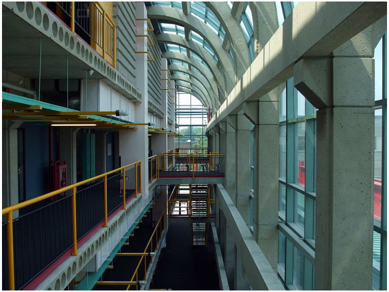
New Results in Additive Number Theory via Automata Theory and Combinatorics on Words

Jeffrey O. Shallit
(Joint work with Jean-Paul Allouche and Jason Bell)

School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1
Canada
shallit@uwaterloo.ca
https://cs.uwaterloo.ca/~shallit/

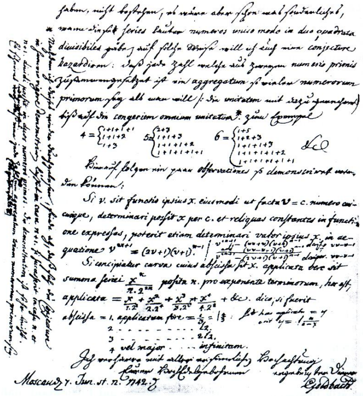


Davis Centre, U. Waterloo

Additive number theory

Additive number theory is the study of the additive properties of integers.

Probably the most famous example is *Goldbach's conjecture* from 1742: every even number ≥ 4 is the sum of two primes.



Goldbach letter to Euler
June 7 1742

Additive number theory

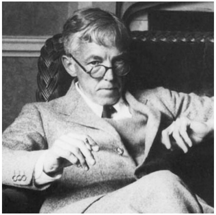
Less famous is the existence of an asymptotic formula that conjecturally predicts the *number* $G_2(n)$ of representations of n as the sum of two primes, due to Hardy and Littlewood in 1923:

$$G_2(n) \approx 2 \cdot \Pi_2 \cdot \left(\prod_{\substack{p|n \\ p \geq 3}} \frac{p-1}{p-2} \right) \frac{n}{(\log n)^2}$$

for n even, where

$$\Pi_2 = \prod_{p \geq 3} \left(1 - \frac{1}{(p-1)^2} \right) \doteq 0.66016$$

is the twin-prime constant.



G. H. Hardy

J. E. Littlewood

Additive number theory

So, given a set S, number theorists are interested in both

- *which* numbers are representable as sums of elements of S, and
- the *number* of such representations.

In this talk I focus on the second: the number of representations.

Let $A \subseteq \mathbb{N} = \{0, 1, 2, \ldots\}$ be a subset of the natural numbers. We define

$$\begin{aligned} r(k, A, n) &:= |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k\}| \\ r_<(k, A, n) &:= |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k, \\ &\qquad a_1 < a_2 < \cdots < a_k\}| \\ r_\leq(k, A, n) &:= |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k, \\ &\qquad a_1 \leq a_2 \leq \cdots \leq a_k\}|. \end{aligned}$$

These functions were originally studied by Erdős, Turán, and co-authors starting in the 1940's.

Motivation for r: powers of power series

$$r(k, A, n) := |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k\}|$$

$r(k, A, n)$ has a nice interpretation in terms of the coefficients of a power series.

Given a set A, we can define its associated *characteristic sequence* $(a(n))_{n \geq 0}$ as follows:

$$a(n) = \begin{cases} 1, & \text{if } n \in A; \\ 0, & \text{otherwise.} \end{cases}$$

And we can define its associated *power series*:

$$A(X) = \sum_{n \geq 0} a(n) X^n.$$

Then $r(k, A, n)$ is just the coefficient of X^n in $A(X)^k$.

Example: Goldbach representations

Take $A = \{2, 3, 5, \ldots\}$ to be the prime numbers.

Then $A(X) = X^2 + X^3 + X^5 + \cdots$ and Goldbach's conjecture can be restated as *the coefficients of* X^{2n} *in*

$$A(X)^2 = X^4 + 2X^5 + X^6 + 2X^7 + 2X^8 + 2X^9 + 3X^{10} + 2X^{12} + \cdots$$

are all positive for $n \geq 2$.

Result of Lambek and Moser

Let $\mathcal{E} = \{0, 3, 5, 6, 9, 10, \ldots\}$ be the evil numbers (number of 1-bits in the binary representation of n is even) and $\mathcal{O} = \{1, 2, 4, 7, 8, 11, \ldots\}$ be the odious numbers (number of 1-bits is odd).

Lambek and Moser (1959) proved the following theorem:

$$r_<(2, \mathcal{E}, n) = r_<(2, \mathcal{O}, n)$$

for $n \geq 0$.

An example of the theorem: the representations of 9 as sums of $\mathcal{E}$ are $(0, 9)$ and $(3, 6)$. The representations as sums of $\mathcal{O}$ are $(1, 8)$ and $(2, 7)$.

This theorem was later proved again by Dombi (2002), Lev (2004), and others.

Joachim "Jim" Lambek

Leo Moser

Detour: linear representations

A *linear representation* for a sequence $(f(n))_{n\geq 0}$ is a triple (v, γ, w), where

- v is a t-element row vector;
- γ is a $t \times t$-matrix-valued morphism;
- w is a t-element column vector

and

$$f(n) = v\,\gamma(x)\,w$$

whenever x is the base-b representation of n.

Here $\gamma(x) = \gamma(a_1)\cdots\gamma(a_i)$ if $x = a_1 \cdots a_i$.

The integer t is called the *rank* of the representation.

Example of a linear representation

Here is a linear representation for the Stern sequence $a(n)$, defined by $a(2n) = a(n)$ and $a(2n+1) = a(n) + a(n+1)$, with initial values $a(0) = 0$ and $a(1) = 1$:

$$v^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad \gamma(0) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}; \quad \gamma(1) = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}; \quad w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For example, let's compute $a(27)$. Express 27 in base 2 as 11011. Then

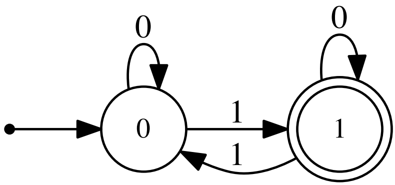
$$a(27) = v\gamma(11011)w = v\gamma(1)\gamma(1)\gamma(0)\gamma(1)\gamma(1)w$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} -5 & 8 \\ -7 & 11 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 8.$$

Automatic sets

A set A is said to be *b-automatic* if there is a finite automaton that recognizes exactly the set of base-b representations of members of A.

For example, the set $\mathcal{O}$ of odious numbers is 2-automatic, and recognized by the following automaton.



To use it, start it state 0, read the representation of n in base 2 and follow the arrows, accept iff you end up at state 1.

Computing linear representations

Theorem

Let $A \subseteq \mathbb{N}$ be an automatic set (i.e., an automaton recognizes representations of A in some representation system, such as base b).

Then $r(k, A, n)$ (resp., $r_<(k, A, n)$; $r_\leq(k, A, n)$) has a linear representation that can be computed directly from the automaton for A.

Proof.

By a theorem of Büchi-Bruyère, it suffices to write first-order logical formulas for $r(k, A, n)$ (resp., $r_<(k, A, n)$; $r_\leq(k, A, n)$). But these are given by the definitions of these functions. □

J. Richard Büchi

Véronique Bruyère

Comparing linear representations

If we have a linear representation (v_f, γ_f, w_f) for $f(n)$ and a linear representation (v_g, γ_g, w_g) for $g(n)$, we can form a linear representation (v, γ, w) for the linear combination

$$\alpha f(n) + \beta g(n)$$

by using block matrices, as follows:

$$v = [\alpha v_f \quad \beta v_g]$$

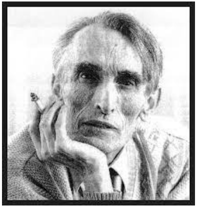
$$\gamma(a) = \left[\begin{array}{cc} \gamma_f(a) & \mathbf{0} \\ \mathbf{0} & \gamma_g(a) \end{array} \right]$$

$$w = \left[\begin{array}{c} w_f \\ w_g \end{array} \right].$$

Comparing linear representations

Furthermore, if we have a linear representation (v, γ, w) there is an algorithm, due to Schützenberger, for finding an equivalent linear representation of minimum rank.

Putting these two ideas together, we have the following theorem:



M.-P. Schützenberger

Theorem

Given a linear representation (v_f, γ_f, w_f) for $f(n)$ and a linear representation (v_g, γ_g, w_g) for $g(n)$, it is decidable if $f(n) = g(n)$ for all n.

Proof.

Form the linear representation for $f(n) - g(n)$, and then minimize it. Then $f(n) = g(n)$ for all n iff the linear representation is of rank 0 computing the 0 function. □

Lambek and Moser: proof via Walnut

To prove the Lambek-Moser result that

$$r_<(2, \mathcal{E}, n) = r_<(2, \mathcal{O}, n)$$

for $n \geq 0$, we just need to find a linear representation for both sides and then use the theorem on the previous slide.

This can be done using the `Walnut` software package as follows:

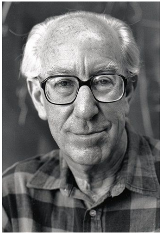
```
def evil_sum n "T[i]=@0 & T[j]=@0 & i<j & n=i+j":
def odious_sum n "T[i]=@1 & T[j]=@1 & i<j & n=i+j":
```

Here `T[i]` is `Walnut`'s way of writing the Thue-Morse sequence, t_i, the parity of the number of 1-bits of i.

These create two linear representations of rank 8, and we can use the ideas above to demonstrate they compute the same function.

The Rudin-Shapiro set

Let $\mathcal{R} = \{3, 6, 11, 12, 13, 15, \ldots\}$ be the Rudin-Shapiro set: the numbers n where the number of 11's (possibly overlapping) in the binary expansion of n is odd.



Walter Rudin

Harold S. Shapiro

Dombi (2002) proved that for $k \geq 5$, the function $r(k, \mathcal{R}, n)$ is an eventually increasing function of n.

He conjectured this is also true for $k = 4$, but no proof is known.

The Rudin-Shapiro set

We can prove that $r(3, \mathcal{R}, n)$ is *not* eventually increasing as follows.

The first step is to create a linear representation for the difference sequence

$$d(n) := r(3, \mathcal{R}, n) - r(3, \mathcal{R}, n-1).$$

We can do that with the following `Walnut` code:

```
def rudin3 n "RS[i]=@1 & RS[j]=@1 & RS[k]=@1 & n=i+j+k":
def rudin3m1 n "RS[i]=@1 & RS[j]=@1 & RS[k]=@1 & n=i+j+k+1":
```

and then combine them with the block matrix trick to get a linear representation (v, γ, w) for $d(n)$.

The goal is to find infinitely many n such that $d(n) < 0$.

Closed forms for linear representations along subsequences

In general a function $f(n)$ given by a linear representation (v, γ, w) will not have a simply-describable behavior.

However, we can always obtain a formula for f evaluated at a *subsequence* $(n_i)_i$ for which the base-b representation is of the form

$$x \overbrace{yy \cdots y}^{i \text{ copies}} z$$

where x, y, z are strings of digits.

This is because

$$v\,\gamma(n_i)\,w = v\,\gamma(x)\,\gamma(y)^i\,\gamma(z)\,w,$$

and each entry of $\gamma(y)^i$ can be expressed as a linear combination of the i'th powers of the zeros of the minimal polynomial of $\gamma(y)$.

We can then solve for the coefficients of this linear combination from the first few values of f, giving an exact closed-form formula for f.

The Rudin-Shapiro set

The n that we choose have a base-2 representation of the form

$$z_t := \overbrace{10\,10\,\cdots\,10}^{t+1 \text{ copies}} = (2^{2t+3} - 2)/3.$$

Now $\gamma(10)$ has minimal polynomial

$$X^2(X-1)(X-2)(X-4)(X^3-5X^2+12X-16)(X^4-13X^3+72X^2-196X+256)$$

and hence there exist constants

$$a, b, c, \alpha, \gamma,\ \alpha_i, \gamma_i\ (1 \le i \le 2), \beta_i, \delta_i\ (1 \le i \le 4)$$

such that

$$d(z_t) = a + b \cdot 2^t + c \cdot 4^t + \alpha_1 \gamma_1^t + \alpha_2 \gamma_2^t + \alpha_\gamma^t + \beta_1 \delta_1^t + \beta_2 \delta_2^t + \zeta_1 \eta_1^t + \zeta_2 \eta_2^t$$

where $\gamma, \gamma_1, \gamma_2$ are the zeros of $X^3 - 5X^2 + 12X - 16$ and the δ_i, η_i are the zeros of $X^4 - 13X^3 + 72X^2 - 196X + 256$.

The Rudin-Shapiro set

Here the α_i are complex conjugates, as are the γ_i, the β_i, the δ_i, the ζ_i, and the η_i.

Using Maple we can find the estimates

$$\text{(zeros of } X^3 - 5X^2 + 12X - 16) \quad \begin{cases} |\gamma_1|, |\gamma_2| & \doteq 2.41114 \\ \gamma & \doteq 2.75217 \end{cases}$$

$$\text{(zeros of } X^4 - 13X^3 + 72X^2 - 196X + 256) \quad \begin{cases} |\delta_1|, |\delta_2| & \doteq 4.88015 \\ |\eta_1|, |\eta_2| & \doteq 3.27859 \end{cases}$$

The dominant roots are clearly the δ_i and the corresponding coefficients are

$$\beta_1 \doteq -.03881 + .00706i$$
$$\beta_2 \doteq -.03881 - .00706i$$

The Rudin-Shapiro set

For t large enough, then, the value of $d(z_t)$ is dominated by

$$\beta_1 \delta_1^t + \beta_2 \delta_2^t = 2\Re(\beta_1 \delta_1^t),$$

which is large and negative when (say)

$$3\pi/4 < \arg(\beta_1 \delta_1^t) = (\arg(\beta_1) + t \arg(\delta_1)) \bmod 2\pi < 5\pi/4.$$

Since $\beta_1/|\beta_1|$ is not a root of unity, this will occur for infinitely many t.

Hence the difference function $d(z_t) < 0$ infinitely often.

Hence $r(3, \mathcal{R}, n)$ is not eventually increasing.

Powers of Thue-Morse power series

We can also study powers of the Thue-Morse power series

$$T(X) := \sum_{n \geq 0} t_n X^n = X + X^2 + X^4 + X^7 + X^8 + \cdots .$$

Allouche recently proved, using complex analysis and following ideas of Dombi, that the coefficients of $T^{10}(X)$ are eventually increasing.

Jean-Paul Allouche

Powers of Thue-Morse power series

More precisely, suppose $(q_n)_{n \geq 0}$ is a sequence of ± 1, and define $Q_n(z) = \sum_{0 \leq j \leq n} q_j z^j$ and $A = \{n \geq 1 \,:\, q_{n-1} = 1\}$.

Theorem (Allouche, 2022)

Suppose there exist constants $C > 0$ and $0 < \alpha < 1$ such that for all complex z with $|z| = 1$ and all $n \geq 1$ one has $|Q_n(z)| \leq Cn^{\alpha}$. Then $(r(k, A, n))_{n \geq 0}$ is eventually strictly increasing for all $k > 2/(1-\alpha)$.

For Thue-Morse we can take $\alpha = (\log 3)/(\log 4) \doteq 0.79248$. Since $10 > 2/(1-\alpha) \doteq 9.63768$, Allouche's result follows.

Powers of Thue-Morse power series

On the other hand, we can prove (just as we did for Rudin-Shapiro) that the coefficients of $T^5(X)$ are *not* eventually increasing.

The status of T^6, T^7, T^8, T^9 is still unknown. It seems likely that T^6 has eventually increasing coefficients.

Dombi's conjecture refuted

Dombi (2002) conjectured that there is no set A such that $\mathbb{N} \setminus A$ is infinite and $r(3, A, n)$ is eventually increasing. But we have:

Theorem (Bell & JOS, 2022)

Let $F = \{3 \cdot 2^n : n \geq 0\} = \{3, 6, 12, 24, \ldots\}$. *Set* $A := \mathbb{N} \setminus F$. *Then* $r(3, A, n)$ *is strictly increasing right from the start.*

Proof.

(Sketch.) Using `Walnut`, we generate a linear representation for $d(n) := r(3, A, n) - r(3, A, n-1)$, guess a closed form for it, and then verify the closed form with `Walnut`. The closed form is strong enough to show that $d(n)$ is always positive. □

A Dombi counterexample of positive density

The example of the previous slide corresponds to a sparse F.

This suggestions the question of whether there is an example where F has positive density.

Indeed there is such an example:

Theorem (JOS, 2023)

Let $F = \{3, 12, 13, 14, 15, 48, 49, 50, \ldots\}$ be the set of natural numbers whose base-2 expansion is of even length and begins with 11.
Then F is of positive lower density and $r(3, \mathbb{N} - F, n)$ is strictly increasing.

Proof.

Like before, using automata and the fact that F is a 2-automatic set. □

Three conjectures

Conjectures

- For the Rudin-Shapiro set $\mathcal{R}$ we have $r(4, \mathcal{R}, n) > r(4, \mathcal{R}, n-1)$ for $n \geq 196$.
- For odious numbers $\mathcal{O}$ we have $r(6, \mathcal{O}, n) > r(6, \mathcal{O}, n-1)$ for $n \geq 6$.
- For the evil numbers we have $r(6, \mathcal{E}, n) > r(6, \mathcal{E}, n-1)$ for $n \geq 38$.

For further reading

- J. Lambek and L. Moser. On some two way classifications of integers. *Canad. Math. Bull.* **2** (1959), 85–89.
- G. Dombi. Additive properties of certain sets. *Acta Arith.* **103** (2002), 137–146.
- J. P. Bell and J. Shallit, Counterexamples to a conjecture of Dombi in additive number theory, arXiv:2212.12473 [math.NT], 2022.
- J.-P. Allouche and J. Shallit, Additive properties of the evil and odious numbers and similar sequences, arXiv:2112.13627 [math.NT], 2022.

Happy birthday to Prof. Simsek!