Automatic sequences
through the lens of higher order Fourier analysis

Jakub Konieczny, University of Oxford & Jagiellonian University

18–22 September, 8. Forum PTM, Lublin

Automatic sequences are among the simplest models of computation. Intuitively speaking, an automatic sequence is one whose n-th term can be computed by a finite device, given the digits on n as input. Perhaps the simplest example of an automatic sequence carries the name of Thue–Morse, and is given by $t(n) = (-1)^{s_2(n)}$, where $s_2(n)$ denotes the sum of digits of n base 2. Somewhat more trivial examples include the periodic sequences, and the sequence $\lfloor \log_2 n \rfloor$.

Various uniformity properties of specific automatic sequences have long been studied. It was already in 1968 that Gelfond obtained quantitative estimates of the Fourier coefficients of the Thue–Morse sequence. As a consequence (observed later by Mauduit and Sarközy), approximately half of the terms in any sufficiently long arithmetic progression have an even sum of binary digits. Later work by Mauduit, Rivat and Drmota, shows likewise that half of the squares and half of the primes have an even sum of binary digits. Similar results were shown and later refined by Müllner, Spiegelhofer and others.

With the advent of higher order Fourier analysis, new notions of uniformity have come to light. Speficically, a sequence a can be construed as uniform — or pseudorandom — if the Gowers uniformity norms $\|a\|_{U^s[N]}$ become small as $N \to \infty$. The usefulness of this notion stems in large part from the fact that Gowers norms control the count of arithmetic progressions, as well as other linear patterns. More precisely, if a set $A \subset [N] = \{1, 2, \ldots, N\}$ is Gowers uniform of order s, meaning that $\|1_A - \alpha 1_{[N]}\|_{U^s[N]}$ is small (where $\alpha = |A|/N$) then A has roughly as many $(s+1)$-term arithmetic progressions as a random set of comparable size.

As it turns out, a simple inductive argument shows that the Gowers norms of the Thue–Morse sequence are very small: $\|t\|_{U^s[N]} = O(N^{-c_s})$, where $c_s > 0$ is a constant. A similar argument shows the same conclusion for another famous automatic sequence by the name of Rudin–Shapiro, and several other examples. For general automatic sequences, jointly with Byszewski and Müllner, we conjecture that an automatic sequence can always be decomposed into two summands, one of which is highly uniform, and the other highly structured. Related results on correlations with polynomial phases follow from joint work with Eisner.

Gowers uniformity norms are intimately tied to nilsequences, that is sequences produced by evaluating a continuous function along the orbit of a point with respect to a rotation on compact quotient of a nilpotent Lie group. Indeed, the celebrated Inverse Theorem of Green, Tao and Ziegler asserts — in rough terms — that sequences with large Gowers norms are precisely those which correlate with a nilsequence of bounded complexity. Nilsequences, in turn, are closely related to bounded generalised polynomials, that is functions obtained from usual polynomials by also allowing the use of the floor function. By a theorem of Bergelson and Leibman, generalised polynomials are precisely the functions obtained by replacing the word "continuous" with the phrase "piecewise polynomial" in the definition of a nilsequence mentioned above. Let us note that there exist non-trivial generalised polynomials which take only finitely many values, such as $\lfloor \varphi(n+1) \rfloor - \lfloor \varphi n \rfloor - 1 \in \{0, 1\}$ (where φ denotes the golden ratio). Interestingly, this last sequence can be computed by a finite automaton reading the Zeckendorf representation of n on input.

Given our interest in Gowers norms of automatic sequence, the following question naturally comes to mind: *"Do there exist any non-trivial examples of automatic sequence which are also given by generalised polynomials?"* (The trivial examples are sequences which are eventually periodic.)

Up to density 0, this question is easily resolved, and the answer is negative. However, interesting phenomena come into view in the sparse setting. As it turns out, any sparse set described by a generalised polynomial formula is very poor in additive structure, which is the opposite of what happens for sets with positive density. It follows from the theory developed by Bergelson and Leibman that a set $A \subset \mathbb{Z}$ whose characteristic function is a generalised polynomial is an IP^* set, provided that the asymptotic density of A is positive. Conversely, if A has density 0 then with Byszewski we show that A not only fails to be IP^*, but also A contains no translate of an IP set. This observation is a key step in showing the following dichotomy: Either there are no non-trivial automatic generalised polynomial sequences; or for some $k \geq 2$, the characteristic function of the set of the powers of k is given by a generalised polynomial.