C. R. Acad. Sci. Paris, Ser. I 353 (2015) 931–936

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Functional analysis/Computer science

New barriers in complexity theory: On the solvability complexity index and the towers of algorithms

Nouvelles barrières en théorie de la complexité : sur l'indice de complexité de la resolubilité et les tours d'algorithmes

Jonathan Ben-Artzi [a], Anders C. Hansen [b,c], Olavi Nevanlinna [d], Markus Seidel [e]

[a] *Department of Mathematics, Imperial College, United Kingdom*
[b] *DAMTP, University of Cambridge, United Kingdom*
[c] *Department of Mathematics, University of Oslo, Norway*
[d] *Department of Mathematics and Systems Analysis, Aalto University, Finland*
[e] *West Saxon University of Applied Sciences, Zwickau, Germany*

A R T I C L E I N F O

Article history:
Received 10 June 2015
Accepted 19 August 2015

Presented by the Editorial Board

A B S T R A C T

We report on new barriers in the theory of computations. These barriers show that the standard theory of computations and complexity theory is insufficient for many core problems in computational theory. Thus we are in need for a new extended complexity theory. The new theory settles the long-standing computational spectral problem and also provides new fundamental algorithms for quantum mechanics.

R É S U M É

On met en evidence de nouvelles barrières en théorie du calcul. Ces barrières montrent que la théorie standard du calcul et, en particulier, la théorie de la complexité ne résolvent pas de nombreux problèmes de base de la théorie du calcul. On se trouve face à la nécessité d'une extension de la théorie de la complexité. Cette nouvelle théorie conduit à la résolution d'un problème ancien concernant le calcul spectral. Elle conduit aussi à l'élaboration de nouveaux algorithmes fondamentaux utiles en mécanique quantique.

1. Introduction

In the classical theory of computations and complexity, we typically have two scenarios: (i) the problem can be computed via an algorithm in finite time given a finite input, or (ii) there is a sequence of approximations produced by an algorithm, where each output is produced in finite time, and the solution to the problem is the limit of this sequence (think about computing an integral of a function, computing eigenvalues of matrices, etc.). It may seem that most, if not all, com-

E-mail addresses: J.Ben-Artzi@imperial.ac.uk (J. Ben-Artzi), A.Hansen@damtp.cam.ac.uk (A.C. Hansen), olavi.nevanlinna@aalto.fi (O. Nevanlinna), markus.seidel@fh-zwickau.de (M. Seidel).

http://dx.doi.org/10.1016/j.crma.2015.08.002
1631-073X/

putational problems would fit into these two scenarios. It is therefore surprising that there are many problems that do not fit into this framework. Moreover, some of these problems are at the heart of computational theory, and include: computing spectra of operators, solving inverse problems, root finding of polynomials with rational maps, and more. The issue is that these problems cannot be computed by simply passing to a limit. However, one can compute the solution by using several limits. Such a phenomenon may be unexpected, but this turns out to be at the heart of the boundaries of computational mathematics.

One of the first rigorous examples of this phenomenon came with the pioneering work of McMullen [4] and Doyle and McMullen [2] on polynomial root finding with rational maps. This example is not exclusive, and in fact there are many problems in this category. In this paper we report on some of the new results concerning these barriers. The main concepts are the Solvability Complexity Index (SCI) and towers of algorithms that merge the frameworks from [3] and [2]. The SCI is the smallest number of limits needed in order to compute a desired quantity given a certain toolbox of allowed mathematical operations.

Main consequences

(I) **Spectral problems.**
It is impossible to compute spectra and essential spectra of infinite matrices in less than three limits. This is universal for all algorithms regardless of the operations allowed (arithmetic operation, radicals, etc.). However, it is possible to compute spectra and essential spectra in three limits when allowing arithmetic operations of complex numbers. It is impossible to compute spectra of self-adjoint infinite matrices in one limit. This is universal regardless of the operations allowed; however, it is possible to compute spectra of matrices with controlled growth on their resolvent in two limits (this then includes normal operators).

(II) **Quantum mechanics.** One can compute spectra of all non-normal Schrödinger operators with bounded potential with bounded local total variation in two limits. If the operator is normal or if the potential blows up at infinity, then one limit suffices.

(III) **Inverse problems.** It is impossible to compute the solution to a general infinite linear system in one limit, yet it is possible in two. For matrices with known/controllable off-diagonal decay, one can compute the solution to a linear system in one limit.

(IV) **Impossibility of error control.** Most problems with SCI > 1 (that is, where more than one limit is necessary) can never be computed with error control, i.e. it is impossible to design an algorithm that can compute an approximation to the solution and know when one is “epsilon” away from the true solution.

(V) **A new complexity theory.** The SCI framework provides a new complexity theory for problems that do not fit into the existing complexity theories. In particular, current complexity theory cannot handle problems that require several limits in the computation.

(VI) **A new classification theory.** If SCI $= k > 1$ for a certain class of problems, the question is: which subclasses of problems will have SCI $= k-1$ or SCI $= k-2$, etc.?

(VII) **Decision problems and Turing machines.** There is a connection between the SCI and the Arithmetical Hierarchy. In particular, the Δ_m sets in the Arithmetical Hierarchy can be equivalently characterised in term of the SCI. Thus, one may view the SCI as a classification tool that is a generalisation of this complexity hierarchy to arbitrary computational problems.

2. The main results

We propose a unified theory for computational problems. The basic objects are: Ω is some set, called the *primary* set, Λ is a set of complex valued functions on Ω, called the *evaluation* set, $\mathcal{M}$ is a metric space, and $\Xi : \Omega \to \mathcal{M}$ is called the *problem* function. The set Ω is the set of objects that give rise to our computational problems. It can be a family of (finite or infinite) matrices, polynomials, Schrödinger or Dirac operators, inverse problems, etc. The problem function $\Xi : \Omega \to \mathcal{M}$ is what we are interested in computing. It could be the set of eigenvalues of an $n \times n$ matrix, root(s) of a polynomial, the spectrum of an operator, etc. Finally, the set Λ is the collection of functions that provide us with the information we are allowed to read, say matrix elements, polynomial coefficients or pointwise values of a potential function of a Schrödinger operator, for example. This leads to the following definition.

Definition 2.1 *(Computational problem).* Given a primary set Ω, an evaluation set Λ, a metric space $\mathcal{M}$ and a problem function $\Xi : \Omega \to \mathcal{M}$ we call the collection $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ a *computational problem*.

The goal of this abstract definition is to allow most of the known computational problems into the framework. However, to make this abstract definition a little more concrete, let us consider the following example.

Example I. Let $\Omega = \mathcal{B}(\mathcal{H})$ be the set of all bounded linear operators on a separable Hilbert space $\mathcal{H}$, and the problem function Ξ be the mapping $A \mapsto \mathrm{sp}(A)$ (the spectrum of A). Here $(\mathcal{M}, d)$ is the set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$. The evaluation functions in Λ could consist of the family of all functions $f_{i,j} : A \mapsto \langle Ae_j, e_i\rangle$,

$i, j \in \mathbb{N}$, which provide the entries of the matrix representation of A w.r.t. an orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$. Of course, Ω could be a strict subset of $\mathcal{B}(\mathcal{H})$, for example the set of self-adjoint or normal operators, and Ξ could have represented the pseudospectrum, the essential spectrum or any other interesting information about the operator.

Our aim is to find and to study families of functions that permit to approximate the function Ξ. The main pillar of our framework is the concept of a tower of algorithms.

Definition 2.2 *(General algorithm).* Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, a *general algorithm* is a mapping $\Gamma : \Omega \to \mathcal{M}$ such that for each $A \in \Omega$

(i) there exists a finite subset of evaluations $\Lambda_\Gamma(A) \subset \Lambda$,
(ii) the action of Γ on A only depends on $\{A_f\}_{f\in\Lambda_\Gamma(A)}$ where $A_f := f(A)$,
(iii) for every $B \in \Omega$ such that $B_f = A_f$ for every $f \in \Lambda_\Gamma(A)$, it holds that $\Lambda_\Gamma(B) = \Lambda_\Gamma(A)$.

A general algorithm has no restrictions on the operations allowed. The only restriction is that it can only take a finite amount of information, though it *is* allowed to *adaptively* choose the finite amount of information it reads depending on the input. Condition (iii) assures that the algorithm reads the information in a consistent way.

Definition 2.3 *(Tower of algorithms).* Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, a *tower of algorithms of height k for* $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a family of sequences of functions

$$\Gamma_{n_k} : \Omega \to \mathcal{M},\ \Gamma_{n_k,n_{k-1}} : \Omega \to \mathcal{M}, \dots,\ \Gamma_{n_k,\dots,n_j} : \Omega \to \mathcal{M},$$

where $n_k, \dots, n_1 \in \mathbb{N}$ and the functions $\Gamma_{n_k,\dots,n_1}$ at the "lowest level" of the tower are general algorithms in the sense of Definition 2.2. Moreover, for every $A \in \Omega$,

$$\Xi(A) = \lim_{n_k\to\infty} \Gamma_{n_k}(A), \quad \Gamma_{n_k,\dots,n_{j+1}}(A) = \lim_{n_j\to\infty} \Gamma_{n_k,\dots,n_j}(A) \quad j = k-1, \dots, 1.$$

In addition to a general tower of algorithms (defined above), we will focus on arithmetic towers.

Definition 2.4 *(Arithmetic tower).* Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, an *Arithmetic Tower of Algorithms* of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a tower of algorithms where the lowest level functions

$$\Gamma = \Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M}$$

satisfy the following: for each $A \in \Omega$, the action of Γ on A consists of only finitely many arithmetic operations on $\{A_f\}_{f\in\Lambda_\Gamma(A)}$, where we remind that $A_f = f(A)$.

Given the definitions above, we can now define the key concept, namely, the Solvability Complexity Index.

Definition 2.5 *(Solvability Complexity Index, SCI).* A given computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is said to have *Solvability Complexity Index* $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = k$ with respect to a tower of algorithms of type α if k is the smallest integer for which there exists a tower of algorithms of type α of height k. If no such tower exists, then $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = \infty$. If there exists a tower $\{\Gamma_n\}_{n\in\mathbb{N}}$ of type α and height one such that $\Xi = \Gamma_{n_1}$ for some $n_1 < \infty$, then we define $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = 0$. The type α may be General, or Arithmetic, denoted respectively G and A. We may sometimes write $\mathrm{SCI}(\Xi, \Omega)_\alpha$ to simplify the notation when $\mathcal{M}$ and Λ are obvious.

We will let $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{A}}$ and $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}}$ denote the SCI with respect to an arithmetic tower and a general tower, respectively. Note that a general tower means just a tower of algorithms as in Definition 2.3, where there are no restrictions on the mathematical operations. Thus, clearly $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{A}} \ge \mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}}$. The evaluation sets Λ and $\Lambda_\Gamma(A)$, given a general algorithm Γ and an element $A \in \Omega$, are crucial when determining the SCI as the following example demonstrates.

Example II. Suppose we want to compute the area of a disc given its radius. Denote the set of discs by Ω. Let f be the evaluation function that assigns to a closed disc D its radius $r = f(D)$. Let $\Lambda_1 = \{f\}$ and let Λ_2 be the union of Λ_1 and the set of all constant functions on Ω. If we allow Λ_2, then the SCI of this problem with respect to an arithmetic tower is obviously zero as the formula πr^2 immediately gives the answer. However, if we only allow Λ_1, then we must have that $\mathrm{SCI} > 0$ (π cannot be obtained in finitely many arithmetic operations from the input r).

Remark I. Motivated by the example above, there are several settings that may be considered when analyzing the SCI. See the following examples. (I): Λ contains all constant functions. (II): Let Γ be a general algorithm, $A, B \in \Omega$ and let $\hat{\Lambda}_\Gamma(A) \subset \Lambda_\Gamma(A)$ denote the set of constant functions. Then $\hat{\Lambda}_\Gamma(A) = \hat{\Lambda}_\Gamma(B)$. In particular, the constant functions are the same for A and B. (III): Λ contains no constant functions. We will specify which of the above conditions are used in the upper bounds of the SCI.

2.1. The main theorems

Spectra of bounded operators. Computing spectra of operators is one of the core problems in the theory of computations and this field has produced a vast amount of research. We refer to [1] and the references therein.

Definition 2.6 *(Dispersion).* We say that the *dispersion* of $A \in \mathcal{B}(\ell^2(\mathbb{N}))$ is bounded by the function $f : \mathbb{N} \to \mathbb{N}$ if $\max\{\|(I - P_{f(m)})AP_m\|, \|P_m A(I - P_{f(m)})\|\} \to 0$ as $m \to \infty$, where P_m denotes the projection onto $\mathrm{span}\{e_1, \dots, e_m\}$ and $\{e_j\}$ is the canonical basis. We denote the set of all operators with the dispersion bounded by f by $\mathcal{B}_f(\ell^2(\mathbb{N}))$.

Definition 2.7 *(Controlled resolvent).* Given some continuous function $g : [0, \infty) \to [0, \infty)$ that vanishes only at $x = 0$ and such that $g(x) \to \infty$ as $x \to \infty$, define $\mathcal{R}_g(\mathcal{H})$ to be the class of all closed linear operators on the Hilbert space $\mathcal{H}$ satisfying $\|(A - zI)^{-1}\|^{-1} \geq g(\mathrm{dist}(z, \mathrm{sp}(A)))$ for every $z \in \mathbb{C}$, where we use the convention $\|B^{-1}\|^{-1} := 0$ if B has no bounded inverse. Note that normal (and self-adjoint) operators satisfy this condition with $g(x) = x$.

Definition 2.8 *(Pseudospectra).* For $N \in \mathbb{Z}_+$ and $\epsilon > 0$, the (N, ϵ)-*pseudospectrum* of a bounded linear operator $A \in \mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ is defined to be the set

$$\mathrm{sp}_{N,\epsilon}(A) := \{z \in \mathbb{C} : \|(A - zI)^{-2^N}\|^{2^{-N}} \geq 1/\epsilon\}.$$

We will consider variants of the computational problem suggested in Example I. Consider the following classes of operators: $\Omega_1 = \mathcal{B}(\ell^2(\mathbb{N}))$, $\Omega_2 = \mathcal{B}_f(\ell^2(\mathbb{N}))$, $\Omega_3 = \mathcal{R}_g(\ell^2(\mathbb{N})) \cap \mathcal{B}(\ell^2(\mathbb{N}))$, $\Omega_4 = \mathcal{K}(\ell^2(\mathbb{N}))$, the set of all compact operators on $\ell^2(\mathbb{N})$, and $\Omega_5 = \Omega_2 \cap \Omega_3$. Define the following problem functions: $\Xi_1(A) = \mathrm{sp}(A)$, for $\epsilon > 0$ and $N \in \mathbb{Z}_+$, let $\Xi_2(A) = \mathrm{sp}_{N,\epsilon}(A)$, as well as $\Xi_3(A) = \mathrm{sp}_{\mathrm{ess}}(A)$ (the essential spectrum). When considering Ω_3 then Λ contains, besides the usual evaluations $f_{i,j} : A \mapsto \langle Ae_j, e_i\rangle$ $(i, j \in \mathbb{N})$ the constant functions $g_{i,j} : A \mapsto g(i/j)$ $(i, j \in \mathbb{N})$, which provide the values of g in all positive rational numbers. When considering Ω_2, the values $f(m)$ $(m \in \mathbb{N})$ shall be available to the algorithms as constant evaluation functions. We then have the following theorem.

Theorem 2.9 *(SCI and spectra). Given the above setup*

$$\begin{aligned}
\textbf{\textit{Spectrum:}}\ & \mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{A}} = 3, \quad \mathrm{SCI}(\Xi_1, \Omega_i)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega_i)_{\mathrm{A}} = 2, \quad i = 2, 3,\\
& \mathrm{SCI}(\Xi_1, \Omega_4)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega_4)_{\mathrm{A}} = 1, \quad \mathrm{SCI}(\Xi_1, \Omega_5)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega_5)_{\mathrm{A}} = 1,\\
\textbf{\textit{Pseudospectrum:}}\ & \mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} = 2, \quad \mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{A}} = 1,\\
& \mathrm{SCI}(\Xi_2, \Omega_4)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2, \Omega_4)_{\mathrm{A}} = 1,\\
\textbf{\textit{Ess-spectrum:}}\ & \mathrm{SCI}(\Xi_3, \Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_3, \Omega_1)_{\mathrm{A}} = 3, \quad \mathrm{SCI}(\Xi_3, \Omega_5)_{\mathrm{G}} = \mathrm{SCI}(\Xi_3, \Omega_5)_{\mathrm{A}} = 2.
\end{aligned}$$

Note that all the lower bounds in the theorem above are valid for a general tower. This implies that regardless of the permissible operations, no algorithm can improve on these lower bounds.

Inverse problems. We seek solutions to problems of the form $Ax = b$, where $A \in \mathcal{B}_{\mathrm{inv}}(\ell^2(\mathbb{N}))$, the class of bounded invertible operators, and $b \in \ell^2(\mathbb{N})$. In particular, $\Xi(A, b) = x$. We define the classes $\Omega_1 = \mathcal{B}_{\mathrm{inv}}(\ell^2(\mathbb{N})) \times \ell^2(\mathbb{N})$, $\Omega_2 = (\mathcal{B}_{\mathrm{inv}}(\ell^2(\mathbb{N})) \cap \mathcal{B}_f(\ell^2(\mathbb{N}))) \times \ell^2(\mathbb{N})$. The metric space $\mathcal{M}$ would simply be $\ell^2(\mathbb{N})$ and Λ the collection of mappings $\{f_{i,j}\}_{i \in \mathbb{N}, j \in \mathbb{Z}_+}$ where $f_{i,j} : (A, b) \mapsto \langle Ae_j, e_i\rangle$ for $j \in \mathbb{N}$ and $f_{i,0} : (A, b) \mapsto \langle b, e_i\rangle$. When considering Ω_2, the values $f(m)$ $(m \in \mathbb{N})$ shall be available to the algorithms as constant evaluation functions. Having defined these computational problems, we have the following.

Theorem 2.10 *(Linear systems). Given the setup above, we have* $\mathrm{SCI}(\Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Omega_1)_{\mathrm{A}} = 2$ *and* $\mathrm{SCI}(\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Omega_2)_{\mathrm{A}} = 1$.

Quantum mechanics. Consider the Schrödinger operator $H = -\Delta + V$ with bounded potentials. Let Ω be some set of bounded potential functions and let $\Xi : V \mapsto \mathrm{sp}(-\Delta + V)$, where the domain $\mathcal{D}(-\Delta + V) = \mathrm{W}^{2,2}(\mathbb{R}^d)$ (the standard Sobolev space). Given that the spectra are unbounded, let $(\mathcal{M}, d_{\mathrm{AW}})$ denote the set of closed subsets of $\mathbb{C}$ equipped with the *Attouch–Wets* metric (see [1]). Also, Λ will be the set of all evaluations $f_x : V \mapsto V(x)$, $x \in \mathbb{R}^d$ and all constant functions. We consider the following potential classes: $\Omega_1 = \{V : V \in \mathrm{L}^\infty(\mathbb{R}^d) \cap \mathrm{BV}_\phi(\mathbb{R}^d)\}$, where $\phi : [0, \infty) \to [0, \infty)$ is some increasing function, $\mathrm{BV}_\phi(\mathbb{R}^d) = \{f : \mathrm{TV}(f_{[-a,a]^d}) \leq \phi(a)\}$, $f_{[-a,a]^d}$ meaning f restricted to the box $[-a, a]^d$, TV being the total variation of a function in the sense of Hardy and Krause (see [1]). Also, $\Omega_2 = \Omega_1 \cap \{V : -\Delta + V \in \mathcal{R}_g(\mathrm{L}^2(\mathbb{R}^d))\}$. We then get the following result.

Theorem 2.11 *(Schrödinger operators). Given the setup above, we have* $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \leq 2$, *and* $\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} = 1$.

Impossibility of error control. A key concept in computations is error control. Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = k$ for some tower of algorithms of type α, and a tower of algorithms of height k, we want to control the convergence $\Gamma_{n_k} \to \Xi, \dots, \Gamma_{n_k,\dots,n_1} \to \Gamma_{n_k,\dots,n_2}$. For $\epsilon > 0$, how big do $n_k, \dots, n_1$ have to be so that $d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) \le \epsilon$, for all $A \in \Omega$. Unfortunately, such choices of $n_k, \dots, n_1$ may be impossible. More precisely, problems with SCI greater than one with respect to a General tower will never have error control.

Theorem 2.12 *(No global error control). Let* $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ *be a computational problem with* $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \ge 2$. *Suppose that there is a general tower of algorithms of height k,* $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ *for* $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$. *Then there do NOT exist integers* $n_k = n_k(m), \dots, n_1 = n_1(m)$ *(depending on m) such that* $d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) \le \frac{1}{m}$, *for all* $A \in \Omega$ *and for all* $m \in \mathbb{N}$.

A weaker concept than global error control is local error control: $\forall A \in \Omega$ and $\forall \epsilon > 0$, $\exists n_k, \dots, n_1$ such that $d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) < \epsilon$. Indeed, the existence of $n_k, \dots, n_1$ is guaranteed by the definition of a tower of algorithms. However, the integers $n_k, \dots, n_1$ cannot be computed as the next theorem demonstrates, and thus local error control is also impossible.

Theorem 2.13 *(Local error control cannot be computed). Given a computational problem* $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ *with* $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \ge 2$, *suppose that there is a general tower of algorithms* $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ *of height k for the computational problem. Then, there does NOT exist a sequence* $\{\tilde{\Gamma}_n\}$ *of general algorithms* $\tilde{\Gamma}_n : \Omega \to \mathbb{N}^k$ *such that for any* $A \in \Omega$, $d(\Gamma_{\tilde{\Gamma}_n(A)_k,\dots,\tilde{\Gamma}_n(A)_1}(A), \Xi(A)) < \frac{1}{n}$.

The SCI, Turing machines and the arithmetical hierarchy. Given a subset $A \subset \mathbb{Z}_+$ with characteristic function χ_A being definable in First-Order Arithmetics, we are interested in the SCI of deciding whether a given number $x \in \mathbb{Z}_+$ belongs to A or not. In other words, we want to determine the value of the characteristic function of A at the point x. Thus, we want to consider towers of algorithms for χ_A where the functions/relations at the lowest level shall be computable, and we again ask for the minimal height. More precisely, we consider the primary set $\Omega := \mathbb{Z}_+$, the evaluation set $\Lambda = \{\lambda\}$ consisting of the function $\lambda : \mathbb{Z}_+ \to \mathbb{C}$, $x \mapsto x$, the metric space $\mathcal{M} := (\{\mathrm{Yes}, \mathrm{No}\}, d_{\mathrm{discr}})$, where d_{discr} denotes the discrete metric, and consider all functions $\Xi : \Omega \to \mathcal{M}$ in the Arithmetical Hierarchy. In honour of Kleene and Shoenfield we call a tower of algorithms that is computable (in the sense of Turing) a Kleene–Shoenfield tower.

Definition 2.14 *(Kleene–Shoenfield tower).* A tower of algorithms given by a family $\{\Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M} : n_k, \dots, n_1 \in \mathbb{N}\}$ of functions at the lowest level is said to be a *Kleene–Shoenfield tower*, if the function

$$\mathbb{N}^k \times \Omega \to \mathcal{M}, \quad (n_k, \dots, n_1, x) \mapsto \Gamma_{n_k,\dots,n_1}(x)$$

is computable. Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ as above, we will write $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{KS}}$ to denote the SCI with respect to a Kleene–Shoenfield tower.

We can now present the main theorem, linking the SCI and the arithmetical hierarchy.

Theorem 2.15 *(The SCI and the arithmetical hierarchy). Let* $m \in \mathbb{Z}_+$ *and recall the classes* Σ_m, Π_m, Δ_m *from the arithmetical hierarchy. If* Ξ *is* Δ_{m+1}, *then there exists a Kleene–Shoenfield tower of algorithms of height m. Conversely, if* $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{KS}} = m$ *then* Ξ *is* Δ_{m+1}, *but not* Δ_m.

This theorem follows from a result by Shoenfield [5] (see [1] for details), and has an immediate corollary.

Corollary 2.16 *(The SCI can become arbitrarily large). For every* $k \in \mathbb{N}$ *there exists a problem function* Ξ *on* Ω *with* $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{KS}} = k$.

Polynomial root finding with rational maps. A *purely iterative algorithm* [6] is a rational map $T : \mathbb{P}_d \to \mathrm{Rat}_m$, $p \mapsto T_p$ which sends any polynomial p of degree $\le d$ to a rational function T_p of a certain degree m. An important example of a purely iterative algorithm is *Newton's method*. Furthermore, a purely iterative algorithm is said to be *generally convergent* if $\lim_{n\to\infty} T_p^n(z)$ exists for (p, z) in an open dense subset of $\mathbb{P}_d \times (\mathbb{C} \cup \{\infty\})$, and the limit is a root of p. Here $T_p^n(z)$ denotes the nth iterate $T_p^n(z) = T_p(T_p^{n-1}(z))$ of T_p. For instance, Newton's method is generally convergent *only* when $d = 2$. This failure of Newton's method prompted S. Smale [6] to ask the following: "*Is there any purely iterative generally convergent algorithm for polynomial zero finding?*" This question led to the definition and theorem below.

Definition 2.17 *(Doyle–McMullen tower).* A (Doyle–McMullen) tower of algorithms is a finite sequence of generally convergent algorithms, linked together serially, so the output of one or more can be used to compute the input to the next. The final output of the tower is a single number, computed rationally from the original input and the outputs of the intermediate generally convergent algorithms.

It can be shown that a Doyle–McMullen tower is a general tower as in Definition 2.3 with the slight change that the convergence holds only for an open dense set, thus the following can be formulated in terms of the SCI.

Theorem 2.18. *(See McMullen [4]; Doyle and McMullen [2].) For $\mathbb{P}_d$ there exists a generally convergent algorithm only for $d \leq 3$. Towers of algorithms exist additionally for $d = 4$ and $d = 5$, but not for $d \geq 6$.*

By the proof of this theorem one gets that the height of the tower is three, and thus the previous theorem can be formulated in terms of the SCI as follows. For $d \leq 3$, the SCI $= 1$, for $d = 4, 5$ one has SCI $\in \{2, 3\}$ while for $d \geq 6$ there is no tower: SCI $= \infty$. See [1] for details.

Acknowledgements

A.C.H. is supported by a Royal Society Fellowship as well as the EPSRC grant EP/L003457/1.

References

[1] J. Ben-Artzi, A.C. Hansen, O. Nevanlinna, M. Seidel, Can everything be computed? – On the solvability complexity index and towers of algorithms, http://www.damtp.cam.ac.uk/research/afha/anders/SCI.pdf, 2015.
[2] P. Doyle, C. McMullen, Solving the quintic by iteration, Acta Math. 163 (3–4) (1989) 151–180.
[3] A.C. Hansen, On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators, J. Amer. Math. Soc. 24 (1) (2011) 81–124.
[4] C. McMullen, Families of rational maps and iterative root-finding algorithms, Ann. Math. (2) 125 (3) (1987) 467–493.
[5] J.R. Shoenfield, On degrees of unsolvability, Ann. Math. (2) 69 (1959) 644–653.
[6] S. Smale, The fundamental theorem of algebra and complexity theory, Bull. Amer. Math. Soc. (N.S.) 4 (1) (1981) 1–36.