Queen Mary
University of London

Queen Mary University of London
School of Mathematical Sciences
Mile End Road
London E1 4NS

https://www.qmul.ac.uk/maths/

Dr Reem Yassawi
Senior Lecturer

Report on the thesis "Dynamical and arithmetic properties of substitutive systems" by Elżbieta Krawczyk

Elżbieta Krawczyk's thesis is mainly concerned with the study of substitutional, and, in particular, automatic sequences. She approaches these sequences from a combinatorial, arithmetic, and dynamical point of view.

The systematic study of automatic sequences has developed over the last 70 years, following Cobham's landmark 1971 paper. Before that individual automatic sequences had been defined and studied, but Cobham gave a wholesale definition of automatic sequences and a thorough study of their properties. Since then, there has developed a rich body of work around automatic sequences, in dynamics, arithmetic and algebra, with connections to logic, computer science algebraic geometry, to name a few areas. There have also been generalisations to the higher dimensional setting, some straightforward, others raising good questions.

Thus Ms Krawczyk's work lies in this rich context, and it is clear that she has been diligent and focused in absorbing much of this background. In this thesis Ms Krawczyk investigates several important questions involving automatic sequences, which build on, extend and improve existing results. She deals with the body of automatic sequences, and she mostly does not restrict to primitive substitutions. This is common when researchers study arithmetic properties of automatic sequences, but it is not common when dynamicists study substitutive sequences. It is a welcome addition that Ms Krawczyk has mastered the theory well enough so that she can work in this general context. Results such as Theorem C (see below) demonstrate that in fact the general statement is usually quite distinct from the specialised statement. Another dimension in which this thesis is impressive is that Ms Krawczyk has developed interests in several aspects of automatic sequences, both dynamical and combinatorial. This is especially seen in her mastery of *effective* results. For, one point about automatic sequences is that while one can ask, and often answer, theoretical questions concerning them, one can also answer these questions constructively. In several cases (such as Theorem A and Theorem C) she answers or improves known results in the literature.

A description of the content of the manuscript, and some of the prominent contributions and results of Elżbieta Krawczyk

The thesis begins with an excellent detailed introduction to the objects of study, approaching the questions of interest, and giving motivational reasons for their study. It also contains good statements of the results proved in this thesis.

The first result stated in the introduction is Theorem A, proved in Chapter 4. (page 5, and Theorem 4.2.1 and Remark 4.2.7), and it is joint work with J. Byszewski and J. Konieczny. It is a generalisation of Cobham's theorem. He showed in 1969 that if k and l are multiplicatively independent, and x is both k– and l–automatic, then x must be eventually periodic. Then, in 1995, I. Fagnot generalised Cobham's result to the following statement. Let k and l be multiplicatively independent natural numbers, and suppose that x is k-automatic and that y is l-automatic. If x and y have the same language, i.e., the set of subwords of x equals the set of subwords of y, then x and y are eventually periodic. Theorem A is the following generalisation and strengthening of Fagnot's result, with the same conditions on k and l. It states that U is the set of common factors of a k-automatic x and an l-automatic y if and only if U is a finite union of languages of eventually periodic sequences. This is a very nice result whose statement and proof is quite subtle. Furthermore, although the statement of this result has been established by Byszewski, Krawczyk and Konieczny, Ms Krawczyk gives another proof which does not depend on Cobham's theorem.

Furthermore, related to Theorem A, Ms Krawczyk proves that the set of common factors can be explicitly computed. A related question, which has been studied by N. Nampersad, J. Shallit and M. Stipulanti, in 2019, is the following. If two sequences are automatic, how much of an initial portion do we need to know that they must agree on in order to conclude that they are actually equal? This is a very natural question in this context. For example, it is similar in flavour to the question, give two power series roots of a polynomial equation, how much of an initial portion do they agree on before we can conclude that they are the same root? (The similarity is pertinent as automatic sequences on fields are the coefficients of an algebraic power series.) Nampersad, Shallit and Stipulanti showed that if the two sequences agree on an initial prefix that is doubly exponential in length (over the sizes of the alphabets), then the sequences are equal. In her Theorem B, proved in Chapter 4, Ms Krawczyk proves that it is possible to reduce this bound to a singly exponential bound, which is moreover the best possible. This result shows that Ms Krawczyk is well versed in the techniques of effective results concerning automatic sequences.

Chapter 5 is concerned with Theorem C, which is Theorem 5.3.10 and Corollary 5.2.5, is concerned with the question of when a fixed point of a substitution is actually k-automatic. It is joint work with C. Muellner, and it gives a complete characterisation of when a substitutional sequence is automatic. It is particularly elegant to state when the substitution is primitive. Namely, an (aperiodic) coding of a fixed point u of a primitive substitution θ is k-automatic iff its *return word* substitution η has vector of substitution word lengths as a left eigenvector for the incidence matrix of η. This is a lovely result which closes this question, a question that has been studied by many. More generally, if θ is not primitive, the authors give a more general necessary and sufficient condition. Roughly speaking, u is automatic if and only if for any letter a, for large enough n, the length of $\theta^n(a)$ is a small fixed linear perturbation of K^n, where K is multiplicatively dependent to k. The substitution needs to be '"non degenerate" but any substitution has a non-degenerate power so this is a mild restriction.

Theorem D is proved in Chapter 2. In Part 1 of Theorem D, Ms Krawczyk proves a result in dynamics. A substitutional sequence generates a language, which generates a shift dynamical system (X, σ), namely the set X of infinite sequences all of whose subwords belong to the language. The dynamics σ is the shift map. A subsystem (Y, σ) of (X, σ) is a closed, shift invariant subset Y of X. Y is *transitive* if is contains a point whose orbit is dense. The first part of Theorem D is joint work with Byszewski and Konieczny, and it shows that any transitive subsystem of a substitutive system is substitutive. In other words, the shift orbit closure of a transitive point in X is substitutional. If the substitution is primitive, then this is trivially true, so the interest of this result is for non primitive substitutions. It should be mentioned that if the substitution is aperiodic, i.e., has no shift periodic points, then I believe that this result may also follow from the description of the Bratteli-Vershik diagrams of these substitutions, in the 2009 paper by S. Bezuglyi, J. Kwiatkowski, and K. Medynets. Certainly, the first part of Theorem D is *not* stated in the paper by Bezuglyi et al. However the authors do give a (Bratteli-Vershik) representation of a substitution shift where points are infinite paths in a stationary Bratteli diagram, and it seems to me accessible with this approach that the closure of such a point under the dynamics will also be a stationary Bratteli-Vershik diagram, i.e., a substitution shift.
Part 2 of Theorem D concerns the question of which sequences in an automatic shift are actually automatic. As automatic shifts generally have uncountable

cardinality, and there are only countably many automatic sequences, this question is relevant. Ms Krawczyk shows that a sequence $y \in X_\theta$ is automatic if and only if $y = \Phi(x)$ (Φ is a factor map) where either $x \in X$ is θ-periodic, or more generally it satisfies a functional equation of the form $x = \sigma^m \theta^j(x)$. Ms Krawczyk calls these points *quasi-fixed* points, and as she has mentioned, they have been studied in the dynamics community. It should be mentioned that Corollary 2.3.3, which is crucial for the proof of Part 2 of theorem D, has already appeared as Corollary 3.7 of the 2015 article
"Computing automorphism groups of shifts using atypical equivalence classes" by E. Coven, A. Quas and R. Yassawi. Although there is no mention of quasi-fixed points, the connection is clear from Proposition 2.1.8(2), which implies that quasi fixed points are precisely those that are mapped to rational expansions in Z_k, the k-adic integers by a factor map $\pi : X_\theta \to Z_k$ which sends fixed points to 0. It should also be mentioned that the techniques of Proposition 3.24 of that same article above will prove results similar to Theorem 2.3.1 in the case of primitive shifts.
Hence at the very least, this paper should be cited and the methods of proof compared. However part 2 of Theorem D is new for non-primitive substitution shifts.

Ms Krawczyk ends the introduction with Theorem E, proved in Chapter 3, which is a technical result concerning an important property which has been very useful in substitutive dynamics. *Recognizability* is the notion that points in a substitution shift can be uniquely de-substituted. This result is originally due to Mosse for primitive, aperiodic (not necessarily constant length) substitutions. Mossés's result states that if θ is aperiodic and primitive, then any long enough word that appears in a fixed point x can essentially be de-substituted in only one way, away from the left and right border. This implies that in a primitive, infinite, two-sided shift dynamical system, any point can be uniquely de-substituted. In fact, the proof of this result for primitive aperiodic constant length substitutions can be extracted from the 1977 work of Dekking. For constant length-k substitutions, it implies that all long enough words can only appear at indices which are in a unique residue class of some power k. This is Theorem 3.23 in Ms Krawczyk's thesis. Her general result on recognizability concerns non primitive substitutions. In this case, theorem E, roughly speaking, says that there is a finite number of residue classes that a difference of indices where w appears can take. This is joint work with C. Muellner.

Opinion and conclusion

To prove her results, Ms Krawczyk has successfully mastered several classical results and generalised them in a theoretical and effective dimension. These results are novel and interesting. They are results that build on important, foundational theorems in the area of automatic sequences, and they push the boundary of what we can ask and answer. Furthermore the fact that Ms Krawczyk has collaborated with two groups of mathematicians is also a very good sign of her effectiveness at communicating and collaborating with other researchers, an important skill in today's academic world. The thesis is a pleasure to read, and it is aptly illustrated with excellent and often simple examples that illuminate the pitfalls and the theory. It is a coherent and cohesive work and contains new, true and truly interesting results that will also be of use to the research community. It demonstrates that the candidate is very well equipped to continue her journey in mathematical research. For these reasons I recommend that Ms. Elżbieta Krawczyk be awarded the degree of Doctor in Mathematics (PhD) at the *Institute of Mathematics, Jagiellonian University in Kraków*, Poland.

Sincerely,

Reem Yassawi