On Stolarsky's conjecture: The sum of digits of n and n^h

Thomas Stoll

(IML, Marseille)

Aussois

(joint work with K. Hare and S. Laishram)

April 7, 2010

1. Introduction
   - Stolarsky’s conjecture
   - A computational example
2. Sum of digits of polynomial values
   - Some context for the conjecture
3. Extremal orders of $s_q(p(n))/s_q(n)$
   - Upper extremal order
   - Lower extremal order
4. A “local” Diophantine problem
5. Open questions

K. Stolarksy, “The binary digits of a power”, Proc. of the AMS, vol. 71, no. 1, 1978.

Let $h \geq 2$ be fixed and denote by $B(n)$ the number of 1’s in the binary expansion of integers.

Is it true that

$$\liminf_{n\to\infty} \frac{B(n^h)}{B(n)} = 0?$$

For $h = 3$ the smallest n such that $B(n) > B(n^3)$ is

$$407182835067 = (1011110110011011111110011111101011\,11011)_2$$

$$407182835067^3 = (1101000000001000000010001010010001\,1100000011000010010011000000000000\,0101000000110000000101000001000001\,000010000000011)_2$$

Hence

$$B(407182835067) = 29, \qquad B(407182835067^3) = 28.$$

Let $q \geq 2$ and denote by $s_q(n)$ the sum of digits in base q of n, i.e.,

$$s_q(n) = \sum_{j \geq 0} \varepsilon_j$$

where $n = \sum_{j\geq 0} \varepsilon_j q^j$ with $\varepsilon_j \in \{0, 1, \ldots, q-1\}$.

What can be said about the distribution of $s_q(p(n))$, where $p(n)$ is some integer-valued polynomial of degree $h \geq 2$?

- mean value and other parameters,
- distribution in arithmetic progressions,
- distribution modulo 1,
- extremal orders,
- limit distributions etc.

Distribution in arithmetic progressions

Theorem (Mauduit/Rivat (2009), Dartyge/Tenenbaum (2006))

Let $r \geq 2$ and put $m = (r, q-1)$.

1. *For all $a \in \mathbb{Z}$,*

$$\#\{n < N : s_q(n^2) \equiv a \bmod r\} = \frac{N}{r} Q(a,m) + o(N),$$

where $Q(a,m) = \#\{0 \leq n < m : n^2 \equiv a \bmod r\}$.

2. *If $m = 1$ then for N sufficiently large,*

$$\#\{n < N : s_q(p(n)) \equiv a \bmod r\} \geq CN^{\min\left(1, \frac{2}{h!}\right)}.$$

Mean value

Theorem (Davenport/Erdős (1952), Delange (1975), Peter (2002) etc.)

As $N \to \infty$,

$$\sum_{n<N} s_q(n) \sim \frac{1}{h} \sum_{n<N} s_q(n^h) \sim \frac{q-1}{2} N \log_q N.$$

In other words, an average power n^h has sum of digits the average digit times the length of the expansion.

Denote

$$p(x) = a_h x^h + a_{h-1} x^{h-1} + \ldots + a_0 \in \mathbb{Z}[x]$$

of degree $h \geq 2$ and $a_h \geq 1$.

Theorem (Lindström (1997))

We have

$$\limsup_{n\to\infty} \frac{s_2(p(n))}{\log_2 n} = h.$$

Other constructions for $p(n) = n^2$ were given by M. Drmota and J. Rivat (2005) with constructions due to J. Cassaigne and G. Baron.

Stolarsky considered the case $q = 2$ and $p(n) = n^h$.

Theorem (Stolarsky (1978))

Let $h \geq 1$. Then for all $n \geq 2$,

$$\frac{s_2(n^h)}{s_2(n)} \leq 2(h \log_2 n)^{1-1/h}.$$

This is best possible in that there is a constant $C > 0$ depending only on h such that

$$\frac{s_2(n^h)}{s_2(n)} > C(\log_2 n)^{1-1/h}$$

infinitely often.

Theorem (K. Hare, S. Laishram, T. Stoll (2010))

(1) *If $p(n)$ has only nonnegative coefficients then there exists C_1 such that for all $n \geq 2$,*

$$\frac{s_q(p(n))}{s_q(n)} \leq C_1 (\log_q n)^{1-1/h}.$$

This is best possible in that there is a constant C_1' such that

$$\frac{s_q(p(n))}{s_q(n)} > C_1' (\log_q n)^{1-1/h}$$

infinitely often.

Theorem (contin.)

(2) *If $p(n)$ has at least one negative coefficient then there exists C_2 and N_0 such that for all $n \geq N_0$,*

$$\frac{s_q(p(n))}{s_q(n)} \leq C_2 \log_q n.$$

This is best possible in that for all $\varepsilon > 0$ we have

$$\frac{s_q(p(n))}{s_q(n)} > (q - 1 - \varepsilon) \log_q n$$

infinitely often.

The proof relies on a notable result in additive number theory.

Theorem (Bose/Chowla (1962/63))

Let $h \geq 2$. Then there are infinitely many integers M for which there exists integers $a_1, \ldots, a_{M+1}$ such that

$$1 \leq a_1 < a_2 < \cdots < a_{M+1} = M^h,$$

while every sum of the form

$$a_{j_1} + \cdots + a_{j_h}, \qquad 1 \leq j_1 \leq \cdots \leq j_h \leq M+1$$

is distinct.

Theorem (Stolarsky (1978))

For all $n > 1$ we have

$$\frac{s_2(n^2)}{s_2(n)} \geq \frac{1}{\lfloor \log_2 n \rfloor + 1}.$$

On the other hand, there are infinitely many integers n such that

$$\frac{s_2(n^2)}{s_2(n)} \leq \frac{4(\log_2 \log_2 n)^2}{\log_2 n}.$$

Conjecture: For every $h \geq 2$ we have

$$\liminf_{n\to\infty} \frac{s_2(n^h)}{s_2(n)} = 0.$$

Theorem (K. Hare, S. Laishram, T. Stoll (2010))

There exist explicitly computable constants C_1 and C_2, dependent only on $p(x)$ and q, such that for all ε with $0 < \varepsilon < h(4h+1)$ there exists an $n < C_1 \cdot C_2^{1/\varepsilon}$ with

$$\frac{s_q(p(n))}{s_q(n)} < \varepsilon.$$

Corollary

There exists a constant C_3, dependent only on $p(x)$ and q, such that there exist infinitely many n with

$$\frac{s_q(p(n))}{s_q(n)} \leq \frac{C_3}{\log n}.$$

Corollary

1. *For any* $\varepsilon > 0$ *there exists an explicitly computable* $\alpha > 0$*, dependent only on* ε*,* $p(x)$ *and* q*, such that*

$$\#\left\{n < N : \quad \frac{s_q(p(n))}{s_q(n)} < \varepsilon\right\} \gg N^{\alpha}.$$

2. *There exists an explicitly computable* $\gamma > 0$*, dependent only on* q *and* $p(x)$*, such that*

$$\#\left\{n < N : \quad |s_q(p(n)) - s_q(n)| \leq \frac{q-1}{2}\right\} \gg N^{\gamma}.$$

3. *We have* $\quad \#\{n < N : \quad s_2(n^2) = s_2(n)\} \gg N^{1/19}.$

Theorem (K. Hare, S. Laishram, T. Stoll (2010))

1. *Let*

$$A_k = \#\{n \text{ odd} : \quad s_2(n^2) = s_2(n) = k\}.$$

   (i) *If* $k \leq 8$ *then* $A_k < \infty$.
   (ii) *If* $k \geq 16$ *or* $k \in \{12, 13\}$ *then* $A_k = \infty$.

2. *Let* $q \geq 3$ *and assume* $k \geq 94(q-1)$. *Then the equation*

$$s_q(n^2) = s_q(n) = k$$

has infinitely many solutions in n *with* $q \nmid n$ *if and only if*

$$k(k-1) \equiv 0 \mod (q-1).$$

1. Let $k \in \{9, 10, 11, 14, 15\}$. Is the set

$$\{n \text{ odd} : \quad s_2(n^2) = s_2(n) = k\}$$

finite or infinite?

2. Is $(s_2(n^2)/s_2(n))_{n\geq 1}$ dense in $\mathbb{R}$?
3. A second conjecture of Stolarsky (1978): As $N \to \infty$,

$$\frac{1}{N}\sum_{n=1}^{N}\frac{s_2(n^h)}{s_2(n)} \to h',$$

where $1 < h' \leq h$.

4. Generalizations: block counting functions, other numeration systems, quasi-polynomials etc.