leqslant cn$, in base 2 ($c > 0$ absolute constant).

Result

Theorem 1 (S. 2020)

For any $g \geqslant 2$, there exist an explicit $c = c(g) \in]0,1[$ and $\delta = \delta(g) > 0$ such that for any $n \geqslant 1$, for any $A \subset \{0, \ldots, n-1\}$ satisfying $\{0, n-1\} \subset A$ and

$$|A| \leqslant cn,$$

for any $(d_j)_{j \in A} \in \{0, \ldots, g-1\}^A$ such that $(d_0, g) = 1$ and $d_{n-1} \geqslant 1$, we have

$$|\{p < g^n : \forall j \in A,\ \varepsilon_j(p) = d_j\}| = \frac{g^{n-|A|}}{\log g^n} \frac{g}{\varphi(g)} \left(1 + O_g\left(n^{-\delta}\right)\right).$$

This generalizes Bourgain's result (2015) to any base.

Theorem 1 holds with $c(g)$ given by

g	2	3	4	5	10	10^3	2^{200}
$c(g) \cdot 10^2$	0.21	0.31	0.36	0.40	0.47	0.68	0.90

Assuming GRH, Theorem 1 holds with $c(g)$ given by

g	2	3	4	5	10	10^3	2^{200}
$c(g) \cdot 10^2$	1.6	2.4	2.9	3.1	3.7	5.2	6.9

Squares with preassigned digits

Squares vs primes

Denote $\mathcal{S} = \{\ell^2, \ell \geqslant 0\}$ the set of squares.

Goal: estimate $|\mathcal{S} \cap \{k < g^n : \forall j \in A, \varepsilon_j(k) = d_j\}|$ as $n \to +\infty$.

- Squares are a priori easier to handle than primes (distribution in short intervals, in arithmetic progressions, ...).

But

- squares are sparser than primes,
- there are algebraic constraints on the digits of squares.

$\to$ New difficulties for squares.

Special case where $A = \{0, n-1\}$

Denoting $N_0 = d_{n-1} g^{n-1}$, $N_1 = (d_{n-1}+1) g^{n-1}$, we have

$$\begin{aligned}
&|\mathcal{S} \cap \{k < g^n : \varepsilon_0(k) = d_0,\ \varepsilon_{n-1}(k) = d_{n-1}\}| \\
&= |\{N_0 \leqslant k < N_1 : k \in \mathcal{S},\ k \equiv d_0 \bmod g\}| \\
&= |\{\sqrt{N_0} \leqslant \ell < \sqrt{N_1} : \ell^2 \equiv d_0 \bmod g\}| \\
&= R(g, d_0) \left(\sqrt{d_{n-1}+1} - \sqrt{d_{n-1}} \right) g^{\frac{n-3}{2}} (1 + o(1)) \quad (n \to +\infty)
\end{aligned}$$

where

$$R(g, d_0) = \text{ number of square roots of } d_0 \text{ modulo } g.$$

A special case where $A = \{\frac{n}{2} - 1\}$

Gross–Vacca (1968): in base 2, for any n divisible by 4, for $j = \frac{n}{2} - 1$,

$$|\mathcal{S} \cap \{k < 2^n : \varepsilon_j(k) = 1\}| = 2^{\frac{n}{2}-1}\left(1 - 2^{-\frac{n}{4}+1}\right) = 2^{\frac{n}{2}-1}(1 + o(1)).$$

Also Prodinger–Wagner (2009) and Preparata–Vacca (2012).

Bassily–Kátai (1996): Let $g \geqslant 2$ and $\delta > 0$.
For any $A \subset \{0, \ldots, n-1\}$ and $\boldsymbol{d} = (d_j)_{j \in A} \in \{0, \ldots, g-1\}^A$ such that

$$|A| \leqslant \log n$$

and

$$n^{1/3} \leqslant \min A \leqslant \max A \leqslant n - n^{1/3},$$

we have

$$|\mathcal{S} \cap \{k < g^n : \forall j \in A,\ \varepsilon_j(k) = d_j\}| = g^{\frac{n}{2} - |A|}(1 + O_{g,\delta}(n^{-\delta})).$$

Hypothesis $\mathcal{H}$ on the preassigned digits

$v_2(g) =$ 2-adic valuation of g.

- If g is odd or $v_2(g) \geqslant 3$,

$$\mathcal{H}(g): \quad \{0\} \subset A,\ (d_0, g) = 1,\ d_0 \text{ square mod } g.$$

- If $v_2(g) = 2$,

$$\mathcal{H}(g): \quad \{0, 1\} \subset A,\ (d_0, g) = 1,\ d_1 g + d_0 \text{ square mod } g^2.$$

- If $v_2(g) = 1$ (e.g. $g = 2$ or $g = 10$),

$$\mathcal{H}(g): \quad \{0, 1, 2\} \subset A,\ (d_0, g) = 1,\ d_2 g^2 + d_1 g + d_0 \text{ square mod } g^3.$$

Hypothesis $\mathcal{H}$ on the preassigned digits

$v_2(g)$ = 2-adic valuation of g.

- If g is odd or $v_2(g) \geqslant 3$,

$$\mathcal{H}(g): \quad \{0\} \subset A,\ (d_0, g) = 1,\ d_0 \text{ square mod } g.$$

- If $v_2(g) = 2$,

$$\mathcal{H}(g): \quad \{0,1\} \subset A,\ (d_0, g) = 1,\ d_1 g + d_0 \text{ square mod } g^2.$$

- If $v_2(g) = 1$ (e.g. $g = 2$ or $g = 10$),

$$\mathcal{H}(g): \quad \{0,1,2\} \subset A,\ (d_0, g) = 1,\ d_2 g^2 + d_1 g + d_0 \text{ square mod } g^3.$$

Under $\mathcal{H}(g)$, we have for any $k \geqslant 0$,

$$(\forall j \in A,\ \varepsilon_j(k) = d_j) \Rightarrow k \text{ is a square modulo any power of } g.$$

Result

Theorem 2 (S. 2023+)

For any $g \geqslant 2$, there exist an explicit $c = c(g) \in]0, 1/2[$ and $\delta = \delta(g) > 0$ such that for any $n \geqslant 3$, for any $A \subset \{0, \dots, n-1\}$ and $\mathbf{d} = (d_j)_{j \in A} \in \{0, \dots, g-1\}^A$ satisfying $\mathcal{H}(g)$, $n-1 \in A$, $d_{n-1} \geqslant 1$ and

$$|A| \leqslant cn,$$

we have

$$|\mathcal{S} \cap \{k < g^n : \forall j \in A,\, \varepsilon_j(k) = d_j\}| = \mathfrak{S}(g, n, A, \mathbf{d}) \left(1 + O_g\left(n^{-\delta}\right)\right)$$

where

$$\mathfrak{S}(g, n, A, \mathbf{d}) = \sum_{\substack{k < g^n \\ \forall j \in A,\, \varepsilon_j(k) = d_j}} \frac{\eta(g)}{2\sqrt{k}}, \quad \eta(g) = \begin{cases} 2^{\omega(g)}, & g \text{ odd}, \\ 2^{\omega(g)+1}, & g \text{ even}. \end{cases}$$

In particular, the order of magnitude of $|\mathcal{S} \cap \{k < g^n : \forall j \in A,\, \varepsilon_j(k) = d_j\}|$ is $g^{\frac{n}{2} - |A|}$.

Explicit values of c

Theorem 2 holds with $c(g)$ given by

g	2	3	4	5	10	16	2^{32}	2^{64}
$c(g) \cdot 10^2$	0.5	0.9	1.1	1.3	1.6	1.8	3.6	4

An example where $(d_0, g) > 1$ $(g = 10,\ d_0 = 5)$

Lemma (S.)

Let m *such that* $\frac{n}{4} - m \to +\infty$ *as* $n \to +\infty$. *Choose*

$$A = \{0, 2, 4, \ldots, 2(m-1), n-1\}.$$

Let s *such that* $s \equiv 1 \bmod 8$ *and* $s \equiv 0 \bmod 5^{2m-1}$ *and let* $d \in \{0, \ldots, 9\}$. *Choose*

$$d_{2i} = \varepsilon_{2i}(s) \text{ for } i = 0, \ldots, m-1, \quad d_{n-1} = d.$$

Then we have

$$|\mathcal{S} \cap \{k < 10^n : \forall j \in A,\ \varepsilon_j(k) = d_j\}| = \frac{C(d)}{2^{|A|}} 10^{\frac{n}{2} - |A|} (1 + o(1))$$

where $C(d) > 0$ *depends only on* d.

So **the order of magnitude may be smaller** than $10^{\frac{n}{2} - |A|}$.

Idea: at the positions $1, 3, \ldots, 2m-3$, the digits of k have to be the digits of s.

Notations for the proof of Theorem 2

- $\mathrm{e}(x) = \exp(2i\pi x)$, $x \in \mathbb{R}$.
- $\mathcal{S} = \{\ell^2,\ \ell \geqslant 0\}$ the set of squares.
- $\mathcal{D}(n, A, \boldsymbol{d}) = \{k < g^n : \forall j \in A,\ \varepsilon_j(k) = d_j\}$.
- $N = g^n$.

We want to estimate

$$\sum_{N_0 \leqslant k < N_1} \mathbf{1}_{\mathcal{S}}(k) \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)$$

where $N_0 = d_{n-1} g^{n-1}$ and $N_1 = (d_{n-1} + 1) g^{n-1}$.

Method

Use the **circle method**:

$$\sum_{N_0 \leqslant k < N_1} \mathbf{1}_{\mathcal{S}}(k)\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k) = \int_0^1 S(\alpha)\overline{R(\alpha)}d\alpha$$

where

$$\underbrace{S(\alpha) = \sum_{N_0 \leqslant k < N_1} \mathbf{1}_{\mathcal{S}}(k)\,\mathrm{e}(k\alpha)}_{\substack{\text{can be large only when } \alpha \text{ is close to} \\ \text{a rational with small denominator} \\ \text{i.e. } \alpha \text{ is in a major arc}}} \quad \text{and} \quad \underbrace{R(\alpha) = \sum_{N_0 \leqslant k < N_1} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)\,\mathrm{e}(k\alpha)}_{\text{depends on the digital conditions}}.$$

- integral over major arcs $\to$ main term ($+$ error term)
- integral over minor arcs $\to$ error term

Fourier transform of $\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}$

$$F_n(\alpha) = \frac{1}{g^{n-|A|}} \sum_{k<g^n} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)\, \mathrm{e}(k\alpha) = \frac{1}{g^{n-|A|}} R(\alpha).$$

By writing k in base g, we obtain:

$$|F_n(\alpha)| = \prod_{\substack{0\leqslant j\leqslant n-1 \\ j\notin A}} \frac{\Phi_g\left(g^j\alpha\right)}{g} \qquad \text{where } \Phi_g(t) = \left|\sum_{v=0}^{g-1} \mathrm{e}(vt)\right| = \left|\frac{\sin \pi g t}{\sin \pi t}\right|.$$

For $g = 2$,

$$|F_n(\alpha)| = \prod_{\substack{0\leqslant j\leqslant n-1 \\ j\notin A}} \left|\cos \pi 2^j \alpha\right|.$$

We need very strong upper bounds for $\|F_n\|_1$ and some (weighted) averages of $|F_n(a/q)|$.

Using a strong bound for some weighted average of $|F_n(a/q)|$, we obtain:

Proposition (S. 2020)

Let $0 < \varepsilon < 1/4$ *and* $0 < c < 2\varepsilon$. *If* $|A| \leqslant cn$ *then*

$$\sum_{\substack{q \leqslant Q \\ (q,g)=1}} \max_{r \in \mathbb{Z}} \left| \sum_{\substack{k < g^n \\ k \equiv r \bmod q}} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k) - \frac{g^{n-|A|}}{q} \right| \ll_{\varepsilon,c} g^{n-|A|}\, n \left(\frac{\log^3 n}{n} \right)^{\frac{2\varepsilon}{c}-1}$$

where $Q = g^{n\left(\frac{1}{4}-\varepsilon\right)}$.

On average over all $q \leqslant Q$ such that $(q,g)=1$, the integers $k < g^n$ such that

$$\forall j \in A,\ \varepsilon_j(k) = d_j$$

are well distributed in arithmetic progressions modulo q (if $|A|$ is small enough).

$B_1 \leqslant B$ "small" powers of $N = g^n$ with $B_1 = o(B)$.

- Major arcs:

$$\mathfrak{M} = \bigcup_{1 \leqslant q \leqslant B_1} \bigcup_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} \mathfrak{M}(q,a)$$

where $\mathfrak{M}(q,a)$ is the interval $\left|\alpha - \frac{a}{q}\right| \leqslant \frac{B}{qN}$ modulo 1.

- Minor arcs:

$$\mathfrak{m} = [0,1[\setminus \mathfrak{M}.$$

Minor arcs contribution

$$\int_{\mathfrak{m}} \left| S(\alpha)\overline{R(\alpha)} \right| d\alpha = g^{n-|A|} \int_{\mathfrak{m}} \left| S(\alpha)\overline{F_n(\alpha)} \right| d\alpha \leqslant g^{n-|A|} \, \|F_n\|_1 \sup_{\alpha\in\mathfrak{m}} |S(\alpha)|$$

- Use the strong upper bound:

$$\|F_n\|_1 \ll N^{\xi-1} \log N \qquad \text{(trivial: } 1\text{)}$$

where ξ is explicit and $\xi \to 0$ as $|A|/n \to 0$.

- Use a classical estimate on Weyl sums to bound $|S(\alpha)|$ over the minor arcs:

$$\sup_{\alpha\in\mathfrak{m}} |S(\alpha)| = \sup_{\alpha\in\mathfrak{m}} \left| \sum_{\sqrt{N_0}\leqslant \ell < \sqrt{N_1}} \mathrm{e}(\ell^2\alpha) \right| \ll \frac{\sqrt{N}}{\sqrt{B_1}}. \qquad \text{(trivial: } \sqrt{N}\text{)}$$

- This gives

$$\int_{\mathfrak{m}} \left| S(\alpha)\overline{R(\alpha)} \right| d\alpha \ll g^{\frac{n}{2}-|A|} \frac{N^{\xi}\log N}{\sqrt{B_1}}.$$

Major arcs contribution

$$\int_{\mathfrak{M}} S(\alpha)\overline{R(\alpha)}d\alpha = \sum_{1\leqslant q\leqslant B_1} \sum_{\substack{1\leqslant a\leqslant q \\ (a,q)=1}} \int_{\left|\alpha-\frac{a}{q}\right|\leqslant\frac{B}{qN}} S(\alpha)\overline{R(\alpha)}d\alpha$$

First step: replace the indicator function of the interval $\left|\alpha - \frac{a}{q}\right| \leqslant \frac{B}{qN}$ by a well chosen smooth function:

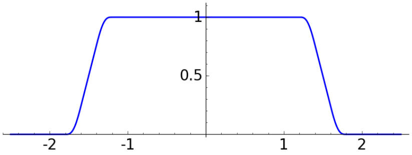
$$\alpha \mapsto w\left(\frac{qN}{B}\left(\alpha - \frac{a}{q}\right)\right).$$

This creates an error term which is bounded by $\int_{\mathfrak{m}} \left|S(\alpha)\overline{R(\alpha)}\right| d\alpha$.

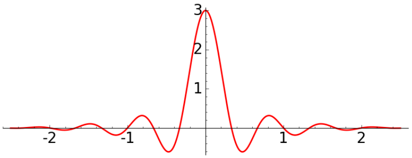
Function w

Using a construction of Ingham or Iwaniec, one can construct a function w such that:

- $0 \leqslant w \leqslant 1$,
- $w = 1$ on $[-1, 1]$,
- $\operatorname{supp} w \subset [-2, 2]$,
- $w \in \mathcal{C}^\infty(\mathbb{R})$,
- $\widehat{w}(y) = O\left(e^{-|y|^{1/2}}\right)$ for any $y \in \mathbb{R}$.



Graph of w



Graph of $\widehat{w}$

Contribution of the major arc around a/q

We want to estimate the "contribution of the major arc around a/q":

$$\int_{\mathbb{R}} w\left(\tfrac{qN}{B}\left(\alpha - \tfrac{a}{q}\right)\right) S(\alpha)\overline{R(\alpha)}\, d\alpha$$

$$= \int_{\mathbb{R}} w\left(\tfrac{qN}{B}\left(\alpha - \tfrac{a}{q}\right)\right) \sum_{N_0 \leqslant k_1 < N_1} \mathbf{1}_{\mathcal{S}}(k_1)\, \mathrm{e}\left(k_1\alpha\right) \sum_{N_0 \leqslant k_2 < N_1} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k_2)\, \mathrm{e}\left(-k_2\alpha\right) d\alpha$$

$$= \sum_{N_0 \leqslant k_2 < N_1} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k_2)\, \mathrm{e}\left(\frac{-k_2 a}{q}\right) \sum_{r=0}^{q-1} \mathrm{e}\left(\frac{ra}{q}\right) \sum_{\substack{N_0 \leqslant k_1 < N_1 \\ k_1 \equiv r \bmod q}} \mathbf{1}_{\mathcal{S}}(k_1)\, \tfrac{B}{qN}\, \widehat{w}\left((k_2 - k_1)\tfrac{B}{qN}\right).$$

Contribution of the major arc around a/q

Up to admissible errors,

$$\sum_{\substack{N_0 \leqslant k_1 < N_1 \\ k_1 \equiv r \bmod q}} \mathbf{1}_{\mathcal{S}}(k_1) \frac{B}{qN} \widehat{w}\left((k_2 - k_1)\frac{B}{qN}\right)$$

$$\downarrow$$

$$\frac{R(q,r)}{q} \int_{N_0}^{N_1} \frac{B}{qN} \widehat{w}\left((k_2 - t)\frac{B}{qN}\right) \frac{dt}{2\sqrt{t}}$$

$$\downarrow$$

$$\frac{R(q,r)}{q} \frac{1}{2\sqrt{k_2}}.$$

- partial summation
- estimate for the number of squares in arithmetic progressions ($R(q,r) =$ number of square roots of r mod q)

- size of $\widehat{w}$ at infinity
- Fourier inversion

Contribution of the major arc around a/q

Up to an admissible error, the contribution of the major arc around a/q is

$$\sum_{N_0 \leqslant k_2 < N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k_2)}{2\sqrt{k_2}} \, \mathrm{e}\left(\frac{-k_2 a}{q}\right) \sum_{r=0}^{q-1} \mathrm{e}\left(\frac{ra}{q}\right) \frac{R(q,r)}{q}$$

$$= \sum_{N_0 \leqslant k < N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)}{2\sqrt{k}} \, \mathrm{e}\left(\frac{-ka}{q}\right) \frac{G(q,a)}{q}$$

where $G(q,a)$ is the quadratic Gauss sum:

$$G(q,a) = \sum_{u=1}^{q} \mathrm{e}\left(\frac{au^2}{q}\right).$$

Contribution of all major arcs around a/q, q fixed

Up to an admissible error, the contribution of all major arcs around a/q (q fixed) is

$$\mathcal{C}(q) := \sum_{N_0 \leqslant k < N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)}{2\sqrt{k}} H(q,k)$$

where

$$H(q,k) = \frac{1}{q} \sum_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} G(q,a)\, \mathrm{e}\left(\frac{-ka}{q}\right) = \sum_{d \mid q} \mu(d) R\left(\frac{q}{d}, k\right) \in \mathbb{Z}.$$

- $q \mapsto H(q,k)$ is multiplicative.
- For any k such that $\left(\frac{k}{p}\right) = 1$, we have

$$H(p,k) = 1, \qquad H(p^\nu, k) = 0 \text{ for any } \nu \geqslant 2.$$

Contribution of all major arcs around a/q, q fixed

For simplicity, we assume here that the base g is a prime $p \geqslant 3$.

Write $q = p^\nu q'$ where $p \nmid q'$.

Three cases depending on ν and q' (under the hypothesis $\mathcal{H}(g)$):

1. If $\nu \geqslant 2$ then $\mathcal{C}(q) = 0$.
2. If $\nu \in \{0,1\}$ and $q' = 1$ (i.e $q = 1$ or $q = p$) then
$$\mathcal{C}(q) = \sum_{N_0 \leqslant k < N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)}{2\sqrt{k}}.$$
This gives the **main term**.
3. If $\nu \in \{0,1\}$ and $q' \geqslant 2$ then
$$\mathcal{C}(q) = \sum_{N_0 \leqslant k < N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)}{2\sqrt{k}} H(q', k).$$
We show that this is small on average over $q' \geqslant 2$ with $(q', g) = 1$ (see below). This gives an **error term**.

Contribution of all major arcs around a/q, q fixed, third case

We want to prove that

$$\sum_{\substack{2\leqslant q'\leqslant B_1\\(q',g)=1}} \left| \sum_{N_0\leqslant k<N_1} \frac{\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)}{\sqrt{k}} H(q',k) \right| = o(g^{\frac{n}{2}-|A|}).$$

After using the upper bound $|G(q',a)| \ll \sqrt{q'}$ and a partial summation, it suffices to show that

$$\sum_{\substack{2\leqslant q'\leqslant B_1\\(q',g)=1}} \frac{1}{\sqrt{q'}} \sum_{\substack{1\leqslant a\leqslant q'\\(a,q')=1}} \max_{0<t\leqslant g^n} \underbrace{\left| \frac{1}{g^{n-|A|}} \sum_{k<t} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)\, \mathrm{e}\left(\frac{ak}{q'}\right) \right|}_{=\begin{cases} |\text{FT of } \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})} \text{ at } a/q'| & \text{if } t = g^n \\ \text{“incomplete sum”} & \text{otherwise} \end{cases}} = o(1).$$

A weighted average of $|F_n(a/q)|$

To handle the "complete sums", we use:

Lemma (S. 2020)

Let $0 < c < \frac{1}{8}$. *If* $|A| \leqslant cn$ *then*

$$\sum_{\substack{2 \leqslant q \leqslant Q \\ (q,g)=1}} \frac{1}{\sqrt{q}} \sum_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} \left| F_n\left(\frac{a}{q}\right) \right| \ll_c \left(\frac{\log^3 n}{n}\right)^{\frac{1}{8c}-1}$$

where $Q = g^{\frac{n}{8}}$ *and* F_n *is the Fourier transform of* $\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}$.

How to handle the "incomplete sums"?

- For a good choice of m, write $[0,t[$ as the disjoint union of intervals of the form $[\ell g^m, (\ell+1)g^m[$ and at most one interval of length $< g^m$.

$$\max_{0<t\leqslant g^n} \left| \frac{1}{g^{n-|A|}} \sum_{k<t} \mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k)\,\mathrm{e}\,(k\alpha) \right| \leqslant \underbrace{\left| \frac{1}{g^{m-|A'|}} \sum_{h<g^m} \mathbf{1}_{\mathcal{D}(m,A',\boldsymbol{d'})}(h)\,\mathrm{e}\,(h\alpha) \right|}_{|\text{FT of } \mathbf{1}_{\mathcal{D}(m,A',\boldsymbol{d'})} \text{ at } \alpha|} + g^{m+|A|-n}$$

where $A' = A \cap \{0,\ldots,m-1\}$ and $\boldsymbol{d'} = (d_j)_{j\in A'} \in \{0,\ldots,g-1\}^{A'}$.

- Apply the previous bound for the Fourier transform.

Conclusion of the proof

With a good choice of the parameters B_1 and B and taking c sufficiently small, we get

$$\sum_{k<g^n} \mathbf{1}_{\mathcal{S}}(k)\mathbf{1}_{\mathcal{D}(n,A,\boldsymbol{d})}(k) = \mathfrak{S}(g,n,A,\boldsymbol{d})\left(1+O_g\left(n^{-\delta}\right)\right)$$

for some $\delta > 0$, where

$$\mathfrak{S}(g,n,A,\boldsymbol{d}) = \sum_{\substack{k<g^n \\ \forall j\in A,\, \varepsilon_j(k)=d_j}} \frac{\eta(g)}{2\sqrt{k}}, \quad \eta(g) = \begin{cases} 2^{\omega(g)}, & g \text{ odd}, \\ 2^{\omega(g)+1}, & g \text{ even}. \end{cases}$$

The main term comes from the major arcs around a/q with

- $q \in \{1, p\}$ if g is a prime $p \geqslant 3$,
- $q \in \{1, 4, 8\}$ if $g = 2$,
- $q \in \{1, 4, 5, 8, 20, 40\}$ if $g = 10$.

Conclusion

- In any base $g \geqslant 2$, we obtain an asymptotic formula for the number of squares with a positive proportion of preassigned digits.
- We give explicit values for the proportion of digits this method allows us to preassign.

Conclusion

- In any base $g \geqslant 2$, we obtain an asymptotic formula for the number of squares with a positive proportion of preassigned digits.
- We give explicit values for the proportion of digits this method allows us to preassign.

Thank you for your attention!