FIFTY YEARS OF THE SPECTRUM PROBLEM: SURVEY AND NEW RESULTS

A. DURAND, N. D. JONES, J. A. MAKOWSKY, AND M. MORE

Abstract. In 1952, Heinrich Scholz published a question in the Journal of Symbolic Logic asking for a characterization of spectra, i.e., sets of natural numbers that are the cardinalities of finite models of first order sentences. Günter Asser asked whether the complement of a spectrum is always a spectrum. These innocent questions turned out to be seminal for the development of finite model theory and descriptive complexity. In this paper we survey developments over the last 50-odd years pertaining to the spectrum problem. Our presentation follows conceptual developments rather than the chronological order. Originally a number theoretic problem, it has been approached in terms of recursion theory, resource bounded complexity theory, classification by complexity of the defining sentences, and finally in terms of structural graph theory. Although Scholz' question was answered in various ways, Asser's question remains open. One appendix paraphrases the contents of several early and not easily accesible papers by G. Asser, A. Mostowski, J. Bennett and S. Mo. Another appendix contains a compendium of questions and conjectures which remain open.

To be submitted to the Bulletin of Symbolic Logic.
July 17, 2009 (version 13.2)

Contents

Sganarelle: Ah! Monsieur, c'est un **spectre**:
je le reconnais au marcher.
Dom Juan: Spectre, fantôme, ou diable,
je veux voir ce que c'est.
J.B. Poquelin, dit Molière, Dom Juan, Acte V, scène V

§1. Introduction. At the Annual Symposium of the European Association of Computer Science Logic, CSL'05, held in Oxford in 2005, Arnaud Durand, Etienne Grandjean and Malika More organized a special workshop dedicated to the spectrum problem. The workshop speakers and the title of their talks where

- Annie Chateau (UQAM, Montreal)
  The Ultra-Weak Ash Conjecture is Equivalent to the Spectrum Conjecture, and Some Relative Results
- Mor Doron (Hebrew University, Jerusalem).
  Weakly Decomposable Classes and Their Spectra (joint work with S. Shelah),
- Aaron Hunter (Simon Fraser University, Burnaby).
  Closure Results for First-Order Spectra: The Model Theoretic Approach
- Neil Immerman (University of Massachusetts, Amherst)
  Recent Progress in Descriptive Complexity
- Neil Jones (University of Copenhagen, Copenhagen)
  Some remarks on the spectrum problem
- Johann A. Makowsky (Technion–Israel Institute of Technology, Haifa)
  50 years of the spectrum problem

The organizers and speakers then decided to use the occasion to expand the survey talk given by J.A. Makowsky into the present survey paper, rather than publish the talks.

Acknowledgements. We would like to thank A. Chateau, M. Doron, A. Esbelin, R. Fagin, E. Fischer, E. Grandjean, A. Hunter, N. Immerman and S. Shelah for their fruitful comments which helped our preparation of this survey.

§2. The Emergence of the Spectrum Problem.

2.1. Scholz's problem. In 1952, H. Scholz published an innocent question in the Journal of Symbolic Logic [116]:

> **1.** Ein ungelöstes Problem in der symbolischen Logik. IK sei der Prädikatenkalkül der ersten Stufe mit der Identität. In IK ist ein Postulatensystem BP für die Boole'sche Algebra mit einer einzigen zweistelligen Prädikatenvariablen formalisierbar. θ sei die Konjunktion der Postulate von PB. Dann ist θ für endliche m m-zahlig erfüllbar genau dann, wennn es ein $n > 0$ gibt, sodaß $m = 2^n$.
>
> Hieraus ergibt sich das folgende Problem. H sei ein Ausdruck des IK. Unter dem *Spectrum von H* soll die Menge der natürlichen Zahlen verstanden sein, für welche H erfüllbar ist. M sei eine beliebige Menge von natürlichen Zahlen. Gesucht ist eine hinreichende [hinrerichende] und notwendige Bedingung dafür, daß es ein H gibt, sodaß M das Spectrum von H ist.
> *(Received September 19, 1951).*

In English:

> **1.** An unsolved problem in symbolic logic. Let IK [Identitätskalkül] be the first order predicate calculus with identity. In IK one can formalize an axiom system BP [Boole'sche Postulate] for Boolean algebras with only one binary relation variable. Let θ be the conjunction of the axioms of BP. Then θ is satisfiable in a finite domain of m elements if and only if there is an $n > 0$ such that $m = 2^n$.
>
> From this results the following problem. Let H be an expression of IK. We call the set of natural numbers, for which H is satisfiable, the *spectrum of H*. Let M be an arbitrary set of natural numbers. We look for a sufficient and necessary condition that ensures that there exists an H, such that M is the spectrum of H.
> *(Received September 19, 1951).*

This question inaugurated a new column of *Problems* to be published in the Journal of Symbolic Logic and edited by L. Henkin. Other questions published in the same issue were authored by G. Kreisel and L. Henkin. They deal with a question about interpretations of non-finitist proofs dealing with recursive ordinals and the no-counter-example interpretation (Kreisel), the provability of formulas asserting the provability or independence of provability assertions (Henkin), and the question whether the ordering principle is equivalent to the axiom of choice (Henkin). All in all 9 problems were published, the last in 1956.

The context in which Scholz's question was formulated is given by the various completeness and incompleteness results for First Order Logic that were the main concern of logicians of the period. An easy consequence of Gödel's classical completeness theorem of 1929 states that validity of first order sentences in all (finite and infinite) structures is recursively enumerable, whereas Church's and Turing's classical theorems state that it is not recursive. In contrast to this, the following was shown in 1950 by B. Trakhtenbrot.

THEOREM 2.1 (Trakhtenbrot 50[126]). *Validity of first order sentences in all finite structures (f-validity) is not recursively enumerable, and hence satisfiability of first order sentences in some finite structure (f-satisfiability) is not decidable, although it is recursively enumerable.*

Heinrich Scholz, a German philosopher, was born 17. December 1884 in Berlin and died 30. December 1956 in Münster. He was a student of Adolf von Harnack. He studied in Berlin and Erlangen philosophy and theology and got his habilitation in 1910 in Berin for the subjects philosophy of religion and systematic theology. He received his Ph.D. in 1913 for his thesis *Schleiermacher and Goethe. A contribution to the history of German thought*. In 1917 he was appointed full professor for philosophy of religion in Breslau (Wroclaw, today Poland). In 1919 he moved to Kiel, and from 1928 on he taught in Münster. From 1924 till 1928 he studied exact sciences and logic and formed in Münster a center for mathematical logic and foundational studies, later to be known as *the school of Münster*. His chair became in 1936 the first chair for *mathematical logic and the foundations of exact sciences*. His seminar underwent several administrative metamorphoses that culminated in 1950 in the creation of the *Institute for mathematical logic and the foundations of exact sciences*, which he led until his untimely death. Among his pupils and collaborators we find W. Ackermann, F. Bachmann, G. Hasenjäger, H. Hermes, K. Schröter and H. Schweitzer. He was also among the founders of the German society bearing the same name (DVMLG). H. Scholz was a Platonist, and he considered mathematical logic as the foundation of epistemology. He is credited for his discovery of Frege's estate, and for making Frege's writing accessible to a wider readership. Together with his pupil Hasenjäger he authored the monograph *Grundzüge der Mathematischen Logik*, published posthumously in 1961.

Thus, H. Scholz really asked whether one could prove anything meaningful about f-satisfiablity besides its undecidability.

2.2. Basic facts and questions. In our notation to be used throughout the paper, H. Scholz introduced the following:

Let τ be a vocabulary, i.e., set of relation and function symbols. Let ϕ be a sentence in some logic with equality over a vocabulary τ. Unless otherwise stated the logic will be first order logic $FOL(\tau)$. Sometimes we shall also discuss second order logic, or a fragment thereof.

DEFINITION 2.2. The *spectrum* $\text{SPEC}(\phi)$ of ϕ is the set of finite cardinalities (viewed as a subset of $\mathbb{N}$), in which ϕ has a model.

We denote by SPEC the set of spectra of first-order sentences, i.e.,

$$\text{SPEC} = \{\text{SPEC}(\phi) \mid \phi \text{ is a first-order formula}\}$$

We shall use S, S_i to denote spectra. For the definition of spectra it does not matter whether we use function symbols or not. So, unless otherwise stated, vocabularies will be without function symbols. However, we shall allow function symbols when dealing with sentences of special forms.

Clearly, $\text{SPEC}(\phi) = \emptyset$ if and only if ϕ is not f-satisfiable. By definition of satisfiability 0 is never part of a spectrum. Very often a spectrum is finite, cofinite or even of the form $\mathbb{N}^+ = \mathbb{N} - \{0\}$.

QUESTION 2.3. *Is it decidable whether, for a given ϕ, $\text{SPEC}(\phi) = \mathbb{N}^+$?*

As H. Scholz noted, (the set of) powers of 2 form a spectrum, because they are the cardinalities of finite Boolean algebras. Similarly, powers of primes form a spectrum, because they are the cardinalities of finite fields. For $a, b \in \mathbb{N}^+$ there are many ways to construct a sentence ϕ with $\text{SPEC}(\phi) = a + b\mathbb{N}$, one of which consists in using one unary function symbol. With a moment of reflection, one sees that spectra have the following closure properties.

PROPOSITION 2.4. *Let S_1 and S_2 be spectra.*

(i) *Then $S_1 \cup S_2, S_1 \cap S_2$ are also spectra.*
(ii) *Let $S_1 + S_2 = \{m + n : m \in S_1, n \in S_2\}$. Then $S_1 + S_2$ is a spectrum.*
(iii) *Let $S_1 \star S_2 = \{m \cdot n : m \in S_1, n \in S_2\}$. Then $S_1 \star S_2$ is a spectrum.*

In the spirit of Question 2.3 we can also ask:

QUESTION 2.5. *Which of the following sets are recursive? The set of sentences ϕ such that*

(i) $\text{SPEC}(\phi)$ *is finite, cofinite.*
(ii) $\text{SPEC}(\phi)$ *is ultimately periodic.*
(iii) $\text{SPEC}(\phi)$ *is, for given $a, b \in \mathbb{N}$ of the form $a + b\mathbb{N}$.*
(iv) $\text{SPEC}(\phi) = S$ *for a given set $S \subseteq \mathbb{N}^+$.*

We shall answer Questions 2.3 and 2.5 in Section 3.4.

2.3. Immediate responses to H. Scholz's problem. The first to publish a paper in response to H. Scholz's problem was G. Asser [7]. A. Robinson's review [112] summarizes it as follows:

> (...) The present paper is concerned with the characterisation of all representable sets [=spectra]. A rather intricate necessary and sufficient condition is stated for arithmetical function $X(n)$ to be the characteristic function of a representable set. The condition shows that

> such a function is elementary in the sense of Kalmar. (...) On the other hand, the author establishes that there exist non-representable sets whose characteristic function is elementary. Examples of representable sets (some of which are by no means obvious) are given without proof and the author suggests that further research in this field is desirable.

Asser also noted that his characterization did not establish whether the complement of a spectrum is a spectrum.

About the same time, A. Mostowski [98] also considered the problem. H. Curry [29] summarizes Mostowski's paper as follows:

> (...) The author proves that for each function $f(n)$ of a class K of functions, which is like the class of primitive functions except that at each step all functions are truncated above at n, there is a formula H that has a model in a set of $n+1$ individuals if and only if $f(n) = 0$. From this he deduces positive solutions to Scholz's problem in a number of special cases.

It is usually considered that A. Mostowski really proved

Theorem 2.6. *All sets of natural numbers, whose characteristic functions are in the second level of the Grzegorzcyk Hierarchy $\mathcal{E}^2$, are first order spectra.*

The detailed definitions and contents of this theorem will be discussed in Section 4.

In the last 50 years a steady stream of papers appeared dealing with spectra of first order and higher order logics. The problem seems not too important at first sight. However, some of these papers had considerable impact on what is now called Finite Model Theory and Descriptive Complexity Theory.

Open Question 1 (Scholz's Problem). *Characterize the sets of natural numbers that are first order spectra.*

Scholz's Problem, as stated, is rather vague. He asks for a characterization of a family of subsets of the natural numbers without specifying, what kind of an answer he had in mind. The answer could be in terms of number theory, recursion theory, it could be algebraic, or in terms of something still to be developed. We shall see in the sequel many solutions to Scholz's Problem, but we consider it still open, because further answers are still possible.

The same question can be asked for any logic, in particular second order logic SO, or fragments thereof, like monadic second order logic MSOL, fixed point logic, etc., as discussed in [36, 84].

Open Question 2 (Asser's Problem). *Is the complement of a first order spectrum a first order spectrum?*

Here the answer should be yes or no.

The corresponding problem for SO has a trivial solution. Let $\phi \in \mathrm{SO}(\tau)$ with $\tau = \{R_1, \ldots, R_k\}$. An integer n is in $\mathrm{SPEC}(\phi)$ iff

$$n \models \exists R_1 \exists R_2 \ldots \exists R_k \phi.$$

Then, the complement of $\mathrm{SPEC}(\phi)$ is easily seen to be the spectrum of the SO sentence $\neg(\exists R_1 \ldots \exists R_k \phi)$. In passing, note that every SO-spectrum is a $\mathrm{SO}(\tau)$ over a language τ containing equality only.

However, for fixed fragments of SO, Asser's Problem remains open. In particular

Open Question 3. *Is the complement of a spectrum of an* MSOL-*sentence again a spectrum of an* MSOL-*sentence?*

2.4. Approaches and themes. In this survey we shall describe the various solutions and attempts to solve Scholz's and Asser's problems, and the developments these attempts triggered. We shall emphasize more the various ways the questions were approached, and focus less on the historical order of the papers.

There are several discernible themes:

Recursion Theory: The early authors H. Asser and A. Mostowski approached the question in the language of the theory of recursive functions i.e. they looked for characterization of spectra in terms of recursion schemes, or hierarchies of recursive functions. Most prominently in terms of Kalmar's elementary functions, the Grzegorczyk hierarchy and hierarchies of arithmetical predicates, in particular *rudimentary relations*. This line of thought culminates in 1962 in the thesis of J. Bennett [9][1]. Although G. Asser already characterized first order spectra in such terms, his characterization was not considered satisfactory even by himself, because it did not use standard terms and was not useful in proving that a given set of integers is (or not) a spectrum. We shall discuss the recursion theoretic approach in detail in Section 4.

Complexity Theory: In the 1970s, D. Rödding and H. Schwichtenberg of the Münster school [113] gave a sufficient but not necessary condition: any set of integers recognizable by a deterministic linear space-bounded Turing machine is a first-order spectrum. (This is also a consequence of results of Bennett and Ritchie [9, 110], obtained before the emergence of complexity theory.) Further, Rödding and Schwichtenberg showed that sets of integers recognisable using

[1]It seems that some of Bennet's unpublished results were rediscovered independently in China in the late 1980ties by Shaokui Mo [95]. We shall discuss his work in Section A.4.

larger space bounds are higher order spectra. C. Christen developed this line further [18, 19], independently obtaining a number of the following results.

At the same time the spectrum problem gained renewed interest in the USA. In 1972 A. Selman and N. Jones found an exact solution to Scholz's original question [79]: a set of integers is a first order spectrum if and only if it is recognizable by a non-deterministic Turing machine in time $O(2^{c \cdot n})$.

This result was independently also obtained by R. Fagin in his thesis [40], which contains an abundance of further results. Most importantly, R. Fagin studies *generalized spectra*, which are the projective classes of Tarski, restricted to finite structures, and really laid the foundations for Finite Model Theory and Descriptive Complexity, as can be seen in the monographs [76, 36, 84]. We shall discuss the complexity theoretic approach in detail in Section 5.

Images and preimages of spectra: From Proposition 2.4 it follows that, if S is a first order spectrum and p is a polynomial with positive coefficents, then $p(S) = \{p(m) : m \in S\}$ is also a spectrum. In J. Bennett's thesis it is essentially proved that there is a first order spectrum S and an integer k such that $\{n : 2^{n^k} \in S\}$ is not a spectrum. It is natural to ask what happens to a spectrum under images and preimages of number theoretic functions. The general line of this type of results states that certain images or preimages of spectra of specific forms of sentences are or are not spectra of other specific forms of sentences.

Spectra of syntactically restricted sentences: Already in a paper by L. Löwenheim from 1915 [87] it is noted that, what later will be called the spectrum of a sentence in monadic second order logic (MSOL) with unary relation symbols only, is finite or cofinite. The set of even numbers is the spectrum of an MSOL sentence with one binary relation symbol, and it is ultimately periodic. Further, *every* ultimately periodic set of positive integers is a spectrum of a first order MSOL sentence with one unary function symbol. Over the last fifty years various papers were written relating restrictions on the use of relation and function symbols, or other syntactic restrictions, to special forms of spectra. R. Fagin, in his thesis, poses the following problem

Open Question 4 (Fagin's Problem for binary relations). *Is every first order spectrum the spectrum of a first order sentence of one binary relation symbol?*

The question is even open, if restricted to any fixed vocabulary that contains at least one binary relation symbol or two unary function symbols.

Much of this line of research is motivated by attempts to solve Fagin's problem.

Transfer theorems: Another way of studying spectra is given by the following result, again from Fagin's thesis: If S is a spectrum of a purely relational sentence where all the predicate symbols have arity bounded by k, then $S^k = \{m^k : m \in S\}$ is a spectrum of a sentence with one binary relation symbol only, or even a spectrum on simple graphs. One can view this an approach combining the study of images and preimages of spectra with either syntactically or semantically restricted spectra. Over the years quite a few results along this line were published. We shall discuss the last three approaches under the common theme of restrictions on vocabularies in detail in Section 6.

Spectra of semantically restricted classes: R. Fagin shows that Asser's problem has a positive answer if and only if it has a positive answer if restricted to the class of simple graphs. Similarly, in order to understand Fagin's problem better, one could consider restricted graph classes K, and study first order spectra restricted to graphs in K. One may think of graphs of bounded degree, planar graphs, trees, graphs of tree-width at most k, etc.

Open Question 5 (Fagin's Problem for simple graphs). *Is every first order spectrum the spectrum of a first order sentence over simple graphs?*

Open Question 6. *Is every first order spectrum the spectrum of a first order sentence over planar graphs?*

For restrictions to graph classes of bounded tree-width, the answer is negative. The reason for this is that spectra of graphs of bounded tree-width are ultimately periodic. In fact, this holds for a much wider class of spectra. E. Fischer and J.A. Makowsky, [49], have analyzed under what conditions MSOL-spectra are ultimately periodic. We shall discuss their results in detail in Section 8.

This line of thought has not been extensively explored, this may well be a fruitful avenue for studying spectra in the future.

In the sequel of this survey we shall summarize what is known about spectra along these themes. Various solutions to Scholz's Problem were offered in the literature, varying with the tastes of the times, but there may be still more to come. Asser's and Fagin's Problems are still open. Both problems are intimately related to our understanding of definability hierarchies in Descriptive Complexity Theory. They may well serve as benchmarks of our understanding.

§3. Understanding Spectra: counting functions and number theory. In this section we formulate various ways to test our understanding of spectra. It will turn out that there still many questions we do not know how to answer.

3.1. Representation of spectra and counting functions. Spectra are sets of positive natural numbers. These sets can be represented in various ways. We shall use the following:

Definition 3.1. Let $M \subseteq \mathbb{N}^+$, and let $m_1, m_2, \ldots$ an enumeration of M ordered by the size of its elements.

(i) $\chi_M(n)$ is the characteristic function of M, i.e.,

$$\chi_M(n) = \begin{cases} 1 & \text{if } n \in M \\ 0 & \text{else .} \end{cases}$$

(ii) $\eta_M(n)$ is the enumeration function of M, i.e.,

$$\eta_M(n) = \begin{cases} m_n & \text{if it exists} \\ 0 & \text{else .} \end{cases}$$

(iii) $\gamma_M(n)$ is the counting function of M, i.e., $\gamma_M(n)$ is the number of elements in M that are strictly smaller than n.

(iv) A *gap* of M is a pair of integers g_1, g_2 such that $g_1, g_2 \in M$ but for each n with $g_1 < n < g_2$ we have that $n \notin M$. Now let $\delta_M(n)$ be the length of the nth gap of M. Clearly, $\delta_M(n) = \eta_M(n+1) - \eta_M(n)$.

Obvious questions are of the following type:

Open Question 7. *Which strictly increasing sequences of positive integers, are enumerating functions of spectra? For instance, how fast can they grow?*

Open Question 8. *If M is a spectrum how can $\delta_M(n)$ behave?*

Coding runs of Turing machines one can easily obtain the following.

Proposition 3.2. *For every recursive monotonically increasing function f there is a first order formula ϕ such that $\delta_\phi(n) = f(n)$.*

Various other partial answers to these questions will appear throughout our narrative.

3.2. Prime numbers. An obvious question is whether the primes form a spectrum. If one gets more ambitious one can ask for special sets of primes such as Fermat primes (of the form $2^{2^n} + 1$), Mersenne primes (of the form $2^p - 1$ with p a prime), or the set of primes p such that $p+2$ is also a prime (twin primes). Even if we do not know whether such a set is finite, which is the case for twin primes, it may still be possible to prove that it is a spectrum. The answer to all these question is yes, because all these sets are easily proved to be rudimentary, see Section 4.

In the sense of the above definitions we have χ_{primes} is the characteristic function of the set of primes, $\eta_{primes}(n) = p_n$, and $\gamma_{primes}(n)$ is the counting function of the primes, usually denoted by $\pi(n)$. $\delta_{primes}(n)$ is usually denoted by d_n. All these functions related to primes are subject to intensive study in the literature, see eg. [108]. As we have said that the primes form a first order spectrum, all the features of these functions observed on primes do occur on spectra.

For instance, $\pi(n)$ is approximated by the integral logarithm $li(n)$, and it was shown by J.E. Littlewood in 1914, cf. [108] that $\pi(n)-li(n)$ changes sign infinitely many often. For logical aspects of Littlewood's theorem, see [82].

Let us define

$$\pi^+ = \{n : \pi(n) - li(n) > 0\}$$
$$\pi^- = \{n : \pi(n) - li(n) \geq 0\}$$

A less obvious question concerning spectra and primes is

Open Question 9. *Are the sets π^+ and π^- spectra?*

3.3. Density functions. Many combinatorial functions are defined by linear or polynomial recurrence relations. Among them we have the powers of 2, factorials, the Fibonacci numbers, Bernoulli numbers, Lucas numbers, Stirling numbers and many more, cf. [57].

Question 3.3. *Are the sets of values of these combinatorial functions first order spectra?*

The answer will be yes in all of these cases. We shall sketch a proof in Section A.3 that is based on the existence of such recurrence relations.

But these functions also allow combinatorial interpretations as counting functions: The powers of 2 count subsets, the factorials count linear orderings, the Stirling numbers are related to counting equivalence relations. We shall see below that in these three examples the combinatorial definitions allow us to give alternative proofs that these sets of numbers are first order spectra.

The spectrum of a sentence ϕ witnesses the existence of models of ϕ of corresponding cardinalities. Instead, one could also ask for the number of ways the set $\{0, 1, \ldots, n-1\} = [n]$ can be made into a model of ϕ. Alternatively one could count models up to isomorphisms or up to some other equivalence relation.

Combinatorial counting functions come in different flavours;

DEFINITION 3.4. Let $\mathcal{C}$ be a class of finite τ-structures. With $\mathcal{C}$ we associate the following counting functions:

(i) $f_{\mathcal{C}}(n)$ is the number of ways one can interpret the relation symbols of τ on the universe $[n]$ such that the resulting structure is in $\mathcal{C}$. This corresponds to counting labeled structures.

(ii) Let $Str(\tau)(n)$ denote the number of labeled τ-structures of size n. We put

$$prob_{\mathcal{C}}(n) = \frac{f_{\mathcal{C}}(n)}{Str(\tau)(n)}$$

which can be interpreted as the probability that a labeled τ-structure of size n is in $\mathcal{C}$.

(iii) $f_{\mathcal{C}}^{iso}(n)$ is the number of non-isomorphic models in $\mathcal{C}$ of size n.

(iv) For an equivalence relation E on $\mathcal{C}$ we denote by $f_{\mathcal{C}}^{E}(n)$ the number of non-E–equivalent models in $\mathcal{C}$ of size n.

(v) If E is the k-equivalence from Ehrenfeucht-Fraïssé games, $f_{\mathcal{C}}^{E}(n)$ is denoted by $N_{\mathcal{C},k}(n)$, and is called an Ash-function, cf. [6] and Section 7.

(vi) If $\mathcal{C}$ consists of all the finite models of a sentence ϕ we write $f_{\phi}^{E}(n)$ instead of $f_{\mathcal{C}}^{E}(n)$. Similarly for $prob_{\phi}(n)$.

Counting labeled and non-labeled structures has a rich literature, cf. [68, 130]. Note that counting non-labeled non-isomorphic structures is in general much harder than the labeled case. The first connection between counting labeled structures and logic is the celebrated 0-1 Law for first order logic:

THEOREM 3.5 (0-1 Laws). *For every first order sentence ϕ over a purely relational vocabulary τ we have:*

(i) *(Y. Glebskii, D. Kogan, M. Liogonki and V. Talanov [53]; R. Fagin [40])*

$$\lim_{n\to\infty} prob_{\phi}(n) = \begin{cases} 0 \\ 1 \end{cases}$$

and the limit always exists.

(ii) *(E. Grandjean [58]) Furthermore, the set of sentences ϕ such that $\lim_{n\to\infty} prob_{\phi}(n) = 1$ is decidable, and in fact* **PSpace***-complete.*

What we are interested in here, is the relationship of such counting functions to spectra. Our example of powers of 2 shows that

$$2^n = f_{\phi}(n) = \eta_{\psi}(n)$$

where ϕ is an always-true first order sentence with one unary relation symbol, and ψ is the conjunction of the axioms of Boolean algebras. Similarly,

$$n! = f_{\phi_{LIN}}(n) = \eta_{\psi}(n)$$

where ϕ_{LIN} are the axioms of linear orders, and ψ describes the following situation:

(i) P is a unary relation and R is an linear order on P.
(ii) E is a ternary relation that is a bijection between the universe (first argument x) and all the linear orderings on P (remaining two arguments y, z).
(iii) First we say that there is an x that corresponds to R; and that for $x \neq x'$ the orderings are different. This says that E is injective. To ensure that we get all the orderings on P we say that for every ordering and every transposition of two elements in this ordering, there is a corresponding ordering.

Hence, the size of the model of ψ is the number of linear orderings on P.

Clearly, if f_ϕ is not strictly increasing, there is no ψ with $f_\phi(n) = \eta_\psi(n)$. For instance, for ϕ which says that some function is a bijection of a part of the universe to its complement, we have

$$f_\phi(n) = \begin{cases} \binom{2m}{m} \cdot m! & \text{if } n = 2m \\ 0 & \text{else} \end{cases}$$

Open Question 10. *Let ϕ a first order sentence, and f_ϕ be the associated labeled counting function that is monotonically increasing. Is there a first order sentence ψ such that for all n*

$$f_\phi(n) = \eta_\psi(n)$$

The converse question seems more complicated. For instance, as we have noted before, the primes p_n are of the form η_ψ for some first order ψ, but we are not aware of any labeled counting function that will produce the primes.

R. Fagin [41] calls ϕ *categorical* if $f^{iso}_\phi(n) \leq 1$ for every n. For instance, ϕ_{LIN} is categorical. The counting function up to isomorphisms can be bounded by any finite number m, using disjunctions of different categorical sentences. So it may be less promising to study for which first order sentences ϕ there is a ψ such that $f^{iso}_\phi(n) = \eta_\psi(n)$, or vice versa.

Surprisingly enough, C. Ash [6] has found a connection between Asser's Problem and the behaviour of the Ash functions defined in Definition 3.4(v). We shall discuss this in Section 7.

3.4. Sentences with prescribed spectra. In the light of Theorem 3.5 we note that if $prob_\phi(n)$ tends to 1 then $spec(\phi)$ is cofinite. Obviously, the converse does not hold, because there are categorical sentences with models in all finite cardinalities.

Trakhtenbrot's Theorem says that it is undecidable whether a spectrum is empty, and Grandjean's Theorem says that it is decidable, whether a sentence is almost always true, i.e. $prob_\phi(n)$ tends to 1. As a partial answer to Questions 2.3 and 2.5 we have:

PROPOSITION 3.6. *Let ϕ be a first order sentence. The following are undecidable:*

(i) $\text{SPEC}(\phi)$ *is finite, cofinite.*
(ii) $\text{SPEC}(\phi)$ *is ultimately periodic.*
(iii) $\text{SPEC}(\phi)$ *is, for given $a, b \in \mathbb{N}$ of the form $a + b\mathbb{N}$.*
(iv) $\text{SPEC}(\phi) = S$ *for a given set $S \subseteq \mathbb{N}^+$.*

SKETCH OF PROOF. Let $\varphi \in$ FO. We describe the construction of FO-sentences ψ_1, ψ_2, ψ_3, ψ_4 and ψ_5 such that $\text{SPEC}(\varphi) = \emptyset$ if and only if:

- $\text{SPEC}(\psi_1)$ is finite.
- $\text{SPEC}(\psi_2) = \mathbb{N}^+$.
- More generally, $\text{SPEC}(\psi_3) = \{f(i) \mid i \in \mathbb{N}^+\}$ for a given function f such that $f(i) \geq i$ for all i and the graph $n = f(i)$ seen as a binary relation is rudimentary (see Section 4.2 for a precise definition).
- $\text{SPEC}(\psi_4)$ is cofinite.
- $\text{SPEC}(\psi_5)$ is ultimately periodic.

Since the problem of emptiness of spectra is undecidable, the announced result follows.

Let 0 and max be two constant symbols, let $\leq$ be a binary predicate symbol and let $+$ and $\times$ be two ternary predicate symbols. Let $Arithm(0, \max, \leq, +, \times)$ denote a first-order sentence axiomatizing the usual arithmetic predicates. Our sentences ψ_i $(i = 1, \ldots, 3)$ consist of the conjunction of $Arithm(0, \max, \leq, +, \times)$ with a specific part ψ'_i that we will describe below. We shall use the fact that the Bit predicate[2] is definable from $+$ and $\times$ in finite structures, as well as the ternary relation $a = b^c$. For simplicity, we will assume w.l.o.g. that the signature of φ consists of a binary relation R only. Let n and y be new variable symbols. Let $\varphi'(n, y)$ be the formula obtained from φ by replacing every quantification $\forall x$ by $\forall x < n$ and $\exists x$ by $\exists x < n$, and every atomic formula $R(x, x')$ by $Bit(y, x + nx')$. The idea is that a graph R on a set of n elements seen as $\{0, \ldots, n-1\}$ is encoded by the number $y < 2^{n^2}$ written in binary with a 1 in position $a + bn$ if and only if $R(a, b)$ holds. Hence for all $n \in \mathbb{N}^+$, we have $\exists y < 2^{n^2} \varphi'(n, y)$ if and only if φ has a model with n elements.

- Let $\psi'_1 \equiv \exists m, n, y < \max(\max = 3^m \times 2^{n^2} \wedge y < 2^{n^2} \wedge \varphi'(n, y))$.

  It is easy to verify that if $\text{SPEC}(\varphi) = \emptyset$, then $\text{SPEC}(\psi_1)$ is also empty (hence finite), and conversely, if $\text{SPEC}(\varphi) \neq \emptyset$, then $\text{SPEC}(\psi_1)$ contains all the integers of the form $3^m \times 2^{n^2}$ for some $m \in \mathbb{N}^+$ and $n \in \text{SPEC}(\varphi)$, i.e. $\text{SPEC}(\psi_1)$ is infinite.

[2] $Bit(a, b)$ is true iff the bit of rank b of a is 1.

- Let $\psi_2' \equiv (\forall n < \max\ \max \neq 2^{n^2}) \vee (\exists n < \max(\max = 2^{n^2} \wedge \forall y < 2^{n^2} \neg\varphi'(n,y)))$.

  If $\text{SPEC}(\varphi) = \emptyset$, then for all n and y, the condition $\exists y < 2^{n^2} \varphi'(n,y)$ is false, hence $\text{SPEC}(\psi_2) = \mathbb{N}^+$. Conversely, if $\text{SPEC}(\varphi) \neq \emptyset$, then the integers of the form 2^{n^2} with $n \in \text{SPEC}(\varphi)$ are not in $\text{SPEC}(\psi_2)$.
- Since the binary relation $y = f(x)$ is rudimentary, it is definable from $+$ and $\times$ in finite structures. Let $\psi_3' \equiv \exists i < \max(\max = f(i) \wedge ((\forall n < i\ i \neq 2^{n^2}) \vee (\exists n < i(i = 2^{n^2} \wedge \forall y < 2^{n^2} \neg\varphi'(n,y)))))$.

  The verification that $\text{SPEC}(\varphi) = \emptyset$ if and only if $\text{SPEC}(\psi_3) = \{f(i) \mid i \in \mathbb{N}^+\}$ is similar to the previous case.
- Let $\psi_4' \equiv \forall n < \max(2^{n^2} \leq \max \longrightarrow \forall y < 2^{n^2} \neg\varphi'(n,y))$.

  If $\text{SPEC}(\varphi) = \emptyset$, then for all n and y, the condition $\exists y < 2^{n^2} \varphi'(n,y)$ is false, hence $\text{SPEC}(\psi_4) = \mathbb{N}^+$ (hence is cofinite). Conversely, if $\text{SPEC}(\varphi) \neq \emptyset$, then the integers greater than 2^{n^2} with $n \in \text{SPEC}(\varphi)$ are not in $\text{SPEC}(\psi_2)$, which is not cofinite.
- Let $\psi_5' \equiv \exists n, m < \max(\max = 2^{n^2} \times m^2 \wedge \exists y < 2^{n^2} \varphi'(n,y) \wedge \forall n' < n \forall y' < 2^{n^2} \neg\varphi'(n',y')))$.

  If $\text{SPEC}(\varphi) = \emptyset$, then $\text{SPEC}(\psi_5) = \emptyset$ (hence is ultimately periodic). Conversely, observe that $\text{SPEC}(\psi_5) = \{n_0 \times m^2 \mid m \in \mathbb{N}+\}$, where $n_0 = \inf(\text{SPEC}(\varphi))$. Hence $\text{SPEC}(\psi_5)$ is not ultimately periodic.

$\dashv$

3.5. Real numbers and spectra. Let $\chi_\phi(n)$ be the characteristic function of the spectrum of a first order sentence ϕ. We can associate with ϕ and $a \in \mathbb{Z}$ the real number $r_\phi = a + \sum_n \chi_\phi(n) 2^{-n}$.

Definition 3.7. A real number is *first order spectral* if it is of the form r_ϕ for some $a \in \mathbb{Z}$ and some first order sentence ϕ.

As we have noted ultimately periodic sets of natural numbers are first order spectra, and correspond to rational numbers. Also every ultimately periodic spectrum can be realized by a formula with one function symbol only. We have

Proposition 3.8. *Every rational number q is first order spectral using a formula with one function symbol only.*

Question 3.9. *Do the the first order spectral reals form a field?*

E. Specker, [122], proved that there is a real x primitive recursive in base 2 such that $3x$, $x + \frac{1}{3}$, x^2 are not primitive recursive in base 2. A modern treatment can be found in [16]. H. Friedman [51] mentioned on an internet discussion site that primitive recursive can be replaced in Specker's Theorem by much lower complexity within the Grzegotczyk Hierarchy. J. Miller kindly provided us, [93], with the more precise statement

THEOREM 3.10 (E. Specker, 1949 and H. Friedman 2003).
There is a real x which is primitive recursive in base 2 (and can even be taken to be in $\mathcal{E}^2$ of the Grzegorczyk Hierarchy), such that $3x$, $x + \frac{1}{3}$, x^2 are not primitive recursive in base 2.

SKETCH OF PROOF[3]. This can be proved by exploiting the fact that none of these functions $3x, x^2, x + 1/3$ are continuous as functions on binary expansions of reals. They can take a number x that is not a binary rational to one that is.

Let us focus on $3x$. So, for example, if $x = 0.0101010101\ldots$, then $3x = 1$. We can exploit this as follows. Say we have built the binary expansion of x up to position $n - 1$ and it looks like $0.b$ (where b is a finite string) and that we want to diagonalize against the ith primitive recursive function p_i. Compute $p_i(n)$, step by step. As long as it does not converge, keep building x to look like $0.b00101010101010101\ldots$. If $p_i(n)$ converges at stage s, then use position $n + 2s$ or position $n + 2s + 1$ to spring the delayed trap. If $p_i(n) = 0$, then let $x = 0.b0010101\ldots01011$. If $p_i(n) = 1$, then let $x = 0.b0010101\ldots0100$. Either way, $p_i(n)$ does not correctly compute the nth bit of $3x$. Note that we spread out the unbounded search, so that each bit is computed by a bounded (primitive recursive) procedure.

In this way we can diagonalize against $3x, x^2$ and $x+1/3$ being primitive recursive while making x primitive recursive. To make x to be in $\mathcal{E}^2$ one uses the fact that $\mathcal{E}^2$ is the same as computable in linear space [110].

In the argument above, when we are trying to figure out bit t of x, we compute $p_i(n)$ (for some i and n determined earlier in the construction of x) for t steps and if it does not halt we output the default bit (alternating between 0 and 1), so it can be made in linear time. Actually, in the case of x^a2, one has to work a little harder to determine the default bit, but this can definitely be done in linear space (and polynomial time). $\dashv$

We shall see in Section 4, Theorem 4.11, that Theorem 3.10 covers all the spectral reals, therefore the spectral reals do not form a field. More precisely we have the following Corollary:

COROLLARY 3.11. *The spectral reals are not closed under addition nor under multiplication. Furthermore, they are closed under the operation $1 - x$ iff the complement of a spectrum is a spectrum.*

We now turn the question of algebraicity and transcendence of spectral reals. Clearly, every first order spectral real is a recursive real in the sense of A. Turing [127]. Using Liouville's Theorem[4], we can see that many transcendental reals are first order spectral.

[4]Liouville's Theorem states, in simplified form, that a real of the form $r = \sum_n 2^{-f(n)}$ where $f(n) \geq n!$ is transcendental.

OPEN QUESTION 11. *Are there any irrational algebraic reals which are spectral?*

One way of analyzing irrational numbers is by counting the number of 1s in their binary representation. For a real $r \in (0,1)$ let $\gamma_r(n)$ be the number of 1s among its first n digits. If $r = r_\phi$ is spectral we have $\gamma_r(n) = \gamma_{phi}(n)$.

In the sequel we follow closely and quote from M. Waldschmidt [129].

THEOREM 3.12 (Bailey, Borwein, Crandall, and Pomerance, 2004, [8]). *Let r be a real algebraic number of degree $d \geq 2$. Then there is a positive number $C_{r,d}$, which depends only on r, such that $\gamma_r(n) \geq C_{r,d} n^{\frac{1}{d}}$.*

In other words, if a spectral number r_ϕ is algebraic of degree $d \geq 2$, then $\gamma_\phi(n) \geq C_{\phi,d} n^{\frac{1}{d}}$, for some positive number $C_{\phi,d}$.

To get more information about irrational numbers r we have to look at the binary string complexity of $r \in (0,1)$. We consider r as an infinite binary word.

DEFINITION 3.13 (Binary string complexity). The *binary string complexity of* r is the function $p_r(m)$ which counts, for each m the number of distinct binary words w of length m occuring in r. Hence we have $1 \leq p_r(m) \leq 2^m$, and the function $p_r(m)$ is non-decreasing.

CONJECTURE 12 (E. Borel 1950, [12]). The binary string complexity of an irrational algebraic number r should be $p_r(m) = 2^m$.

DEFINITION 3.14. We call a real number $r \in (0,1)$ *automatic* if the n-th bit of its binary expansion can be generated by a finite automaton from the binary representation of n.

Clearly, the binary string complexity of an automatic real is $O(m)$.

OPEN QUESTION 13. *Is every automatic real a spectral real?*

In 1968 A. Cobham, [20] conjectured that automatic numbers are transcendental. This was proven in 2007 by B. Adamczewski and Y. Bugeaud, [1]. They actually proved a stronger theorem.

THEOREM 3.15 (B. Adamczewski and Y. Bugeaud, 2007). *The binary string complexity $p_r(m)$ of a real irrational algebraic number r satisfies*

$$\liminf_{m\to\infty} \frac{p_r(m)}{m} = +\infty$$

Borel's Conjecture would imply that the binary string complexity $p_r(m)$ of a real irrational algebraic number r satisfies

$$\liminf_{m\to\infty} \frac{p_r(m)}{2^m} = 1$$

Open Question 14. *Does the binary string complexity $p_r(m)$ of a spectral real r satisfy*

$$\liminf_{m\to\infty} \frac{p_r(m)}{2^m} < 1$$

or even

$$\liminf_{m\to\infty} \frac{p_r(m)}{2^m} = 0?$$

From Theorem 3.15 one gets that the Fibonacci numbers, which will be shown to form a spectrum in Corollary 4.12 of Section 4.3, give us a transcendental spectral number. More generally, we get the following:

Proposition 3.16. *Let r_ϕ be a spectral real such that the gap function $\delta_\phi(n)$ is monotonically increasing and grows exponentially. Then $p_\phi(n) = O(n)$. Therefore, r_ϕ is transcendental.*

The analysis of computable reals in binary or b-adic presentation is tricky because of the behaviour of the carry, cf. [17]. Let $\mathcal{F}$ be a class of functions $f : \mathbb{N} \to \mathbb{N}$.

Definition 3.17.
A real number α is called *$\mathcal{F}$-Cauchy computable* if there are functions $f, g, h \in \mathcal{F}$ such that for

$$r_n = \frac{f(n) - g(n)}{h(n) + 1}$$

we have that for all $n \in \mathbb{N}$

$$\mid r_n - \alpha \mid \leq \frac{1}{n+1}.$$

A real number $\alpha \in [0, 1]$ in b-adic presentation is called *$\mathcal{F}$-computable* if there is $f : \mathbb{N} \to \{0, \dots, b-1\}$ such that

$$\alpha = \sum_{n\in\mathbb{N}} f(n) b^{-n}$$

Note that it is not clear at all how to define spectral Cauchy reals. If $\mathcal{F}$ contains the function 2^n then the $\mathcal{F}$ computable reals in b-adic presentation are also $\mathcal{F}$-Cauchy computable. In particular, this is true for $\mathcal{F} = \mathcal{E}^i$ and $i \geq 3$.

Open Question 15. *Are the b-adic $\mathcal{E}^2$-computable reals $\mathcal{E}^2$-Cauchy computable?*

Recently, $\mathcal{E}^2$-Cauchy computable reals have received quite a bit of attention, cf. [119, 120]. The following summarizes what is known.

Proposition 3.18 (D. Skordev). (i) *The $\mathcal{E}^2$-Cauchy computable reals form a real closed field.*

(ii) *The transcendental numbers e and* π*, and the Euler constant* γ *and the Liouville number* $\sum_{n\in\mathbb{N}} 10^{a}-n!$ *are* $\mathcal{E}^2$*-Cauchy computable.*
(iii) *There are* $\mathcal{E}^3$*-Cauchy computable reals which are not* $\mathcal{E}^2$*-Cauchy computable.*

Let $\mathcal{F}_{low}$ be the smallest class of functions in $\mathcal{E}^2$ which contains the constant functions, projections, successor, modified difference, and which is closed under composition and bounded summation. A real α is *low* if α is $\mathcal{F}_{low}$-Cauchy computable. The low reals also form a real closed field. In [125] low reals are studied and some very deep theorems about low transcendental numbers are obtained, the discussion of which would take too much space.

Open Question 16. *Is the inclusion* $\mathcal{F}_{low} \subseteq \mathcal{E}^2$ *proper?*

§4. Approach I: Recursion Theory.

This approach has generated all in all four papers (namely [7] by G. Asser in 1955, [98] by A. Mostowski in 1956, [110] by R. Ritchie, and [95] by S. Mo in 1991) and two Ph.D. dissertations, namely [109] by R. Ritchie in 1960 and [9] by J. Bennett in 1962. These works share the common feature of being hardly available for many readers on various grounds: Asser's and Mostowski's papers are difficult to read because they are more than fifty years old and Asser's paper is in German. Bennett's thesis, cited in many papers, is almost equally old and in addition has remained unpublished. Finally, Mo's paper, though more recent, is in Chinese. This is the reason why we propose in Section A a detailed review of these references, including several sketches of proofs in modern language. In the present section, after some background material, we present a synthetic survey of the recursive approach of the spectrum problem.

4.1. Grzegorczyk's Hierarchy. For a detailed presentation of the material in this subsection, see eg. [114]. A. Grzegorczyk's seminal paper [65] about classification of primitive recursive functions was published in 1953, one year after Scholz's question, and two years before Asser's paper. Hence, Grzegorczyk's Hierarchy was not the standard way to consider primitive recursive functions in the mid-fifties. And actually, G. Asser and A. Mostowski deal with recursive aspects of spectra, but not explicitly with Grzegorczyk's classes, though it is the usual framework in which their results are presented. It is only in J. Bennett's thesis in 1962 and especially in S. Mo's paper in 1991 that one finds an explicit study of spectra in terms of Grzegorczyk's classes.

In the sequel a function is always intended to be a function from some $\mathbb{N}^k$ to $\mathbb{N}$ (total, unless otherwise specified).

Definition 4.1 (Elementary functions). The class $\mathcal{E}$ of elementary functions is the smallest class of functions containing the zero, successor, projections, addition, multiplication and modified subtraction functions and which is closed under composition and bounded sum and product (i.e. $f(n,\vec{x}) = \sum_{i=0}^{n} g(i,\vec{x})$ and $f(n,\vec{x}) = \prod_{i=0}^{n} g(i,\vec{x})$, with previously defined g). We denote by $\mathcal{E}_\star$ the elementary relations, i.e. the class of relations whose characteristic functions are elementary.

The class $\mathcal{E}$ was introduced by Kalmár [81] and Csillag [28] in the forties, and contains most usual number-theoretic functions. It also corresponds to Grzegorczyk's class $\mathcal{E}^3$, that we define below.

DEFINITION 4.2 (Primitive recursion). Let f, g, h be functions. We say that f is defined from g and h by primitive recursion when it obeys a schema: $\begin{cases} f(0,\vec{x}) & = & g(\vec{x}) \\ f(n+1,\vec{x}) & = & h(n,\vec{x},f(n,\vec{x})) \end{cases}$.

The class of primitive recursive functions, denoted by $\mathcal{P}R$, is the smallest class of functions containing the zero function, the successor function, the projection functions, and which is closed under composition and primitive recursion.

For instance, elementary functions are primitive recursive. The following binary function Ack, known as Ackermann's function, is provably not primitive recursive, whereas all unary specialised functions $Ack_x : y \mapsto Ack(x,y)$ are primitive recursive:

$$\begin{cases} Ack(0,y) = y+1 \\ Ack(x+1,0) = Ack(x,1) \\ Ack(x+1,y+1) = Ack(x, Ack(x+1,y)) \end{cases}$$

In order to introduce Grzegorczyk's hierarchy, we need a weaker version of primitive recursion, in which the newly defined functions have to be bounded by some previously defined function.

DEFINITION 4.3 (Bounded recursion). Let f, g, h, j be functions. We say that f is defined from g, h and j by bounded recursion when it obeys a schema: $\begin{cases} f(0,\vec{x}) & = & g(\vec{x}) \\ f(n+1,\vec{x}) & = & h(n,\vec{x},f(n,\vec{x})) \\ f(n,\vec{x}) & \leq & j(n,\vec{x}) \end{cases}$

Let f_n ($n = 0, 1, 2, \dots$) be the following sequence of primitive recursive functions :

- $f_0(x,y) = y+1$,
- $f_1(x,y) = x+y$,
- $f_2(x,y) = (x+1)\cdot(y+1)$,
- and for $k \geq 0$ $\begin{cases} f_{k+3}(0,y) & = f_{k+2}(y+1,y+1) \\ f_{k+3}(x+1,y) & = f_{k+3}(x, f_{k+3}(x,y)) \end{cases}$

Roughly speaking, the important feature is that the functions f_n are more and more rapidly growing. Several other similar sequences of increasingly growing functions can be used to define Grzegorczyk's classes.

DEFINITION 4.4 (Grzegorczyk's hierarchy). The Grzegorczyk's class $\mathcal{E}^n$ is the smallest class of functions containing the zero function, the projections functions and f_n and which is closed under composition and bounded recursion. The associated classes of relations $\mathcal{E}^n_\star$ are defined as the class of relations on integers with a characteristic function in $\mathcal{E}^n$.

Note that, for sake of simplicity, we use the same notation for a class of relations of various arities (eg. $\mathcal{E}^3_\star$) and the class of unary relations (i.e.

sets) it contains. Which one is intended will always be clear from the context.

The main features of Grzegorczyk's classes were studied by A. Grzegorczyk in [65] and by R. Ritchie in [110].

THEOREM 4.5 (A. Grzegorczyk (1953)).

- *The functional hierarchy is strict for $n \geq 0$, i.e. we have $\mathcal{E}^n \subsetneq \mathcal{E}^{n+1}$.*
- *The relational hierarchy is strict for $n \geq 3$, i.e. we have $\mathcal{E}^n_\star \subsetneq \mathcal{E}^{n+1}_\star$.*
- *For the initial levels of the relational hierarchy, we have $\mathcal{E}^0_\star \subseteq \mathcal{E}^1_\star \subseteq \mathcal{E}^2_\star \subseteq \mathcal{E}^3_\star$.*
- *The Kalmár-Csillag class of elementary functions $\mathcal{E}$ is equal to $\mathcal{E}^3$.*
- *Finally, the full hierarchy corresponds to primitive recursion, i.e. $\mathcal{P}R = \bigcup_{n=0}^{+\infty} \mathcal{E}^n$.*

THEOREM 4.6 (R. Ritchie (1963)). *We have $\mathcal{E}^2_\star \neq \mathcal{E}^3_\star$ [110].*

Note that the possible separation of the relational classes $\mathcal{E}^0_\star, \mathcal{E}^1_\star, \mathcal{E}^2_\star$ is still an open question.

OPEN QUESTION 17. *Are the inclusions in*

$$\mathcal{E}^0_\star \subseteq \mathcal{E}^1_\star \subseteq \mathcal{E}^2_\star$$

proper?

An important point is that the functional hierarchy deals with the rate at which functions may grow: intuitively, functions in the low level of the hierarchy grow very slowly, while functions higher up in the hierarchy grow very rapidly. However, this feature does not hold for the relational hierarchy, because characteristic functions do not grow at all (they are $0-1$ valued). For instance, the ternary relations $z = x+y$, $z = x \times y$, $z = x^y$ as well as $z = Ack(x, y)$ all belong to $\mathcal{E}^0_\star$, whereas the corresponding functions provably do not lie in $\mathcal{E}^0$.

4.2. Rudimentary relations and strictly rudimentary relations. In addition to primitive recursive classes of relations, we also introduce two new classes of relations with an arithmetical flavour, namely the rudimentary and strictly rudimentary relations. These classes were originally introduced by R. Smullyan [121], and a major reference about rudimentary relations and subclasses is J. Bennett's thesis [9].

DEFINITION 4.7 (Rudimentary relations). Denote by RUD the smallest class of relations over integers containing the graphs of addition and multiplication (seen as ternary relations) and closed under Boolean operations ($\neg$, $\wedge$, $\vee$) and bounded quantifications ($\forall x < y \ldots$ and $\exists x < y \ldots$).

In spite of its very restricted definition, the class RUD is surprisingly robust (eg. it has several equivalent definitions in the fields of computational complexity, recursion theory, formal languages etc.) and large. For instance, the following formula defines the set of prime numbers:

$$x > 1 \ \wedge \ \forall y < x \ \forall z < x \ \neg(\ x = y.z \ )$$

More (sometimes VERY) sophisticated formulas prove that the ternary relation $z = x^y$ is rudimentary (Bennett [9]), as well as the graph $z = Ack(x, y)$ of Ackermann's function (Calude [13]), which is not primitive recursive (as a function), or the four-ary relation $x^y \equiv z \ [mod \ t]$ (Hesse, Allender, Barrington [71]). Actually, we are not aware of a natural number theoretic relation which is provably not rudimentary.

The following is easy to see.

PROPOSITION 4.8. RUD $\subseteq \mathcal{E}^0_\star \subseteq \mathcal{E}^1_\star \subseteq \mathcal{E}^2_\star$.

However, the equality remains an open question (and would imply RUD $= \mathcal{E}^2_\star$ as well, since the closure of $\mathcal{E}^0_\star$ by polynomial substitution is $\mathcal{E}^2_\star$ whereas RUD is closed under polynomial substitution).

OPEN QUESTION 18. *Are the inclusions in* RUD $\subseteq \mathcal{E}^0_\star \subseteq \mathcal{E}^1_\star \subseteq \mathcal{E}^2_\star$ *proper?*

It remains to introduce the strictly rudimentary relations. Let us consider the dyadic representation of integers, i.e. $n \in \mathbb{N}$ is represented by a word in $\{1, 2\}^*$. Compared to binary notation, dyadic notation avoids the problem of leading 0s and yields a bijection between integers and words. When integers are seen as words, it is natural to consider subword quantifications instead of ordinary bounded quantification. We say that $w = w_1 \dots w_k$ is a subword of $v = v_1 \dots v_p$ and we denote $w \upharpoonright v$ when there exists $1 \leq i \leq p$ such that $w_1 = v_i, \dots, w_k = v_{i+k-1}$. Of course, if $x \upharpoonright y$, then $x \leq y$.

DEFINITION 4.9 (Strictly rudimentary relations). Denote by SRUD the smallest class of relations over integers containing the graphs of dyadic concatenation (seen as a ternary relation) and closed under Boolean operations ($\neg$, $\wedge$, $\vee$) and subword quantifications ($\forall x \upharpoonright y \dots$ and $\exists x \upharpoonright y \dots$).

There are only few examples of strictly rudimentary relations, e.g. x begins (or ends or is a part of) y (as dyadic words), $x = y$, the dyadic representation of x is a single symbol, the dyadic representation of x contains only one type of symbol. On the other hand, several relations are provably not strictly rudimentary such as $x \leq y$, $x = y + 1$, x and y have the same dyadic length and the dyadic representation of x is of the form $1^n 2^n$ for some n (V. Nepomnjascii 1978, see [101]).

Note that rudimentary relations were originally (and equivalently) defined by Smullyan [121] using dyadic concatenation as a basis relation instead of addition and multiplication. Clearly, we have SRUD $\subsetneq$ RUD.

4.3. Recursive and arithmetical characterizations of spectra. In the fifties and sixties, following the tastes of their time, logicians aim at characterizing SPEC via recursion and arithmetics. Typically, they wished to obtain the characteristic functions of spectra as the 0-1-valued functions in a class defined by closure of a certain set of simple functions under certain operators (such as composition or various recursion schemas). From this point of view, their results are not totally satisfactory because they are either partial, or somehow cumbersome or unnatural. However, these studies show that the class of spectra is very broad, and that most classical arithmetical sets are spectra.

The class SPEC is set within Grzegorczyk's hierarchy (by G. Asser in [7] and A. Mostowski in [98]), from which we can deduce that all rudimentary sets are spectra.

THEOREM 4.10 (G. Asser (1955)). SPEC $\subsetneq \mathcal{E}^3_\star$

Asser's theorem is based on a rather complicated and artificial arithmetical characterization of spectra (see Subsection A.1). In particular, Asser's construction is of no help in proving that a particular set is (or not) spectrum.

Though he actually uses a slightly different class (see Subsection A.2), the following result is usually attributed to Mostowski:

THEOREM 4.11 (A. Mostowski (1956)). $\mathcal{E}^2_\star \subseteq$ SPEC

Note that equality in Mostowski's theorem remains an open question.

OPEN QUESTION 19. *Is the inclusion in $\mathcal{E}^2_\star \subseteq$ SPEC proper?*

The following corollary is not stated by A. Mostowski, but can be found in J. Bennett's thesis. It is worth noting because one of the most fruitful ways in proving that various arithmetical sets are spectra is to prove that they are actually rudimentary.

COROLLARY 4.12. RUD $\subseteq$ SPEC

For instance, any set defined by a linear or polynomial recurrence condition, such as the Fibonacci numbers (i.e. those numbers appearing in the sequence defined by $u_0 = u_1 = 1$ and $u_{n+2} = u_n + u_{n+1}$), is rudimentary (see [39]). From Corollary 4.12, we deduce that such sets are spectra, as announced in Section 3. Similarly, using the fact that the set of prime numbers is rudimentary and the exponentiation has a rudimentary graph, one proves that the sets of Fermat primes (of the form $2^{2^n}+1$), Mersenne primes (of the form $2^p - 1$ with p a prime), or twin primes (p prime such that $p + 2$ is also a prime) are rudimentary (hence also spectra).

Note that the question of whether the inclusion in Corollary 4.12 is proper is still open.

Open Question 20. *Do we have* Rud = Spec*?*

This problem is further investigated in Subsubsection 6.3.3.

An arithmetic characterization of Spec in terms of strictly rudimentary relations is also given, among many other results (see Subsection A.3), by J. Bennett in his thesis.

Theorem 4.13 (J. Bennett (1962)).
A set $S \subseteq \mathbb{N}$ is in Spec *iff it can be defined by a formula of the form $\exists y \leq 2^{x^j} R(x, y)$ for some $j \geq 1$, where R is in* Srud*. i.e.,*

$$S = \{x \in \mathbb{N} \mid \exists y \leq 2^{x^j} R(x, y)\}$$

for some $R \in$ Srud *and $j \geq 1$.*

J. Bennett also characterizes spectra of higher order logics and shows that the union of spectra of various orders equals the class of elementary relations $\mathcal{E}^3_\star$.

The characterization of spectra stated in Theorem 4.13 is rather simple and elegant. However, once again, it is not really useful in proving that a given set is a spectrum, now because Srud is very restrictive. A somehow similar characterization of Spec using Rud instead of Srud would have been more powerful - but, one gets this way second-order spectra.

Finally, let us note a late paper on the recursive aspect of spectra, namely [95], due to the Chinese logician Mo Shaokui in 1991 (see Subsection A.4). The solution to Scholz's problem proposed there is of the same type as Bennett's characterization. However, the only bibliographic references in Mo's paper are Scholz's question [116] and Grzegorczyk's paper [65], so that it can be considered completely independent from all other contributions about spectra. Section A summarises this paper's results.

§5. Approach II: Complexity Theory. The spectrum problem, formulated in the early 1950s, predates complexity theory since the notions of time or space bounded Turing machines first emerged in the 1960s (see [70, 83]). However, the first results about complexity of spectra appeared very soon (see Subsection 5.3), and computational complexity characterisations of spectra were found, in at least three independent early contexts (see Subsection 5.4). Later on, several refinements and developments of these seminal results have been published (see Subsections 5.7 and 5.8).

Turing machines and other standard models of computation operate on words, not on numbers. Let $L \subseteq \Sigma^*$ be a set of finite words over a fixed finite alphabet Σ. Without loss of generality we assume $\Sigma = \{0, 1\}$, and that input words have no leading zeros.

The archetypical task, given a language L, is to study the complexity of deciding membership in L of a word x as a function of the length $|x|$, i.e., asymptotic growth rate of the time, space or other computational resources needed to decide whether $x \in L$.

5.1. Complexity and spectra. In this section, for a fixed sentence ϕ, the set of natural numbers $\mathrm{SPEC}(\phi)$ is seen as the set of positive instances of a decision problem (given a number n, is there a model of ϕ with n elements?).

When dealing with computational complexity, we convert spectra (sets of natural numbers) into languages (over alphabet $\{0, 1\}$). The spectrum problem can thus be rephrased as: What is the computational complexity of the decision problems for spectra?

Complexity classes. Denote by $\mathrm{NTIME}(f(n))$ (resp. $\mathrm{DTIME}(f(n))$) the class of binary languages accepted in time $O(f(n))$ by some non-deterministic (resp. deterministic) Turing machine, where n is the length of the input. Similarly, let us denote by $\mathrm{DSPACE}(f(n))$ the class of languages accepted in space $O(f(n))$ by some deterministic Turing machine. Some well-known complexity classes which concern us here are:

$$\mathrm{L} = \mathrm{DSPACE}(\log n) \subseteq \mathrm{NL} = \mathrm{NSPACE}(\log n)$$

$$\mathrm{LINSPACE} = \mathrm{DSPACE}(n) \subseteq \mathrm{NLINSPACE} = \mathrm{NSPACE}(n)$$

$$\mathrm{P} = \bigcup_{c \geq 1} \mathrm{DTIME}(n^c) \subseteq \mathrm{NP} = \bigcup_{c \geq 1} \mathrm{NTIME}(n^c)$$

$$\mathrm{E} = \bigcup_{c \geq 1} \mathrm{DTIME}(2^{c \cdot n}) \subseteq \mathrm{NE} = \bigcup_{c \geq 1} \mathrm{NTIME}(2^{c \cdot n})$$

Finally, if C denotes a complexity class, we denote its complement class, i.e. the class of binary languages L such that $\Sigma^* - L \in \mathrm{C}$, by coC.

Of course, the perennial open questions are:

OPEN QUESTION 21. (i) *Are any of the inclusions*
$L \subseteq NL$, LINSPACE $\subseteq$ NLINSPACE, $P \subseteq NP$ *and* $E \subseteq NE$ *proper?*
(ii) *Do any of the equalities* NP = coNP *and* NE = coNE *hold?*

Surprisingly, the following was shown independently by N. Immermann and R. Szelepcźenyi in 1982, cf. [76]:

THEOREM 5.1 (Immermann, Szelepcźenyi 1982).
NL = coNL *and* NLINSPACE = coNLINSPACE.

In Section 6 we shall also make use of the polynomial time hierarchy PH and its linear analogue LTH.

The class RUD lies between L and LINSPACE, and must be different from one of them.

PROPOSITION 5.2.
(i) (Nepomnjascii 1970, [100]) $L \subseteq$ RUD
(ii) (Wrathall 1978, [132]) RUD = LTH
(iii) (Myhill 1960, [99]) LTH $\subseteq$ LINSPACE

OPEN QUESTION 22.
Are the inclusions $L \subseteq$ RUD = LTH $\subseteq$ LINSPACE *proper?*

Number representations by binary or unary words. It is natural to use binary notation for natural numbers (an alternative without leading zeros is Smullyan's *dyadic* notation [121]). The shortest binary length and dyadic length of the natural number n are very close to $\lceil \log_2 n \rceil$, whereas its unary length is of course n, and we have $n = 2^{\log_2 n}$. Consequently, the same (mathematical) computation that is performed by some Turing machine in time eg. $O(2^{c \cdot |n|})$ when $|n|$ is the binary length of the natural number input, is also performed (by a slightly different Turing machine) in time $O(n^c)$ when n is the (unary length of the) natural number input.

Unary notation (also called *tally* notation, i.e. the number n is represented by the word $1 \ldots 1$ composed of n ones) also has its fans, for reasons explained in the description of Fagin's work. Most results in this section may be stated in either notation, but for sake of simplicity, and unless explicitly stated otherwise, we use binary notation. The length of a binary or unary word x is written $|x|$.

Recall that SPEC denotes the set of spectra of first-order sentences, i.e.,

$$\text{SPEC} = \{\text{spec}(\phi) \mid \phi \text{ is a first-order sentence}\}$$

5.2. Spectra, formal languages, and complexity theory. Formal language theory was much studied in the early 1960s, cf. [69], in particular the Chomsky hierarchy. While the regular and context-free language classes were well-understood, several questions remained open for larger classes. We need here the following:

THEOREM 5.3.
(i) (Ritchie 1963, [110]) $\mathcal{E}^2_\star = \text{LINSPACE}$
(ii) (Kuroda 1964, [83]) *A language* L *is* context sensitive *iff* $L \in \text{NLINSPACE}$.

For our discussion one should remember that at that time it was then (as now) unknown whether LINSPACE = NLINSPACE and also unknown whether NLINSPACE was closed under complementation. The latter was only resolved positively more than 20 years later, see Theorem 5.1.

These open questions showed a tantalising similarity to Scholz' and Asser's questions. If we identify characteristic functions with sets, then Bennett's 1962 thesis combined with Asser, Mostowski and Ritchie's results, yield

$$\text{LINSPACE} \subseteq \textsc{Spec} \subseteq \mathcal{E}^3_\star \text{ and } \text{LINSPACE} \subseteq \text{NLINSPACE} \subseteq \mathcal{E}^3_\star.$$

This led to a conjecture $\textsc{Spec} \stackrel{?}{=} \text{NLINSPACE}$, that spectra might be coextensive to the context sensitive languages. The analogy fails, though, since more than n "bits of storage" are needed to store an n-element model $\mathcal{M}$ of a sentence ϕ.

5.3. An early paper. One of the first papers explicitly relating spectra to bounded resource machine models of computation is [113] (in German), due to Rödding and Schwichtenberg from Münster in 1972. This switch from recursion theory to complexity theory had been prepared ten years before by Bennett and Ritchie, and Rödding and Schwichtenberg made a step further. The model of computation they use is not Turing machines, but register machines. As Bennett does, Rödding and Schwichtenberg not only consider spectra of first-order sentences, but also higher order spectra, namely spectra of sentences using i-th order variables. Let us denote by $\textsc{ho}-\textsc{spec}_i$ the class of spectra of sentences using i-th order variables. Let us define the following sequence of functions : let $exp_0(n) = n$, and $exp_{i+1}(n) = 2^{exp_i(n)}$. Along with other results in the field of recursion theory, Rödding and Schwichtenberg prove the following theorem.

THEOREM 5.4 (Rödding and Schwichtenberg 1972 [113]). *For all* $i \in \mathbb{N}$, *we have* $\text{DSPACE}(exp_i(n)) \subseteq \textsc{ho}-\textsc{spec}_{i+1}$.

In particular, taking $i = 0$, first-order spectra are thereby shown to contain DSPACE(n). Let us finally remark that Rödding and Schwichtenberg did not consider non-deterministic complexity classes.

5.4. First-order spectra and non-deterministic exponential time. Scholz's original question (see [116]) was finally answered after twenty years, when Jones and Selman related first-order spectra to non-deterministic time bounded Turing machines. Their result was first published in a conference version in 1972 (see [79]), and the journal version appeared in 1974 (see [80]). The following theorem holds.

THEOREM 5.5 (Jones and Selman 1972 [79]). SPEC = NE.

This leads to a complexity theory counterpart of Asser's question:

COROLLARY 5.6. SPEC = coSPEC *if and only if* NE = coNE.

They note that this does not answer Asser's question, but it shows the link with a wide range of closure under complement questions, in complexity theory. Presently, we know that many of them are very difficult questions.

PROOF IDEAS. To see that SPEC $\subseteq$ NE, consider SPEC(ϕ) $\in$ SPEC. Since the sentence ϕ is fixed, satisfaction $\mathcal{M} \models \phi$ can be decided in time that is at most polynomial in the size of model $\mathcal{M}$, where the polynomial's degree depends on the quantifier nesting depth in ϕ. A simple guess-and-verify algorithm is: given number x, non-deterministically guess an x-element model $\mathcal{M}$, then decide whether $\mathcal{M} \models \phi$ is true. Time and space $2^{c \cdot n}$ suffice to store an x-element model and check $\mathcal{M} \models \phi$, where constant c is independent of x and n is the length of x's binary notation. Thus the algorithm works in non-deterministic exponential time (as a function of input length n).

To show SPEC $\supseteq$ NE, let Z be a nondeterministic time-bounded Turing machine that runs in time $2^{c \cdot n}$ on inputs of length n. Here the input is a word x of length n. We think of x as a binary-coded natural number. Computation C can be coded as a word $C = \mathit{config}_0 \mathit{config}_1 \ldots \mathit{config}_{2^{c \cdot n}}$ where config_0 contains the Turing machine's input, and each config_t encodes the tape contents and control point at its t-th computational step.

Now we have to find a first-order sentence ϕ such that:

(i) For every input-accepting Z computation, $\mathcal{M} \models \phi$ for some model $\mathcal{M}$ of cardinality x

(ii) If Z has no computation that accepts its input, then ϕ has no model of cardinality x

Each config_t length is at most $2^{c \cdot n}$, so C has length bound $2^{c' \cdot n} = x^{c'}$ for c' independent of n. A model $\mathcal{M}$ of cardinality x contains, for each k-ary predicate symbol P of sentence ϕ, a relation $\overline{P} \subseteq \{0, 1, \ldots, x-1\}^k$. Thus a model can in principle "have enough bits" $x^k = 2^{k \log x} = 2^{O(|x|)}$ to encode all the symbols of computation C.

The remaining task is to actually construct ϕ so it has a model $\mathcal{M}$ of cardinality x if and only if Z has a well-formed computation C that accepts input x. In effect, the task is to use predicate logic to check that C is well-formed and accepts x. The technical details are omitted from this survey paper; some approaches may be seen in [80, 40, 18] $\dashv$

5.5. Relationship to the question $\mathrm{P} \stackrel{?}{=} \mathrm{NP}$**.** Let $\mathrm{UN} = \{L | L \subseteq 1^*\}$ be the set of *tally languages* (each is a set of unary words over the one-letter alphabet $\{1\}$) and let $\mathrm{NP}_1 = \mathrm{NP} \cap \mathrm{UN}$. Since there is a natural identification between NP_1 and NE, we can deduce that if $\mathrm{P} = \mathrm{NP}$, then $\mathrm{NP} = \mathrm{coNP}$ and $\mathrm{NE} = \mathrm{coNE}$, i.e. the complement of a spectrum is a spectrum. Of course, it also holds that if there is a spectrum whose complement is not a spectrum, i.e. if $\mathrm{NE} \neq \mathrm{coNE}$, then $\mathrm{NP} \neq \mathrm{coNP}$ and $\mathrm{P} \neq \mathrm{NP}$. The converse implication remains open.

5.6. Independent solutions to Scholz's problem. The characterization of spectra via non-deterministic complexity classes was independently found also by Christen on the one hand and Fagin on the other hand during their PhD studies.

Claude Christen's thesis[5] [18] (1974, ETH Zürich, E. Specker) remains unpublished, and only a small part was published in German [19]. Christen discovered all his results independently, and only in the late stage of his work his attention was drawn to Bennett's work [9] and the paper of Jones and Selman [79]. It turned out that most of his independently found results were already in print or published by Fagin after completion of Christen's thesis.

Ronald Fagin's thesis (1973, UC Berkeley, R. Vaught) is treasure of results introducing *projective classes of finite structures*, which he called *generalized spectra* (see Subsection 5.7) that had wide impact on what is now called descriptive complexity and finite model theory. Most of our knowledge about spectra till about 1985 and, to some extent far beyond that, is contained in the published papers (see [41, 43, 42]) emanating from Fagin's thesis [40]. In this survey, these papers are pervasive. Right now, let us begin with reviewing what is said in [41] about the consequences of the complexity characterization of spectra *per se*.

Recall that $\mathrm{E} = \bigcup_{c \geq 1} \mathrm{DTIME}(2^{c \cdot n}) \subseteq \mathrm{NE}$ and let us examine the closure under complementation problem. Since $\mathrm{E} = \mathrm{coE}$, it is clear that if a first-order spectrum is in E, then its complement is also a first-order spectrum. Of course, the question $\mathrm{E} \stackrel{?}{=} \mathrm{NE}$ is still open. Fagin notes that E contains the spectra of categorical sentences, i.e. sentences that have at most one model for every cardinality. Thus, one obtains a model theoretic question closely related to Asser's question.

Open Question 23. *Is every spectrum the spectrum of a categorical sentence ?*

Besides reviewing many natural sets of numbers that are spectra, Fagin also proves by a complexity argument the existence of a spectrum S such

[5] Claude Christen, born 1943, joined the faculty of CS at the University of Montreal in 1976 and died there, a full professor, prematurely, April 10, 1994.

that $\{n \mid 2^n \in S\}$ is *not* a spectrum (see also [74] for a recent proof by diagonalization).

5.7. Generalized first-order spectra and NP: Fagin's result. Let us spend some time on what is called *generalized first-order spectra* by Fagin in his 1974 paper [41], and is nowadays more usually refered to as (classes of finite structures definable in) *existential second-order logic*. Our main goal is to clarify the differences and the connections with ordinary first-order spectra.

In this subsection, we are no longer interested in the *size* of the finite models of some given sentence, but in the models themselves. Hence, let σ and τ be two disjoint vocabularies, and let ϕ be a first-order $\sigma \cup \tau$-sentence. The *generalized spectrum* of ϕ is the class of finite τ-structures that can be expanded to models of ϕ. In other words, it is the class of finite models of the existential second-order sentence $\exists\sigma\phi$ with vocabulary τ. The vocabulary τ is usually refered to as the *built-in* vocabulary, whereas σ is often called the *extra* vocabulary. Note that generalized spectra are finite counterparts to Tarski's *projective classes* (see [124]). Fagin's theorem states the equivalence between generalized spectra and classes of finite structures accepted in NP.

THEOREM 5.7 (Fagin 1974 [41]). *Let τ be a non-empty vocabulary. A class of finite τ-structures K is a generalized spectrum iff $K \in \mathrm{NP}$.*

If the built-in vocabulary τ is empty, then the τ-structures are merely sets. From a computational point of view, it is natural to see such empty structures as unary representations of natural numbers. From a logical point of view, one obtains ordinary spectra. Hence, Fagin rephrases Jones and Selman's complexity characterization of first-order spectra as follows:

PROPOSITION 5.8. *A set S, if regarded as a set of unary words, is a first-order spectrum if and only if $S \in \mathrm{NP}_1$.*

Concerning the complement problem for generalized spectra, in view of Fagin's theorem, it is not surprising that the general case remains open. However, the following is known. If σ consists of unary predicates only, it is called unary. It has been proved in several occasions that unary generalized spectra are not closed under complement (see Fagin 1975 [42], Hajek 1975 [67], Ajtai and Fagin 1990 [3]). For instance, it is shown in [42] by a game argument that the set of connected simple graphs is not a unary generalized spectrum. In contrast, it is easy to design a monadic existential second-order sentence defining the class of non-connected simple graphs.

Since our survey deals with spectra and not with descriptive complexity as a whole, we will not say any more on this subject. However, let us note that *descriptive complexity* [76] emerged as a specific field of research out of Fagin's paper about generalized spectra.

5.8. Further results and refinements. During the late 1970s and the 1990s, several results were published that generalize Jones and Selman's result to higher order spectra on the one hand, and that refine this result, in order to obtain correspondences between certain complexity classes and the spectra of certain types of sentences.

In 1977, Lovász and Gács [86], it is shown essentially that there are generalized first order spectra such that their complement cannot be expressed with a smaller number of variables. To do this they introduced first order reductions, which became a very important tool in finite model theory and descriptive complexity. In fact they were the first to show the existence of decision problems which are **NP**-complete with respect to first order reductions.

First order reductions were used in (un)decidability results early on, [123], and more explicitely in [107]. For a systematic survey, see [21, 89]. In the context of generalized spectra they were rediscovered independently also by Immerman in [75], Vardi in [128] and Dahlhaus in [30]. First order reductions are of very low complexity, essentially they are uniform $\mathbf{AC}^0$ transductions. The first use of low complexity reduction techniques seems to be Jones [78] who termed them log-rudimentary. Allender and Gore [4] showed them equivalent to uniform $\mathbf{AC}^0$ reductions.

Open Question 24. *Is there a universal (complete) spectrum S_0 and a suitable notion of reduction such that every spectrum S is reducible to S_0?*

Note that this question has two flavors, one where we look at spectra in terms of sets of natural numbers, and one where we look at sets of (cardinalities of) finite models and their defining sentences. First order reductions may be appropriate in the latter case.

In 1982, Lynch [88] relates the computation time needed to decide property on set of integers to the maximal arity of symbols required in the sentence to define this property. Below, we refine the definition of classes of spectra in order to take into account some specific syntactic restrictions on sentences.

Definition 5.9. Let $d \in \mathbb{N}^*$, the following classes are defined as follows.

(i) SPEC_d (resp. F-SPEC_d) is the class of spectra of first-order sentences over arbitrary predicate (resp. predicate and function) symbols of arity at most d.

(ii) $\text{SPEC}(k\forall)$ (resp. $\text{SPEC}_d(k\forall)$) is the class of spectra of prenex first-order sentences involving at most k variables that are all universally quantified (resp. and involving predicate symbols of arity at most d). The classes $\text{F-SPEC}(k\forall)$ and $\text{F-SPEC}_d(k\forall)$ are analogously defined.

(iii) Finally, let $\text{SPEC}_d(+)$ be the class of spectra of first-order sentences over the language containing one ternary relation interpreted as the addition relation over finite segments of $\mathbb{N}$ and predicate symbols of arity at most d

Some inclusions between these classes are easy to obtain: more resources (in terms of arity or number of variables) means more expressive power. Hence, for example, for all $i, j \in \mathbb{N}$ such that $i < j$:

$$\text{SPEC}_i \subseteq \text{SPEC}_j,\ \text{F-SPEC}_i \subseteq \text{F-SPEC}_j \text{ and } \text{SPEC}_i \subseteq \text{F-SPEC}_i.$$

The following results proves a first relationship between time complexity of computation and "syntactic" complexity of definition.

Proposition 5.10 (Lynch 1982 [88]).
For all $d \geq 1$, $\text{NTIME}(2^{d \cdot n}) \subseteq \text{SPEC}_d(+)$.

The converse of this result remains open. It refines the complexity characterization of first-order spectra and has many further developments that we present in Section 6. From a technical point of view, note that Lynch works with so-called "word-models". Namely, a binary word w with length n is seen as a structure with universe $\{0, \dots, n-1\}$, equipped with some arithmetics (eg. successor predicate or addition predicate) and with a unary predicate that indicates the positions of the digits 1 of w. The methods developed in this paper are re-used later on by several authors.

Open Question 25. *Is the inclusion $d \geq 1$,* $\text{NTIME}(2^{d \cdot n}) \subseteq \text{SPEC}_d(+)$ *proper?*

Finally, Lynch explains that, from an attentive reading of Fagin's proof, one can only deduce that if some language L is in $\text{NTIME}(2^{d \cdot n})$, then L is in SPEC_{2d} i.e. is a the spectrum of a first-order sentence involving predicates of arity at most $2d$. Even though Lynch's result is not an exact characterization, but only an inclusion, it has been very influential to other researchers.

In a series of papers published between 1983 and 1990 (see [58, 60, 59, 61, 62, 64]), Grandjean proposes two fruitful ideas. The first one is to use RAM machines as a natural model of computation for general logical structures instead of Turing machines, which are best fitted for languages (or word structures). The second idea is to remark that the time complexity seems closely related to the syntactical form of the sentences (and more specifically in this case with the number of universally quantified variables). Let $\text{NRAM}(f(n))$ be the class of binary languages accepted in time $O(f(n))$ by a non-deterministic RAM (with successor), where n is the length of the input.

THEOREM 5.11 (Grandjean 1983 [60, 61, 62, 64]).
For all $d \geq 1$*, we have* $\mathrm{NRAM}(2^{d \cdot n}) = \text{F-SPEC}(d\forall) = \text{F-SPEC}_d(d\forall)$.

The case $d = 1$ i.e. the case of sentences with one universally quantified variable, is more involved: it requires to encode arithmetic predicates such as linear order or addition that appear intrinsically in the characterization of computation by sentences with one variable. It is developed in the papers [62, 64]). In passing, this implies that the presence of the addition relation is not mandatory in Proposition 5.10 provided (unary) functions are allowed in the language.

An interesting corollary of the latter characterization is that when the number of (universally quantified) variables is fixed, restricting the language to contain function or relation symbols of arity bounded by d only does not weaken the expressive power of sentences and define the same class of spectra. In other words, the following holds.

COROLLARY 5.12 (Grandjean 1983 [60, 61, 62, 64]). *For all* $d \geq 1$*, it holds that* $\text{F-SPEC}(d\forall) = \text{F-SPEC}_d(d\forall)$.

The original proof of this result relies on complexity arguments based on the characterization of $\text{F-SPEC}(d\forall)$. We give here a purely logical proof [6].

PROOF OF COROLLARY 5.12. For simplicity of notation, we give the proof in the case $d = 1$. Let $\varphi \equiv \forall t \Psi$ where Ψ is quantifier-free and whose vocabulary is composed of function symbols of various arities. Let $\mathsf{Term}(\Psi)$ be the set of terms and subterms of Ψ. The first idea is to associate with each element τ of $\mathsf{Term}(\Psi)$ a new unary function f_τ. The definition of f_τ is as follows:

(i) if $\tau = t$ or τ is a constant symbol, then $f_\tau(t) = \tau$,
(ii) if $\tau = f(\tau_1(t), \dots, \tau_k(t))$ for some function symbol f of arity k, then $f_\tau(t) = f(f_{\tau_1}(t), \dots, f_{\tau_k}(t))$.

One obtains a new sentence Ψ' instead of Ψ by replacing each term $\tau \in \mathsf{Term}(\Psi)$ by $f_\tau(t)$ in conjunction with the definition of each function symbol f_τ. Let us explain the transformation on some example.

Let Ψ be the following very simple sentence with f of arity 2 and g of arity 1:

$$\underbrace{E(f(\overbrace{f(t, \underbrace{g(t)}_{\tau_1})}^{\tau_2}, t), t)}_{\tau_3}$$

Then, Ψ' corresponds to:

[6] We thank Étienne Grandjean for kindly giving us this proof.

$$E(f_{\tau_3}(t), f_t(t)) \wedge f_{\tau_3}(t) = f(f_{\tau_2}(t), f_t(t))$$
$$\wedge f_{\tau_2}(t) = f(f_t(t), f_{\tau_1}(t)) \wedge f_{\tau_1}(t) = g(f_t(t)) \wedge f_t(t) = t$$

It is easily seen that the only non unary symbols (here f) appear (if they do at all) only as an outermost symbol in atomic formula. Let now $f(\overline{\sigma_1}(t)), \dots, f(\overline{\sigma_h}(t))$ be the list of terms in Ψ' involving f. The idea is now to replace in Ψ' each $f(\overline{\sigma_i}(t))$ by some new term $F_i(t)$ where F_i is of arity one (let's call this new sentence Ψ'') and to write down the relations between each pair $F_i(t)$ and $F_j(t)$ for $i, j \leq h$. This provides a new sentence $\varphi' = \forall t \forall t' \ (\Psi'' \wedge \Delta)$ where

$$\Delta \equiv \bigwedge_{i,j \leq h} (\overline{\sigma}_i(t) = \overline{\sigma}_j(t') \rightarrow F_i(t) = F_j(t'))$$

The above method shows, when the number of variables is $d = 1$ how to replace h-ary functions by unary functions. However, in order to control the definition of the $F_i s$ we introduce one additional quantified variable. To get rid of this additional variable one can proceed as follows. First, the vocabulary is enriched with a binary predicate $<$ interpreted as a linear order on the domain, and h unary functions $\overline{N}_j$ for $j \leq h$. Let Δ' be the following sentence.

$$\begin{aligned}
&\Delta' \equiv \\
&\forall(i,t)\exists(j,x)\ \overline{N}(j,x) = (\overline{\sigma}_i(t), i, t) \wedge \\
&\forall(j,x) \neq (h, max)\exists(j',x')\ (j',x') = (j,x)+1 \wedge \overline{N}(j,x) < \overline{N}(j',x') \wedge \\
&\forall(j,x) \neq (h, max)\exists(j',x')\exists(i,t)\exists(i',t') \\
&\qquad (j',x') = (j,x)+1 \wedge \overline{N}(j,x) = (\overline{\sigma}_i(t), i, t) \wedge \overline{N}(j',x') = (\overline{\sigma}_{i'}(t'), i', t') \\
&\qquad \wedge(\overline{\sigma}_i(t) = \overline{\sigma}_{i'}(t') \rightarrow F_i(t) = F_{i'}(t')).
\end{aligned}$$

where $\forall(i,t)$ stands for $\bigwedge_{1 \leq i \leq h} \forall t$ and $\exists(j,x)$ for $\bigvee_{1 \leq i \leq h} \exists x$; $N(j,x)$ stands for $N_j(x)$. Similarly, $(j,x)+1$ represents the successor of pair (j,x) in the lexicographic ordering of pairs (j,x), $j \in \{1, \dots, h\}$ and $x \in D$. The above sentence expresses the fact that the function $\overline{N}$ (in fact the union of functions $\overline{N}_j$, $j \leq h$) is an increasing bijection from the set $\{1, \dots, h\} \times D$ to the set $\{(\overline{\sigma}_i(t), i, t) | i \leq h, t \in D\}$. This sentence plays the same role as the sentence Δ but this time tuples $(\overline{\sigma}_i(t), i, t)$ with the same first component $\overline{\sigma}_i(t)$ are contiguous in the numbering $\overline{N}$. Using a result of Grandjean [64], one can replace the linear ordering $<$ by additional unary functions. ⊣

To be complete, one should also mention the earlier (and weaker) result obtained by Pudlák [105] by purely logical argument at that time.

Proposition 5.13 (Pudlák 75 [105]).
F-SPEC$(d\forall) \subseteq$ F-SPEC$_d(2d\forall)$ *for all* $d \geq 1$.

In the next section, we will examine more closely the expressive power of spectra on restricted vocabulary. The results of this section show that a tight connection exists between nondeterministic complexity classes and classes of spectra defined by limiting the number of universally quantified variables in sentences. A natural question is whether such a connection exists when the language itself is limited. In particular

OPEN QUESTION 26. *Is there a characterization as a complexity class of the classes* F-SPEC_d *for all* $d \geq 1$ *?*

This question has also some connections with problems addressed in Section 6.4.

§6. Approach III: Restricted vocabularies.

6.1. Spectra for monadic predicates. Maybe the simplest way to restrict vocabularies is by limiting the arity of the symbols. In that direction, the smallest restricted class of spectra that can be studied is that of sentences involving only relation symbols of arity one (so-called monadic in the literature). In this case, the following can be proved:

PROPOSITION 6.1 (Löwenheim 1915 [87], Fagin 1975 [42]). *Let τ be a vocabulary consisting of unary relation symbols only and $\phi \in \mathrm{MSOL}(\tau)$. Then the spectrum of ϕ is finite or co-finite.*

PROOF. Use quantifier elimination or Ehrenfeucht-Fraïssé games. $\dashv$

Remark that the even numbers are a spectrum of the following sentence with one unary function:

$$\forall x\ f(x) \neq x \wedge f^2(x) = x.$$

Hence, the most trivial extension of monadic relational vocabulary already provide a spectrum which is neither finite nor co-finite. Then, a natural question is whether the converse of Proposition 6.1 is true or not. The following observation can be made by remarking that one can express the cardinality of a finite domain set by an existential first-order formula.

OBSERVATION 6.2. Every finite or co-finite set $X \subseteq \mathbb{N}$ is a first-order spectrum for a sentence with equality only (i.e., no relation or function symbols).

This contrast with the fact that every SO-spectrum is also an SO-spectrum over equality only. This allows to conclude.

PROPOSITION 6.3. *If τ consists of a finite (possibly empty) set of unary relation symbols, the* $\mathrm{MSOL}(\tau)$*-spectra are exactly all finite and cofinite subsets of* $\mathbb{N}$.

6.2. Spectra for one unary function. As remarked above, one unary function is enough to define nontrivial spectra. It turns out, however, that a complete characterization of spectra for one unary function (with additional unary relations) is possible.

DEFINITION 6.4. A set $X \subseteq \mathbb{N}$ is *ultimately periodic* if there are $a, p \in \mathbb{N}$ such that for each $n \geq a$ we have that $n \in X$ iff $n + p \in X$.

The set of even numbers is ultimately periodic with $a = p = 2$. Again, one may observe

OBSERVATION 6.5. Every ultimately periodic set $X \subseteq \mathbb{N}$ is a first order spectrum for a sentence with one unary function and equality only (this is already true if the function is restricted to be a permutation).

Surprisingly, ultimately periodic sets are precisely the spectra of sentences with one unary function [34, 66].

THEOREM 6.6 (Durand, Fagin, Loescher 1997, Gurevich, Shelah 2003). *Let ϕ be a sentence of* MSOL(τ) *where τ consists of*

- *finitely many unary relation symbols,*
- *one unary function and equality only.*

Then SPEC(ϕ) *is ultimately periodic,*

PROOF. The proof of [34] uses Ehrenfeucht-Fraïssé game argument and is restricted to the EMSOL(τ) case. The generalization of [66] is an application of the Feferman-Vaught-Shelah decomposition method. $\dashv$

There exists alternative ways to characterize ultimately periodic sets. Among others, they can also be seen as sets of integers definable in Presburger Arithmetic. Also, since ultimately periodic sets are closed under complementation, one have:

COROLLARY 6.7. *Spectra involving a single unary function symbol are closed under complement.*

6.3. Beyond one unary function and transfer theorems.

There exist several ways to extend a vocabulary with one unary function: one may choose to add one (or several) new unary function(s) or, in the opposite direction, one may consider vocabularies with only one unrestricted binary relation symbol. It turns out that, up to what will be called "transfer theorems" in the sequel, both kinds of extension lead to very expressive formulas.

Before going further, notations about classes of spectra need to be refined to take into account the number of symbols of distinct arities. Again, we will distinguish in the sequel *whether function symbols are allowed or not*. We write F-SPEC with various indices when function symbols are allowed, and SPEC when function symbols are not allowed.

DEFINITION 6.8. A set S of integers is in F-SPEC$_i^{\alpha,\beta}$ if there exists a first-order sentence ϕ such that $S =$ SPEC(ϕ) and the vocabulary of φ contains only

- α function symbols of arity i and
- at most β function symbols of arity less than i, or relation symbols of arity less or equal to i.

Said another way: a set of integers is in SPEC$_i^{\alpha,\beta}$ if it can be defined by a first order formula over the language of α relation symbols of arity i and β relation symbols of arity less than i.

When the number of symbols of arity less than i is not restricted, the respective class of spectra are denoted by SPEC$_i^{\beta,\omega}$ and F-SPEC$_i^{\beta,\omega}$. For example, the class of first order spectra over one unary function and an arbitrary number of monadic relation and constant symbols (studied in Section 6.2) is denoted $\text{F-SPEC}_1^{1,\omega}$. Similarly SPEC_i^{α} and F-SPEC_i^{α} are abbreviations for $\text{SPEC}_i^{\alpha,0}$ and $\text{F-SPEC}_i^{\alpha,0}$ Finally, in this setting SPEC_i abbreviates for $\text{SPEC}_i^{\omega,\omega}$ (the same holds for F-SPEC_i).

Let us examine what are the relations between these different classes of spectra. Recall that for all $i, j \in \mathbb{N}$ such that $i < j$:

$$\text{SPEC}_i \subseteq \text{SPEC}_j,\ \ \text{F-SPEC}_i \subseteq \text{F-SPEC}_j \text{ and } \text{SPEC}_i \subseteq \text{F-SPEC}_i.$$

The following inclusions are also easy to see. For all $i, \beta \in \mathbb{N}$,

$$\text{SPEC}_i^{\alpha,\omega} \subseteq \text{SPEC}_i^{\alpha+1} \text{ and } \text{F-SPEC}_i^{\alpha,\omega} \subseteq \text{F-SPEC}_i^{\alpha+1}.$$

The relationships between spectra of i-ary functions and spectra of $i+1$-ary relations can be made more precise. In [35], it is shown that a spectrum of a first-order formula involving any number of unary function symbols is also the spectrum of a formula using only one binary relation. This can be generalized to any arity.

Proposition 6.9 ([35, 33]). *For every integer $i \geq 1$,* $\text{F-SPEC}_i \subseteq \text{SPEC}_{i+1}^1$

The converse is not known not be true. All in once, the following chain of inclusions holds.

$$\text{SPEC}_i \subseteq \text{F-SPEC}_i^1 \subseteq \text{F-SPEC}_i \subseteq \text{SPEC}_{i+1}^1 \subseteq \text{SPEC}_{i+1}$$

It seems difficult to prove the converse of any of the inclusions given above. Then, a natural way to study the expressive power of languages relatively to spectrum definition is to reduce them through functional (here polynomial) transformation.

Definition 6.10. If $f : \mathbb{N} \to \mathbb{N}$ and S is a set of integers, then $f(S) = \{f(n) : n \in S\}$ and $S^i = \{n^i : n \in S\}$.

Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two classes of spectra and $f : \mathbb{N} \to \mathbb{N}$ be a function. A natural question is then the following:

Let S be a spectrum in $\mathcal{C}_1$ and f be a $\mathbb{N} \to \mathbb{N}$ function, is $f(S)$ a spectrum in $\mathcal{C}_2$?

In [43], Fagin showed an interesting equivalence between spectra defined by different relational languages. Such results have been called transfer theorems since then. The proof can be seen as an extension of the well-known interpretation method of Rabin [106] with an additional constraint that describes how the domain size of the structure needs to change.

THEOREM 6.11 (Fagin 1975 [43]). *For every $i \geq 1$, $S \in \text{SPEC}_{2i} \iff S^i \in \text{SPEC}_2$.*

Since relational spectra of arity one are finite or cofinite the "transfer" theorem above cannot be extended to SPEC_1. Not too surprisingly, if function symbols are allowed, a similar and more uniform equivalence can be proved.

PROPOSITION 6.12. *For every $i \geq 1$, $S \in \text{F-SPEC}_i \iff S^i \in \text{F-SPEC}_1$.*

The flexibility of unary functions as well as their expressive power are well emphasized by the following result which show that the image of any spectrum under any polynomial transformation is a spectrum involving unary functions only, provided the polynomial is "big enough".

THEOREM 6.13 (Durand, Ranaivoson 1996 [35]). *Let $P(X) \in \mathbb{Q}[X]$ of degree $m \geq i$ and with a strictly positive dominating coefficient. Then,*

$$S \in \text{F-SPEC}_i \Rightarrow \lceil P(S) \rceil = \{\lceil P(n) \rceil,\ n \in S\} \in \text{F-SPEC}_1.$$

6.3.1. *Spectra for one binary relation symbol.* An easy consequence of Theorem 6.11 is that for every spectrum S, there exists $i \in \mathbb{N}$, such that $S^i \in \text{SPEC}_2^1$. This result underlines the great expressive power of sentences involving exactly one binary relation symbol. Up to polynomial transformation, any spectrum is a spectrum of such a sentence.

Let $BIN \subseteq \text{SPEC}_2^1$ be the set of spectra of a symmetric, irreflexive relation (simple graphs). The whole complexity of the spectrum problem is already contained in the apparently weaker question of whether BIN is closed under complement.

THEOREM 6.14 (Fagin 1974, [41]). *BIN is closed under complement iff the complement of every first order spectrum is also a spectrum.*

In fact, one can prove the following stronger result: For all $X \in BIN$ the complement $\mathbb{N} - X \in \text{SPEC}_2^1$ iff the complement of every first order spectrum is also a spectrum.

OPEN QUESTION 27 (Fagin 1974, [41]). *Is every first order spectrum in SPEC_2^1?*

6.3.2. *Spectra for two unary functions and more.* Here again, Proposition 6.12 implies that for any spectrum S, there exists an integer i such that S^i is a spectrum involving unary functions only. This should be compared with the very weak expressive power of sentences involving unary predicates only. We also know that one unary function leads to the very specific class of ultimately periodic sets. The question now is: how many unary function symbols are necessary to obtain an expressive (in the spectrum framework) fragment of first-order logic.

Recall that F-SPEC_1^i denotes the set of first order spectra using at most i unary function symbols. Obviously, for all positive integer i, $\text{F-SPEC}_1^i \subseteq \text{F-SPEC}_1^{i+1}$

The inclusion between the two first levels is strict, as shown in [85].

THEOREM 6.15 (Loescher 1997 [85]). *The set* $\{n^2 : n \in \mathbb{N}\}$ *belongs to the class* $\text{F-SPEC}_1^2 \backslash \text{F-SPEC}_1^1$, *hence the inclusion* $\text{F-SPEC}_1^1 \subset \text{F-SPEC}_1^2$ *is proper.*

In fact, more can be proved on the expressive power of sentences with two unary functions.

THEOREM 6.16 (Durand, Fagin and Loescher, 1998, [34]). *Given* k *and a spectrum* S *in* F-SPEC_1^k. *Then* $kS = \{kn : n \in S\} \in \text{F-SPEC}_1^2$.

Combined with Proposition 6.12 or with Theorem 6.13, this implies that there is a first order spectrum over two unary function symbols which, written in unary, is **NP**-complete. This also implies that there is a transfer result that maps every spectrum to a spectrum over two unary functions.

6.3.3. *Rudimentary sets and spectra of restricted vocabularies.* The relation between rudimentary sets and spectra have been investigated. It has been first observed that the set of primes is in F-SPEC_1 (Grandjean 1988 [63]). More generaly, it holds that:

$$\text{RUD} \subseteq \text{F-SPEC}_1 \text{ (Olive 1996 [102]).}$$

Due to the closure of RUD by polynomial transformation and to the existence of the above described transfer results for spectra, it is clear that, not only, one can improve:

$$\text{RUD} \subseteq \text{F-SPEC}_1^2$$

but also

$$\text{RUD} = \text{SPEC if and only if RUD} = \text{F-SPEC}_1^2$$

However, in view of the following surprising result, there are evidences that none of these equalities hold.

THEOREM 6.17 (Woods, 1981, [131]). *If* RUD = SPEC *then* NP $\neq$ coNP *and* NE = coNE

Since the proof is hardly available and given in a different framework in Woods's thesis, we sketch it below.

PROOF. Let LTH the linear time analog of the polynomial hierarchy PH. Celia Wrathall gave in [132] a precise complexity characterization of rudimentary set by proving that LTH = RUD.

The following facts are easy to prove.

(i) LTH $\subseteq$ DSPACE(n) $\subseteq$ NE
(ii) If LTH = DSPACE(n) then LTH collapses to some level k.
(iii) If LTH collapses to some level k then PH collapses to some level k'.
(iv) LTH = DSPACE(n) implies PH = PSPACE

If RUD = SPEC then, since SPEC = NE, it holds LTH = DSPACE(n) = NE. Hence, both LTH and PH collapse.

For the other consequence, suppose NP = coNP. In this case, the polynomial hierarchy collapses to NP i.e. NP = PH. If again RUD = SPEC one knows than LTH = DSPACE(n) = NE and PH = PSPACE. Then, since NP $\subseteq$ NE one obtains PSPACE = PH $\subseteq$ DSPACE(n) which contradict the well-known results DSPACE(n) $\subsetneq$ PSPACE. ⊣

Hence, although the equality RUD = F-SPEC_1^2 might seem realistic at first sight, its consequences makes it probably hard to prove.

6.4. The unary and the arity hierarchies. The results of the preceding section show that for any spectrum S, there is a polynomial P such that $P(S)$ is the spectrum of a first order sentence with two unary functions. This underlines the expressive power of this latter class of spectra *up to polynomial transformation*. However, as we know *equality* between particular classes of spectra defined, for example, by restriction on the number or the arity of the predicates in the languages is often an open problem. Taking a very particular case, it is even not known whether three unary functions "say" more than two as far as spectra are concerned. This leads to the following open problems about spectra of unary functions.

OPEN QUESTION 28 (The unary hierarchy). *Is the following hierarchy proper:*

$$\text{F-SPEC}_1^1 \subset \text{F-SPEC}_1^2 \subseteq \text{F-SPEC}_1^3 \subseteq \ldots \subseteq \text{F-SPEC}_1^k \subseteq \ldots ?$$

OPEN QUESTION 29. *Is* RUD = F-SPEC_1^2*?*

OPEN QUESTION 30. *Is* $\text{SPEC}_2^1 = \text{F-SPEC}_1$ *or even* $\text{SPEC}_2^1 = \text{F-SPEC}_1^2$*?*

Both positive or negative answers to these questions would have non-trivial consequences. For example, proving RUD $\neq$ F-SPEC_1^2 would separate the classes RUD and SPEC. In the opposite, we already know that RUD = F-SPEC_1^2 implies NP $\neq$ coNP and NE = coNE.

Proving that $\text{F-SPEC}_1^k \subsetneq \text{F-SPEC}_1^{k+1}$ for some integer k would also separate SPEC from RUD. Similarly, a collapse of the unary hierarchy to some level k would have strong consequences. It is easily seen that, testing if a number n (as input i.e. in unary) is in the spectrum of a first order sentence over k unary functions can be decided by a deterministic polynomial time RAM algorithm that uses $k \cdot n$ additional non deterministic steps. Since from Theorem 5.11, $\text{NRAM}(2^n) \subseteq \text{F-SPEC}_1$ then, the inclusion $\text{F-SPEC}_1 \subseteq \text{F-SPEC}_1^k$ would imply immediately the following "trade-off" result on nondeterministic RAM computations (see [34]): *for any arbitrary constant c, any nondeterministic RAM, which given a number n as input, runs in time $c \cdot n$ can be simulated by a RAM which runs in $k \cdot n$ nondeterministic steps and in a polynomial number of deterministic steps.* For c greater than k (which is fixed) such a result is rather unexpected and would strongly modify our understanding of the relationships between determinism and nondeterminism.

Another natural question concerns the relative expressive power of first order sentences defined by restriction on the arity of the symbols involved in the language. It is open whether any spectrum is the spectrum of a sentence over one binary relation only. The question may be refined as follows (See Fagin [43, 44]).

Open Question 31 (The arity hierarchy). *Is the following hierarchy proper:*

$$\text{SPEC}_1 \subseteq \text{SPEC}_2 \subseteq \text{SPEC}_3 \subseteq \ldots \subseteq \text{SPEC}_k \subseteq \ldots ?$$

The same question could be asked for spectra over i-ary functions.

Although the above problem is still open, Fagin proved the following partial result.

Theorem 6.18 (Fagin 1975 [43]). *If $\text{SPEC}_k = \text{SPEC}_{k+1}$ for some integer k then, the arity hierarchy collapses and $\text{SPEC}_k = \text{SPEC}_m$ for every $m \geq k$.*

6.4.1. *Collapse of hierarchies and closure under functions.* Let $\mathcal{C}$ be a class of spectra and f be a $\mathbb{N} \to \mathbb{N}$ function. Class $\mathcal{C}$ is *closed* under f if for any spectrum S in $\mathcal{C}$, $f(S)$ is in $\mathcal{C}$. In the spirit of Theorems 6.11 and 6.13 and Proposition 6.12 one can relate the collapse of hierarchies to the possible closure of class of spectra under some function. In [73], such problems are studied and several related results are given.

Theorem 6.19.

(i) *(Hunter, [73]) The arity hierarchy collapses to SPEC_2^1 if and only if SPEC_2^1 is closed under function $f : n \mapsto \lceil \sqrt{n} \rceil$.*

(ii) *The unary hierarchy collapses to F-SPEC_1^2 if and only if F-SPEC_1^2 is closed under function $f : n \mapsto \lceil \frac{n}{2} \rceil$.*

PROOF OF (II). In [73], a result similar to (ii) is proved about the number of binary predicates instead of the number of unary functions. It is not hard to see that his result extends to the case of (ii). ⊣

6.5. Higher order spectra. In [88], Lynch also proposes a generalization of the characterization given in Proposition 5.10 to higher order spectra. It is shown that the polynomial time hierarchy corresponds to second-order logic. In other words, let $\mathbf{PH}_1 = \mathbf{PH} \cap \mathrm{UN}$ then, a set $X \subseteq \mathbb{N}$ is a second order spectrum if and only if $\{1^n : n \in X\} \in \mathbf{PH}_1$.

Consider SO_k the class of second-order formulas where all the second order variables are of arity at most k. Let $2^{X^k} = \{2^{n^k} \in \mathbb{N} : n \in X\}$ for $X \subseteq \mathbb{N}$. The following precise characterization can be obtained.

THEOREM 6.20 (More and Olive 1997). *[96] A set X is a spectrum of a sentence in SO_k iff 2^{X^k} is rudimentary.*

6.6. Spectra of finite variable logic $\mathbf{FOL}^k$. We denote by FOL^k first order logic with only k distinct variables (bound or free), and by $Spec\mathrm{FOL}^k$ the set of spectra of sentences of FOL^k. FOL^k has been studied extensively in finite model theory, [103, 56], but somehow the spectrum problem was not adressed for FOL^k in the literature[7].

With the same proof as for Proposition 6.1 one gets easily the following.

PROPOSITION 6.21. *The spectrum of a sentence in FOL^1 is finite or cofinite.*

Let M be a set of natural numbers. Recall that a *gap* of M is a pair of integers g_1, g_2 such that $g_1, g_2 \in M$ but for each n with $g_1 < n < g_2$ we have that $n \notin M$. We define $Tow(k,n)$ inductively: $Tow(0,n) = n$ and $Tow(k+1,n) = 2^{Tow(k,n)}$.

By coding structures which model iterated powersets one gets immediately the following.

PROPOSITION 6.22. *For every k there is a $\phi \in FO^4$ such that the spectrum $sp(\phi)$ contains gaps of size $Tow(k,n)$.*

Four variables are used here to express extensionality of the membership relation, and closure under unions with singletons.

Working a bit harder one can prove the following[8].

THEOREM 6.23 (Grohe). (i) *For every Turing machine T there is a sentence $\phi_T \in \mathrm{FOL}^3$ such that*

$$sp_{\phi_T} = \{t^2 : T \ \ \textit{has an accepting run of length} \ \ t\}$$

[7] While the third author was lecturing in Chennai in January 2009 on the spectrum problem, Dr. S. P. Suresh asked about it. The results below are the fruit of discussions with Amaldev Manuel and Martin Grohe.

[8] M. Grohe, personal communication

(ii) *For every recursive function f there is a sentence ϕ_f in* FOL^3 *such that the gaps in the spectrum of ϕ_f grow faster than f.*

The key idea is to encode Turing machines on grids, hence the t^2 in the statement of the Theorem.

In contrast to the above, it follows from the proof that FOL^2 has the finite model property and hence is decidable, [56], that the gaps are bounded.

THEOREM 6.24. *For every $\phi \in FOL^2$ the spectrum $sp(\phi)$ has gaps of size at most $2^{poly(n)}$.*

COROLLARY 6.25. *The following inclusions are proper:*

$$Spec\mathrm{FOL}^1 \subset Spec\mathrm{FOL}^2 \subset Spec\mathrm{FOL}^3$$

OPEN QUESTION 32 (Finite Variable Hierarchy). *Does the hierarchy $Sp\mathrm{FOL}^k$ collapse at level 3?*

§7. The Ash conjectures. The notion of k-equivalence of structures was introduced by Fraïssé [50] in 1954 and the game presentation is due to Ehrenfeucht [37] in 1961.

DEFINITION 7.1. Let σ be a vocabulary, let $\mathcal{A}$ and $\mathcal{B}$ be σ-structures and let k be an integer. The two structures $\mathcal{A}$ and $\mathcal{B}$ are k-equivalent if and only if they satisfy the same σ-sentences with quantifier depth $\leq k$.

For a detailed presentation of k-equivalence, we refer the reader to [36, 72]. For our purpose, the two most important features of the above notion are the following. For each finite vocabulary σ and for each quantifier depth k, the number of k-equivalence classes of σ-structures is finite and we denote it by $M_{\sigma,k}$. For each finite vocabulary σ, for each quantifier depth k and for each k-equivalence class of σ-structures $\mathcal{C}$, there exists a σ-sentence $\Psi_{\mathcal{C}}$ of quantifier depth k such that, for all σ-structure $\mathcal{A}$, we have $\mathcal{A} \in \mathcal{C}$ if and only if $\mathcal{A} \models \Psi_{\mathcal{C}}$.

In 1994, Ash [6] introduces a counting function relative to the k-equivalence classes:

DEFINITION 7.2. Let σ be a finite relational vocabulary, and let k be a positive integer. Ash's function $N_{\sigma,k}$ counts, for each positive integer n, the number of k-equivalence classes of σ-structures of size n.

This function is obviously bounded by the total number of classes, $M_{\sigma,k}$, and Ash's conjecture deals with its asymptotic behavior.

OPEN QUESTION 33 (Ash's constant conjecture). *Is it true that for any finite relational vocabulary σ and any positive integer k, the Ash function $N_{\sigma,k}$ is eventually constant?*

A weaker version of his conjecture is also proposed:

OPEN QUESTION 34 (Ash's periodic conjecture). *Is it true that for any finite relational vocabulary σ and any positive integer k, the Ash function $N_{\sigma,k}$ is eventually periodic?*

Ash shows by a very neat proof that both conjectures imply the spectrum conjecture (i.e. SPEC = coSPEC). Let us present the idea of the proof. Fix σ and k and assume for simplicity that $N_{\sigma,k}(n) = a$ for all n. Take a quantifier depth k first-order σ-sentence ϕ. We exhibit a quantifier depth k first-order σ'-sentence θ_a such that $\mathrm{SPEC}(\theta_a) = \mathbb{N}^+ \setminus \mathrm{SPEC}(\phi)$, where σ' consists in a disjoint copies of σ. Let $\mathcal{C}_1, \ldots, \mathcal{C}_{M_{\sigma,k}}$ be the classes of k-equivalence of σ-structures. All the structures in a given class $\mathcal{C}_i$ agree on ϕ, i.e. either $\forall \mathcal{A} \in \mathcal{C}_i \; \mathcal{A} \models \phi$ or $\forall \mathcal{A} \in \mathcal{C}_i \; \mathcal{A} \models \neg\phi$, because ϕ has quantifier depth k. Form the set X_a consisting of all sets of a distinct $\mathcal{C}_i$s such that $\forall \mathcal{A} \in \mathcal{C}_i \; \mathcal{A} \models \neg\phi$. Take $\theta_a \equiv \bigvee_{Y \in X_a} \bigwedge_{\mathcal{C} \in Y} \Psi'_{\mathcal{C}}$, where $\Psi'_{\mathcal{C}}$ characterizes $\mathcal{C}$ and is written using a distinct copy of σ. The sentence θ_a has quantifier depth k and $n \in \mathrm{SPEC}(\theta_a)$ if and only if there are (at least)

a distinct k-equivalence classes containing a structure $\mathcal{A}$ with size n and $\mathcal{A} \models \neg\phi$. Since the total number of k-equivalence classes containing a structure $\mathcal{A}$ with size n is exactly a, there is no class left for a model of ϕ, and the anounced result follows.

These ideas of Ash's have not been exploited afterwards, and his paper has remained isolated until recently. In 2006, Chateau and More [15] published a second paper related to Ash's counting functions. For all $i \in \mathbb{N}^+$, note $N^{-1}_{\sigma,k}(i) = \{n \in \mathbb{N}^+ | N_{\sigma,k}(n) = i\}$, the inverse image of the positive integer i under the function $N_{\sigma,k}$. Both Ash's conjectures can be rephrased in terms of the sets $N^{-1}_{\sigma,k}(i)$ and are subsumed under the following condition:

Open Question 35 (Ultra-weak Ash conjecture). *Is it true that for any finite relational vocabulary σ, for any positive integer k and for all $i \in \mathbb{N}^+$, the set $N^{-1}_{\sigma,k}(i)$ is a spectrum?*

This last conjecture is proved to be a necessary and sufficient condition for the complement of a spectrum to be a spectrum.

Theorem 7.3. *Let σ be a finite relational vocabulary, and let k be a positive integer. For all $i \in \mathbb{N}^+$, the set $N^{-1}_{\sigma,k}(i)$ is a spectrum if and only if for every σ-sentence φ of quantifier depth $\leq k$, the set $\mathbb{N}^+ \setminus \textsc{spec}(\varphi)$ is a spectrum.*

All in all, the spectrum conjecture is true if and only if the ultra-weak Ash conjecture is true.

Note that, in some cases, one gets interesting additional information about complements of spectra. Eg., if the Ash function $N_{\sigma,k}$ is eventually constant, then for every σ-sentence φ of quantifier depth $\leq k$, the set $\mathbb{N}^+ \setminus \textsc{spec}(\varphi)$ is the spectrum of a sentence of the same quantifier depth as φ, over a vocabulary with the same arities as σ.

In order to make some progress, particular cases of Ash conjectures may be considered, i.e., by restricting the the sets of pairs (σ, k). In his original paper, Ash [6] already did so. For instance, he shows that if σ is a unary vocabulary, then for all $k \geq 1$, Ash's function $N_{\sigma,k}$ is eventually constant. In [15], it is also proved that for all finite relational vocabulary σ, Ash's function $N_{\sigma,2}$ is eventually constant. However, these results are of very limited interest because unary vocabularies or quantifier depth two (see [97]) only allow to define finite and cofinite spectra.

In other cases, solving restricted versions of Ash conjectures happens to be as difficult as solving the full conjectures. Let $\mathcal{G}$ be a vocabulary restricted to a single binary relation symbol (i.e. the language of graphs). As already said, Fagin [43] shows that, up to a polynomial padding, every spectrum is a spectrum in $\textsc{spec}^1_2$. Since spectra are closed under inverse images of polynomial, one obtains:

Proposition 7.4. *If for all $k \geq 1$ and for all $i \in \mathbb{N}^+$, the set $N^{-1}_{\mathcal{G},k}(i)$ is a spectrum, then the (full) spectrum conjecture holds.*

It is less known that, as an easy corollary of a result of Grandjean in [64], up to a polynomial padding, every spectrum is the spectrum of a quantifier depth 3 sentence, using an unbounded number of binary relations. Once again, it follows:

Proposition 7.5. *If for all binary vocabulary σ and for all $i \in \mathbb{N}^+$, the set $N^{-1}_{\sigma,3}(i)$ is a spectrum, then the spectrum conjecture holds.*

Since there is no known padding result that uses a finite set of pairs (σ, k), these results are presently the best possible.

In order to make further progress in solving particular cases of Ash conjectures, Chateau and More introduce a new type of restriction, concerning the semantics of the vocabularies.

Definition 7.6. Let σ be a finite relational vocabulary, and let $\mathcal{C}$ be a class of finite σ-structures. For any positive integer k, the Ash function for the class $\mathcal{C}$, denoted by $N_{\mathcal{C},k}$, counts the number ($\leq M_{\sigma,k}$) of non k-equivalent structures in $\mathcal{C}$ of size n for all $n \geq 1$.

Let $\mathcal{B}$ be the set of finite Boolean algebras. Clearly, we have: $N^{-1}_{\mathcal{B},k}(1) = \{2^\alpha \in \mathbb{N} \mid \alpha \geq 2\}$, $N^{-1}_{\mathcal{B},k}(0) = \mathbb{N}^+ \setminus \{2^\alpha \in \mathbb{N} \mid \alpha \geq 2\}$ and $N^{-1}_{\mathcal{B},k}(i) = \emptyset$ for $i \notin \{0, 1\}$. From this example, it follows that we cannot expect Ash's constant or periodic conjectures to hold for classes of structures. A natural question arises: "Which functions can be Ash's function for some class of structures?". Let $f : \mathbb{N}^+ \mapsto \mathbb{N}$ a function bounded by some constant M_f and computable in E. Then, it is proved that there exist a vocabulary σ, a σ-sentence φ and a quantifier depth K such that $f = N_{Mod_f(\varphi),K}$, where $Mod_f(\varphi)$ denotes the class of finite models of φ.

Now, let us turn to the ultra-weak Ash conjecture for classes of structures. Consider the class $Mod_f(\Psi)$ of finite models of a first-order sentence Ψ.

Open Question 36 (Ultra-weak Ash conjecture for classes of structures). *Is it true that for any finite relational vocabulary σ, for any first-order σ-sentence Ψ, for any positive integer k and for all $i \in \mathbb{N}$, the set $N^{-1}_{Mod_f(\Psi),k}(i)$ is a spectrum?*

Only somehow expressive classes of structures are interesting, and in particular it is natural to require that $Mod_f(\Psi)$ contains at least one structure of size n, for all positive integer n.

Theorem 7.7. *Let σ be a finite relational vocabulary, let Ψ be a first-order σ-sentence such that $Mod_f(\Psi)$ contains at least one structure of size n, for all positive integer n and let k be a positive integer. For all*

$i \in \mathbb{N}^+$, the set $N^{-1}_{Mod_f(\Psi),k}(i)$ is a spectrum if and only if for every σ-sentence φ of quantifier depth $\leq k$ which implies Ψ (i.e. every model of φ is also a model of Ψ), the set $\mathbb{N}^+ \setminus \textsc{spec}(\varphi)$ is a spectrum.

In some particular cases, we obtain more information about complements of spectra. For instance, if Ψ expresses that all binary relations in σ are functional (which can be done using a quantifier depth 3 sentence), and if $k \geq 3$, then all binary relations in the sentence defining $\mathbb{N}^+ \setminus \textsc{spec}(\varphi)$ are functional too.

Let us turn to classes of structures for which the study of Ash's functions is as difficult as the general case. More precisely, let us consider the following classes of structures: let $\mathcal{SG}$ be the class of simple graphs; let $2\mathcal{E}$ be the class of two equivalence relations and let $2\mathcal{F}$ be the class of two functions.

Proposition 7.8. *Let $\mathcal{C} \in \{\mathcal{SG}, 2\mathcal{E}, 2\mathcal{F}\}$. If for all positive integer k and for all $i \in \mathbb{N}^+$, the set $N^{-1}_{\mathcal{C},k}(i)$ is a spectrum, then the spectrum conjecture holds.*

Once again, this is a consequence of various padding results. Fagin actually proves in [43], that simple graphs allow a polynomial padding for all spectra. It is also the case for two unary functions, as proved by Durand, Fagin and Loescher in [34]. The last result concerning two equivalence relations is proved in [14, 38].

§8. Approach IV: Structures of bounded width.

8.1. From restricted vocabularies to bounded tree-width. In this section we look again at spectra of sentences with one unary function and a fixed set of unary predicates. The finite structures which have only one unary function consist of disjoint unions of components of the form given in Figure 1. They look like directed forests where the roots

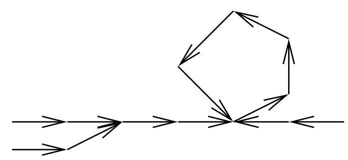

FIGURE 1. Models of one unary function

are replaced by a directed cycle. The unary predicates are just colours attached to the nodes. The similarity of labeled graphs to labeled trees can be measured by the notion of *tree-width*, and in fact, these structures have tree width at most 2. Inspired by Theorem 6.6, E. Fischer and J.A. Makowsky [49] generalized Theorem 6.6 by replacing the restriction on the vocabulary by a purely model theoretic condition involving the width of a relational structure. In this section we discuss their results.

8.1.1. *Tree-width.* In the eighties the notion of tree-width of a graph became a central focus of research in graph theory through the monumental work of Robertson and Seymour on graph minor closed classes of graphs, and its algorithmic consequences [111]. The literature is very rich, but good references and orientation may be found in [31, 10, 11]. Tree-width is a parameter that measures to what extent a graph is similar to a tree. Additional unary predicates do not affect the tree-width. Tree-width of directed graphs is defined as the tree-width of the underlying undirected graph[9].

DEFINITION 8.1 (Tree-width). A k-tree decomposition of a graph $G = (V, E)$ is a pair $(\{X_i \mid i \in I\}, T = (I, F))$ with $\{X_i \mid i \in I\}$ a family of subsets of V, one for each node of T, and T a tree such that

(i) $\bigcup_{i \in I} X_i = V$.
(ii) for all edges $(v, w) \in E$ there exists an $i \in I$ with $v \in X_i$ and $w \in X_i$.
(iii) for all $i, j, k \in I$: if j is on the path from i to k in T, then $X_i \cap X_k \subseteq X_j$ in other words, the subset $\{t \mid v \in X_t\}$ is connected for all v.

[9] In [77] a different definition is given, which attempts to capture the specific situation of directed graphs. But the original definition is the one which is used when dealing with hypergraphs and general relational structures.

(iv) for all $i \in I$, $|X_i| \leq k+1$.

A graph G is of *tree-width at most* k if there exists a k-tree decomposition of G. A class of graphs K is a $TW(k)$-class iff all its members have tree width at most k.

Given a graph G and $k \in \mathbb{N}$ there are polynomial time, even linear time, algorithms, which determine whether G has tree-width k, and if the answer is yes, produce a tree decomposition, cf. [11]. However, if k is part of the input, the problem is NP-complete [5]

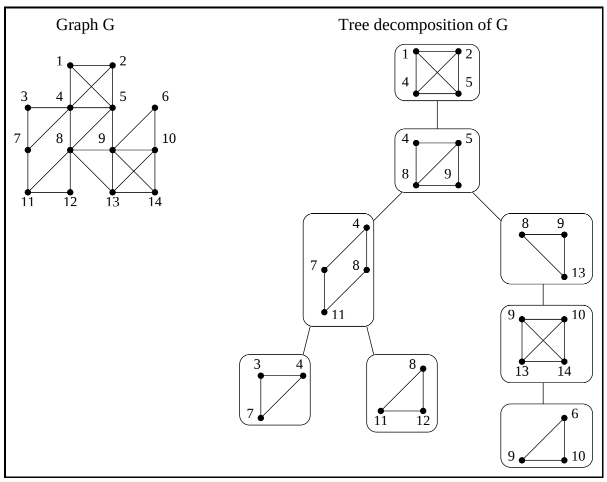

FIGURE 2. A graph and one of its tree-decompositions

Trees have tree-width 1. The clique K_n has tree-width $n-1$. Furthermore, for fixed k, the class of finite graphs of tree-width at most k denoted by $TW(k)$, is MSOL-definable.

EXAMPLE 8.2. The following graph classes are of tree-width at most k:
(i) Planar graphs of radius r with $k = 3r$.
(ii) Chordal graphs with maximal clique of size c with $k = c-1$.
(iii) Interval graphs with maximal clique of size c with $k = c-1$.

EXAMPLE 8.3. The following graph classes have unbounded tree-width and are all MSOL-definable.
(i) All planar graphs and the class of all planar grids $G_{m,n}$.
Note that if $n \leq n_0$ for some fixed $n_0 \in \mathbb{N}$, then the tree-width of the grids $G_{m,n}, n \leq n_0$, is bounded by $2n_0$.
(ii) The regular graphs of degree $r, r \geq 3$ have unbounded tree-width.

Tree-width for labeled graphs can be generalized to arbitrary relational structures in a straightforward way. Clause (ii) in the above definition is replaced by

(ii-rel): For each r-ary relation R, if $\bar{v} \in R$, there exists an $i \in I$ with $\bar{v} \in X_i^r$.

This was first used in [45].

It is now natural to ask, whether Theorem 6.6 can be generalized to arbitrary vocabularies, provided one restricts the spectrum to structures of fixed tree-width k. Indeed, E. Fischer and J. Makowsky [49] have established the following:

THEOREM 8.4 (E. Fischer and J.A. Makowsky (2004)).
Let ϕ be an MSOL *sentence and $k \in \mathbb{N}$. Assume that all the models of ϕ are in $TW(k)$. Then $spec(\phi)$ is ultimately periodic.*

8.2. Extending the logic. First one observes that the logic MSOL can be extended by modular counting quantifiers $C_{k,m}$, where $C_{k,m}x\ \phi(x)$ is interpreted as "there are, modulo m, exactly k elements satisfying $\phi(x)$". We denote the extension of MSOL obtained by adding, for all $k, m \in \mathbb{N}$ the quantifiers $C_{k,m}$, by CMSOL.

Typical graph theoretic concepts expressible in FOL are the presence or absence (up to isomorphism) of a fixed (induced) subgraph H, and fixed lower or upper bounds on the degree of the vertices (hence also r-regularity). Typical graph theoretic concepts expressible in MSOL but not in FOL are connectivity, k-connectivity, reachability, k-colorability (of the vertices), and the presence or absence of a fixed (topological) minor. The latter includes planarity, and more generally, graphs of a fixed genus g. Typical graph theoretic concepts expressible in CMSOL but not in MSOL are the existence of an eulerian circuit (path), the size of a connected component being a multiple of k, and the number of connected components is a multiple of k. All the non-definability statements above can be proved using Ehrenfeucht-Fraïssé games. The definability statements are straightforward.

Second, one observes that very little of the definition of tree-width is used in the proof. The techniques used in the proof of Theorem 8.4 can be adapted to other notions of width of relational structures, such as *clique-width*, which was introduced first in [25] and studied more systematically in [27], and to *patch-width*, introduced in [49].

8.3. Ingredients of the proof of Theorem 8.4. To prove Theorem 8.4, and its generalizations below, the authors use three tools:

(i) A generalization of the Feferman-Vaught Theorem for k-sums of labeled graphs to the logic CMSOL, due to B. Courcelle, [23], and further refined by J.A. Makowsky in [89].
(ii) A reduction of the problem to spectra of labeled trees by a technique first used by B. Courcelle in [24] in his study of graph grammars.
(iii) An adaptation of the Pumping Lemma for labeled trees, cf. [52].

The proof of Theorem 8.4 is quite general. Its proof can be adapted to stronger logics and other notions of width of relational structures. However, the details are rather technical.

8.4. Clique-width. A k-coloured τ-structures is a $\tau_k = \tau \cup \{P_1, \ldots, P_k\}$-structure where $P_i, i \leq k$ are unary predicate symbols the interpretation of which are disjoint (but can be empty).

Definition 8.5. Let $\mathfrak{A}$ be a k-coloured τ-structure.

(i) (Adding hyper-edges) Let $R \in \tau$ be an r-ary relation symbol. $\eta_{R,P_{j_1},\ldots,P_{j_r}}(\mathfrak{A})$ denotes the k-coloured τ structure $\mathfrak{B}$ with the same universe as $\mathfrak{A}$, and for each $S \in \tau_k$, $S \neq R$ the interpretation is also unchanged. Only for R we put

$$R^B = R^A \cup \{\bar{a} \in A^r : a_i \in P^A_{j_i}\}.$$

We call the operation η *hyper edge creation*, or simply *edge creation* in the case of directed graphs. In the case of undirected graphs we denote by $\eta_{P_{j_1},P_{j_2}}$ the operation of adding the corresponding undirected edges.

(ii) (Recolouring) $\rho_{i,j}(\mathfrak{A})$ denotes the k-coloured τ structure $\mathfrak{B}$ with the same universe as $\mathfrak{A}$, and all the relations unchanged but for P^A_i and P^A_j. We put

$$P^B_i = \emptyset \text{ and } P^B_j = P^A_j \cup P^A_i.$$

We call this operation *recolouring*.

(iii) (modification via quantifier free translation) More generally, for $S \in \tau_k$ of arity r and $B(x_1, \ldots, x_r)$ a quantifier free τ_k-formula, $\delta_{S,B}(\mathfrak{A})$ denotes the k-coloured τ structure $\mathfrak{B}$ with the same universe as $\mathfrak{A}$, and for each $S' \in \tau_k$, $S' \neq S$ the interpretation is also unchanged. Only for S we put

$$S^B = \{\bar{a} \in A^r : \bar{a} \in B^A\}.$$

where B^A denotes the interpretation of B in $\mathfrak{A}$.

Note that the operations of type ρ and η are special cases of the operation of type δ.

Definition 8.6 (Clique-Width, [27, 89]).

(i) Here $\tau = \{R_E\}$ consist of the symbol for the edge relation. Given a graph $G = (V, E)$, the *clique-width of* G *(*cwd(G)) is the minimal number of colours required to obtain the given graph as an $\{R_E\}$-reduct from a k-coloured graph constructed inductively from coloured singletons and closure under the following operations:
  - (i.a) disjoint union ($\sqcup$)
  - (i.b) recolouring ($\rho_{i \to j}$)
  - (i.c) edge creation (η_{E,P_i,P_j})

(ii) For τ containing more than one binary relation symbol, we replace the edge creation by the corresponding hyper edge creation $\eta_{R,P_{j_1},\dots,P_{j_r}}$ for each $R \in \tau$.
(iii) A class of τ-structures is a $CW(k)$-class if all its members have clique-width at most k.

If τ contains a unary predicate symbol U, the interpretation of U is not affected by the operations recoloring or edge creation. Only the disjoint union affects it.

A description of a graph or a structure using these operations is called a *clique-width parse term* (or *parse term*, if no confusion arises). Every structure of size n has clique-width at most n. The simplest class of graphs of unbounded tree-width but of clique-width at most 2 are the cliques. Given a graph G and $k \in \mathbb{N}$, determining whether G has clique-width k is in **NP**. A polynomial time algorithm was presented for $k \leq 3$ in [22].

It was shown in [47, 46] that for fixed $k \geq 4$ the problem is **NP**-complete. The recognition problem for clique-width of relational structures has not been studied so far even for $k = 2$. The relationship between tree-width and clique-width was studied in [27, 54, 2].

THEOREM 8.7 (Courcelle and Olariu (2000)). *Let K be a $TW(k)$-class of graphs. Then K is a $CW(m)$-class of graphs with $m \leq 2^{k+1} + 1$.*

THEOREM 8.8 (Adler and Adler (2008)). *For every non-negative integer n there is a structure $\mathfrak{A}$ with only one ternary relation symbol such that $\mathfrak{A} \in TW(2)$ and $\mathfrak{A} \notin CW(n)$.*

The following examples are from [92, 55].

EXAMPLE 8.9 (Classes of clique-width at most k).
(i) The cographs with $k = 2$.
(ii) The distance-hereditary graphs with $k = 3$.
(iii) The cycles C_n with $k = 4$.
(iv) The complement graphs $\bar{C}_n$ of the cycles C_n with $k = 4$.

The cycles C_n have tree-width at most 2, but the other examples have unbounded tree-width.

EXAMPLE 8.10 (Classes of unbounded clique-width).
(i) The class of all finite graphs.
(ii) The class of unit interval graphs.
(iii) The class of permutation graphs.
(iv) The regular graphs of degree 4 have unbounded clique-width.
(v) The class grids *Grid*, consisting of the graphs $Grid_{n\times n}$.

For more non-trivial examples, cf. [92, 55]. In contrast to $TW(k)$, we do not know whether the class of all $CW(k)$-graphs is MSOL-definable.

To find more examples it is useful to note, cf. [90]:

Proposition 8.11. *If a graph is of clique-width at most k and G' is an induced subgraph of G, then the clique-width of G' is at most k.*

In [49] the following is shown:

Theorem 8.12 (E. Fischer and J.A. Makowsky (2004)).
Let $\phi \in \mathrm{CMSOL}(\tau)$ be such that all its finite models have clique-width at most k. Then there are $m_0, n_0 \in \mathbb{N}$ such that if ϕ has a model of size $n \geq n_0$ then ϕ has also a model of size $n + m_0$.

From this we get immediately a further generalization of Theorem 8.4.

Corollary 8.13. *Let $\phi \in \mathrm{CMSOL}(\tau)$ be such that all its finite models have clique-width at most k. Then $spec(\phi)$ is ultimately periodic.*

8.4.1. *Patch-width.* Here is a further generalization of clique-width for which our theorem still works. The choice of operation is discussed in detail in [26].

Definition 8.14. Given a τ-structure $\mathfrak{A}$, the *patch-width of G (*pwd(G)) is the minimal number of colours required to obtain $\mathfrak{G}$ as a $\{\tau\}$-reduct from a k-coloured τ-structure inductively from fixed finite number of τ_k-structures and closure under the following operations:
- (i) disjoint union ($\sqcup$),
- (ii) recoloring ($\rho_{i \to j}$) and
- (iii) modifications ($\delta_{S,B}$).

A class of τ-structures is a $PW_\tau(k)$-class if all its members have patch-width at most k.

A description of a τ-structure using these operations is called a *patch term.*

Example 8.15.
- (i) In [27] it is shown that if a graph G has clique-width at most k then its complement graph $\bar{G}$ has clique-width at most $2k$. However, its patch-width is also k as $\bar{G}$ can be obtained from G by $\delta_{E,\neg E}$.
- (ii) The clique K_n has clique-width 2. However if we consider graphs as structures on a two-sorted universe (respectively for vertices and edges), then K_n has clique-width $c(n)$ and patch-width $p(n)$ where $c(n)$ and $p(n)$ are functions which tend to infinity. This will easily follow from Theorem 8.20. For the clique-width of K_n as a two-sorted-structure this was already shown in [115].

Remark 8.16. In [26] it is shown that a class of graphs of patch-width at most k is of clique-width at most $f(k)$ for some function f. It is shown in [48] that this is not true for relational structures in general.

In the definition of patch-width we allowed only unary predicates as auxiliary predicates (colours). We could also allow r-ary predicates and speak of r-ary patch-width. The theorems where bounded patch-width is required are also true for this more general case. The relative strength of clique-width and the various forms of patch-width are discussed in [48].

In [49] the following is shown:

THEOREM 8.17 (E. Fischer and J.A. Makowsky (2004)).
Let $\phi \in \mathrm{CMSOL}(\tau)$ *be such that all its finite models have patch-width at most* k. *Then there are* $m_0, n_0 \in \mathbb{N}$ *such that if* ϕ *has a model of size* $n \geq n_0$ *then* ϕ *has also a model of size* $n + m_0$.

From this we get yet another generalization of Theorem 8.4.

COROLLARY 8.18. *Let* $\phi \in \mathrm{CMSOL}(\tau)$ *be such that all its finite models have patch-width at most* k. *Then* $spec(\phi)$ *is ultimately periodic.*

More recent work on spectra and patch-width may be found in [117, 32].

8.5. Classes of unbounded patch-width. Theorem 8.17 gives a new method to show that certain classes K of graphs have unbounded tree-width, clique-width or patch-width.

To see this we look at the class *Grid* of all grids $Grid_{n \times n}$. They are known to have unbounded tree-width, cf. [31], and in fact, every minor closed class of graphs of unbounded tree-width contains these grids. They were shown to have unbounded clique-width in [55]. However, for patch-width these arguments do not work. On the other hand *Grid* is MSOL-definable, and its spectrum consists of the numbers n^2, so by Theorems 8.12 and 8.17, the unboundedness follows directly.

In particular, as this is also true for every $K' \supseteq K$, the class of all graphs is of unbounded patch-width.

Without Theorem 8.17, there was only a conditional proof of unbounded patch-width available. It depends on the assumption that the polynomial hierarchy Σ^{P} does not collapse to **NP**. The argument then procceds as follows:

(i) Checking patch-width at most k of a structure $\mathfrak{A}$, for k fixed, is in **NP**. Given a structure $\mathfrak{A}$, one just has to guess a patch-term of size polynomial in the size of $\mathfrak{A}$.
(ii) Using the results of [89] one gets that checking a $\mathrm{CMSOL}(\tau)$-property ϕ on the class $PW_\tau(k)$ is in **NP**, whereas, by [91], there are Σ^{P}_n-hard problems definable in MSOL for every level Σ^{P}_n of the polynomial hierarchy.
(iii) Hence, if the polynomial hierarchy does not collapse to **NP**, the class of all τ-structures is of unbounded patch-width, provided τ is large enough.

Open Question 37. *What is the complexity of checking whether a* τ*-structure* $\mathfrak{A}$ *has patch-width at most* k*, for a fixed* k*?*

8.6. Parikh's Theorem. R. Parikh's celebrated theorem, first proved in [104], counts the number of occurences of letters in k-letter words of context-free languages. For a given word w, the numbers of these occurences is denoted by a vector $n(w) \in \mathbb{N}^k$, and the theorem states

Theorem 8.19 (Parikh 1966). *For a context-free language* L*, the set* $Par(L) = \{n(w) \in \mathbb{N}^k : w \in L\}$ *is semilinear.*

A set $X \subseteq \mathbb{N}^s$ is *linear in* $\mathbb{N}^s$ iff there is vector $\bar{a} \in \mathbb{N}^s$ and a matrix $M \in \mathbb{N}^{s \times r}$ such that

$$X = A_{\bar{a},\bar{M}} = \{\bar{b} \in \mathbb{N}^s : \text{ there is } \bar{u} \in \mathbb{N}^r \text{ with } \bar{b} = \bar{a} + M \cdot \bar{u}\}$$

Singletons are linear sets with $M = 0$. If $M \neq 0$ the series is *nontrivial.* $X \subseteq \mathbb{N}^s$ is *semilinear in* $\mathbb{N}^s$ iff X is a finite union of linear sets $A_i \subseteq \mathbb{N}^s$. The terminology is from [104], and has since become standard terminology in formal language theory. We note that for unary languages, Parikh's Theorem looks at the spectrum of context-free languages.

B. Courcelle has generalized Theorem 8.19 further to context-free vertex replacement graph grammars, [24]. We want to generalize Theorem 8.19 to spectra. Rather than counting occurences of letters, we look at many-sorted structures and the sizes of the different sorts, which we call many-sorted spectra. In [49] the following theorem is proved:

Theorem 8.20 (E. Fischer and J.A. Makowsky (2006)).
Let K *be a class of* CMSOL*-definable many-sorted relational structures which are of patch-width at most* k*. Then the many-sorted spectrum of* K *forms a semilinear set.*

Appendix A. A review of some hardly accessible references.

This section contains a detailed presentation of the material of Subsection 4.3. Note that the proofs sketched here do not necessarily correspond to the original proofs.

A.1. Asser's paper. In chronological ordering, the first paper related to spectra is [7], in German, due to Asser in 1955. Though it does not use the name "spectrum", nor refer to Scholz in its title or in the text, the long introduction clarifies the context in which the concept of spectrum was born. The author addresses the general question of classes of cardinal numbers (not only natural numbers) so-called "representable" by a sentence of first-order logic with equality, both in the framework of satisfiability theory and validity theory. Here a first-order sentence φ represents a given (finite or infinite) cardinality m regarding satisfiablity if there is a structure which domain has cardinality m which is a model of φ (i.e. for finite m, it means $m \in \textsc{spec}(\varphi)$), and regarding validity, if φ holds in every structure with cardinality m. Asser first notices that φ represents m regarding satisfiability if and only if $\neg\varphi$ does not represent m regarding validity, so that validity reduces to satisfiability via complement. Then, he remarks that, from Löwenheim-Skolem Theorem [87, 118], the representation question in satisfiability theory for infinite cardinalities is trivial: the first-order sentence φ either has no infinite model (and in this case it has finite models in finitely many finite cardinalities only) or has models in every infinite cardinality. Hence, the problem actually is about exactly which sets of natural numbers are the set of cardinalities of finite models of first-order sentences, i.e. what we would call spectra. In a footnote, one reads "this question was also asked by Scholz as a problem in [116]".

With this background, Asser's aim is to give a purely arithmetical characterization of spectra. This is done via an arithmetical encoding of finite structures, first-order sentences and satisfiability. Let us make precise Asser's construction.

Note that in the sequel "characteristic functions" (of sets or relations) are not taken in the usual way: a unary function f is said to be the characteristic function of the set of integers n such that $f(n) = 0$. It is only a technical matter to come back to the usual definition with little machinery, for instance use $\chi(n) = 1 \dot{-} f(n)$ (so-called *modified substraction* i.e. $x \dot{-} y = x - y$ if $x \geq y$ and 0 otherwise). Using this alternative definition, characteristic functions are not required to be 0-1-valued.

W.l.o.g., let φ be a sentence in relational Skolem normal form, i.e. $\varphi \equiv \forall x_1 \dots \forall x_r \exists x_{r+1} \dots \exists x_s \psi(x_1, \dots x_s)$, where $\psi(x_1, \dots x_s)$ is a Boolean combination of atomic formulas $R_i^{(a_i)}(x_{j_1}, \dots x_{j_{a_i}})$ with $i = 1, \dots, t$ and of atoms $x_{l_1} = x_{l_2}$. Assume that ψ contains u different atoms of type $R_i^{(a_i)}(x_{j_1}, \dots x_{j_{a_i}})$ and v different atoms of type $x_{l_1} = x_{l_2}$. Let Ψ :

$\{0,1\}^{u+v} \longrightarrow \{0,1\}$ be the Boolean function associated to the propositional version of ψ (using the convention that 0 encodes true and 1 encodes false).

Denote by $Bit_k(y, z_1, \dots, z_k, n)$ the binary digit of y of rank $\sum_{l=1}^{k} z_l \cdot n^{l-1}$, assuming $y < 2^{n^k}, z_1 < n, \dots, z_k < n$. Encode the k-ary relation R on the domain $\{0, \dots, n-1\}$ by the number $y < 2^{n^k}$ such that $Bit_k(y, z_1, \dots, z_k, n) = 0$ if and only if $R(z_1, \dots, z_k)$ holds. Let $\delta(z_1, z_2) = 0$ if $z_1 = z_2$ and 1 otherwise. Let $\Psi^*(y_1, \dots, y_t, x_1, \dots, x_s, n)$ be obtained from Ψ by replacing each atom $R_i^{(a_i)}(x_{j_1}, \dots x_{j_{a_i}})$ by $Bit_{a_i}(y_i, x_{j_1}, \dots, x_{j_{a_i}}, n)$ and every atoms $x_{l_1} = x_{l_2}$ by $\delta(x_{l_1}, x_{l_2})$. The first-order quantifiers $\forall x_1 \dots \forall x_r \exists x_{r+1} \dots \exists x_s$ are dealt with by defining

$$\Psi^{**}(y_1, \dots, y_t, n) = \sum_{x_1=0}^{n-1} \dots \sum_{x_r=0}^{n-1} \prod_{x_{r+1}=0}^{n-1} \dots \prod_{x_s=0}^{n-1} \Psi^*(\overline{y}, \overline{x}, n).$$

Note the non-standard use of $\sum$ for $\forall$ and $\prod$ for $\exists$, due to the fact that 0 encodes true and 1 encodes false. Finally the characteristic function of the spectrum of the sentence $\varphi \equiv \forall x_1 \dots \forall x_r \exists x_{r+1} \dots \exists x_s \psi(x_1, \dots x_s)$ is $\chi(n) = \prod_{y_1=0}^{2^{n^{a_1}}-1} \dots \prod_{y_t=0}^{2^{n^{a_t}}-1} \Psi^{**}(y_1, \dots, y_t, n)$. This construction is clearly elementary. Conversely, it is also easy to verify that any function defined as $\chi(n) = \prod_{y_1=0}^{2^{n^{a_1}}-1} \dots \prod_{y_t=0}^{2^{n^{a_t}}-1} \Psi^{**}(y_1, \dots, y_t, n)$, where Ψ^{**} is obtained from some Boolean function Ψ by the same type of construction, is the characteristic function of the spectrum of the corresponding first-order sentence. Hence we have the following result.

Theorem A.1. *A set $\mathcal{S}$ is a spectrum iff its characteristic function χ has the form $\chi(n) = \prod_{y_1=0}^{2^{n^{a_1}}-1} \dots \prod_{y_t=0}^{2^{n^{a_t}}-1} \Psi^{**}(y_1, \dots, y_t, n)$, where Ψ^{**} is obtained from some Boolean function Ψ by the above construction.*

Note that Asser judges his result "non satisfactory", in particular because this paraphrastic characterization is of no help in proving that a given set is or not a spectrum, or in providing any concrete spectrum. However, Asser's characterization is enough to prove Theorem 4.10, that we restate here for sake of self-containment.

Theorem 4.10 Spec $\subsetneq \mathcal{E}^3_\star$

The inclusion follows from the fact that Theorem A.1 provides elementary characteristic functions for spectra. The properness is obtained by diagonalization.

As a conclusion, Asser asks some questions, that have essentially remained open up to now. First, he asked for a recursive characterization of spectra. He notes that there are actually two different problems. The first one asks for a recursively defined class of functions, i.e. a class of functions defined via some basis functions and closure under some functional operations, such that the unary functions in this class are exactly the characteristic functions of spectra. Second, he asks for a recursively defined class of functions, but now such that the unary functions in it enumerate exactly the spectra, i.e. a set S is a spectrum if and only if $S = f(\mathbb{N})$ for some f in the class. Note that this is not the most commonly admitted meaning for enumeration, because the enumeration functions are usually required to be strictly increasing, which is not the case here.

Next, Asser refers to "work in progress" that proves that a large class of unary functions are characteristic functions of spectra, among which the following arithmetically defined sets: prime numbers, multiples of a given integer k, powers of a given k, k-th powers, composite numbers.a

Finally, the third and most famous open question proposed in this paper is usually known as Asser's Problem (Open Question 2) and asks whether spectra are closed under complement.

A.2. Mostowski's paper. A paper almost simultaneous with Asser's is [98], due to A. Mostowski in 1956. It also adresses recursive characterization of spectra, and explicitly uses the name "spectrum". It is noticed that "The results of Asser overlap in part with results which I have found in 1953 while attempting (unsuccessfully) to solve Scholz's problem (cf. Roczniki Polskiego Towarzystwa Matematycznego, series I, vol. 1 (1955), p.427). I shall give here proofs of my results which do not overlap with Asser's." [10].

Here, A. Mostowski defines a class of functions denoted by K as follows.

Definition A.2. The class K is the least class

- containing the functions Z_k, U_k^i, S, C respectively defined by:
  - $Z_k(x_1, \dots, x_k, n) = 0$
  - $U_k^i(x_1, \dots, x_k, n) = min(x_i, n)$, for $i = 1, \dots, k$
  - $S(x, n) = min(x + 1, n)$
  - $C(x) = n$
- closed under composition: $f(x_1, \dots, x_{j-1}, g(y_1, \dots, y_p, n), x_{j+1}, \dots, x_k, n)$
- closed under recursion: $\begin{cases} f(0, \vec{x}, n) = g(\vec{x}, n) \\ f(x+1, \vec{x}, n) = min(h(x, f(x, \vec{x}, n), \vec{x}, n), n) \end{cases}$

The basis functions Z_k, U_k^i and S are intended as the classical zero, projections and successor functions, but the special variable n always bounds their values. The function C is intended as a maximum function.

[10]Thanks to J. Tomasik, we have seen a translation of the Polish reference. It is the abstract of a seminar given by Andrzej Mostowski on October, 16. 1953. In addition to the following material, it is also stated that spectra form a strict subclass of primitive recursive sets, a result which indeed overlaps with Asser's.

The functional operations composition and recursion are also bounded by n. The main result of Mostowski's paper is the following theorem.

THEOREM A.3. *For any unary function $f \in K$, the set $\{n+1 \mid f(n) = 0\}$ is a spectrum.*

Let us give an idea of the proof via an example. Consider the functions f, g and h defined as follows:

- $f(x,n) = 1$ if $x = 0$ and $f(x,n) = 0$ otherwise.
  I.e. $\begin{cases} f(0,n) = S(Z(C(n),n),n) \\ f(x+1,n) = min(Z(f(x,n),n),n) \end{cases}$
- $g(x,n) = 0$ if x is even and $g(x,n) = 1$ otherwise.
  I.e. $\begin{cases} g(0,n) = Z(C(n),n) \\ g(x+1,n) = min(f(g(x,n),n),n) \end{cases}$
- $h(n) = g(C(n),n)$.

Clearly we have $h \in K$ and $h(n) = 0$ if and only if n is even. Let us derive from the definition of h a sentence ψ in the vocabulary

$$\sigma = \{\leq, min, max, Succ^{(2)}, R_f^{(3)}, R_g^{(3)}, R_h^{(2)}\}$$

such that ψ has a model with $n+1$ elements if and only if $h(n) = 0$ (i.e. SPEC(ψ) is the set of odd numbers). The key point of the construction is that the functions in the class K can be interpreted as functions on finite structures (eg. from $\{0,\ldots,n\}^k$ to $\{0,\ldots,n\}$) without loss of generality, because of the special variable n that bounds all their values.

The sentence ψ first expresses the fact that $\leq$ is a linear ordering, min and max are its first and last elements and $Succ$ its successor relation. Then, ψ describes the behavior of the predicates R_f, R_g and R_h corresponding to the graphs of the functions f, g and h. For instance, R_g obeys the conjunction of the following sentences:

- g is functional in its first variable:
  $\forall x, y\,(y = max \longrightarrow \exists! z R_g(x,y,z))$
- the second variable in g is always n:
  $\forall x, y, z\,(R_g(x,y,z) \longrightarrow y = max)$
- description of the base case of the definition of g:
  $\forall x, y, z\,[(R_g(x,y,z) \wedge x = min) \longrightarrow z = min]$
- description of the recursive recursive case of the definition of g:
  $\forall x, y, z[(R_g(x,y,z) \wedge \neg x = min)$
  $\longrightarrow \exists t, u(Succ(t,x) \wedge R_g(t,y,u) \wedge R_f(u,y,z))]$

Our goal is then achieved by adding to ψ the following condition:
$\forall x, y\,(R_h(x,y) \longrightarrow (x = max \wedge y = min))$
Finally, it is clear that SPEC(ψ) is the set of odd numbers as required.

Mostowski asks if the converse is true, i.e.

OPEN QUESTION 38. *Is every spectrum representable as $\{n+1 \mid f(n) = 0\}$ for some function $f \in K$?*

No answer is known up to now.

As a conclusion, new examples of spectra are presented: the set of integers having the form $n!$ for some n, and the set $\{n \mid n^2+1 \text{ is prime }\}$. Also, Mostowski asks whether Fermat's prime numbers, i.e. primes of the form $2^{2^n}+1$, form a spectrum. This question can be understood in two different ways, as noticed by Bennett: which one of the sets $A = \{p \mid p \text{ is prime and } p = 2^{2^n}+1 \text{ for some integer } n\}$ and $B = \{n \mid 2^{2^n} + 1 \text{ is prime }\}$ is intended ? Using rudimentary relations, the set A is easily proved to be a spectrum, whereas it is still not known for the set B.

Finally, let us remark that it is ordinarily considered that what Mostowski proved is that the unary relations in $\mathcal{E}^2_\star$ are spectra. This is not exactly the case, but the legend is most probably due to the fact that Bennett attributes this result to Mostowski. However, Bennett also notes that, even if it is easy to prove that $K \subseteq \mathcal{E}^2$, it is not clear that the bounded version of any function in $\mathcal{E}^2$ (i.e. $f_b(x_1,\dots,x_k,n) = min(f(x_1,\dots,x_k),n))$ is in K. Mostowski's construction crucially relies on the fact that the functions in K are bounded by their last variable, and does not generalize to functions in $\mathcal{E}^2$. In contrast, it is not difficult to verify that the bounded versions of addition and multiplication are in K, and consequently that the rudimentary relations have their characteristic functions in K. Whatever, it is true that the unary relations in $\mathcal{E}^2_\star$ are indeed spectra, see Corollary A.5.

A.3. Bennett's thesis. This is a huge work titled "On spectra" [9], but which also deals with a lot of other subjects. Bennett's thesis is unpublished, and only available via library services. It is one of the remarkable early texts anticipating later developments in finite model theory, definability theory and complexity theory. It contains a characterization (and various definitions) of rudimentary sets and already relates spectra to space bounded Turing machines, thus catching a glimpse of many of the results concerning spectra that were formulated and proved in more modern language after 1970.

Not only first-order spectra are considered by Bennett, but also spectra of higher order logics, and not only sets, but also many-sorted sets, all in all spectra of the whole theory of types. This full generality makes the notations quite clumsy. The use of many-sorted structures corresponds to relations of arity greater than one, and the use of higher order logics provides more complicated relations.

We shall limit ourselves with the cases of one-sorted (i.e. ordinary) spectra of orders one and two. Note that the first item of Theorem A.4 is also stated as Theorem 4.13 in Subsection 4.3.

Theorem A.4 (Bennett, 1962 [9]).

(i) *A set $S \subseteq \mathbb{N}$ is a first-order spectrum iff it can be defined by a formula of the form $\exists y \leq 2^{x^j} R(x, y)$ for some $j \geq 1$, where R is strictly rudimentary.*
(ii) *A set $S \subseteq \mathbb{N}$ is a second-order spectrum iff it can be defined by a formula of the form $\exists y \leq 2^{x^j} R(x, y)$ for some $j \geq 1$, where R is rudimentary.*

Spectra of higher order are characterized by similar features: spectra of order $2n$ correspond to rudimentary relations prefixed by an existential quantifier bounded by an iterated exponential $2^{\cdot^{2^{x^j}}}$ with height n, and spectra of order $2n-1$ correspond to strictly rudimentary relations prefixed by an existential quantifier bounded by an iterated exponential with height n. Spectra of sentences over a d-sorted universe have the same types of characterizations, using $\exists y \leq 2^{\max(x_1,\dots,x_d)^j} R(x_1, \dots, x_d, y)$. Finally, the spectra of the whole type theory are characterized as the elementary relations.

Bennett also introduces several other subrudimentary classes, respectively called "strongly", "positive" and "extended" rudimentary relations, which yield a bunch of slightly different characterizations of spectra, which may witness various unsuccessful attempts to design a truly satisfactory characterization. In this survey, we shall limit ourselves to Rud and Srud.

Some consequences of the characterization theorem (not all of them are immediate):

Corollary A.5.

(i) *For each $n \geq 1$, the class of spectra of order n is closed under $\wedge$, $\vee$, bounded quantifications, substitution of rudimentary functions, explicit transformations and finite modifications.*
(ii) *For each $n \geq 1$, the class of spectra of order $2n$ is closed under $\neg$.*
(iii) *The class of first-order spectra contains the rudimentary relations and $\mathcal{E}^2_\star$.*
(iv) *The class of second-order spectra strictly contains the rudimentary relations.*
(v) *For each $n \geq 1$, spectra of order n form a subset of spectra of order $n+1$ and a strict subset of spectra of order $n+2$.*

We propose below a proof of Bennett's theorem.

Proof of Theorem A.4. (ii) We first present the second-order case, because it has less technical difficulties.

- First inclusion: $\{\text{SPEC}(\varphi) \mid \varphi \in \text{SO}\} \subseteq \{\exists y \leq 2^{x^j}\ R(x,y) \mid j \geq 1 \text{ and } R \in \text{RUD}\}$ i.e. φ has a model with x elements iff $\exists y \leq 2^{x^j}\ R(x,y)$ is true.

W.l.o.g. we may assume that φ has no first-order or second-order free variable (just quantify existentially in case there are any). Assume the second-order variables appearing in φ have arities strictly less than j. Then we take $y = 2^{x^j}$. We encode a second-order variable Z with arity $a < j$ by the number $z < 2^{x^a} < y$ in the usual way. Hence, every second-order quantification $QZ^{(a)}$ in φ is translated into the first-order bounded quantification $Qz < 2^{x^a} < y$. Recall that $Bit(a,b)$ is true iff the bit of rank b of a is 1. Now, every atomic formula $Z(z_1,\dots,z_a)$ is translated into $Bit(z, z_1 + z_2 \cdot x + \dots + z_a \cdot x^{a-1})$. Every first-order quantification qz in φ is translated into the bounded quantification $qz < x$. The atomic formulas $z = z'$ in φ remain unchanged. Let φ' denote the obtained formula. Finally, let $R \equiv (y = 2^{x^j}) \wedge \varphi'$.

- Second inclusion: $\{\text{SPEC}(\varphi) \mid \varphi \in SO\} \supseteq \{\exists y \leq 2^{x^j}\ R(x,y) \mid j \geq 1 \text{ and } R \in \text{RUD}\}$ i.e. $\exists y \leq 2^{x^j}\ R(x,y)$ is true iff φ has a model with x elements.

First, we use three existentially quantified relations, namely $\leq^{(2j)}$ which is bound to be a linear ordering over the j-tuples of vertices, $+^{(3j)}$ which is bound to be the associated addition and $\times^{(3j)}$ which is bound to be the associated multiplication. Let us denote by $Arithm(\leq,+,\times)$ the first-order sentence expressing this requirement. Note that we may now use for free any usual arithmetic predicate on numbers bounded by x^j (written in x-ary notation, i.e. seen as j-tuples of integers in $\{0,\dots,x-1\}$). Next, all variables in R, including x and y, are encoded by j-ary second-order variables in φ in the usual way. For instance if $x = \sum_{l=0}^{p} 2^{i_l}$, we let $X = \{(i_0,0,\dots,0),\dots,(i_p,0,\dots,0)\}$.

W.l.o.g. we may assume that all the atomic formulas in R are of type $u \cdot v = w$ (concatenation), which we translate into

$$\begin{aligned}
&Concat(U,V,W) \equiv \\
&\quad \exists \bar{t} V(\bar{t}) \wedge \forall \bar{z}(V(\bar{z}) \longrightarrow \bar{z} \leq \bar{t}) \wedge \forall \bar{z}(U(\bar{z}) \longrightarrow \bar{z} \leq (\overline{max} - \bar{t})) \\
&\quad \wedge \forall \bar{z} \left(W(\bar{z}) \longleftrightarrow \left(\left(\bar{z} \leq \bar{t} \wedge V(\bar{z})\right) \vee \left(\bar{z} > \bar{t} \wedge U(\bar{z} - \bar{t})\right)\right)\right).
\end{aligned}$$

Note that this sentence would be cleaner in dyadic than it is in binary, but the whole encoding would also be more complicated because two unary relations are needed to encode an integer in dyadic (the set of 1s and the set of 2s) because its length is fixed.

In order to translate the bounded quantifications in R, we also need the following first-order sentence, which expresses the fact that the integers u and v respectively encoded by U and V are such that $u < v$.

$$\begin{array}{l} Smaller(U,V) \equiv \\ \exists \overline{z}\left(V(\overline{z}) \wedge \neg U(\overline{z}) \wedge \forall \overline{z'} > \overline{z} \neg U(\overline{z'})\right) \\ \vee \exists \overline{z}\left(V(\overline{z}) \wedge U(\overline{z}) \wedge \forall \overline{z'} > \overline{z}\left(\neg V(\overline{z'}) \wedge \neg U(\overline{z'})\right) \wedge \exists \overline{z'} < \overline{z}\left(V(\overline{z'}) \wedge \neg U(\overline{z'})\right)\right). \end{array}$$

Now, let R' be obtained from R by applying the following rules: every bounded first-order quantification $\forall z < z' \dots$ is translated into the second order quantification $\forall Z^{(j)} Smaller(Z,Z') \longrightarrow \dots$, and accordingly for $\exists z < z' \dots$; and every atomic formula $u \cdot v = w$ is translated into $Concat(U,V,W)$.

It remains to express that X encodes the size x of the domain, which is done using the binary notation of $(max, 0, \dots, 0)$, which represents the largest element of the domain. More precisely, we have $max + 1 = x$, which translates in binary as follows:

$$\begin{array}{l} Dom(X) \equiv \\ \forall \overline{z}(((X(\overline{z}) \wedge \forall \overline{z'} < z \neg X(\overline{z'})) \longrightarrow \neg Bit((max,0,\dots,0),\overline{z})) \\ \wedge((\exists \overline{z'} < z X(\overline{z'})) \longrightarrow (X(\overline{z}) \longleftrightarrow Bit((max,0,\dots,0),\overline{z}))) \\ \wedge((\exists \overline{z'} > \overline{z}(X(\overline{z'}) \wedge \forall \overline{z''} < \overline{z'} \neg X(\overline{z''}))) \longrightarrow Bit((max,0,\dots,0),\overline{z}))) \end{array}$$

Finally, φ is $\exists \leq^{(2j)} \exists +^{(3j)} \exists \times^{(3j)} \exists Y^{(j)} \exists X^{(j)} (Arithm(\leq,+,\times) \wedge Dom(X) \wedge R')$.

(i) Next we turn to the first-order case. We consider the proof of the second-order case and show how it has to be modified in order to fit to the first-order case. Note that the proof is now more tricky, and we use dyadic notation because we have to be more precise.

- First inclusion: $\{\text{SPEC}(\varphi) \mid \varphi \in FO\} \subseteq \{\exists y \leq 2^{x^j}\ R(x,y) \mid j \geq 1 \text{ and } R \in \text{SRUD}\}$

The main difference concerning φ is that it contains no second-order quantifications. Concerning R, we have to deal with two differences: bounded quantifications are now replaced by part-of quantifications ($\forall z_1 \upharpoonright z_2$ and $\exists z_1 \upharpoonright z_2$) on the one hand and we have to use concatenation instead of arithmetic on the other hand.

However, φ does contain free second-order variables, say $Z_1^{(a_1)}, \dots, Z_k^{(a_k)}$, which we do not encode in the usual way because SRUD does not allow to use arithmetical predicates, hence the Bit predicate is not available. Instead, we assume for now that the alphabet is $\{1, 2, *, \star, \bullet\}$ and we first define a provisional predicate $R'(x,y)$. We shall explain later how to get rid of the extra symbols $*$, $\star$ and $\bullet$ to obtain the expected $R(x,y)$.

We use the following encoding: if $Z = \{(x_1^1, \dots, x_a^1), \dots, (x_1^p, \dots, x_a^p)\}$, with $p \leq x^a$, then let $z = \star * x_1^1 * \dots * x_a^1 * \star \dots \star * x_1^p * \dots * x_a^p * \star$. Note that we have $|z| \leq x^a \cdot a \cdot (|x| + 2)$.

Let us define $x_0 = *x*(x-1)*\ldots*1*$, i.e. the dyadic representation of x_0 is the concatenation of the dyadic representations of all integers in $\{1,\ldots,x\}$, separated by $*$s. Note that $|x_0| \leq x\cdot(|x+2|) < x^2$. Finally, let $y = \bullet z_1 \bullet \ldots \bullet z_k \bullet x_0 \bullet$. Clearly we have $y \leq 2^{x^j}$ for some $j \geq 1$.

Now, $R'(x,y)$ will begin with $\exists z_1 \restriction y\ \ldots\ \exists z_k \restriction y\ \exists x_0 \restriction y\ ((y = \bullet z_1 \bullet \ldots \bullet z_k \bullet x_0 \bullet) \wedge \neg(\bullet \restriction z_1) \wedge \ldots \wedge \neg(\bullet \restriction z_k) \wedge \neg(\bullet \restriction x_0) \wedge \ldots)$, in order to retrieve the significant parts of y.

We use x_0 to replace every first-order quantification $\forall u \ldots$ appearing in φ by a part-of quantification $\forall u \restriction x_0\,(Int(u,x_0) \longrightarrow \ldots)$ in R', and similarly for $\exists u \ldots$, where $Int(u,x_0)$ means that u is a maximal non-empty string of 1s and 2s in x_0. The most technical part of the proof is to write a strictly rudimentary formula $Dom(x_0,x)$ which is true iff x_0 has the expected form, but for sake of brevity, we do not explicit this formula. In particular, note that we now consider the domain as $\{1,\ldots,x\}$ instead of $\{0,\ldots,x-1\}$ as we did previously. Finally it is not difficult to write a formula $Verif(x_0,z)$ expressing the fact that z has the expected form $\star * x_1^1 * \ldots * x_a^1 * \star \ldots \star * x_1^p * \ldots * x_a^p * \star$. Namely, take

$$\begin{aligned}
&Verif(x_0,z) \equiv \exists u \restriction z\\
&(\star u \star \restriction z) \wedge \forall u \restriction z(((\star u \star \restriction z) \wedge \neg(\star \restriction u) \wedge u \neq \epsilon) \longrightarrow \exists v_1 \restriction u \ldots \exists v_a \restriction u\\
&(Int(v_1,x_0) \wedge \ldots \wedge Int(v_a,x_0) \wedge u = *v_1 * \ldots * v_a *))\\
&\wedge \forall u_1,u_2,\alpha,\beta,\gamma \restriction z(((z = \alpha \star u_1 \star \beta \star u_2 \star \gamma \vee z = \alpha \star u_1 \star u_2 \star \gamma)\\
&\wedge \neg(\star \restriction u_1) \wedge \neg(\star \restriction u_2)) \longrightarrow u_1 \neq u_2).
\end{aligned}$$

There are two types of atomic formulas in φ: equalities $z_1 = z_2$ and atoms $Z(z_1,\ldots,z_a)$. Equalities remain unchanged and $Z(z_1,\ldots,z_a)$ is changed into $\star * z_1 * \ldots * z_a * \star \restriction z$. These operations lead to the strictly rudimentary formula φ'.

Finally, $R'(x,y)$ is $\exists z_1 \restriction y \ldots \exists z_k \restriction y \exists x_0 \restriction y((y = \bullet z_1 \bullet \ldots \bullet z_k \bullet x_0 \bullet) \wedge \neg(\bullet \restriction z_1) \wedge \ldots \wedge \neg(\bullet \restriction z_k) \wedge \neg(\bullet \restriction x_0) \wedge Dom(x_0,x) \wedge Verif(x_0,z_1) \wedge \ldots \wedge Verif(x_0,z_k) \wedge \varphi')$.

To obtain R, it remains to get rid of the alphabet $\{1,2,*,\star,\bullet\}$. Let $*$ be a string of 1s which is not a subword of $x, x-1,\ldots,2$ and 1. For instance, $*$ could be of length $|x|+1$. Let $\star = 2*2$ and $\bullet = 22*22$. The final length of y is polynomially longer than it used to be, which remains acceptable. Finally, take $R(x,y) \equiv \exists * \restriction y\ \exists \star \restriction y\ \exists \bullet \restriction y\ ((\forall u \restriction *\ (u=1)) \wedge * \neq \epsilon \wedge \star = 2*2 \wedge \bullet = 22*22 \wedge R')$. Note that strictly rudimentary relations do not define predicates referring to the length of integers, so that $*$ cannot be bound to be some specific word like $1^{|x|+1}$.

- Second inclusion: $\{\text{SPEC}(\varphi) \mid \varphi \in FO\} \supseteq \{\exists y \leq 2^{x^j}\ R(x,y) \mid j \geq 1 \text{ and } R \in \text{SRUD}\}$

The main difference with the second-order case concerning φ is that it only contains first-order quantifications. However, we are still free to choose as many free second-order variables as we may need. In particular, we still use usual arithmetic predicates on the (j-tuples of) elements of the domain, and the previous first-order sentence $Arithm(\leq, +, \times)$ is still required to hold for this purpose. In addition, we introduce the second-order variables X_1, X_2 and Y_1, Y_2, both of arity j, respectively representing the set of positions where x and y have 1s and 2s and no other second-order variables are introduced. Let $Word(X_1, X_2)$ be the sentence expressing the fact that X_1 and X_2 (and similarly Y_1, Y_2) do represent a dyadic word, namely

$$\begin{aligned}&Word(X_1, X_2) \equiv \forall \overline{z} \neg (X_1(\overline{z}) \wedge X_2(\overline{z}))\\ &\wedge \exists \overline{z} \forall \overline{t}((\overline{t} > \overline{z} \longrightarrow (\neg X_1(\overline{t}) \wedge \neg X_2(\overline{t}))) \wedge (\overline{t} \leq \overline{z} \longrightarrow (X_1(\overline{t}) \vee X_2(\overline{t})))).\end{aligned}$$

Concerning R, we may assume w.l.o.g. that it only contains part-of quantifications $qz \restriction x$ and $qz \restriction y$ and no $qz \restriction z'$ for $z' \notin \{x, y\}$.

The main trick is that a part-of quantification $\exists z \restriction y \dots$ (for instance) will be replaced by $2j$ first-order quantifications $\exists \overline{z_1} \exists \overline{z_2} \, (\overline{z_1} \leq \overline{z_2} \wedge \dots)$, where $\overline{z_1}$ and $\overline{z_2}$ encode the positions where z begins and ends, as a subword of y.

We have to translate the atomic formulas $u \cdot v = w$. W.l.o.g. we may rewrite R in an equivalent formula by replacing everywhere $u \cdot v = w$ with $(u \cdot v = w \wedge u \restriction y \wedge v \restriction y \wedge w \restriction y) \vee (u \cdot v = w \wedge u \restriction y \wedge v \restriction y \wedge w \restriction x) \vee \dots \vee (u \cdot v = w \wedge u \restriction x \wedge v \restriction x \wedge w \restriction x)$. Hence, there are 8 slightly different cases to be taken care of. We limit ourselves with the case $u \cdot v \cdot w \wedge u \restriction y \wedge v \restriction y \wedge w \restriction y$. The corresponding formula $Concat_{yyy}(\overline{u_1}, \overline{u_2}, \overline{v_1}, \overline{v_2}, \overline{w_1}, \overline{w_2})$ is as follows:

$$\begin{aligned}&\overline{w_2} = \overline{w_1} + \overline{u_2} - \overline{u_1} + \overline{v_2} - \overline{v_1}\\ &\wedge \forall \overline{z}(\overline{w_1} \leq \overline{z} < \overline{w_1} + \overline{u_2} - \overline{u_1} \longrightarrow\\ &((Y_1(\overline{z}) \longleftrightarrow Y_1(\overline{z} - \overline{w_1})) \wedge (Y_2(\overline{z}) \longleftrightarrow Y_2(\overline{z} - \overline{w_1}))))\\ &\wedge \forall \overline{z}(\overline{w_1} + \overline{u_2} - \overline{u_1} \leq \overline{z} < \overline{w_2} \longrightarrow\\ &((Y_1(\overline{z}) \longleftrightarrow Y_1(\overline{z} - \overline{w_1} - \overline{u_2} + \overline{u_1})) \wedge (Y_2(\overline{z}) \longleftrightarrow Y_2(\overline{z} - \overline{w_1} - \overline{u_2} + \overline{u_1})))).\end{aligned}$$

Let us denote by R' the obtained sentence.

The last remaining part is to write out a sentence $Dom'(X_1, X_2)$ expressing the fact that X_1, X_2 encodes (in dyadic) the cardinality of the domain, i.e. the successor of the j-tuple $(max, 0, \dots, 0)$. This is a bit more technical than the sentence $Dom(X)$ we used for the binary notation and we do not spell it out here. Finally, take $\varphi \equiv Arithm \wedge Word(X_1, X_2) \wedge Word(Y_1, Y_2) \wedge Dom(X_1, X_2) \wedge R'$. $\dashv$

Connections with complexity classes. At the beginning of complexity theory, the usual compexity classes such as the polynomial hierarchy had not emerged yet. So the classes used by Bennett are not standard ones. He considers two hierarchies based on space-bounded deterministic Turing machines defined in a recursive fashion: the base class is of type

$FDSpace(f(n))$, and the next class has a space bound which is a function in the previous class.

Let us denote by $(\mathcal{R}^i)_{i\geq 1}$ the first hierarchy, introduced in Ritchie's 1963 paper [110], which comes from from his Ph.D. thesis [109].

Definition A.6 (Ritchie's classes).

- Let $\mathcal{R}^1$ be the class of functions computable by some (deterministic) Turing machine in space bounded by $b \cdot \max(\overrightarrow{x})$ on input $\overrightarrow{x}$, where $b \geq 1$ is some integer fixed for each machine, i.e. $\mathcal{R}^1 = FDSpace(\mathcal{O}(2^n))$ in modern notation.
- For each $i \geq 1$, let us denote by $\mathcal{R}^{i+1}$ the class of functions computable by a Turing machine in space bounded by $B(\overrightarrow{x})$, where B is some function in $\mathcal{R}^i$, fixed for each machine.
- For each $i \geq 1$, let us denote by $\mathcal{R}^i_\star$ the class of relations whose characteristic functions are in $\mathcal{R}^i$.

It is proved in [110] that this hierarchy $(\mathcal{R}^i_\star)_{i\geq 1}$ is strict and that its union corresponds to elementary relations.

Using the same pattern, Bennett introduces a second hierarchy, that we denote by $(\mathcal{B}^i)_{i\geq 1}$.

Definition A.7 (Bennett's classes).

- Let $\mathcal{B}^1$ be the class of functions computable by some (deterministic) Turing machine in space bounded by $P(\overrightarrow{x})$ on input $\overrightarrow{x}$, where P is some arithmetical polynomial fixed for each machine, i.e. $\mathcal{B}^1 = FDSpace((2^{\mathcal{O}(n)}))$ in modern notation.
- For each $i \geq 1$, let us denote by $\mathcal{B}^{i+1}$ the class of functions computable by a Turing machine in space bounded by $B(\overrightarrow{x})$, where B is some function in $\mathcal{B}^i$, fixed for each machine.
- For each $i \geq 1$, let us denote by $\mathcal{B}^i_\star$ the class of relations whose characteristic functions are in $\mathcal{B}^i$.

Bennett shows that Ritchie's classes $\mathcal{R}^i_\star$ come in between spectra of various orders, but not in a very nice way. In contrast, he proves nice closure properties and an exact intercalation between the classes of spectra of consecutive orders for the classes $\mathcal{B}^i_\star$. However, all these classes are too big to be informative concerning relationship between first-order spectra and complexity classes.

In order to state the next theorem, let us denote by $\mathcal{S}^i$ the class of (many-sorted) spectra of formulas of order i. For instance Spec is the class of unary relations in $\mathcal{S}^1$.

Theorem A.8.

(i) *$\mathcal{R}^1_\star \subseteq \mathcal{S}^3$ and for each $i \geq 2$, $\mathcal{S}^{2i-2} \subseteq \mathcal{R}^i_\star \subseteq \mathcal{S}^{2i+1}$. Moreover, for no $i, j \geq 1$ does $\mathcal{R}^i_\star = \mathcal{S}^j$.*

(ii) *For each* $i \geq 1$, $\mathcal{S}^{2i} \subseteq \mathcal{B}^i_\star \subseteq \mathcal{S}^{2i+1}$ *(equality or strictness is unknown) and* $\mathcal{R}^i_\star \subsetneq \mathcal{B}^i_\star \subsetneq \mathcal{R}^{i+1}_\star$. *Moreover,* $\mathcal{B}^i_\star$ *is closed with respect to union, intersection, bounded quantifications, substitution of rudimentary functions, explicit transformations and finite modifications.*

The proof of item (i) is based on recursive characterizations of the classes $\mathcal{R}^i_\star$, whereas item (ii) is stated without proof.

A.4. Mo's paper.

There is a late paper on the recursive aspect of spectra, namely [95], due to the Chinese logician Mo Shaokui in 1991, only available in Chinese (see the author's English abstract in Mathematics Abstracts of Zentralblatt [94]). With the help of Zhu Ping [133], we have been able to state Mo's result, and we propose a proof sketch.

Definition A.9. Let $x, x_1, x_2, \dots$ and $y, y_1, y_2, \dots$ be two disjoint sets of variables. Let Mo be the smallest class of predicates over integers containing the relations $x_1 + x_2 = x_3$, $x_1 \times x_2 = x_3$ (both for variables of type x only) and $Bit(y, x)$ (where the first variable is of type y, the second of type x) and closed under Boolean operations and (polynomially) bounded quantifications for variables of type x only.

Note that a predicate in Mo has two types of variables, which do not play similar roles, and that Mo extends the rudimentary relations by the use of $Bit(y, x)$ atoms, which are not definable because y variables are not allowed in the atomic formulas for addition and multiplication.

Theorem A.10. $\{\mathrm{SPEC}(\varphi) \mid \varphi \in FO\} = \{\exists y_1 \leq 2^{x^{j_1}} \dots \exists y_k \leq 2^{x^{j_k}} R(x, y_1, \dots, y_k) \mid k, j_1, \dots, j_k \geq 1$ *and* $R \in \mathrm{Mo}\}$

Proof. It is a slightly modified version of the proof of Bennett's theorem for second-order spectra.

- First inclusion: φ has a model with x elements iff $\exists y_1 \leq 2^{x^{j_1}} \dots \exists y_k \leq 2^{x^{j_k}}\ R(x, y_1, \dots, y_k)$ is true.

We encode a predicate symbol Y with arity j by the number $y < 2^{x^j}$ in the usual way. Hence, every atomic formula $Y(x_1, \dots, x_j)$ is translated into $\exists x' < x^j (Bit(y, x') \wedge x' = x_1 + x_2 \cdot x + \dots + x_j \cdot x^{j-1})$. Every first-order quantification qx_i in φ is translated into the bounded quantification $qx_i < x$. The atomic formulas $x_1 = x_2$ in φ remain unchanged. Let R denote the obtained formula with free variables $x, y_1, \dots, y_k$.

- Second inclusion: $\exists y_1 \leq 2^{x^{j_1}} \dots \exists y_k \leq 2^{x^{j_k}} R(x, y_1, \dots, y_k)$ is true iff φ has a model with x elements.

First, we use three predicate symbols, namely $\leq^{(2)}$ which is bound to be a linear ordering on the vertices, $+^{(3)}$ which is bound to be the associated addition and $\times^{(3)}$ which is bound to be the associated multiplication. Let us denote by $Arithm_1(\leq, +, \times)$ the first-order sentence expressing this requirement. Note that we may now use for free any usual arithmetic predicate on numbers bounded by x.

Next, every free variable of type y in R and bounded by 2^{x^j} is translated into a predicate symbol Y of arity j.

W.l.o.g., we may assume that all the bounded quantifications in R are of type $qx' < x^i$ for some $i \geq 1$. The bounded quantification $qx' < x^i$ in R is simply translated into $qx'_1 \dots qx'_i$ and x' is represented by the i-tuple $(x'_1, \dots, x'_i)$.

There are three types of atomic formulas in R. Let us first consider formulas $x_1 + x_2 = x_3$ and $x_1 \times x_2 = x_3$. Assume we have $x_1 < x^{i_1} \leq x^j, x_2 < x^{i_2} \leq x^j, x_3 < x^{i_3} \leq x^j$, with $j = \max(i_1, i_2, i_3)$. The variables x_1, x_2, x_3 correspond to the tuples $(x_1^1, \dots, x_1^j), (x_2^1, \dots, x_2^j), (x_3^1, \dots, x_3^j)$ (padding with as many 0s as necessary). This includes the case $x < x^2$ so that x corresponds to $(0, 1, 0, \dots, 0)$. Then $x_1 + x_2 = x_3$ is changed into $Add_j(x_1^1, \dots, x_1^j, x_2^1, \dots, x_2^j, x_3^1, \dots, x_3^j)$ and $x_1 \times x_2 = x_3$ is changed into $Mult_j(x_1^1, \dots, x_1^j, x_2^1, \dots, x_2^j, x_3^1, \dots, x_3^j)$, where the formulas Add_j and $Mult_j$ express addition and multiplication on j-tuples in x-ary notation. The case of atomic formulas $Bit(y, x')$ is dealt with similarly. Assume we have $y < 2^{x^j}$, and $x' < x^i$ then there are three possibilities. If $i < j$, then $Bit(y, x')$ is changed into $Y(x'_1, \dots, x'_i, 0, \dots, 0)$. If $i = j$, then $Bit(y, x')$ is changed into $Y(x'_1, \dots, x'_j)$. If $i > j$, then $Bit(y, x')$ is changed into $Y(x'_1, \dots, x'_j) \wedge x_{j+1} = 0 \wedge \dots \wedge x_i = 0$. Similarly, $Bit(y, x)$ (which may also occur in R because x is a free variable of type x) translates into $Y(0, 1, 0, \dots, 0)$ if Y has arity 2 at least and into $0 \neq 0$ (false) if Y is unary. Let us denote by R' the first-order sentence thus obtained.

Finally, φ is $Arithm_1(\leq, +, \times) \wedge R'$. $\dashv$

Note that, in order to uniformize the proofs with that of Bennett's theorem and help comparison, we have slightly modified the original statement in two points. First, Mo uses functional vocabularies, which yields bounds of type x^{x^j} for y type variables and the use of atoms $Digit_x(y, x') = x''$ (meaning "the digit of rank x' of y in x-ary notation is x''") instead of $Bit(y, x)$. Second, the relation R is originally described using Grzegorczyk's classes $\mathcal{E}^0_\star$, $\mathcal{E}^1_\star$ or $\mathcal{E}^2_\star$ instead of RUD.

Finally, concerning Asser's problem (so-called second Scholz problem here), the author's abstract [94] asserts that:

> It is also shown that if all the functions in $\mathcal{E}^0$ can be enumerated by a function in $\mathcal{E}^2$, then the complement of a certain finite spectrum cannot be any finite spectrum. Hence, under such a condition, the answer to the second Scholz problem is negative.

Hence, the conditional negative solution proposed here seems to be linked to some separation of $\mathcal{E}^0$ and $\mathcal{E}^2$ via diagonalization, which seems unlikely (the classical proof of separation of $\mathcal{E}^i$ and $\mathcal{E}^{i+1}$ uses the bound on the growth of the functions in $\mathcal{E}^i$).

Appendix B. A compendium of questions and conjectures.

In this appendix we list, for convenience, all the Open Questions (OQ)and stated in our survey.

From Section 2.

- **OQ 1:** (Scholz) Characterize the sets of natural numbers that are first order spectra.
- **OQ 2:** (Asser) Is the complement of a first order spectrum a first order spectrum?
- **OQ 3:** Is the complement of a spectrum of an MSOL-sentence again a spectrum of an MSOL-sentence?
- **OQ 4:** (Fagin) Is every first order spectrum the spectrum of a first order sentence of one binary relation symbol?
- **OQ 5:** (Fagin) Is every first order spectrum the spectrum of a first order sentence over simple graphs?
- **OQ 6:** (Fagin) Is every first order spectrum the spectrum of a first order sentence over planar graphs?

From Section 3.

We recall a few definitions: Let $M \subseteq \mathbb{N}^+$, and let $m_1, m_2, \ldots$ an enumeration of M ordered by the size of its elements. $\chi_M(n)$ is the characteristic function of M. $\eta_M(n)$ is the enumeration function of M, i.e., $\eta_M(n) = m_n$ if it exists, and $\eta_M(n) = 0$ otherwise. Finally, $\delta_M(n) = \eta_M(n+1) - \eta_M(n)$.

- **OQ 7:** Which strictly increasing sequences of positive integers, are enumerating functions of spectra? For instance, how fast can they grow?
- **OQ 8:** If M is a spectrum how can $\delta_M(n)$ behave?
- **OQ 9:** Let $\pi(n)$ be the counting function of the primes, and let $li(n)$ be its approximation by the integral logarithm. Define
  $\pi^+ = \{n : \pi(n) - li(n) > 0\}$ and $\pi^- = \{n : \pi(n) - li(n) \geq 0\}$.
  Are the sets π^+ and π^- spectra?
- **OQ 10:** Let ϕ a first order sentence, and f_ϕ be the associated labeled counting function that is monotonically increasing. Is there a first order sentence ψ such that for all n we have $f_\phi(n) = \eta_\psi(n)$?
- **OQ 11:** Are there any irrational algebraic reals which are spectral?
- **OQ 13:** Is every automatic real a spectral real?

The binary string complexity of a real in binary presentation is the function $p_r(m)$ which counts, for each m the number of distinct binary words w of length m occurring in r.

OQ 14: Does the binary string complexity $p_r(m)$ of a spectral real r satisfy

$$\liminf_{m\to\infty} \frac{p_r(m)}{2^m} < 1$$

or even

$$\liminf_{m\to\infty} \frac{p_r(m)}{2^m} = 1?$$

OQ 15: Are the b-adic $\mathcal{E}^2$-computable reals $\mathcal{E}^2$-Cauchy computable?
OQ 16: Is the inclusion $\mathcal{F}_{low} \subseteq \mathcal{E}^2$ proper?

From Section 4.

OQ 17: Are the inclusions in $\mathcal{E}^0_\star \subseteq \mathcal{E}^1_\star \subseteq \mathcal{E}^2_\star$ proper?
OQ 18: Additionally, is the inclusion $\textsc{Rud} \subseteq \mathcal{E}^0_\star$ proper?
OQ 19: Is the inclusion in $\mathcal{E}^2_\star \subseteq \textsc{Spec}$ proper?
OQ 20: Is the inclusion $\textsc{Rud} \subseteq \textsc{Spec}$ proper?

From Section 5.

OQ 21:
(i) Are any of the inclusions
L $\subseteq$ NL, LINSPACE $\subseteq$ NLINSPACE, P $\subseteq$ NP and E $\subseteq$ NE proper?
(ii) Do any of the equalities NP = coNP and NE = coNE hold?

OQ 22: Are the inclusions L $\subseteq$ $\textsc{Rud}$ = LTH $\subseteq$ LINSPACE proper?
OQ 23: Is every spectrum the spectrum of a categorical sentence ?
OQ 24: Is there a universal (complete) spectrum S_0 and a suitable notion of reduction such that every spectrum S is reducible to S_0?
OQ 25: Is the inclusion $d \geq 1$, NTIME$(2^{d\cdot n}) \subseteq \textsc{spec}_d(+)$ proper?
OQ 26: Is there a characterization as a complexity class of the classes $\textsc{f-spec}_d$ for all $d \geq 1$?

From Section 6.

OQ 27: Is every first order spectrum in $\textsc{spec}^1_2$?
OQ 28: Is the following hierarchy proper:

$$\textsc{f-spec}^1_1 \subset \textsc{f-spec}^2_1 \subseteq \textsc{f-spec}^3_1 \subseteq \ldots \subseteq \textsc{f-spec}^k_1 \subseteq \ldots ?$$

OQ 29: Is $\textsc{Rud} = \textsc{f-spec}^2_1$?
OQ 30: Is $\textsc{spec}^1_2 = \textsc{f-spec}_1$ or even $\textsc{spec}^1_2 = \textsc{f-spec}^2_1$?
OQ 31: Is the following hierarchy proper:

$$\textsc{spec}_1 \subseteq \textsc{spec}_2 \subseteq \textsc{spec}_3 \subseteq \ldots \subseteq \textsc{spec}_k \subseteq \ldots ?$$

The same question may be asked for spectra over i-ary functions.

OQ 32: Does the hierarchy $Sp\mathrm{FOL}^k$ collapse at level 3?

From Section 7.

OQ 33: (Ash's constant conjecture)
Is it true that for any finite relational vocabulary σ and any positive integer k, the Ash function $N_{\sigma,k}$ is eventually constant?

OQ 34: (Ash's periodic conjecture)
Is it true that for any finite relational vocabulary σ and any positive integer k, the Ash function $N_{\sigma,k}$ is eventually periodic?

OQ 35: (Ultra-weak Ash conjecture)
Is it true that for any finite relational vocabulary σ, for any positive integer k and for all $i \in \mathbb{N}^+$, the set $N^{-1}_{\sigma,k}(i)$ is a spectrum?

OQ 36: (Ultra-weak Ash conjecture for classes of structures)
Is it true that for any finite relational vocabulary σ, for any first-order σ-sentence Ψ, for any positive integer k and for all $i \in \mathbb{N}$, the set $N^{-1}_{Mod_f(\Psi),k}(i)$ is a spectrum?

From Section 8.

OQ 37: What is the complexity of checking whether a τ-structure $\mathfrak{A}$ has patch-width at most k, for a fixed k?

From Section 8.

OQ38: Is every spectrum representable as $\{n+1 \mid f(n) = 0\}$ for some function $f \in K$?

Recall that K is the class of functions defined in Section A Definition A.2.

REFERENCES

[1] B. Adamczewski and Y. Bugeaud, *On the complexity of algebraic numbers I: Expansion in integer bases*, ***Ann. of Math.***, vol. 165 (2007), pp. 547–565.

[2] H. Adler and I. Adler, *A note on clique-width and tree-width on structures*, Arxiv preprint arXiv:0806.0103v2, [cs.LO], 2008.

[3] Miklos Ajtai and Ronald Fagin, *Reachability is harder for directed than for undirected finite graphs*, ***The Journal of Symbolic Logic***, vol. 55 (1990), no. 1, pp. 113–150.

[4] Eric Allender and Vivek Gore, *Rudimentary reductions revisited*, ***Inf. Process. Lett.***, vol. 40 (1991), no. 2, pp. 89–95.

[5] S. Arnborg, D.G. Corneil, and A. Proskurowski, *Complexity of finding embedding in a k–tree*, ***SIAM. J. Algebraic Discrete Methods***, vol. 8 (1987), pp. 277–284.

[6] Christopher J. Ash, *A conjecture concerning the spectrum of a sentence*, ***Mathematical Logic Quartely***, vol. 40 (1994), pp. 393–397.

[7] Günter Asser, *Das Repräsentenproblem in Prädikatenkalkül der ersten Stufe mit Identität*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 1 (1955), pp. 252–263.

[8] D.H. Bailey, J.M. Borwein, R.E. Crandall, and C. Pomerance, *On the binary expansions of algebraic numbers*, ***J. Théor. Nombres Bordeaux***, vol. 16 (2004), pp. 487–518.

[9] James H. Bennett, *On spectra*, ***Ph.D. thesis***, Princeton University, Princeton, New Jersey, USA, 1962.

[10] H. Bodlaender, *A tourist guide through tree width*, ***Acta Cybernetica***, vol. 11 (1993), pp. 1–23.

[11] ———, *A partial k-arboretum of graphs with bounded tree width (tutorial)*, ***Theoretical Computer Science***, vol. 208 (1998), pp. 1–45.

[12] E. Borel, *Sur les chiffres décimaux de $\sqrt{2}$ et divers problèmes de probabilités en chaînes*, ***C. R. Acad. Sci. Paris***, vol. 230 (1950), pp. 591–593.

[13] Cristian Calude, *Super-exponentials nonprimitive recursive, but rudimentary.*, ***Inf. Process. Lett.***, vol. 25 (1987), no. 5, pp. 311–316.

[14] A. Chateau, *Utilisation des destinées pour la décision et sa complexité dans le cas de formules à profondeur de quantification bornée sur des structures logiques finies et infinies*, ***Ph.D. thesis***, Université d'Auvergne, 2003.

[15] Annie Chateau and Malika More, *The ultraweak Ash conjecture and some particular cases*, ***Mathematical Logic Quarterly***, vol. 52 (2006), no. 1, pp. 4–13.

[16] Qingliang Chen, Kaile Su, and Xizhong Zheng, *Primitive recursive real numbers*, ***Mathematical Logic Quarterly***, vol. 53 (2007), no. 4-5, pp. 365–380.

[17] ———, *Primitive recursiveness of real numbers under different representations*, ***Electronic Notes in Theoretical Computer Science***, vol. 167 (2007), pp. 303–324.

[18] Claude A. Christen, *Spektralproblem und Komplexitätstheorie*, ***Ph.D. thesis***, Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland, 1974.

[19] ———, *Spektralproblem und Komplexitätstheorie*, ***Komplexität von entscheidungsproblemen: ein seminar*** (Ernst Specker and Volker Strassen, editors), Lecture Notes in Computer Science, vol. 43, Springer, 1976, pp. 102–126.

[20] A. Cobham, *On the Hartmanis-Stearns problem for a class of tag machine*, ***Ieee conference record of 1968 ninth annual symposium on switching and automata theory, schenectady, new york***, IEEE Computer Society Press, 1968, pp. 51–60.

[21] K.J. Compton and C.W. Henson, *A uniform method for proving lower bounds on the computational complexity of logical theories*, ***Annals of Pure and Applied***

Logic, vol. 48 (1990), pp. 1–79.

[22] D. G. Corneil, M. Habib, J.-M. Lanlignel, B. Reed, and U. Rotics, *Polynomial time recognition of clique-width ≤ 3 graphs*, ***Proceedings of latin'2000***, Lecture Notes in Computer Science, vol. 1776, Springer, 2000, pp. 126–134.

[23] B. Courcelle, *The monadic second–order logic of graphs I: Recognizable sets of finite graphs*, ***Information and Computation***, vol. 85 (1990), pp. 12–75.

[24] ———, *Structural properties of context-free sets of graphs generated by vertex replacement*, ***ic***, vol. 116 (1995), pp. 275–293.

[25] B. Courcelle, J. Engelfriet, and G. Rozenberg, *Handle-rewriting hypergraph grammars*, ***J. Comput. System Sci.***, vol. 46 (1993), pp. 218–270.

[26] B. Courcelle and J.A. Makowsky, *Fusion on relational structures and the verification of monadic second order properties*, ***Mathematical Structures in Computer Science***, vol. 12.2 (2002), pp. 203–235.

[27] B. Courcelle and S. Olariu, *Upper bounds to the clique–width of graphs*, ***Discrete Applied Mathematics***, vol. 101 (2000), pp. 77–114.

[28] Paul Csillag, *Eine Bemerkung zur Auflösung der eingeschachtelten Rekursion.*, ***Acta Univ. Szeged., Acta Sci. Math.***, vol. 11 (1947), pp. 169–173 (German).

[29] Haskell B. Curry, *Review of [98]*, MR0086766 (19,240c).

[30] E. Dahlhaus, *Combinatorial and logical properties of reductions to some complete problems in NP and NL*, ***Ph.D. thesis***, Technische Universität Berlin, Germany, 1982.

[31] R. Diestel, ***Graph theory***, Graduate Texts in Mathematics, Springer, 1996.

[32] M. Doron and S. Shelah, *Bounded m-ary patch-width are equivalent for* $m > 2$, Electronically available at arXiv:math/0607375v1, 2006.

[33] Arnaud Durand, *Hiérarchies de définissabilité logique au second ordre*, ***Ph.D. thesis***, Université de Caen, Caen, France, 1996.

[34] Arnaud Durand, Ronald Fagin, and Bernd Loescher, *Spectra with only unary function symbols*, ***Computer science logic, 11th international workshop, csl'97, annual conference of the eacsl, aarhus, denmark, august 23-29, 1997, selected papers*** (Mogens Nielsen and Wolfgang Thomas, editors), Lecture Notes in Computer Science, vol. 1414, Springer, 1998, pp. 189–202.

[35] Arnaud Durand and Solomompionona Ranaivoson, *First-order spectra with one binary predicate*, ***Theoretical Computer Science***, vol. 160 (1996), no. 1&2, pp. 189–202.

[36] Heinz-Dieter Ebbinghaus and Jörg Flum, ***Finite model theory***, Perspectives in Mathematical Logic, Springer, 1995.

[37] A. Ehrenfeucht, *An application of games to the completeness problem for formalized theories*, ***Fondamenta Mathematicae***, vol. 49 (1961), pp. 129–141.

[38] Y. Ershov, I. Lavrov, A. Taimanov, and M. Taitslin, *Elementary theories*, ***Russian Mathematical Surveys***, vol. 20 (1965), pp. 35–105.

[39] Henri-Alex Esbelin and Malika More, *Rudimentary relations and primitive recursion : a toolbox*, ***Theoretical Computer Science***, vol. 193 (1998), no. 1-2, pp. 129–148.

[40] Ronald Fagin, *Contributions to the model theory of finite structures*, ***Ph.D. thesis***, University of California, Berkeley, California, 1973.

[41] ———, *Generalized first-order spectra and polynomial-time recognizable sets*, ***Complexity of computation (proc. siam-ams sympos. appl. math., new york, 1973)*** (Providence, R.I.) (Richard M. Karp, editor), SIAM-AMS Proceedings, vol. 7, American Mathematical Society, 1974, pp. 43–73.

[42] ———, *Monadic generalized spectra*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 21 (1975), pp. 89–96.

[43] ———, *A spectrum hierarchy*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 21 (1975), pp. 123–134.

[44] ———, *Finite-model theory - a personal perspective*, ***Theoretical Computer Science***, (1993), no. 116, pp. pp.3–31.

[45] T. Feder and M. Vardi, *The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory*, ***SIAM Journal on Computing***, vol. 28 (1999), pp. 57–104.

[46] Michael R. Fellows, Frances A. Rosamond, Udi Rotics, and Stefan Szeider, *Clique-width minimization is np-hard*, ***Stoc '06: Proceedings of the thirty-eighth annual acm symposium on theory of computing***, ACM, 2006, pp. 354–362.

[47] M.R. Fellows, F.A. Rosamond, U. Rotics, and S. Szeider, *Proving NP-hardness for clique width ii: Non-approximability of clique-width*, Electronic Colloquium on Computational Complexity, 2005.

[48] E. Fischer and J.A. Makowsky, *Patch-width, a generalization of clique-width for relational structures*, in preparation.

[49] Eldar Fischer and Johann A. Makowsky, *On spectra of sentences of monadic second order logic with counting*, ***The Journal of Symbolic Logic***, vol. 69 (2004), no. 3, pp. 617–640.

[50] R. Fraïssé, *Sur quelques classifications des systèmes de relations*, ***Publ. Sci. Univ. Alger, Sér. A***, vol. 1 (1954), pp. 35–182.

[51] Harvey Friedman, *Primitive recursive reals,* ,
http://www.cs.nyu.edu/pipermail/fom/2006-April/010452.html, April 2006.

[52] F. Gécseg and M. Steinby, *Tree languages*, ***Handbook of formal languages, vol. 3 : Beyond words*** (G. Rozenberg and A. Salomaa, editors), Springer Verlag, Berlin, 1997, pp. 1–68.

[53] Ju. V. Glebskiĭ, D. I. Kogan, M. I. Liogon′kiĭ, and V. A. Talanov, *Volume and fraction of satisfiability of formulas of the lower predicate calculus*, ***Otdelenie Matematiki, Mekhaniki i Kibernetiki Akademii Nauk Ukrainskoĭ SSR. Kibernetika***, (1969), no. 2, pp. 17–27.

[54] A. Glikson and J.A. Makowsky, *NCE graph grammars and clique-width*, ***Proceedings of the 29th international workshop on graph-theoretic concepts in computer science (wg 2003), elspeet (the netherlands)*** (H.L. Bodlaender, editor), Lecture Notes in Computer Science, vol. 2880, Springer, 2003, pp. 237–248.

[55] M. C. Golumbic and U. Rotics, *On the clique-width of some perfect graph classes*, ***Internation Journal of Foundations of Computer Science***, vol. 11 (2000), pp. 423–443.

[56] E. Grädel, P. Kolaitis, and M. Vardi, *On the decision problem for two-variable first-order logic*, this Bulletin, vol. 3 (1997), pp. 53–69.

[57] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, ***Concrete mathematics. a foundation for computer science***, Addison Wesley, Reading, Massachusetts, 1989.

[58] Étienne Grandjean, *Complexity of the first-order theory of almost all finite structures*, ***Information and Control***, vol. 57 (1983), no. 2&3, pp. 180–204.

[59] ———, *The spectra of first-order sentences and computational complexity*, ***SIAM Journal on Computing***, vol. 13 (1984), no. 2, pp. 356–373.

[60] ———, *Universal quantifiers and time complexity of random access machines*, ***Logic and machines: Decision problems and complexity, proceedings of the symposium "Rekursive Kombinatorik" held from may 23-28, 1983 at the institut für mathematische logik und grundlagenforschung der universität münster/westfahlen*** (Egon Börger, Gisbert Hasenjaeger, and Dieter Rödding, editors), Lecture Notes in Computer Science, vol. 171, Springer, 1984, pp. 366–379.

[61] ———, *Universal quantifiers and time complexity of random access machines*, ***Mathematical Systems Theory***, vol. 18 (1985), no. 2, pp. 171–187.

[62] ———, *First-order spectra with one variable*, ***Computation theory and logic, in memory of dieter rödding*** (Egon Börger, editor), Lecture Notes in Computer Science, vol. 270, Springer, 1987, pp. 166–180.

[63] ———, *A natural NP-complete problem with a nontrivial lower bound*, ***SIAM Journal on Computing***, vol. 17 (1988), no. 4, pp. 786–809.

[64] ———, *First-order spectra with one variable*, ***Journal of Computer and System Sciences***, vol. 40 (1990), no. 2, pp. 136–153.

[65] Andrzej Grzegorczyk, *Some classes of recursive functions*, ***Rosprawy Matematyczne***, vol. 4 (1953), pp. 1–46.

[66] Yuri Gurevich and Saharon Shelah, *Spectra of monadic second-order formulas with one unary function*, ***18th ieee symposium on logic in computer science (lics 2003), 22-25 june 2003, ottawa, canada, proceedings*** (Los Alamitos, California, USA), IEEE Computer Society, 2003, pp. 291–300.

[67] Petr Hájek, *On logics of discovery*, ***Mathematical foundations of computer science 1975, 4th symposium, mariánské lázne, czechoslovakia, september 1-5, 1975, proceedings*** (Jirí Becvár, editor), Lecture Notes in Computer Science, vol. 32, Springer, 1975, pp. 30–45.

[68] Frank Harary and Edgar M. Palmer, ***Graphical enumeration***, Academic Press, 1973.

[69] M.A. Harrison, ***Introduction to formal language theory***, Addison Wesley, 1978.

[70] J. Hartmanis and R.E. Stearns, *On the complexity of algorithms*, ***Transactions of the American Mathematical Society***, vol. 117 (1965), pp. 285–306.

[71] William Hesse, Eric Allender, and David A. Mix Barrington, *Uniform constant-depth threshold circuits for division and iterated multiplication.*, ***J. Comput. Syst. Sci.***, vol. 65 (2002), no. 4, pp. 695–716.

[72] W. Hodges, *Model Theory*, ***Encyclopedia of Mathematics and its Applications***, vol. 42, Cambridge University Press, 1993.

[73] Aaron Hunter, *Spectrum hierarchies and subdiagonal functions*, ***18th international symposium on logic in computer science (lics'03)***, IEEE Press, 2003, pp. 281–290.

[74] ———, *Limiting cases for spectrum closure results*, ***Australasian Journal of Logic***, vol. 2 (2004), pp. 70–90.

[75] N. Immerman, *Languages that capture complexity classes*, ***SIAM Journal on Computing***, vol. 16 (1987), no. 4, pp. 760–778.

[76] Neil Immerman, ***Descriptive complexity***, Graduate Texts in Computer Science, Springer, 1999.

[77] T. Johnson, N. Robertson, P. Seymour, and R. Thomas, *Directed tree–width*, ***Journal of Combinatorial Theory, Serie B***, vol. 81 (2001), no. 1, pp. 138–154.

[78] Neil D. Jones, *Space-bounded reducibility among combinatorial problems*, ***J. Comput. Syst. Sci.***, vol. 11 (1975), no. 1, pp. 68–85.

[79] Neil D. Jones and Alan L. Selman, *Turing machines and the spectra of first-order formulas with equality*, ***Conference record, fourth annual acm symposium on theory of computing, 1-3 may 1972, denver, colorado, usa*** (New York, NY, USA), ACM Press, 1972, pp. 157–167.

[80] ———, *Turing machines and the spectra of first-order formulas*, ***The Journal of Symbolic Logic***, vol. 39 (1974), pp. 139–150.

[81] Laszlo Kalmár, *Egyszerü példa eldönthetetlen aritmetikai problémára. (Ein einfaches Beispiel für ein unentscheidbares arithmetisches Problem.)*, ***Mate és Fizikai Lapok***, vol. 50 (1943), pp. 1–23 (Hungarian. German summary).

[82] Georg Kreisel, *On the interpretation of non-finitist proofs. II. Interpretation of number theory. Applications*, ***The Journal of Symbolic Logic***, vol. 17 (1952), no. 2, pp. 43–58.

[83] S.-Y. Kuroda, *Classes of languages and linear-bounded automata*, ***Information and Control***, vol. 7 (1964), no. 2, pp. 207–223.

[84] Leonid Libkin, ***Elements of finite model theory***, Texts in Theoretical Computer Science, Springer, 2004.

[85] Bernd Loescher, *One unary function says less than two in existential second order logic*, ***Information Processing Letters***, vol. 61 (1997), no. 2, pp. 69–75.

[86] L. Lovász and P. Gács, *Some remarks on generalized spectra*, ***Zeitschr. f. math. Logik und Grundlagen d. Math.***, vol. 23 (1977), pp. 547–554.

[87] Leopold Löwenheim, *Über möglichkeiten im relativkalkül*, ***Mathematische Annalen***, vol. 76 (1915), no. 4, pp. 447–470.

[88] James F. Lynch, *Complexity classes and theories of finite models*, ***Mathematical Systems Theory***, vol. 15 (1982), no. 2, pp. 127–144.

[89] J.A. Makowsky, *Algorithmic uses of the Feferman-Vaught theorem*, ***Annals of Pure and Applied Logic***, vol. 126.1-3 (2004), pp. 159–213.

[90] J.A. Makowsky and J.P. Mariño, *Tree-width and the monadic quantifier hierarchy*, ***Theoretical Computer Science***, vol. 303 (2003), pp. 157–170.

[91] J.A. Makowsky and Y. Pnueli, *Arity vs. alternation in second order logic*, ***Annals of Pure and Applied Logic***, vol. 78 (1996), no. 2, pp. 189–202.

[92] J.A. Makowsky and U. Rotics, *On the cliquewidth of graphs with few P_4's*, ***International Journal on Foundations of Computer Science***, vol. 10 (1999), pp. 329–348.

[93] J. Miller, *Statement of result on primitive recursive reals*, E-mail communication.

[94] Shaokui Mo, *Abstract of [95]*, Zbl 0729.03004.

[95] ———, *The solution of Scholz problems.*, ***Chin. Ann. Math., Ser. A***, vol. 12 (1991), no. 1, pp. 89–97 (Chinese).

[96] Malika More and Frédéric Olive, *Rudimentary languages and second order logic*, ***Mathematical Logic Quarterly***, vol. 43 (1997), pp. 419–426.

[97] M. Mortimer, *On Languages with Two Variables*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 21 (1975), pp. 135–140.

[98] Andrzej Mostowski, *Concerning a problem of H.Scholz*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 12 (1956), pp. 210–214.

[99] John Myhill, *Linear bounded automata*, ***Technical Note 60–165***, Wright-Patterson Air Force Base, Wright Air Development Division, Ohio, 1960.

[100] V. A. Nepomnjaščiĭ, *Rudimentary predicates and Turing computations*, ***Dokl. Akad. Nauk SSSR***, vol. 195 (1970), pp. 282–284.

[101] V. A. Nepomnjascii, *Examples of predicates inexpressible by s-rudimentary formulas*, ***Kibernetika (Kiev)***, vol. 2 (1978), pp. 44–46, Russian, English summary.

[102] Frédéric Olive, *A conjunctive logical characterization of nondeterministic linear time*, ***Computer science logic, 11th international workshop, csl'97, annual conference of the eacsl, aarhus, denmark, august 23-29, 1997, selected papers*** (Mogens Nielsen and Wolfgang Thomas, editors), Lecture Notes in Computer Science, vol. 1414, Springer, 1998, pp. 360–372.

[103] M. Otto, ***Bounded variable logics and counting — A study in finite models***, vol. 9, Springer-Verlag, 1997, IX+183 pages.

[104] R. Parikh, *On context-free languages*, ***jacm***, vol. 13 (1966), pp. 570–581.

[105] Pavel Pudlák, *The observational predicate calculus and complexity of computations*, ***Commentationes Mathematicae Universitatis Carolinae***, vol. 16 (1975), pp. 395–398.

[106] Michael O. Rabin, *A simple method of undecidability proofs and some applications*, ***Logic, methodology and phylosophy of science (proceedings 1964 international congress)*** (Amsterdam) (Bar-Hillel, editor), vol. 1, North-Holland, 1965, pp. 58–68.

[107] M.O. Rabin, *A simple method for undecidability proofs and some applications*, ***Logic, methodology and philosophy of science II*** (Y. Bar Hillel, editor), Studies in Logic, North Holland, 1965, pp. 58–68.

[108] Paulo Ribenboim, ***The book of prime number records***, second ed., Springer, 1989.

[109] Robert W. Ritchie, *Classes of recursive functions of predictable complexity*, ***Ph.D. thesis***, Princeton University, 1960.

[110] ———, *Classes of predictably computable functions*, ***Transactions of the American Mathematical Society***, vol. 106 (1963), pp. 139–173.

[111] N. Robertson and P. D. Seymour, *Graph minors. ii. algorithmic aspects of tree–width*, ***Journal of Algorithms***, vol. 7 (1986), pp. 309–322.

[112] Abraham Robinson, *Review of [7]*, MR0077468 (17,1038c).

[113] Dieter Rödding and Helmut Schwichtenberg, *Bemerkungen zum Spektralproblem*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 18 (1972), pp. 1–12.

[114] H.E. Rose, ***Subrecursion. Functions and hierarchies.***, Oxford Logic Guides, 9. Oxford: Clarendon Press. XIII, 191 p., 1984 (English).

[115] U. Rotics, *Efficient algorithms for generally intractable graph problems restricted to specific classes of graphs*, ***Ph.D. thesis***, Technion- Israel Institute of Technology, 1998.

[116] Heinrich Scholz, *Ein ungelöstes Problem in der symbolischen Logik*, ***The Journal of Symbolic Logic***, vol. 17 (1952), p. 160.

[117] S. Shelah, *Spectra of monadic second order sentences*, Paper No. 817, Electronically available at arXiv:math/0405158, 2004.

[118] Thoralf Skolem, *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über die dichte Mengen*, ***Videnskapsselskapets skrifter, I. Matematisk-naturvidenskabelig klasse***, vol. 4 (1919), pp. 1–36.

[119] D. Skordev, *Computability of real numbers by using a given class of functions in the set of natural numbers*, ***Mathematical Logic Quarterly***, vol. 48 (2002), no. Suppl.1, pp. 91–106.

[120] ———, *On the subrecursive computability of several famous constants*, ***Journal of Universal Computer Science***, vol. 14 (2008), no. 6, pp. 861–875.

[121] Raymond M. Smullyan, ***Theory of formal systems***, Annals of Mathematical Studies, vol. 47, Princeton University Press, Princeton, New Jersey, 1961.

[122] E. Specker, *Nicht konstruktiv beweisbare Sätze in der Analysis*, ***The Journal of Symbolic Logic***, vol. 14 (1949), pp. 145–158.

[123] A. Tarski, A. Mostowski, and R. M. Robinson, ***Undecidable theories***, North-Holland Publishing Company, 1953.

[124] Alfred Tarski, *Contribution to the theory of models, I,II*, ***Indagationes Mathematicae***, vol. 16 (1954), pp. 572–588.

[125] K. Tent and M. Ziegler, *Low functions of reals*, Manuscript, 2009.

[126] Boris A. Trakhtenbrot, *Impossibility of an algorithm for the decision problem in finite classes*, ***Doklady Akademii Nauk SSSR***, vol. 70 (1950), pp. 569–572.

[127] Alan M. Turing, *On computable numbers, with an application to the entscheidungsproblem*, ***Procedings of the London Mathematical Society***, vol. 42 (1936), no. 2, pp. 230–265.

[128] M. Vardi, *The complexity of relational query languages*, ***Stoc'82***, ACM, 1982, pp. 137–146.

[129] M. Waldschmidt, *Words and transcendence*, ***Analytic number theory - essays in honour of Klaus Roth***, Cambridge University Press, 2008, pp. 449–470.

[130] Herbert S. Wilf, ***Generatingfunctionology***, Academic Press, 1990.

[131] A.R. Woods, *Some problems in logic and number theory and their connections*, ***Ph.D. thesis***, University of Manchester, 1981.

[132] Celia Wrathall, *Rudimentary predicates and relative computation*, ***SIAM J. Comput.***, vol. 7 (1978), no. 2, pp. 194–209.

[133] Ping Zhu, *Translation of [95]*, Personal communication.

ÉQUIPE DE LOGIQUE, UNIVERSITÉ PARIS 7, FRANCE
E-mail: durand@logique.jussieu.fr

DIKU, UNIVERSITY OF COPENHAGEN, DENMARK
E-mail: neil@diku.dk

FACULTY OF COMPUTER SCIENCE, TECHNION, HAIFA, ISRAEL
E-mail: janos@cs.technion.ac.il

UNIV CLERMONT 1, LAIC, FRANCE
E-mail: more@laic.u-clermont1.fr