Primes and squares with preassigned digits

Cathy Swaenepoel

IMJ-PRG, Université Paris Cité.

Numeration 2023,
Liège, May 25.
Representation of an integer $k \geq 0$ in base $g \geq 2$:

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

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

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.
\( s_g(k) = \text{sum of digits of } k \text{ in base } g. \)

**Gelfond’s problems** (1968): estimate, as \( x \to +\infty, \)

\[
|\{ p \leq x, \ s_g(p) \equiv a \mod m \}| \quad \text{and} \quad |\{ n \leq x, \ s_g(P(n)) \equiv a \mod m \}| \quad (\deg P \geq 2).
\]

- For polynomials \( P \) of degree \( \geq 3 \):
  - 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.
Let $d \in \{0, \ldots, g - 1\}$.

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

- 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 $\geq 2$: solved by Maynard (2021) in any large enough base.
Integers with preassigned digits

Representation of an integer \( k \in [0, g^n] \) in base \( g \geq 2 \):

\[
k = \sum_{j=0}^{n-1} \varepsilon_j(k) g^j, \quad 0 \leq \varepsilon_j \leq g - 1.
\]

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

\[
\begin{array}{cccccc}
  & d_{n-2} &  &  & d_6 & d_4 & d_1 \\
\hline
n-1 & n-2 &  &  & 6 & 4 & 1 \\
\end{array}
\]

\[
|\{k < g^n : \forall j \in A, \varepsilon_j(k) = d_j\}| = g^{n-|A|}
\]
Integers with preassigned digits

Representation of an integer \( k \in [0, g^n] \) in base \( g \geq 2 \):

\[
k = \sum_{j=0}^{n-1} \varepsilon_j(k)g^j, \quad 0 \leq \varepsilon_j \leq g - 1.
\]

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

|  |  |  |  |  |  |
|---|---|---|---|---|
|  |  |  |  |  |
| \( n-1 \) | \( n-2 \) | 6 | 4 | 1 0 |

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

sparse set
if \( |A| \to +\infty \) as \( n \to +\infty \)
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

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

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

Cathy Swaenepoel
Primes and squares with preassigned digits

6 / 37
Prime numbers with preassigned digits
Goal: estimate $|\{p < g^n : \forall j \in A, \varepsilon_j(p) = d_j\}|$ as $n \to +\infty$.

- **Kátai (1986).**

- **Wolke (2005):** asymptotic, $|A| \leq 2$ \hspace{1cm} ($|A| \leq (1 - \varepsilon)\sqrt{n}$ under GRH).

- **Harman (2006):** lower bound, $|A| \leq \text{constant}$.

- **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, $|A| \leq cn$, in base 2 ($c > 0$ absolute constant).
Theorem 1 (S. 2020)

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

\[
|A| \leq cn,
\]

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

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

This generalizes Bourgain’s result (2015) to any base.
Explicit values of $c$

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

<table>
<thead>
<tr>
<th>$g$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>$10^3$</th>
<th>$2^{200}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c(g) \cdot 10^2$</td>
<td>0.21</td>
<td>0.31</td>
<td>0.36</td>
<td>0.40</td>
<td>0.47</td>
<td>0.68</td>
<td>0.90</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>$g$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>$10^3$</th>
<th>$2^{200}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c(g) \cdot 10^2$</td>
<td>1.6</td>
<td>2.4</td>
<td>2.9</td>
<td>3.1</td>
<td>3.7</td>
<td>5.2</td>
<td>6.9</td>
</tr>
</tbody>
</table>
Squares with preassigned digits
Denote $S = \{ \ell^2, \ell \geq 0 \}$ the set of squares.

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

→ New difficulties for squares.
Denoting $N_0 = d_{n-1} g^{n-1}$, $N_1 = (d_{n-1} + 1) g^{n-1}$, we have

$$|\mathcal{S} \cap \{ k < g^n : \varepsilon_0(k) = d_0, \varepsilon_{n-1}(k) = d_{n-1} \}|$$

$$= |\{ N_0 \leq k < N_1 : k \in \mathcal{S}, \; k \equiv d_0 \mod g \}|$$

$$= |\{ \sqrt{N_0} \leq \ell < \sqrt{N_1} : \ell^2 \equiv d_0 \mod g \}|$$

$$= R(g, d_0) \left( \sqrt{d_{n-1} + 1} - \sqrt{d_{n-1}} \right) g^{\frac{n-3}{2}} (1 + o(1)) \quad (n \to +\infty)$$

where

$$R(g, d_0) = \text{number of square roots of } d_0 \mod g.$$
A special case where $A = \left\{ \frac{n}{2} - 1 \right\}$

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

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

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

$$|A| \leq \log n$$

and

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

we have

$$|S \cap \{k < g^n : \forall j \in A, \varepsilon_j(k) = d_j\}| = g^{n-|A|/2}(1 + O_{g,\delta}(n^{-\delta})).$$
Hypothesis $\mathcal{H}$ on the preassigned digits

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

- If $g$ is odd or $v_2(g) \geq 3$,
  
  $\mathcal{H}(g) : \{0\} \subset A, (d_0, g) = 1, d_0$ square mod $g$.

- If $v_2(g) = 2$,
  
  $\mathcal{H}(g) : \{0, 1\} \subset A, (d_0, g) = 1, d_1g + d_0$ square mod $g^2$.

- If $v_2(g) = 1$ (e.g. $g = 2$ or $g = 10$),
  
  $\mathcal{H}(g) : \{0, 1, 2\} \subset A, (d_0, g) = 1, d_2g^2 + d_1g + d_0$ square mod $g^3$. 
Hypothesis $\mathcal{H}$ on the preassigned digits

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

- If $g$ is odd or $v_2(g) \geq 3$,
  \[ \mathcal{H}(g) : \{0\} \subset A, (d_0, g) = 1, d_0 \text{ square mod } g. \]

- If $v_2(g) = 2$,
  \[ \mathcal{H}(g) : \{0, 1\} \subset A, (d_0, g) = 1, d_1g + d_0 \text{ square mod } g^2. \]

- If $v_2(g) = 1$ (e.g. $g = 2$ or $g = 10$),
  \[ \mathcal{H}(g) : \{0, 1, 2\} \subset A, (d_0, g) = 1, d_2g^2 + d_1g + d_0 \text{ square mod } g^3. \]

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

\[ (\forall j \in A, \varepsilon_j(k) = d_j) \Rightarrow k \text{ is a square modulo any power of } g. \]
Theorem 2 (S. 2023+)

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

$$|A| \leq cn,$$

we have

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

where

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

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

<table>
<thead>
<tr>
<th>$g$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>16</th>
<th>$2^{32}$</th>
<th>$2^{64}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c(g) \cdot 10^2$</td>
<td>0.5</td>
<td>0.9</td>
<td>1.1</td>
<td>1.3</td>
<td>1.6</td>
<td>1.8</td>
<td>3.6</td>
<td>4</td>
</tr>
</tbody>
</table>
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 \(s \equiv 1 \mod 8\) and \(s \equiv 0 \mod 5^{2m-1}\) 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

- \( e(x) = \exp(2i\pi x), \ x \in \mathbb{R} \).

- \( S = \{\ell^2, \ \ell \geq 0\} \) the set of squares.

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

- \( N = g^n \).

We want to estimate

\[
\sum_{N_0 \leq k < N_1} 1_S(k) 1_{D(n, A, d)}(k)
\]

where \( N_0 = d_{n-1} g^{n-1} \) and \( N_1 = (d_{n-1} + 1) g^{n-1} \).
Use the **circle method**:

\[ \sum_{N_0 \leq k < N_1} 1_{S(k)} 1_{\mathcal{D}(n,A,d)}(k) = \int_0^1 S(\alpha) \overline{R(\alpha)} d\alpha \]

where

\[ S(\alpha) = \sum_{N_0 \leq k < N_1} 1_{S(k)} e(k\alpha) \]  
\[ R(\alpha) = \sum_{N_0 \leq k < N_1} 1_{\mathcal{D}(n,A,d)}(k) e(k\alpha) \]

- Can be large only when \( \alpha \) is close to a rational with small denominator, i.e. \( \alpha \) is in a major arc.
- Depends on the digital conditions.

- Integral over major arcs \( \rightarrow \) main term (+ error term)
- Integral over minor arcs \( \rightarrow \) error term
Fourier transform of $1_{D(n,A,d)}$

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

By writing $k$ in base $g$, we obtain:

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

For $g = 2$,

$$|F_n(\alpha)| = \prod_{0 \leq j \leq n-1 \atop j \notin A} |\cos \pi 2^j \alpha|.$$ 

We need very strong upper bounds for $\|F_n\|_1$ and some (weighted) averages of $|F_n(a/q)|$. 
Proposition (S. 2020)

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

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

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

On average over all $q \leq Q$ 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).
$B_1 \leq B$ “small” powers of $N = g^n$ with $B_1 = o(B)$.

- **Major arcs:**

\[
M = \bigcup_{1 \leq q \leq B_1} \bigcup_{1 \leq a \leq q} M(q, a)
\]

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

- **Minor arcs:**

\[
m = [0, 1] \setminus M.
\]
Minor arcs contribution

\[ \int_m \left| S(\alpha)R(\alpha) \right| \, d\alpha = g^{n-|A|} \int_m \left| S(\alpha)F_n(\alpha) \right| \, d\alpha \leq g^{n-|A|} \| F_n \|_1 \sup_{\alpha \in m} \left| S(\alpha) \right| \]

- Use the strong upper bound:
  \[ \| F_n \|_1 \ll N^{\xi - 1} \log N \quad \text{(trivial: 1)} \]
  where \( \xi \) 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 m} |S(\alpha)| = \sup_{\alpha \in m} \left| \sum_{\sqrt{N_0} \leq \ell < \sqrt{N_1}} e(\ell^2 \alpha) \right| \ll \frac{\sqrt{N}}{\sqrt{B_1}}. \quad \text{(trivial: \( \sqrt{N} \))} \]

- This gives
  \[ \int_m \left| S(\alpha)R(\alpha) \right| \, d\alpha \ll g^{\frac{n}{2}-|A|} \frac{N^\xi \log N}{\sqrt{B_1}}. \]
Major arcs contribution

\[ \int_{\mathfrak{m}} S(\alpha) \overline{R(\alpha)} d\alpha = \sum_{1 \leq q \leq B_1} \sum_{1 \leq a \leq q \atop (a,q)=1} \int_{|\alpha - \frac{a}{q}| \leq \frac{B}{qN}} S(\alpha) \overline{R(\alpha)} d\alpha \]

First step: replace the indicator function of the interval \( |\alpha - \frac{a}{q}| \leq \frac{B}{qN} \) 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 \( \int_{\mathfrak{m}} |S(\alpha) \overline{R(\alpha)}| \, d\alpha \).
Using a construction of Ingham or Iwaniec, one can construct a function $w$ such that:

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

Graph of $w$

Graph of $\hat{w}$
We want to estimate the "contribution of the major arc around $a/q$":

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

$$= \int_{\mathbb{R}} w \left( \frac{qN}{B} \left( \alpha - \frac{a}{q} \right) \right) \sum_{N_0 \leq k_1 < N_1} 1_S(k_1) \, e(k_1 \alpha) \sum_{N_0 \leq k_2 < N_1} 1_{D(n,A,d)}(k_2) \, e(-k_2 \alpha) \, d\alpha$$

$$= \sum_{N_0 \leq k_2 < N_1} 1_{D(n,A,d)}(k_2) \, e \left( \frac{-k_2 a}{q} \right) \sum_{r=0}^{q-1} e \left( \frac{ra}{q} \right) \sum_{N_0 \leq k_1 < N_1} 1_S(k_1) \frac{B}{qN} \hat{w} \left( (k_2 - k_1) \frac{B}{qN} \right).$$
Contribution of the major arc around $a/q$

Up to admissible errors,

$$
\sum_{N_0 \leq k_1 < N_1} 1_{S(k_1)} \frac{B}{qN} \hat{w} \left( (k_2 - k_1) \frac{B}{qN} \right)
$$

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

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

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

$$
\frac{R(q, r)}{q} \frac{1}{2\sqrt{k_2}}.
$$
Contribution of the major arc around $a/q$

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

$$\sum_{N_0 \leq k_2 < N_1} \frac{1_{D(n,A,d)}(k_2)}{2\sqrt{k_2}} e\left(\frac{-k_2a}{q}\right) \sum_{r=0}^{q-1} e\left(\frac{ra}{q}\right) \frac{R(q,r)}{q}$$

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

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

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

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

where

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

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

$$H(p,k) = 1, \quad H(p^\nu, k) = 0 \text{ for any } \nu \geq 2.$$
Contribution of all major arcs around $\alpha/q$, $q$ fixed

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

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

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

1. If $\nu \geq 2$ then $C(q) = 0$.
2. If $\nu \in \{0, 1\}$ and $q' = 1$ (i.e. $q = 1$ or $q = p$) then
   \[
   C(q) = \sum_{N_0 \leq k < N_1} \frac{1_{D(n,A,d)}(k)}{2\sqrt{k}}.
   \]
   This gives the **main term**.
3. If $\nu \in \{0, 1\}$ and $q' \geq 2$ then
   \[
   C(q) = \sum_{N_0 \leq k < N_1} \frac{1_{D(n,A,d)}(k)}{2\sqrt{k}} H(q', k).
   \]
   We show that this is small on average over $q' \geq 2$ with $(q', g) = 1$ (see below).
   This gives an **error term**.
We want to prove that
\[
\sum_{2 \leq q' \leq B_1 \atop (q',g)=1} \sum_{N_0 \leq k < N_1} \frac{1_{\mathcal{D}(n,A,d)}(k)}{\sqrt{k}} H(q',k) \right|_{(q',g)=1} = o(g^{\frac{n}{2}} - |A|).
\]

After using the upper bound $|G(q',a)| \ll \sqrt{q'}$ and a partial summation, it suffices to show that
\[
\sum_{2 \leq q' \leq B_1 \atop (q',g)=1} \frac{1}{\sqrt{q'}} \sum_{1 \leq a \leq q' \atop (a,q')=1} \max_{0 < t \leq g^n} \left| \frac{1}{g^{n-|A|}} \sum_{k < t} 1_{\mathcal{D}(n,A,d)}(k) e\left(\frac{ak}{q'}\right) \right|_{(a,q')=1} = o(1).
\]

\[
= \begin{cases} 
|\text{FT of } 1_{\mathcal{D}(n,A,d)} \text{ at } a/q'| & \text{if } t = g^n \\
\text{“incomplete sum”} & \text{otherwise}
\end{cases}
\]
To handle the “complete sums”, we use:

**Lemma (S. 2020)**

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

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

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

\[
\max_{0 < t \leq g^n} \left| \frac{1}{g^n - |A|} \sum_{k < t} 1_D(n, A, d)(k) e(k\alpha) \right|
\]

\[
\leq \left| \frac{1}{g^m - |A'|} \sum_{h < g^m} 1_D(m, A', d')(h) e(h\alpha) \right| + g^m + |A| - n
\]

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

- 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

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

for some $\delta > 0$, where

$$\mathcal{S}(g, n, A, d) = \sum_{k<g^n} \frac{\eta(g)}{2 \sqrt{k}}, \quad \eta(g) = \begin{cases} 2 \omega(g), & g \text{ odd}, \\ 2 \omega(g)+1, & g \text{ even}. \end{cases}$$

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

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