European Mathematical Society Magazine

ISSN 2747-7894 / Issue 122 / December 2021

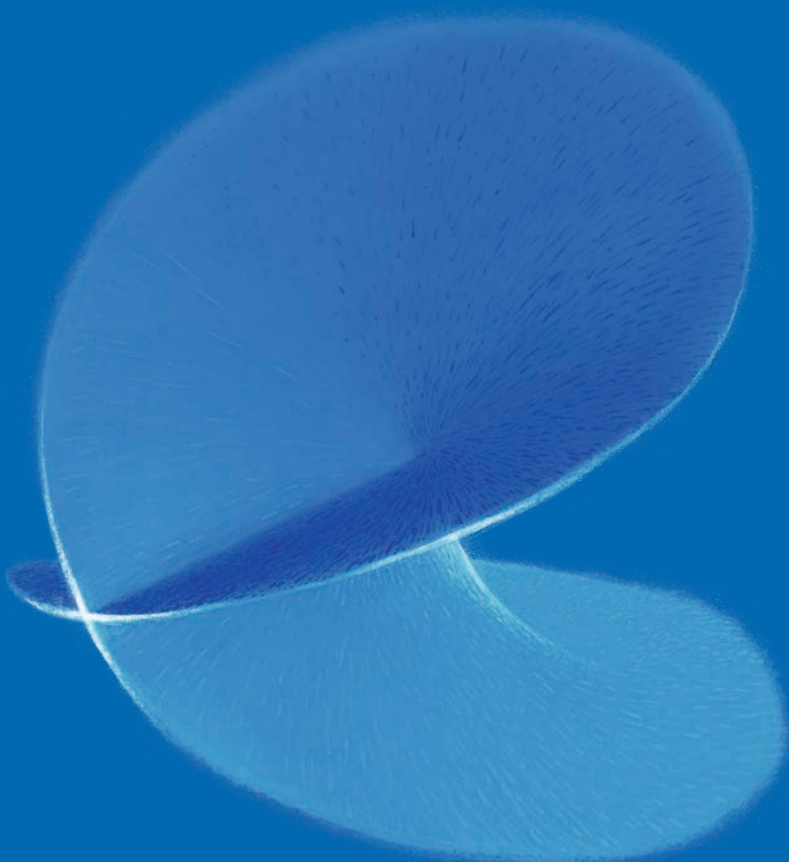
EMS Magazine

Bjørn Ian Dundas and Christian F. Skau
Abel interview 2021:
László Lovász and Avi Wigderson

Zinovy Reichstein
From Hilbert's 13th problem to
essential dimension and back

Joana Grah, Tamara Grossmann and
Julia Kroos
Her Maths Story – Sharing stories
of women in mathematics in all
walks of life

Muharem Avdispahić
The Bosnian Mathematical Society
and the mathematical life in
Bosnia and Herzegovina



EUROPEAN
MATHEMATICAL
SOCIETY

100 years Unione Matematica Italiana
800 years Università di Padova

Padova – May 23 – 27, 2022

AIM AND SCOPE Year 2022 marks two great anniversaries for the Italian mathematical community: 100 years of Unione Matematica Italiana (UMI) and 800 years of University of Padova. On this unique occasion, UMI and the Department of Mathematics "Tullio Levi-Civita" of University of Padova will organize an international conference.

PLENARY SPEAKERS

Alberto Bressan, Penn State University
Camillo De Lellis, IAS Princeton
Luca Dell'Aglio, Università della Calabria
Laura DeMarco, Harvard University
Daniele Di Pietro, Université de Montpellier
Cynthia Dwork, Harvard University
Livia Giacardi, Università di Torino
Alessandro Giuliani, Università Roma 3
Martin Hairer, Imperial College London
Andrea Mondino, University of Oxford
Giulia Saccà, Columbia University
Peter Scholze, Universität Bonn
Claire Voisin, CNRS Paris

PUBLIC LECTURE

Alessio Figalli, ETH Zurich

ROUND TABLES

Mathematical challenges in an AI driven world
Pierre Baldi, University of California, Irvine
Gitta Kutyniok, LMU München
Yann LeCun, Facebook and New York University
Tomaso Poggio, MIT Boston

The Usefulness of Useless Knowledge
Ingrid Daubechies, Duke University
Robbert Dijkgraaf, IAS Princeton
Alberto Sangiovanni-Vincentelli, Univ California, Berkeley

11^2 years after Volterra:
applying math to biological and social sciences
Ilaria Dorigatti, Imperial College London
Massimo Fornasier, TU München
Benedetto Piccoli, Rutgers University–Camden

www.100umi800unipd.it

DEADLINES

December 31, 2021: Abstracts for contributed talks – only Italian PhD awarded within 2019–2022
March 14, 2022: Early bird reduced fees

800 1222·2022 ANNI

Università degli Studi di Padova

ORGANIZERS

Unione Matematica Italiana

European Mathematical Society Magazine

The views expressed in the *European Mathematical Society Magazine* are those of the authors and do not necessarily represent those of the European Mathematical Society (EMS) or the Editorial Team.

For advertisements and reprint permission requests please contact magazine@ems.press.

Published by EMS Press, an imprint of the
European Mathematical Society – EMS – Publishing House GmbH
Technische Universität Berlin, Institut für Mathematik,
Straße des 17. Juni 136, 10623 Berlin, Germany

https://ems.press

Typesetting: WisSat Publishing + Consulting GmbH, Fürstenwalde, Germany
Printing: Beltz Bad Langensalza GmbH, Bad Langensalza, Germany

ISSN (print) 2747-7894; ISSN (online) 2747-7908

The cover illustration is a drawing by António B. Araújo of the Riemann surface of the complex square root, rendered as if dimly glimpsed through a mist. It is dedicated to the memory of Mudumbai Seshachalu Narasimhan (1932–2021) whose obituary is published in the present issue.

A message from the president

Dear EMS members,

Even though hybrid and virtual events are still ubiquitous, the increasing return of personal contact in academic life through lectures, conferences, and workshops means that the mathematical community needs to start thinking carefully about the future, starting now. I am fairly certain that we will never completely return to the in-person-only procedures that existed before the pandemic, but at the same time, having attended numerous virtual and hybrid meetings and lectures, I have found that the technological facilities for setting up well-functioning hybrid conferences still often lack professionalism. There is also the fact that particularly for the younger generation, but also for established scientists, personal contact is really essential for the purposes of networking, becoming known in the community, and efficiently learning about new developments. In fact, if we are honest with ourselves, we must admit that given the huge number of publications coming out every day, and despite the fact that almost all of them are freely available on the arXiv or via subscribe-to-open publishing like that of EMS Press, reading research papers is becoming less common, and we rely increasingly on conferences and lectures to find out about new research.

All in all, this means that in fact we are facing another crisis, which the pandemic has only brought more clearly into view.

I already addressed these issues in my last editorial, together with some observations on the 8ECM congress, which was almost entirely virtual with very few people actually physically present. I received quite a lot of supportive feedback on that editorial.

Since these issues concern the future of our profession in a fundamental way, I believe that we need to begin this discussion here and now. Please write to me what you think about the following questions concerning the future:

1. After the pandemic, will we and should we return to the exact same style of meetings, conferences, lectures, and publications we had before the pandemic?
2. Does the community still need big international congresses that cover the broad spectrum of all fields of mathematics?
3. Are there better ways to spread mathematical knowledge than hundreds of new publications every day?
4. Do we need so many scientific prizes, and if we do, should they exist for all age groups or mainly for the young?

Any further suggestions of important questions to add to this list are also welcome. I plan to set up a discussion forum devoted to these topics, and perhaps even to organise an online meeting.

With best wishes for a healthy fall and winter.

Volker Mehrmann
President of the EMS

Brief words from the editor-in-chief

Dear readers of the EMS Magazine,

With this issue, the first year of publication of the EMS Magazine comes to an end. Of course, our Magazine is a continuation of the "old" Newsletter (as implied by the numbering), but it is also in some ways a brand-new publication, with a new look and some new content. Even though some parts are still in the process of fine-tuning, I believe that the challenge of renewing the journal of the European Mathematical Society has been largely successfully met.

In this issue, you can read many interesting articles, of which, and in no way disparaging the other contributions, I would like to highlight the interview with the 2021 Abel prize winners Avi Wigderson and László Lovász by Bjorn Dundas and Christian Skau.

A regular feature of the Magazine (and of the Newsletter before it) that is almost coming to an end is the series of articles about the history and current activities of EMS full member national societies. Indeed, with the article about the Bosnian Mathematical Society in the current issue, almost all of these societies have had the opportunity to present themselves to readers of the Magazine. I therefore take advantage of this particular moment to publicly thank all the national societies for their continuing support of the European Mathematical Society in general, and of the Newsletter/Magazine in particular.

Fernando Pestana da Costa
Editor-in-chief

From Hilbert's 13th problem to essential dimension and back

Zinovy Reichstein

1 Introduction

The problem of solving polynomial equations in one variable, i.e., equations of the form

$$f(x) = 0, \quad \text{where} \quad f(x) = x^n + a_1 x^{n-1} + \cdots + a_n, \tag{1}$$

goes back to ancient times. Here by "solving" I mean finding a procedure or a formula which produces a solution x for a given set of coefficients $a_1, \ldots, a_n$. The terms "procedure" and "formula" are ambiguous; to get a well-posed problem, we need to specify what kinds of operations we are allowed to perform to obtain x from $a_1, \ldots, a_n$. In the simplest setting, we are only allowed to perform the four arithmetic operations: addition, subtraction, multiplication and division. In other words, we are asking if the polynomial (1) has a root x which is expressible as a rational function of $a_1, \ldots, a_n$. For a general polynomial of degree $n \geqslant 2$, the answer is clearly "no"; this was already known to the ancient Greeks. The focus then shifted to the problem of "solving polynomials in radicals", where one is allowed to use the four arithmetic operations and radicals of any degree. Here the mth radical (or root) of t is a solution to

$$x^m - t = 0. \tag{2}$$

Mathematicians attempted to solve polynomial equations this way for centuries, but only succeeded for $n = 1, 2, 3$ and 4. It was shown by Ruffini, Abel and Galois in the early 19th century that a general polynomial of degree $n \geqslant 5$ cannot be solved in radicals. This was a ground-breaking discovery. However, the story does not end there.

Suppose we allow one additional operation, namely solving

$$x^5 + tx + t = 0. \tag{3}$$

That is, we start with $a_1, \ldots, a_n$, and at each step, we are allowed to enlarge this collection by adding one new number, which is the sum, difference, product or quotient of two numbers in our collection, or a solution to (2) or (3) for any t in our collection. In 1786, Bring [16] showed that every polynomial equation of degree 5 can be solved using these operations.

Note that the coefficients of (2) and (3) only depend on one parameter t. Thus roots of these equations can be thought of as "algebraic functions" of one variable. By contrast, the coefficients of the general polynomial equation (1) depend on n independent parameters $a_1, \ldots, a_n$. With this in mind, we define the resolvent degree $\mathrm{rd}(f)$ of a polynomial $f(x)$ in (1) as the smallest positive integer r such that every root of $f(x)$ can be obtained from $a_1, \ldots, a_n$ in a finite number of steps, assuming that at each step we are allowed to perform the four arithmetic operations and evaluate algebraic functions of r variables. Let us denote the largest possible value of $\mathrm{rd}(f)$ by $\mathrm{rd}(n)$, as $f(x)$ ranges over all polynomials of degree n. The algebraic form of Hilbert's 13th problem asks for the value of $\mathrm{rd}(n)$.

The actual wording of the 13th problem is a little different: Hilbert asked for the minimal integer r one needs to solve every polynomial equation of degree n, assuming that at each step one is allowed to perform the four arithmetic operations and apply any continuous (rather than algebraic) function in r variables. Let us denote the maximal possible resolvent degree in this setting by $\mathrm{crd}(n)$, where "c" stands for "continuous". Specifically, Hilbert asked whether or not $\mathrm{crd}(7) = 3$. In this form, Hilbert's 13th problem was solved by Kolmogorov [37] and Arnold [1] in 1957.[1] They showed that, contrary to Hilbert's expectation, $\mathrm{crd}(n) = 1$ for every n. In other words, continuous functions in 1 variable are enough to solve any polynomial equation of any degree. Moreover, any continuous function in n variables can be expressed as a composition of functions of one variable and addition.

In spite of this achievement, Wikipedia lists the 13th problem as "unresolved". While this designation is subjective, it reflects the view of many mathematicians that Hilbert's true intention was to ask about $\mathrm{rd}(n)$, not $\mathrm{crd}(n)$. They point to the body of work on $\mathrm{rd}(n)$ going back centuries before Hilbert (see, e.g., [21]) and to Hilbert's own 20th century writings, where only $\mathrm{rd}(n)$ was considered (see, e.g., [31]). Arnold himself was a strong proponent of this point of view [13, pp. 45–46], [2].

Progress on the algebraic form of Hilbert's 13th problem has been slow. From what I said above, $\mathrm{rd}(n) = 1$ when $n \leqslant 5$; this was

[1] Arnold was a 19 year old undergraduate student in 1957. He later said that all of his numerous subsequent contributions to mathematics were, in one way or another, motivated by Hilbert's 13th problem; see [2].

known before Hilbert and even before Galois. The value of $\mathrm{rd}(n)$ remains open for every $n \geqslant 6$, and the possibility that $\mathrm{rd}(n) = 1$ for every n has not been ruled out. The best known upper bounds on $\mathrm{rd}(n)$ are of the form $\mathrm{rd}(n) \leqslant n - \alpha(n)$, where $\alpha(n)$ is an unbounded but very slow growing function of n. The list of people who have proved inequalities of this form includes some of the leading mathematicians of the past two centuries: Hamilton, Sylvester, Klein, Hilbert, Chebotarev, Segre, Brauer. Recently, their methods have been refined and their bounds sharpened by Wolfson [63], Sutherland [60] and Heberle–Sutherland [30].

There is another reading of the 13th problem, to the effect that Hilbert meant to allow global multi-valued continuous functions; see [2, p. 613]. These behave in many ways like algebraic functions. If we denote the resolvent degree in this sense by $\mathrm{Crd}(n)$, where "C" stands for "global continuous", then

$$1 = \mathrm{crd}(n) \leqslant \mathrm{Crd}(n) \leqslant \mathrm{rd}(n) \leqslant n - \alpha(n).$$

As far as I am aware, nothing else is known about $\mathrm{Crd}(n)$ or $\mathrm{rd}(n)$ for $n \geqslant 6$.

On the other hand, in recent decades, considerable progress has been made in studying a related but different invariant, the essential dimension.[2] Joe Buhler and I [14] introduced this notion in the late 1990s. In special instances, it came up earlier, e.g., in the work of Kronecker [38], Klein [35], Chebotarev [15], Procesi [48][3] and Kawamata [34][4]. Our focus in [14] was on polynomials and field extensions. It later became clear that the notion of essential dimension is of interest in other contexts: quadratic forms, central simple algebras, torsors, moduli stacks, representations of groups and algebras, etc. In each case, it poses new questions about the underlying objects and occasionally leads to solutions of preexisting open problems.

This paper has two goals. The first is to survey some of the research on essential dimension in Sections 2–7. This survey is not comprehensive; it is only intended to convey the flavor of the subject and sample some of its highlights. My second goal for this paper is to define the notion of resolvent degree of an algebraic group in Section 8, building on the work of Farb and Wolfson [25] but focusing on connected, rather than finite groups. The quantity $\mathrm{rd}(n)$ defined above is recovered in this setting as $\mathrm{rd}(\mathrm{S}_n)$. For more comprehensive surveys of essential dimension and resolvent degree, see [41, 51] and [25], respectively.

[2] The term "essential dimension" was coined by Joe Buhler. The term "resolvent degree" was introduced by Richard Brauer in [8].

[3] Procesi asked about the minimal number of independent parameters required to define a generic division algebra of degree n. In modern terminology, this number is the essential dimension of the projective linear group PGL_n.

[4] Kawamata defined an invariant $\mathrm{Var}(f)$ of an algebraic fiber space $f\colon X \to S$, which he informally described as "the number of moduli of fibers of f in the sense of birational geometry". In modern terminology, $\mathrm{Var}(f)$ is the essential dimension of f.

2 Essential dimension of a polynomial

Let k be a base field, K be a field containing k and L be a finite-dimensional K-algebra (not necessarily commutative, associative or unital). We say that L descends to an intermediate field $k \subset K_0 \subset K$ if there exists a finite-dimensional K_0-algebra L_0 such that $L = L_0 \otimes_{K_0} K$. Equivalently, recall that, for any choice of an K-vector space basis $e_1, \dots, e_n$ of L, one can encode multiplication in L into the n^3 structure constants $c_{ij}^h \in K$ given by $e_i e_j = \sum_{h=1}^n c_{ij}^h e_h$. Then L descends to $K_0 \subset K$ if and only if there exists a basis $e_1, \dots, e_n$ such that all of the structure constants e_{ij}^h with respect to this basis lie in K_0. The essential dimension $\mathrm{ed}_k(L/K)$ is defined as the minimal value of the transcendence degree $\mathrm{trdeg}_k(K_0)$, where L descends to K_0. If the reference to the base field k is clear from the context, we will write ed in place of ed_k.

If $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$ is a polynomial over K, for some $a_1, \dots, a_n$, as in (1), we define $\mathrm{ed}_k(f)$ as $\mathrm{ed}_k(L/K)$, where $L = K[x]/(f(x))$. Note that if $f(x)$ (or equivalently, L) is separable over K, then L descends to K_0 if and only if there exists an element $\bar{y} \in L$ which generates L as an K-algebra and such that the minimal polynomial $g(y) = y^n + b_1 y^{n-1} + \dots + b_n$ of $\bar{y}$ lies in $K_0[y]$.

In classical language, the passage from $f(x)$ to $g(y)$ is called a Tschirnhaus transformation. Note that

$$\bar{y} = c_0 + c_1 \bar{x} + \dots + c_{n-1} \bar{x}^{n-1} \tag{4}$$

for some $c_0, c_1, \dots, c_{n-1} \in K$. Here $\bar{x} \in L$ is x modulo $(f(x))$. Tschirnhaus' strategy for solving polynomial equations in radicals by induction on degree was to transform $f(x)$ to a simpler polynomial $g(y)$, find a root of $g(y)$ and then recover a root of $f(x)$ from (4) by solving a polynomial equation of degree $\leqslant n - 1$. In his 1683 paper [62], Tschirnhaus successfully implemented this strategy for $n = 3$ but made a mistake in implementing it for higher n. Tschirnhaus did not know that a general polynomial of degree $\geqslant 5$ cannot be solved in radicals or that his method for solving cubic polynomials had been discovered by Cardano a century earlier.

Let us denote the maximal value of $\mathrm{ed}(f)$ taken over all field extensions K/k and all separable polynomials $f(x) \in K[x]$ of degree n by $\mathrm{ed}_k(n)$. Kronecker [38] and Klein [35] showed that

$$\mathrm{ed}_{\mathbb{C}}(5) = 2. \tag{5}$$

This classical result is strengthened in [14] as follows.

Theorem 1. *Assume* $\mathrm{char}(k) \neq 2$. *Then* $\mathrm{ed}_k(1) = 0$,

$$\mathrm{ed}_k(2) = \mathrm{ed}_k(3) = 1, \quad \mathrm{ed}_k(4) = \mathrm{ed}_k(5) = 2, \quad \mathrm{ed}_k(6) = 3$$

and $\mathrm{ed}_k(n+2) \geqslant \mathrm{ed}_k(n) + 1$ *for every* $n \geqslant 1$. *In particular,*

$$\left\lfloor \frac{n}{2} \right\rfloor \leqslant \mathrm{ed}_k(n) \leqslant n - 3 \tag{6}$$

for every $n \geqslant 5$.

I recently learned that a variant of the inequality $\mathrm{ed}_{\mathbb{C}}(n) \geqslant \lfloor \frac{n}{2} \rfloor$ was known to Chebotarev [15] as far back as 1943.

The problem of finding the exact value of $\mathrm{ed}(n)$ may be viewed as being analogous to Hilbert's 13th problem with $\mathrm{rd}(n)$, $\mathrm{crd}(n)$ or $\mathrm{Crd}(n)$ replaced by $\mathrm{ed}(n)$. Since Hilbert specifically asked about $\mathrm{rd}(7)$, the case where $n = 7$ is of particular interest.

Theorem 2 (Duncan [23]). *If* $\mathrm{char}(k) = 0$, *then* $\mathrm{ed}_k(7) = 4$.

The proof of Theorem 2 relies on the same general strategy as Klein's proof of (5); I will discuss it further it in Section 6. Combining Theorem 2 with the inequality $\mathrm{ed}_k(n+2) \geqslant \mathrm{ed}_k(n) + 1$ from Theorem 1, we can slightly strengthen (6) in characteristic 0 as follows:

$$\left\lfloor \frac{n+1}{2} \right\rfloor \leqslant \mathrm{ed}(n) \leqslant n - 3 \quad \text{for every } n \geqslant 7. \tag{7}$$

Beyond (7), nothing is known about $\mathrm{ed}_{\mathbb{C}}(n)$ for any $n \geqslant 8$. I will explain where I think the difficulty lies in Section 5.

Analogous questions can be asked about polynomials that are not separable, assuming $\mathrm{char}(k) = p > 0$. In this setting, the role of the degree is played by the "generalized degree" $(n, \mathbf{e})$. Here $n = [S : K]$, where S is the separable closure of K in $L = K[x]/(f(x))$ and $\mathbf{e} = (e_1, \dots, e_r)$ is the so-called type of the purely inseparable algebra L/S defined as follows. Given $x \in L$, let us define the exponent $\exp(x, S)$ to be the smallest integer e such that $x^{p^e} \in S$. Then e_1 is the largest value of $\exp(x, S)$ as x ranges over L. Choose an $x_1 \in L$ of exponent e_1, and define e_2 as the largest value of $\exp(x, S[x_1])$. Now choose $x_2 \in L$ of exponent e_2, and define e_3 as the largest value of $\exp(x, S[x_1, x_2])$, etc. We stop when $S[x_1, \dots, x_r] = L$. By a theorem of Pickert, the resulting integer sequence $e_1, \dots, e_r$ satisfies $e_1 \geqslant \dots \geqslant e_r \geqslant 1$ and does not depend on the choice of the elements $x_1, \dots, x_r$. One can now define $\mathrm{ed}_k(n, \mathbf{e})$ by analogy with $\mathrm{ed}_k(n)$: $\mathrm{ed}_k(n, \mathbf{e})$ is the maximal value of $\mathrm{ed}_k(f)$, as K ranges over all field extension of k and $f(x) \in K[x]$ ranges over all polynomials of generalized degree $(n, \mathbf{e})$. Surprisingly, the case where $\mathbf{e} \neq \emptyset$ (i.e., the polynomials $f(x)$ in question are not separable) turns out to be easier. We refer the reader to [53], where an exact formula for $\mathrm{ed}(n, \mathbf{e})$ is obtained.

3 Essential dimension of a functor

Following Merkurjev [6], we will now define essential dimension for a broader class of objects, beyond polynomials or finite-dimensional algebras. Let k be a base field, which we assume to be fixed throughout, and $\mathcal{F}$ be a covariant functor from the category of field extensions K/k to the category of sets. Any object $\alpha \in \mathcal{F}(K)$ in the image of the natural ("base change") map $\mathcal{F}(K_0) \to \mathcal{F}(K)$ is said to *descend* to K_0. The essential dimension $\mathrm{ed}_k(\alpha)$ is defined as the minimal value of $\mathrm{trdeg}_k(K_0)$, where the minimum is taken over all intermediate fields $k \subset K_0 \subset K$ such that α descends to K_0.

For example, consider the functor Ass_n of n-dimensional associative algebras given by

$$\mathrm{Ass}_n(K) = \{n\text{-dimensional associative } K\text{-algebras, up to } K\text{-isomorphism}\}.$$

For $A \in \mathrm{Ass}_n(K)$, the new definition of $\mathrm{ed}_k(A)$ is the same as the definition in the previous section. Recall that, after choosing a K-basis for A, we can describe A completely in terms of the n^3 structure constants c_{ij}^h. In particular, A descends to the subfield $K_0 = k(c_{ij}^h)$ of K, and consequently, $\mathrm{ed}_k(A) \leqslant n^3$.

Another interesting example is the functor of non-degenerate n-dimensional quadratic forms,

$$\mathrm{Quad}_n(K) = \{\text{non-degenerate quadratic forms on } K^n, \text{ up to } K\text{-isomorphism}\}.$$

For simplicity, let us assume that the base field k is of characteristic different from 2. Under this assumption, a quadratic form q on K^n is the same thing as a symmetric bilinear form b. One passes back and forth between q and b using the formulas

$$q(v) = b(v,v) \quad \text{and} \quad b(v,w) = \frac{q(v+w) - q(v) - q(w)}{2}$$

for any $v, w \in K^n$. The form q (or equivalently, b) is called degenerate if the linear form $b(v, *)$ is identically zero for some $0 \neq v \in K^n$. A variant of the Gram–Schmidt process shows that there exists an orthogonal basis of K^n with respect to b. In other words, in some basis $e_1, \dots, e_n$ of K^n, q can be written as

$$q(x_1 e_1 + \dots + x_n e_n) = a_1 x_1^2 + \dots + a_n x_n^2$$

for some $a_1, \dots, a_n$ in K. In particular, we have that q descends to $K_0 = k(a_1, \dots, a_n)$, and thus $\mathrm{ed}_k(q) \leqslant n$. Note that q is non-degenerate if and only if $a_1, \dots, a_n \neq 0$.

Yet another interesting example is provided by the functor of elliptic curves,

$$\mathrm{Ell}(K) = \{\text{elliptic curves over } K, \text{ up to } K\text{-isomorphism}\}.$$

For simplicity, assume that $\mathrm{char}(k) \neq 2$ or 3. Then every elliptic curve X over K is isomorphic to the plane curve cut out by a Weierstrass equation $y^2 = x^3 + ax + b$ for some $a, b \in K$. Hence, X descends to $K_0 = k(a, b)$ and $\mathrm{ed}(X) \leqslant 2$.

Informally, we think of $\mathcal{F}$ as specifying the type of algebraic object under consideration (e.g., algebras or quadratic forms or elliptic curves), $\mathcal{F}(K)$ as the set of objects of this type defined over K, and $\mathrm{ed}_k(\alpha)$ as the minimal number of parameters required to define α. In most cases, essential dimension varies from object to object, and it is natural to consider what happens under a "worst case scenario", i.e., how many parameters are needed to define the most general object of a given type. This number is called the essential dimension of the functor $\mathcal{F}$. That is,

$$\mathrm{ed}_k(\mathcal{F}) = \sup_{K, \alpha} \mathrm{ed}_k(\alpha),$$

as K varies over all fields containing k and α varies over $\mathcal{F}(K)$. Note that $\mathrm{ed}_k(\mathcal{F})$ can be either a non-negative integer or ∞. In particular, the arguments above yield

$$\mathrm{ed}(\mathrm{Ass}_n) \leqslant n^3, \quad \mathrm{ed}(\mathrm{Quad}_n) \leqslant n \quad \text{and} \quad \mathrm{ed}(\mathrm{Ell}) \leqslant 2.$$

One can show that the last two of these inequalities are, in fact, sharp. The exact value of $\mathrm{ed}(\mathrm{Ass}_n)$ is unknown for most n; however, for large n,

$$\mathrm{ed}(\mathrm{Ass}_n) = 4n^3/27 + O(n^{8/3}).$$

Similarly,

$$\mathrm{ed}(\mathrm{Lie}_n) = 2n^3/27 + O(n^{8/3}),$$
$$\mathrm{ed}(\mathrm{Comm}_n) = 2n^3/27 + O(n^{8/3}),$$

where Lie_n and Comm_n are the functors of n-dimensional Lie algebras and commutative algebras, respectively. These formulas are deduced from the formulas for the dimensions of the varieties of structure constants for n-dimensional associative, Lie and commutative algebras due to Neretin [44].[5]

This brings us to the functor $H^1(*, G)$, where G is an algebraic group defined over k. The essential dimension of this functor is a numerical invariant of G. This invariant has been extensively studied; it will be our main focus in the next section. The functor $H^1(*, G)$ associates to a field K/k, the set $H^1(K, G)$ of isomorphism classes of G-torsors T over K. Recall that a G-torsor over T over K is an algebraic variety with a G-action defined over K such that, over the algebraic closure $\overline{K}$, T becomes equivariantly isomorphic to G acting on itself by left translations. If T has a K-point x, then $G \to T$ taking g to $g \cdot x$ is, in fact, an isomorphism over K. In this case, the torsor T is called "trivial" or "split". The interesting (non-trivial) torsors over K have no K-points. For example, if $G = C_2$ is a cyclic group of order 2 and $\mathrm{char}(k) \neq 2$, then every C_2-torsor is of the form T_a, where T_a is the subvariety of $\mathbb{A}^1$ cut out by the quadratic equation $x^2 - a = 0$ for some $a \in K$. Informally, T_a is a pair of points (roots of this equation) permuted by C_2; it is split if and only if these points are defined over K (i.e., a is a complete square in K). In fact, $H^1(K, C_2)$ is in bijective correspondence with $K^*/(K^*)^2$ given by $T_a \mapsto a \bmod (K^*)^2$, where K^* is the multiplicative group of K. Note that, in this example, $H^1(K, G)$ is, in fact, a group. This is the case whenever G is abelian. For a non-abelian algebraic group G, $H^1(K, G)$ carries no natural group structure; it is only a set with a marked element (the trivial torsor).

For many linear algebraic groups G, the functor $H^1(*, G)$ parametrizes interesting algebraic objects. For example, when G is the orthogonal group O_n, $H^1(*, \mathrm{O}_n)$ is the functor Quad_n we considered above. When G is the projective linear group PGL_n, $H^1(K, \mathrm{PGL}_n)$ is the set of isomorphism classes of central simple algebras of degree n over K. When G is the exceptional group of type G_2, $H^1(K, G_2)$ is the set of isomorphism classes of octonion algebras over K.

[5] Note the resemblance of these asymptotic formulas to the classical theorem of Higman and Sims, which assert that the number of finite p-groups of order p^n (up to isomorphism) is asymptotically $p^{2n^3/27 + O(n^{8/3})}$. This is not an accident; see [45].

4 Essential dimension of an algebraic group

The essential dimension of the functor $H^1(*, G)$ is abbreviated as $\mathrm{ed}_k(G)$. Here G is an algebraic group defined over k. This number is always finite if G is linear but may be infinite if G is an abelian variety [12]. If G is the symmetric group S_n, then

$$\mathrm{ed}_k(\mathrm{S}_n) = \mathrm{ed}_k(n), \tag{8}$$

where $\mathrm{ed}_k(n)$ is the quantity we defined and studied in Section 2. Indeed, $H^1(K, \mathrm{S}_n)$ is the set of étale algebras L/K of degree n. Étale algebras of degree n are precisely the algebras of the form $K[x]/(f(x))$, where $f(x)$ is a separable (but not necessarily irreducible) polynomial of degree n over K. Thus (8) is just a restatement of the definition of $\mathrm{ed}_k(n)$.

Another interesting example is the general linear group $G = \mathrm{GL}_n$. Elements of $H^1(K, \mathrm{GL}_n)$ are the n-dimensional vector spaces over K. Since there is only one n-dimensional K-vector space up to K-isomorphism, we see that $H^1(K, \mathrm{GL}_n) = \{1\}$. In particular, every object of $H^1(K, \mathrm{GL}_n)$ descends to k, and we conclude that $\mathrm{ed}_k(\mathrm{GL}_n) = 0$. I will now give a brief summary of three methods for proving lower bounds on $\mathrm{ed}_k(G)$ for various linear algebraic groups G.

4.1 Cohomological invariants

Let $\mathcal{F}$ be a covariant functor from the category of field extensions K/k to the category of sets, as in the previous section. A cohomological invariant of degree d for $\mathcal{F}$ is a morphism of functors

$$\mathcal{F} \to H^d(*, M)$$

for some discrete $\mathrm{Gal}(k)$-module M. In many interesting examples, $M = \mu_m$ is the module of mth roots of unity with a natural $\mathrm{Gal}(k)$-action (trivial if k contains a primitive m-th root of unity). The following observation is due to J.-P. Serre.

Theorem 3. *Assume that the base field k is algebraically closed. If $\mathcal{F}$ has a non-trivial cohomological invariant $\mathcal{F} \to H^d(*, M)$, then $\mathrm{ed}_k(\mathcal{F}) \geqslant d$.*

The proof is an immediate consequence of the Serre vanishing theorem. Cohomological invariants of an algebraic group G (or equivalently, of the functor $H^1(*, G)$) were introduced by Serre and Rost in the early 1990s, and have been extensively studied since then; see [57]. These invariants give rise to a number of interesting lower bounds on $\mathrm{ed}_k(G)$ for various groups G; in particular,

(i) $\mathrm{ed}(\mathrm{O}_n) \geqslant n$,

(ii) $\mathrm{ed}(\mathrm{SO}_n) \geqslant n-1$ for every $n \geqslant 3$,

(iii) $\mathrm{ed}(G_2) \geqslant 3$,

(iv) $\mathrm{ed}(F_4) \geqslant 5$,

(v) $\mathrm{ed}(\mathrm{S}_n) \geqslant \lfloor \frac{n}{2} \rfloor$.

Inequalities (i), (ii) and (iii) turn out to be exact; (iv) is best known, and (v) is best known for even n; see (7).

4.2 Finite abelian subgroups

Theorem 4. *Let G be a reductive group over k and A be a finite abelian subgroup of G of rank r.*

(a) [55] *Assume* $\mathrm{char}(k) = 0$*. If the centralizer $C_G(A)$ is finite, then* $\mathrm{ed}(G) \geqslant r$*.*

(b) [29] *Assume* $\mathrm{char}(k)$ *does not divide* $|A|$*. If G is connected and the dimension of the maximal torus of $C_G(A)$ is d, then* $\mathrm{ed}(G) \geqslant r - d$*.*

Note that both parts are vacuous if A lies in a maximal torus T of G. Indeed, in this case, the centralizer $C_G(A)$ contains T, so $d \geqslant r$. In other words, only non-toral finite abelian subgroups A of linear algebraic groups are of interest here. These have been much studied and catalogued, starting with the work of Borel in the 1950s. Theorem 4 yields the best known lower bound on $\mathrm{ed}(G)$ in many cases, such as $\mathrm{ed}(E_7) \geqslant 7$ and $\mathrm{ed}(E_8) \geqslant 9$, where E_7 denotes the split simply connected exceptional group of type E_7 and similarly for E_8.

4.3 The Brauer class bound

Consider a linear algebraic group G defined over our base field k. Suppose G fits into a central exact sequence of algebraic groups (again, defined over k)

$$1 \to D \to G \to \overline{G} \to 1,$$

where D is diagonalizable over k. For every field extension K/k, this sequence gives rise to the exact sequence of pointed sets

$$H^1(K,G) \to H^1(K,\overline{G}) \xrightarrow{\partial} H^2(K,D).$$

Every element $\alpha \in H^2(K,D)$ has an index, $\mathrm{ind}(\alpha)$, defined as follows. If $D \simeq \mathbb{G}_m$, then α is a Brauer class over K, and $\mathrm{ind}(\alpha)$ denotes the Schur index of α, as usual. In general, we consider the character group $X(D)$ whose elements are homomorphisms $x\colon D \to \mathbb{G}_m$. Note that $X(D)$ is a finitely generated abelian group and each character $x \in X(D)$ induces a homomorphism

$$x_*\colon H^2(K,D) \to H^2(K,\mathbb{G}_m).$$

The index of $\alpha \in H^2(K,D)$ is defined as the minimal value of

$$\mathrm{ind}(x_1)_*(\alpha) + \cdots + \mathrm{ind}(x_r)_*(\alpha)$$

as $\{x_1, \ldots, x_r\}$ ranges over generating sets of $X(D)$. Here each $(x_i)_*(\alpha)$ lies in $H^1(K,\mathbb{G}_m)$, and $\mathrm{ind}(x_i)_*(\alpha)$ denotes its Schur index, as above. We now define $\mathrm{ind}(G,D)$ as the maximal index of $\alpha \in \partial(H^1(K,\overline{G})) \subset H^2(K,D)$, where the maximum is taken over all field extensions K/k, as α ranges over the image $H^1(K,\overline{G})$ in $H^2(K,D)$.

Theorem 5.

(a) $\mathrm{ind}(G,D)$ *is the greatest common divisor of* $\dim(\rho)$*, where ρ ranges over the linear representations of G over k such that the restriction $\rho_{|D}$ is faithful.*

(b) *Let p be a prime different from* $\mathrm{char}(k)$*. Assume that the exponent of every element of $H^2(K,D)$ in the image of*

$$\partial\colon H^1(K,\overline{G}) \to H^2(K,D)$$

is a power of p for every field extension K/k. (This is automatic if D is a p-group.) Then $\mathrm{ed}_k(G) \geqslant \mathrm{ind}(G,D) - \dim(G)$*.*

Part (a) is known as Merkurjev's index formula. The inequality of part (b) is based on Karpenko's incompressibility theorem. Part (b) first appeared in [9] in the special case where $D = \mathbb{G}_m$ or μ_{p^r} and in [26] in an even more special case, where $D = \mu_p$. It was proved in full generality in [33].

Theorem 5 is responsible for some of the strongest results in this theory, including the exact formulas for the essential dimension of a finite p-group (Theorem 6 below), the essential p-dimension of an algebraic torus, and the essential dimension of spinor groups Spin_n. The latter turned out to increase exponentially in n:

$$\mathrm{ed}(\mathrm{Spin}_n) \geqslant 2^{\lfloor (n-1)/2 \rfloor} - \frac{n(n-1)}{2}. \tag{9}$$

This inequality was first proved in [9]. The exact value of $\mathrm{ed}(\mathrm{Spin}_n)$ subsequently got pinned down in [10, 18] in characteristic 0, [28] in characteristic $p \neq 2$ and [61] in characteristic 2. When $n \geqslant 15$, inequality (9) is sharp for $n \not\equiv 0$ modulo 4, and is off by $2^{\nu_2(n)}$ otherwise. Here $2^{\nu_2(n)}$ is the largest power of 2 dividing n.

The exponential growth of $\mathrm{ed}(\mathrm{Spin}_n)$ came as a surprise. Prior to [9], the best known lower bounds on $\mathrm{ed}(\mathrm{Spin}_n)$ were linear (see [19, Section 7]), on the order of $\frac{n}{2}$. Moreover, the exact values of $\mathrm{ed}(\mathrm{Spin}_n)$ for $n \leqslant 14$ obtained by Rost and Garibaldi [27] appeared to suggest that these linear bounds should be sharp. The fact that $\mathrm{ed}(\mathrm{Spin}_n)$ increases exponentially in n has found interesting applications in the theory of quadratic forms. For details, see [10, 18].

5 Essential dimension at p

Once again, fix a base field k, and let $\mathcal{F}$ be a covariant functor from the category of field extensions K/k to the category of sets. The essential dimension $\mathrm{ed}_k(\alpha; p)$ of an object $\alpha \in \mathcal{F}(K)$ at a prime p

is defined as the minimal value of $\mathrm{ed}_k(\alpha'; p)$, where the minimum ranges over all finite field extensions K'/K of degree prime to p and α' denotes the image of α under the natural map $\mathcal{F}(K) \to \mathcal{F}(K')$. Finally, the essential dimension $\mathrm{ed}_k(\mathcal{F}; p)$ of $\mathcal{F}$ at p is the maximal value of $\mathrm{ed}_k(\alpha)$, as K ranges over all fields containing k and α ranges over $\mathcal{F}(K)$. When $\mathcal{F} = H^1(*, G)$ for an algebraic group G, we write $\mathrm{ed}_k(G; p)$ in place of $\mathrm{ed}_k(\mathcal{F}; p)$. Once again, if the reference to the base field is clear from the context, we will abbreviate ed_k as ed. By definition, $\mathrm{ed}(\alpha; p) \leqslant \mathrm{ed}(\alpha)$ and $\mathrm{ed}(\mathcal{F}; p) \leqslant \mathrm{ed}(\mathcal{F})$.

The reason to consider $\mathrm{ed}(\mathcal{F}; p)$ in place of $\mathrm{ed}(\mathcal{F})$ is that the former is often more accessible. In fact, most of the methods we have for proving a lower bound on $\mathrm{ed}_k(\alpha)$ (respectively, $\mathrm{ed}_k(\mathcal{F})$) turn out to produce a lower bound on $\mathrm{ed}_k(\alpha; p)$ (respectively, $\mathrm{ed}_k(\mathcal{F}; p)$) for some prime p. For example, the lower bound in Theorem 5 (b) is really $\mathrm{ed}_k(G; p) \geqslant \mathrm{ind}(G, D) - \dim(G)$. In Theorem 4, one can usually choose A to be a p-group, in which case the conclusion can be strengthened to $\mathrm{ed}(G; p) \geqslant r$ in part (a) and $\mathrm{ed}(G; p) \geqslant r - d$ in part (b). In Theorem 3, if M is p-torsion (which can often be arranged), then $\mathrm{ed}(G; p) \geqslant d$.

This is a special case of a general meta-mathematical phenomenon: many problems concerning algebraic objects (such as finite-dimensional algebras or polynomials or algebraic varieties) over fields K can be subdivided into two types. In type 1 problems, we are allowed to pass from K to a finite extension K'/K of degree prime to p, for one prime p, whereas in type 2 problems this is not allowed. For example, given an algebraic variety X defined over K, deciding whether or not X has a 0-cycle of degree 1 is a type 1 problem (it is equivalent to showing that there is a 0-cycle of degree prime to p, for every prime p), whereas deciding whether or not X has a K-point is a type 2 problem. As I observed in [51, Section 5], most of the technical tools we have are tailor-made for type 1 problems, whereas many long-standing open questions across several areas of algebra and algebraic geometry are of type 2.

In the context of essential dimension, the problem of computing $\mathrm{ed}(G; p)$ for a given algebraic group G and a given prime p is of type 1, whereas the problem of computing $\mathrm{ed}(G)$ is of type 2. For simplicity, let us assume that G is a finite group. In this case, $\mathrm{ed}_k(G; p) = \mathrm{ed}_k(G_p; p)$, where G_p is the Sylow p-subgroup of G. In other words, the problem of computing $\mathrm{ed}_k(G; p)$ reduces to the case where G is a p-group. In this case, we have the following remarkable theorem of Karpenko and Merkurjev [32].

Theorem 6. *Let p be a prime and k be a field containing a primitive pth root of unity. Then, for any finite p-group P,*

$$\mathrm{ed}_k(P) = \mathrm{ed}_k(P; p) = \mathrm{rdim}_k(P),$$

where $\mathrm{rdim}_k(P)$ *denotes the minimal dimension of a faithful representation of P defined over k.*

Theorem 6 reduces the computation of $\mathrm{ed}_k(G; p)$ to $\mathrm{rdim}_k(G_p)$. For a given finite p-group P, one can often (though not always) compute $\mathrm{rdim}_k(P)$ in closed form using the machinery of character theory; see, e.g., [3, 36, 42, 43].

The situation is quite different when computing $\mathrm{ed}_k(G)$ for an arbitrary finite group G. Clearly, $\mathrm{ed}_k(G) \geqslant \max_p \mathrm{ed}_k(G; p)$, where p ranges over the prime integers. In those cases, where $\mathrm{ed}_k(G)$ is strictly larger than $\max_p \mathrm{ed}_k(G; p)$, the exact value of $\mathrm{ed}_k(G)$ is usually difficult to establish. The only approach that has been successful to date relies on classification results in algebraic geometry, which are currently only available in low dimensions. I will return to this topic in the next section.

To illustrate the distinction between type 1 and type 2 problems, consider the symmetric group $G = \mathrm{S}_n$. For simplicity, assume that $k = \mathbb{C}$ is the field of complex numbers. Here the type 1 problem is solved completely: $\mathrm{ed}_{\mathbb{C}}(\mathrm{S}_n; p) = \lfloor \frac{n}{p} \rfloor$ for every prime p. Thus $\max_p \mathrm{ed}_{\mathbb{C}}(\mathrm{S}_n; p) = \lfloor \frac{n}{2} \rfloor$, and (7) tells us that

$$\mathrm{ed}_{\mathbb{C}}(\mathrm{S}_n) > \max_p \mathrm{ed}_{\mathbb{C}}(G; p) \quad \text{for every odd } n \geqslant 7.$$

The remaining type 2 problem is to bridge the gap between $\lfloor \frac{n}{2} \rfloor$ and the true value of $\mathrm{ed}_{\mathbb{C}}(\mathrm{S}_n)$. This problem has only been solved for $n \leqslant 7$; see Theorems 1, 2 and (8).

Note that the algebraic form of Hilbert's 13th problem is also of type 2 in the sense that

$$\mathrm{rd}(f; p) \leqslant 1 \tag{10}$$

for any prime p, every field K and every separable polynomial $f(x) \in K[x]$.[6] Indeed, denote the Galois group of $f(x)$ by G. Then, after passing from K to a finite extension K'/K whose degree $[K' : K] = [G : G_p]$ is prime to p, we may replace G by its p-Sylow subgroup G_p. Since every p-group is solvable, this means that $f(x)$ becomes solvable in radicals over K', and hence its resolvent degree becomes $\leqslant 1$, as desired.

Inequality (10) accounts, at least in part, for the difficulty of showing that $\mathrm{rd}(n) \geqslant 2$ for any n. The methods used to prove lower bounds on the essential dimension of algebraic groups in Section 4, and anything resembling these methods, cannot possibly work here; otherwise, we would also be able to prove that $\mathrm{rd}(f; p) \geqslant 2$ for some prime p, contradicting (10).

A similar situation arises in computing the essential dimension of a finite p-group G over a field k of characteristic p. Superficially this problem looks very different from Hilbert's 13th problem (where one usually works over $k = \mathbb{C}$); the common feature is that both are type 2 problems. Indeed, it is shown in [54] that $\mathrm{ed}_k(G; p) = 1$ for every non-trivial p-group G. Using the method described in the next section, one can often show that $\mathrm{ed}_k(G) \geqslant 2$, but we are not able to prove that $\mathrm{ed}_k(G) > 2$ for any p-group G and any field k of characteristic p. On the other hand, Ledet [39]

[6] For the precise definitions of $\mathrm{rd}(f)$ and $\mathrm{rd}(f; p)$, see Section 8.

conjectured that

$$\mathrm{ed}_k(C_{p^n}) = n \tag{11}$$

for any prime p and any infinite field k of characteristic p. Here C_{p^n} denotes the cyclic group of order p^n. Ledet showed that $\mathrm{ed}_k(C_{p^n}) \leqslant n$ for every $n \geqslant 1$ and that equality holds when $n \leqslant 2$.

My general feeling is that type 2 problems arising in different contexts are linked in some way, and that solving one of them (e.g., proving Ledet's conjecture) can shed light on the others (e.g., Hilbert's 13th problem). The only bit of evidence I have in this direction is the following theorem from [11] linking a priori unrelated type 2 problems in characteristic p and in characteristic 0.

Theorem 7. *Let p be a prime and G be a finite group satisfying the following conditions:*

(i) *G does not have a non-trivial normal p-subgroup, and*

(ii) *G has an element of order p^n.*

If Ledet's conjecture (11) *holds, then* $\mathrm{ed}_{\mathbb{C}}(G) \geqslant n$.

The following family of examples is particularly striking. Let p be a prime and n a positive integer. Choose a positive integer m such that $q = mp^n + 1$ is a prime. Note that, by Dirichlet's theorem on primes in arithmetic progressions, there are infinitely many such m. Let C_q be a cyclic group of order q. Then $\mathrm{Aut}(C_q)$ is cyclic of order mp^n; let $C_{p^n} \subseteq \mathrm{Aut}(C_q)$ denote the unique subgroup of order p^n. Applying Theorem 7 to $G = C_q \rtimes C_{p^n}$, we obtain the following.

Corollary 8. *If Ledet's conjecture* (11) *holds, then*

$$\mathrm{ed}_{\mathbb{C}}(C_q \rtimes C_{p^n}) \geqslant n.$$

Note that, since the Sylow subgroups of $C_q \rtimes C_{p^n}$ are all cyclic,

$$\mathrm{ed}_{\mathbb{C}}(C_q \rtimes C_{p^n}; l) \leqslant 1$$

for every prime l, so the inequality of Corollary 8 is a type 2 result. An unconditional proof of this inequality or even of the weaker inequality $\mathrm{ed}_{\mathbb{C}}(C_q \rtimes C_{p^n}) > 3$ is currently out of reach for any specific choice of q and p^n.

6 Essential dimension and the Jordan property

An alternative (equivalent) definition of essential dimension of a finite group G is as follows. An action of G on an algebraic variety X is said to be linearizable if there exists a G-equivariant dominant rational map $V \dashrightarrow X$ for some linear representation $G \to \mathrm{GL}(V)$. Then $\mathrm{ed}_k(G)$ is the minimal value of $\dim(X)$, as X ranges over all linearizable varieties with a faithful G-action defined over k. In particular, $\mathrm{ed}_k(G) \leqslant \mathrm{rdim}_k(G)$, where $\mathrm{rdim}_k(G)$ is the minimal dimension of a faithful linear representation of G over k, as in Theorem 6.

This geometric interpretation of $\mathrm{ed}_k(G)$ can sometimes be used to prove lower bounds on $\mathrm{ed}(G)$ by narrowing the possibilities for X and ruling them out one by one using Theorem 4 (a). For the remainder of this section, I will assume that G is a finite group and the base field k is the field of complex numbers and will write ed in place of $\mathrm{ed}_{\mathbb{C}}$.

Suppose $\mathrm{ed}(G) = 0$. Then X is a single point, and only the trivial group can act faithfully on a point. Thus $\mathrm{ed}(G) = 0$ if and only if G is the trivial group.

Now suppose $\mathrm{ed}(G) = 1$. Then X is a curve with a dominant map $V \dashrightarrow X$. By Lüroth's theorem, X is birationally isomorphic to $\mathbb{P}^1$ and thus G is a subgroup of PGL_2. Finite subgroups of PGL_2 were classified by Klein [35]. Here is a complete list: cyclic groups C_n and dihedral groups D_n for every n, A_4, S_4 and A_5. Theorem 4 (a) rules out the groups on this list which contain $A = C_2 \times C_2$. We thus obtain the following.

Theorem 9 ([14, Theorem 6.2])**.** *Let G be a finite group. Then* $\mathrm{ed}(G) = 1$ *if and only if G is either cyclic or odd dihedral.*

To classify groups of essential dimension d (or more realistically, show that $\mathrm{ed}(G) > d$ for a particular finite group G) in a similar manner, we need a classification of finite subgroups of $\mathrm{Bir}(X)$, extending Klein's classification of finite subgroups in $\mathrm{Bir}(\mathbb{P}^1)$. Here X ranges over the unirational complex varieties of dimension d, and $\mathrm{Bir}(X)$ denotes the groups of birational automorphisms of X. In dimension 2, every unirational variety is rational, so we are talking about classifying finite subgroups of the Cremona group $\mathrm{Bir}(\mathbb{P}^2)$. Such a classification exists, though it is rather complicated; see [22]. Serre used this approach to show that $\mathrm{ed}(\mathrm{A}_6) = 3$ (see [59, Theorem 3.6]). Again, this is a type 2 phenomenon since $\max_p \mathrm{ed}(\mathrm{A}_6; p) = 2$. Duncan [24] subsequently extended Serre's argument to a full classification of finite groups of essential dimension 2.

In dimension 3, there is the additional complication that unirational complex varieties do not need to be rational. Here only a partial analogue of Klein's classification exists, namely the classification of rationally connected 3-folds with the action of a finite simple group G, due to Prokhorov [49]. Duncan used this classification to prove Theorem 2. More specifically, he showed that $\mathrm{ed}(\mathrm{S}_7) = \mathrm{ed}(\mathrm{A}_7) = 4$; see (8). Subsequently, Beauville [4] showed that the only finite simple groups of essential dimension 3 are A_6 and possibly $\mathrm{PSL}_2(\mathbb{F}_{11})$.[7]

In dimension $d \geqslant 4$, even a partial analogue of Klein's classification of finite subgroups of $\mathrm{Bir}(\mathbb{P}^1)$ is out of reach. However, a recent break-through in Mori theory gives us a new insight into the asymptotic behavior of $\mathrm{ed}(G_n)$ for certain infinite sequences $G_1, G_2, \ldots$ of finite groups. Recall that an abstract group Γ is called *Jordan* if there exists an integer j (called a Jordan constant of Γ) such that every finite subgroup $G \subset \Gamma$ has a normal abelian sub-

[7] It is not known whether the essential dimension of $\mathrm{PSL}_2(\mathbb{F}_{11})$ is 3 or 4.

group A of index $[G : A] \leqslant j$. This definition, due to Popov [46], was motivated by the classical theorem of Camille Jordan which asserts that $\mathrm{GL}_n(\mathbb{C})$ is Jordan, and by a theorem of Serre [58] which asserts that the Cremona group $\mathrm{Bir}(\mathbb{P}^2)$ is also Jordan. The following result, due to Prokhorov, Shramov and Birkar[8], is a far-reaching generalization of Serre's theorem.

I will say that a collection of abstract groups is uniformly Jordan if they are all Jordan with the same constant.

Theorem 10. *Fix $d \geqslant 1$. Then the groups $\mathrm{Bir}(X)$ are uniformly Jordan, as X ranges over d-dimensional rationally connected complex varieties.*

Unirational varieties are rationally connected. The converse is not known, though it is generally believed to be false. Rationally connected varieties naturally arise in the context of Mori theory, and we are forced to consider them even if we are only really interested in unirational varieties. Note that Theorem 10 does not become any easier to prove if one requires X to be unirational. In fact, prior to Birkar [7] (respectively, prior to Prokhorov–Shramov [49]), it was an open question, due to Serre [58, Section 6], whether for each $d \geqslant 4$ (respectively, for $d = 3$) there exists even a single finite group which does not embed into $\mathrm{Bir}(\mathbb{P}^d)$.[9]

Now observe that, while every finite group is obviously Jordan, being uniformly Jordan is a strong condition on a sequence of finite groups

$$G_1, G_2, G_3, \ldots. \tag{12}$$

Suppose sequence (12) is chosen so that no infinite subsequence is uniformly Jordan. Then we claim that

$$\lim_{n \to \infty} \mathrm{ed}(G_n) = \infty. \tag{13}$$

Indeed, if $\mathrm{ed}(G_n) = d$, then there exists a d-dimensional linearizable variety X with a faithful G_n-action. In particular, G_n is contained in $\mathrm{Bir}(X)$. Since X is linearizable, it is unirational and hence rationally connected. On the other hand, since no infinite subsequence of (12) is uniformly Jordan, Theorem 10 tells us that there are at most finitely many groups G_n with $\mathrm{ed}(G_n) = d$, and (13) follows. Here is an interesting family of examples.

Theorem 11. *For each positive integer n, let C_n be a cyclic group of order n and H_n be a subgroup of $\mathrm{Aut}(C_n)$. If $\lim_{n\to\infty} |H_n| = \infty$, then $\lim_{n\to\infty} \mathrm{ed}(C_n \rtimes H_n) = \infty$.*

[8] Prokhorov and Shramov [50] proved this theorem assuming the Borisov–Alexeev–Borisov (BAB) conjecture. The BAB conjecture was subsequently proved by Birkar [7].

[9] For the current status of Serre's questions from [58, Section 6], see [47, Section 3].

Note that this method does not give us any information about $\mathrm{ed}(C_n \rtimes H_n)$ for any particular choice of n and of $H_n \subset \mathrm{Aut}(C_n)$. For example, while Theorem 11 tells us that

$$\mathrm{ed}_{\mathbb{C}}(C_p \rtimes \mathrm{Aut}(C_p)) > 10$$

for all but finitely many primes p, it does not allow us to exhibit a specific prime for which this inequality holds. The reason is that, when $d > 3$, a specific Jordan constant for the family of groups $\mathrm{Bir}(X)$ in Theorem 10 is out of reach. In particular, an unconditional proof of Corollary 8 along these lines does not appear feasible. Nevertheless, Theorem 11 represents a big step forward: previously, it was not even known that $\mathrm{ed}_{\mathbb{C}}((C_p) \rtimes \mathrm{Aut}(C_p)) > 3$ for any prime p.

A classification of the subgroups of $\mathrm{Bir}(X)$, as X ranges over the unirational varieties of dimension d is a rather blunt instrument. It would be preferable to find some topological or algebro-geometric obstruction to the existence of a linearization map $V \dashrightarrow X$, which can be read off from the G-variety X without enumerating all the possibilities for X. Unfortunately, all known obstructions of this sort are of type 1: they do not distinguish between dominant rational maps $V \dashrightarrow X$ and correspondences $V \rightsquigarrow X$ of degree prime to p, for a suitable prime p and thus cannot help us if $\mathrm{ed}(G) > \max_p \mathrm{ed}(G; p)$.

Another draw-back of this method is that, as we mentioned in the previous section, beyond dimension 1,[10] none of the classification theorems we need are available in prime characteristic.

7 Essential dimension of a representation

7.1 *Representations of finite groups in characteristic* 0

Let G be a finite group of exponent e, k be a field of characteristic 0, K/k be a field extension, $\rho : G \to \mathrm{GL}_n(K)$ be a representation of G, and $\chi : G \to K$ be the character of ρ. Can we realize ρ over k? In other words, is there a representation $\rho' : G \to \mathrm{GL}(k)$ such that ρ and ρ' are equivalent over K? A celebrated theorem of Richard Brauer asserts that the answer is "yes" as long as k contains a primitive root of unity of degree e. If it does not, there is a classical way to quantify how far ρ is from being definable over k via the Schur index, at least in the case where ρ is absolutely irreducible and the character value $\chi(g)$ lies in k for every $g \in G$. The Schur index of ρ is defined as the index of the envelope

$$\mathrm{Env}_k(\rho) := \mathrm{Span}_k\{\rho(g) \mid g \in G\} \subset \mathrm{Mat}_n(K)$$

which, under our assumptions on ρ, is a central simple algebra of degree n over k. The Schur index of ρ is equal to the minimal degree $[l : k]$ of a field extension l/k such that ρ can be realized over l.

[10] Groups of essential dimension 1 have been classified over an arbitrary field k; see [20, 40]. Recall that Theorem 9 assumes that $k = \mathbb{C}$.

The essential dimension $\mathrm{ed}_k(\rho)$ gives us a different way to quantify how far ρ is from being definable over k. Here we do not need to assume that ρ is irreducible or that its character values lie in k. We simply think of ρ as an object of the functor

$$\mathrm{Rep}_G \colon K \mapsto \{K\text{-representations of } G, \text{ up to } K\text{-isomorphism}\}.$$

The naive upper bound on $\mathrm{ed}(\rho)$ is rn^2, where n is the dimension of ρ and r is the minimal number of generators of G. Indeed, if G is generated by r elements $g_1, \dots, g_r$ and $\rho(g_h)$ is the $n \times n$ matrix (a_{ij}^h), then ρ descends to the field

$$K_0 = k(a_{ij}^h \mid i, j = 1, \dots, n;\ h = 1, \dots, r)$$

of transcendence degree at most rn^2 over k. It is shown in [32] that, in fact, $\mathrm{ed}(\rho) \leqslant n^2/4$ and, moreover, $\mathrm{ed}(\mathrm{Rep}_G) \leqslant |G|/4$. We have also proved lower bounds on $\mathrm{ed}_k(\rho)$ in many cases (for details, see [32]). Note that these are quite delicate: by Brauer's theorem, $\mathrm{ed}_k(G) = 0$ as long as k contains suitable roots of unity.

7.2 Representations of finite groups in positive characteristic

Here the situation is entirely different.

Theorem 12 ([5, 32]). *Let G be a finite group, k be a field of characteristic $p > 0$ and G_p be the Sylow p-subgroup of G. Then*

$$\mathrm{ed}_k(\mathrm{Rep}_G) = \begin{cases} 0 & \text{if } G_p \text{ is cyclic,} \\ \infty & \text{otherwise.} \end{cases}$$

Note that, by a theorem of Higman, in characteristic p, G_p is cyclic if and only if the group algebra kG is of finite representation type, i.e., if and only if kG (or equivalently, G) has only finitely many indecomposable representations. Since kG is always of finite representation type in characteristic 0, we obtain the following.

Corollary 13. *Let G be a finite group and k be a field of arbitrary characteristic. Then*

- $\mathrm{ed}_k(\mathrm{Rep}_G) < \infty$ *if kG is of finite representation type, and*
- $\mathrm{ed}_k(\mathrm{Rep}_G) = \infty$ *otherwise.*

7.3 Representations of algebras

For simplicity, let us assume that the base field k is algebraically closed. A celebrated theorem of Drozd asserts that every finite-dimensional k-algebra A falls into one of three categories: (a) finite representation type, (b) tame and (c) wild.

Informally speaking, A is of tame representation type if, for every positive integer n, the n-dimensional indecomposable A-modules occur in (at most) a finite number of one-parameter families. On the other hand, A is of wild representation type if the representation theory of A contains that of the free k-algebra on two generators.

We can define the functor of representations Rep_A in the same way as before: to a field K/k, it associates isomorphism classes of finite-dimensional $A \otimes_k K$-modules. Corollary 13 tells us that, when $A = kG$ is a group ring, the essential dimension of the functor Rep_A distinguishes between algebras A of finite representation type and algebras of the other two types. It does not distinguish between tame and wild representations types since $\mathrm{ed}(\mathrm{Rep}_A) = \infty$ in both cases. Benson suggested that it may be possible to distinguish between these two types of algebras by considering the rate of growth of $r_A(n) = \mathrm{ed}(\mathrm{Rep}_A[n])$, where $\mathrm{Rep}_A[n](K)$ is the set of isomorphism classes of K-representations of A of dimension $\leqslant n$. This is confirmed by the following theorem of Scavia [56].

Theorem 14.

(a) *If A is of finite representation type, then $r_A(n)$ is bounded from above as $n \to \infty$.*

(b) *If A is tame, then there exists a constant $c > 0$ such that $cn - 1 \leqslant r_A(n) \leqslant 2n - 1$ for every $n \geqslant 1$.*

(c) *If A is wild, then there exist constants $0 < c_1 < c_2$ such that $c_1 n^2 - 1 \leqslant r_A(n) \leqslant c_2 n^2$ for every $n \geqslant 1$.*

This gives us three new invariants of finite-dimensional algebras, $a_i(A) = \limsup_{n \to \infty} r_A(n)/n^i$ for $i = 0, 1, 2$. Informally, $a_2(A)$ (respectively, $a_1(A)$) quantifies "how wild" (respectively, "how tame") A is. Scavia [56] computed $a_1(A)$ and $a_2(A)$ explicitly in combinatorial terms in the case, where A is a quiver algebra.

8 Back to resolvent degree

8.1 The level of a field extension

Let k be a base field, K be a field containing k, and L/K be a field extension of finite degree. I will say that L/K is of level $\leqslant d$ if there exists a finite tower of subfields

$$K = K_0 \subset K_1 \subset \cdots \subset K_n \tag{14}$$

such that $L \subset K_n$ and $\mathrm{ed}_k(K_{i+1}/K_i) \leqslant d$ for every i. The level of L/K is the smallest such d; I will denote it by $\mathrm{lev}_k(L/K)$. Clearly,

$$\mathrm{lev}_k(L/K) \leqslant \mathrm{ed}_k(L/K).$$

If K is a field of rational functions on some algebraic variety X defined over k, then it is natural to think of elements of K_1 as algebraic (multi-valued) functions on X in at most $\mathrm{ed}_k(K_1/K)$ variables, and elements of L as compositions of algebraic functions in at most $\mathrm{lev}_k(L/K)$ variables.

Example 15. If the field extension L/K is solvable, then we claim that $\mathrm{lev}_k(L/K) \leqslant 1$. Indeed, here we can choose the tower (14) so that each K_{i+1} is obtained from K_i by adjoining a single radical. Then $\mathrm{ed}_k(K_{i+1}/K_i) \leqslant 1$ for each i, and hence, $\mathrm{lev}_k(L/K) \leqslant 1$, as claimed.

8.2 *The resolvent degree of a functor*

Let $\mathcal{F}$ be a functor from the category of field extensions K/k to the category of sets with a marked element. We will denote the marked element in $\mathcal{F}(K)$ by 1 and will refer to it as being "split". We will say that a field extension L/K splits an object $\alpha \in \mathcal{F}(K)$ if $\alpha_L = 1$. Here, as usual, α_L denotes the image of α under the natural map $\mathcal{F}(K) \to \mathcal{F}(L)$. Let us assume that

$$\text{for every field } K/k \text{ and every } \alpha \in \mathcal{F}(K), \quad \alpha \text{ can be split by a field extension } L/K \text{ of finite degree.} \tag{15}$$

This is a strong condition of $\mathcal{F}$; in particular, it implies that $\mathcal{F}(K) = \{1\}$ whenever K is algebraically closed.

I will now define the resolvent degrees $\mathrm{rd}_k(\alpha)$ of $\alpha \in \mathcal{F}(K)$ and $\mathrm{rd}_k(\mathcal{F})$ of the functor $\mathcal{F}$ satisfying condition (15) by analogy with the definitions of $\mathrm{ed}_k(\alpha)$ and $\mathrm{ed}_k(\mathcal{F})$ in Section 3. The resolvent degree $\mathrm{rd}_k(\alpha)$ is the minimal integer $d \geqslant 0$ such that α is split by a field extension L/K of level d (or equivalently, of level $\leqslant d$). The resolvent degree $\mathrm{rd}_k(\mathcal{F})$ is the maximal value of $\mathrm{rd}_k(\alpha)$, as K ranges over all fields containing k and α ranges over $\mathcal{F}(K)$.

Example 16. Let $n \geqslant 2$ be an integer not divisible by $\mathrm{char}(k)$. Then the functor $H^2(*, \mu_n)$ satisfies condition (15). I claim that this functor has resolvent degree 1. Indeed, let $\alpha \in H^2(K, \mu_n)$, and let ζ be a primitive nth root of unity in $\bar{k}$. By the Merkurjev–Suslin theorem, over $K(\zeta)$, we can write

$$\alpha = (a_1) \cup (b_1) + (a_2) \cup (b_2) + \cdots + (a_r) \cup (b_r)$$

for some $0 \neq a_i, b_i \in K(\zeta)$. Now $L = K(\zeta, a_1^{1/n}, \ldots, a_r^{1/n})$ splits α. By our construction, L is solvable over K. Thus, as we saw in Example 15, $\mathrm{lev}_k(L/K) \leqslant 1$. This shows that $\mathrm{rd}_k(\alpha) \leqslant 1$, as claimed. Using the norm residue isomorphism theorem (formerly known as the Bloch–Kato conjecture) in place of Merkurjev–Suslin, one shows in the same manner that $H^d(*, \mu_n)$ has resolvent degree 1 for every $d \geqslant 1$.

The resolvent degrees $\mathrm{rd}_k(\alpha; p)$ and $\mathrm{rd}_k(\mathcal{F}; p)$ at a prime p are defined in the same way as $\mathrm{ed}_k(\alpha; p)$ and $\mathrm{ed}_k(\mathcal{F}; p)$. Here $\mathcal{F}$ is a functor satisfying (15), $\alpha \in \mathcal{F}(K)$ is an object of $\mathcal{F}$. That is, $\mathrm{rd}_k(\alpha; p)$ is the minimal value of $\mathrm{rd}_k(\alpha_{K'})$, as K' ranges over all field extension of K such that $[K' : K]$ is finite and prime to p, and $\mathrm{rd}_k(\mathcal{F}; p)$ is the maximal value of $\mathrm{rd}_k(\alpha; p)$, where K ranges over all field containing k, and α ranges over $\mathcal{F}(K)$. A variant of the argument we used to prove (10) shows that $\mathrm{rd}_k(\mathcal{F}; p) \leqslant 1$ for every base field k, every functor $\mathcal{F}$ satisfying (15) and every prime p.

8.3 *The resolvent degree of an algebraic group*

The functor $\mathcal{F} = H^1(*, G)$ whose objects over K are G-torsors over $\mathrm{Spec}(K)$ satisfies condition (15) for every algebraic group G defined over k. I will write $\mathrm{rd}_k(G)$ for the resolvent degree of this functor. For simplicity, let us assume that $k = \mathbb{C}$ for the remainder of this section. I will write rd in place of $\mathrm{rd}_{\mathbb{C}}$.

Note that the quantity $\mathrm{rd}(n)$ we defined in the introduction can be recovered in this setting as $\mathrm{rd}(\mathrm{S}_n)$; cf. (8). Moreover, for a finite group G, our definition of $\mathrm{rd}(G)$ coincides with the definition given by Farb and Wolfson in [25].

Recall that, for a polynomial $f(x) \in K[x]$, our definition of $\mathrm{rd}(f)$ was motivated by wanting to express a root of $f(x)$ as a composition of algebraic functions in $\leqslant d$ variables applied to the coefficients. Equivalently, we wanted to find the smallest d such that the 0-cycle in $\mathbb{A}^1_K$ given by $f(x) = 0$ has an L-point for some field extension L/K of level $\leqslant d$. If G is a linear algebraic group and $T \to \mathrm{Spec}(K)$ is a G-torsor, then our more general definition of $\mathrm{rd}(T)$ retains this flavor. Indeed, T is an affine variety defined over K, and saying that T is split by L is the same as saying that T has an L-point.

While little is known about $\mathrm{rd}(n) = \mathrm{rd}(\mathrm{S}_n)$, it is natural to ask what $\mathrm{rd}(G)$ is for other algebraic groups G. Such questions can be thought of as variants of Hilbert's 13th problem. Let us now take a closer look at the case where G is linear and connected. The following folklore conjecture is implicit in the work of Tits.

Conjecture 17. *Let G be a connected complex linear algebraic group and K be a field containing $\mathbb{C}$. Then every $\alpha \in H^1(K, G)$ is split by some solvable field extension L/K.*

Since solvable extensions are of level $\leqslant 1$, this conjecture implies that $\mathrm{rd}(G) \leqslant 1$ for every connected linear algebraic group G.[11] I can prove the following weaker inequality unconditionally [52].

Theorem 18. *Let G be a connected complex linear algebraic group. Then $\mathrm{rd}(G) \leqslant 7$.*

Note that if we knew that $\mathrm{rd}(\mathrm{S}_n) \leqslant d$ for every n, we would be able to conclude that $\mathrm{lev}(L/K) \leqslant d$ for every field extension L/K of finite degree. This would, in turn, imply that $\mathrm{rd}(\mathcal{F}) \leqslant d$ for every functor $\mathcal{F}$ satisfying (15). Setting $\mathcal{F} = H^1(*, G)$, we obtain $\mathrm{rd}(G) \leqslant d$ for every algebraic group G. In particular, if we were able to show that $\mathrm{rd}(\mathcal{F}) > 1$ for some functor $\mathcal{F}$ satisfying (15), we would be able to conclude that $\mathrm{rd}(\mathrm{S}_n) > 1$ for some n. This would constitute major progress on Hilbert's 13th problem. I do not see how to reverse this implication though: an upper bound on $\mathrm{rd}(G)$ for every connected group G (such as the inequality $\mathrm{rd}(G) \leqslant 7$ of Theorem 18) does not appear to tell us anything about $\mathrm{rd}(\mathrm{S}_n)$. However, Conjecture 17 and Theorem 18 make me take more seriously the possibility that $\mathrm{rd}(\mathrm{S}_n)$ may be identically 1 or at least bounded as $n \to \infty$.

[11] Other interesting consequences of Conjecture 17 are discussed in [17].

Acknowledgements. I am grateful to EMS Magazine's editor Vladimir L. Popov for asking me to write a survey on essential dimension and resolvent degree, and for his help with the exposition. I would also like to thank Mikhail Borovoi for stimulating discussions concerning Conjecture 17, and to Patrick Brosnan, Danny Ofek, Federico Scavia and Jesse Wolfson for constructive comments.

References

[1] V. I. Arnol'd, On functions of three variables. *Dokl. Akad. Nauk SSSR* **114**, 679–681 (1957)

[2] V. I. Arnol'd, From Hilbert's superposition problem to dynamical systems. *Amer. Math. Monthly* **111**, 608–624 (2004)

[3] M. Bardestani, K. Mallahi-Karai and H. Salmasian, Kirillov's orbit method and polynomiality of the faithful dimension of p-groups. *Compos. Math.* **155**, 1618–1654 (2019)

[4] A. Beauville, Finite simple groups of small essential dimension. In *Trends in contemporary mathematics*, Springer INdAM Ser. 8, Springer, Cham, 221–228 (2014)

[5] D. Benson and Z. Reichstein, Fields of definition for representations of associative algebras. *Proc. Edinb. Math. Soc. (2)* **62**, 291–304 (2019)

[6] G. Berhuy and G. Favi, Essential dimension: a functorial point of view (after A. Merkurjev). *Doc. Math.* **8**, 279–330 (2003)

[7] C. Birkar, Singularities of linear systems and boundedness of Fano varieties. *Ann. of Math. (2)* **193**, 347–405 (2021)

[8] R. Brauer, On the resolvent problem. *Ann. Mat. Pura Appl. (4)* **102**, 45–55 (1975)

[9] P. Brosnan, Z. Reichstein and A. Vistoli, Essential dimension and algebraic stacks. arXiv:math/0701903 (2007)

[10] P. Brosnan, Z. Reichstein and A. Vistoli, Essential dimension, spinor groups, and quadratic forms. *Ann. of Math. (2)* **171**, 533–544 (2010)

[11] P. Brosnan, Z. Reichstein and A. Vistoli, Essential dimension in mixed characteristic. *Doc. Math.* **23**, 1587–1600 (2018)

[12] P. Brosnan and R. Sreekantan, Essential dimension of abelian varieties over number fields. *C. R. Math. Acad. Sci. Paris* **346**, 417–420 (2008)

[13] F. E. Browder (ed.), *Mathematical developments arising from Hilbert problems*, Proceedings of Symposia in Pure Mathematics, Vol. XXVIII, American Mathematical Society, Providence (1976)

[14] J. Buhler and Z. Reichstein, On the essential dimension of a finite group. *Compositio Math.* **106**, 159–179 (1997)

[15] N. G. Chebotarev, The problem of resolvents and critical manifolds. *Bull. Acad. Sci. URSS. Sér. Math. [Izvestia Akad. Nauk SSSR]* **7**, 123–146 (1943)

[16] A. Chen, Y.-H. He and J. McKay, Erland Samuel Bring's "Transformation of Algebraic Equations", arXiv:1711.09253 (2017)

[17] V. Chernousov, P. Gille and Z. Reichstein, Resolving G-torsors by abelian base extensions. *J. Algebra* **296**, 561–581 (2006)

[18] V. Chernousov and A. Merkurjev, Essential dimension of spinor and Clifford groups. *Algebra Number Theory* **8**, 457–472 (2014)

[19] V. Chernousov and J.-P. Serre, Lower bounds for essential dimensions via orthogonal representations. *J. Algebra* **305**, 1055–1070 (2006)

[20] H. Chu, S.-J. Hu, M.-C. Kang and J. Zhang, Groups with essential dimension one. *Asian J. Math.* **12**, 177–191 (2008)

[21] J. Dixmier, Histoire du 13e problème de Hilbert. In *Analyse diophantienne et géométrie algébrique*, Cahiers Sém. Hist. Math. Sér. 2, Univ. Paris VI, Paris, 85–94 (1993)

[22] I. V. Dolgachev and V. A. Iskovskikh, Finite subgroups of the plane Cremona group. In *Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. I*, Progr. Math. 269, Birkhäuser, Boston, 443–548 (2009)

[23] A. Duncan, Essential dimensions of A_7 and S_7. *Math. Res. Lett.* **17**, 263–266 (2010)

[24] A. Duncan, Finite groups of essential dimension 2. *Comment. Math. Helv.* **88**, 555–585 (2013)

[25] B. Farb and J. Wolfson, Resolvent degree, Hilbert's 13th problem and geometry. *Enseign. Math.* **65**, 303–376 (2019)

[26] M. Florence, On the essential dimension of cyclic p-groups. *Invent. Math.* **171**, 175–189 (2008)

[27] S. Garibaldi, Cohomological invariants: exceptional groups and spin groups. *Mem. Amer. Math. Soc.* **200**, xii+81 (2009)

[28] S. Garibaldi and R. M. Guralnick, Spinors and essential dimension. *Compos. Math.* **153**, 535–556 (2017)

[29] P. Gille and Z. Reichstein, A lower bound on the essential dimension of a connected linear group. *Comment. Math. Helv.* **84**, 189–212 (2009)

[30] C. Heberle and A. J. Sutherland, Upper bounds on resolvent degree via sylvester's obliteration algorithm. arXiv:2110.08670 (2021)

[31] D. Hilbert, Über die Gleichung neunten Grades. *Math. Ann.* **97**, 243–250 (1927)

[32] N. Karpenko and Z. Reichstein, A numerical invariant for linear representations of finite groups. *Comment. Math. Helv.* **90**, 667–701 (2015)

[33] N. A. Karpenko and A. S. Merkurjev, Essential dimension of finite p-groups. *Invent. Math.* **172**, 491–508 (2008)

[34] Y. Kawamata, Minimal models and the Kodaira dimension of algebraic fiber spaces. *J. Reine Angew. Math.* **363**, 1–46 (1985)

[35] F. Klein, *Lectures on the icosahedron and the solution of equations of the fifth degree*. Revised ed., Dover Publications, New York (1956)

[36] H. Knight, The essential p-dimension of finite simple groups of Lie type. arXiv:2109.02698 (2021)

[37] A. N. Kolmogorov, On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. *Dokl. Akad. Nauk SSSR (N.S.)* **108**, 179–182 (1956)

[38] L. Kronecker, Ueber die Gleichungen fünften Grades. *J. Reine Angew. Math.* **59**, 306–310 (1861)

[39] A. Ledet, On the essential dimension of p-groups. In *Galois theory and modular forms*, Dev. Math. 11, Kluwer Acad. Publ., Boston, MA, 159–172 (2004)

[40] A. Ledet, Finite groups of essential dimension one. *J. Algebra* **311**, 31–37 (2007)

[41] A. S. Merkurjev, Essential dimension: a survey. *Transform. Groups* **18**, 415–481 (2013)

[42] A. Meyer and Z. Reichstein, Some consequences of the Karpenko–Merkurjev theorem. *Doc. Math.* 445–457 (2010)

[43] A. Moreto, On the minimal dimension of a faithful linear representation of a finite group. arXiv:2102.01463 (2021)

[44] Y. A. Neretin, An estimate for the number of parameters defining an n-dimensional algebra. *Izv. Akad. Nauk SSSR Ser. Mat.* **51**, 306–318, 447 (1987)

[45] B. Poonen, The moduli space of commutative algebras of finite rank. *J. Eur. Math. Soc. (JEMS)* **10**, 817–836 (2008)

[46] V. L. Popov, On the Makar-Limanov, Derksen invariants, and finite automorphism groups of algebraic varieties. In *Affine algebraic geometry*, CRM Proc. Lecture Notes 54, American Mathematical Society, Providence, 289–311 (2011)

[47] V. L. Popov, Three plots about Cremona groups. *Izv. Ross. Akad. Nauk Ser. Mat.* **83**, 194–225 (2019)

[48] C. Procesi, Non-commutative affine rings. *Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. Ia (8)* **8**, 237–255 (1967)

[49] Y. Prokhorov, Simple finite subgroups of the Cremona group of rank 3. *J. Algebraic Geom.* **21**, 563–600 (2012)

[50] Y. Prokhorov and C. Shramov, Jordan property for Cremona groups. *Amer. J. Math.* **138**, 403–418 (2016)

[51] Z. Reichstein, Essential dimension. In *Proceedings of the International Congress of Mathematicians, Vol. II (Hyderabad, India, 2010)*, Hindustan Book Agency, New Delhi, 162–188 (2011)

[52] Z. Reichstein, Hilbert's 13th problem for connected groups. In preparation

[53] Z. Reichstein and A. K. Shukla, Essential dimension of inseparable field extensions. *Algebra Number Theory* **13**, 513–530 (2019)

[54] Z. Reichstein and A. Vistoli, Essential dimension of finite groups in prime characteristic. *C. R. Math. Acad. Sci. Paris* **356**, 463–467 (2018)

[55] Z. Reichstein and B. Youssin, Essential dimensions of algebraic groups and a resolution theorem for G-varieties. *Canad. J. Math.* **52**, 1018–1056 (2000)

[56] F. Scavia, Essential dimension of representations of algebras. *Comment. Math. Helv.* **95**, 661–702 (2020)

[57] J.-P. Serre, Cohomological invariants, Witt invariants, and trace forms. In *Cohomological invariants in Galois cohomology*, Univ. Lecture Ser. 28, American Mathematical Society, Providence, 1–100 (2003)

[58] J.-P. Serre, A Minkowski-style bound for the orders of the finite subgroups of the Cremona group of rank 2 over an arbitrary field. *Mosc. Math. J.* **9**, 193–208, back matter (2009)

[59] J.-P. Serre, Le groupe de Cremona et ses sous-groupes finis. *Astérisque*, Exp. No. 1000, vii, 75–100 (2010)

[60] A. J. Sutherland, Upper bounds on resolvent degree and its growth rate. arXiv:2107.08139 (2021)

[61] B. Totaro, Essential dimension of the spin groups in characteristic 2. *Comment. Math. Helv.* **94**, 1–20 (2019)

[62] E. W. Tschirnhaus, A method for removing all intermediate terms from a given equation. *ACM SIGSAM Bulletin*, **37**, 1–3 (2003)

[63] J. Wolfson, Tschirnhaus transformations after Hilbert. *Enseign. Math.* **66**, 489–540 (2020)

Zinovy Reichstein is a mathematician at the University of British Columbia in Vancouver, Canada. He was an invited ICM speaker in 2010 and was awarded the Jeffery-Williams Prize by the Canadian Mathematical Society in 2013.

reichst@math.ubc.ca

Abel interview 2021: László Lovász and Avi Wigderson

Bjørn Ian Dundas and Christian F. Skau

Professor Lovász and Professor Wigderson. First, we want to congratulate you on being the Abel Prize recipients for 2021. We cite the Abel committee:

> For their foundational contributions to theoretical computer science and discrete mathematics, and their leading role in shaping them into central fields of modern mathematics.

We would first like you to comment on the remarkable change that has occurred over the last few decades in the attitude of, say, mainstream mathematics towards discrete mathematics and theoretical computer science. As you are fully aware of, not that many years ago it was quite common among many first class mathematicians to have a sceptical, if not condescending opinion, of this type of mathematics. Please, could you start, Professor Lovász?

Lovász. I think that is true. It took time before two things were realized about theoretical computer science that are relevant for mathematics.

One is simply that it is a source of exciting problems. When I finished the university, together with some other young researchers, we started a group to study computing and computer science, because we realized that it's such a huge unexplored field; questions about what can be computed, how fast and how well and so on.

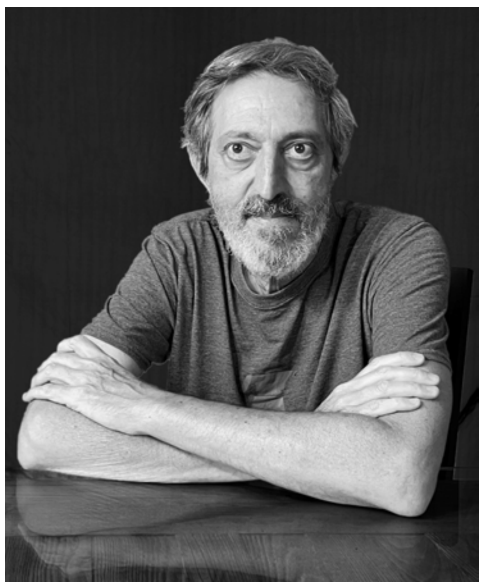
The second thing is that when answers began to come, in particular, when the notions of NP and P, i.e., nondeterministic polynomial time and polynomial time became central, we realized that the whole of mathematics can be viewed in a completely different way through these notions, through effective computation and through short proofs of existence.

For us young people these two things were so inspiring that we started to make connections with the rest of mathematics. I think it took time until other areas of mathematics also realized the significance of this, but gradually it came about. In number theory it turned out to be very important, and also in group theory these notions became important, and then slowly in a lot of other areas of mathematics.

Wigderson. Yeah, I completely agree. In fact, it's true that there was a condescending attitude among some mathematicians towards discrete mathematics. This was perhaps less so in theoretical computer science, because it existed in the realm of computer science as it was developing, and maybe people were less aware of it directly? I think that Lovász is right in that the very idea of efficient algorithms and the notions of computational complexity that were introduced in theoretical computer science are fundamental to mathematics, and it took time to realize that.

However, the real truth is that all mathematicians of all ages, they all used algorithms. They needed to compute things. Gauss' famous challenge to the mathematical community to find fast methods to test whether a number is prime and to factor integers is extremely eloquent, given the time it was written. It's really calling for fast algorithms to be developed.

Parts of discrete mathematics were viewed by some as trivial in the sense that there are only finite number of possibilities that



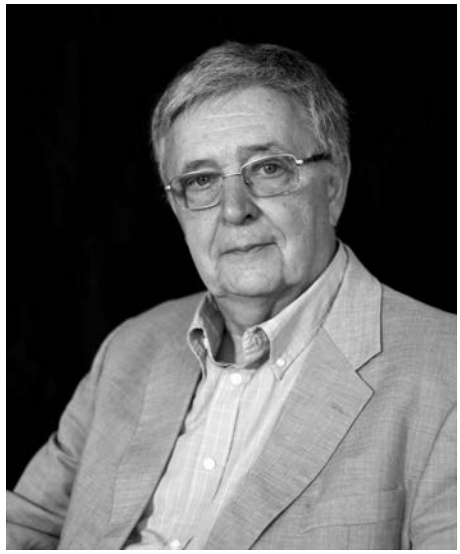
Avi Wigderson
© Peter Badge/Abel Prize

EMS MAGAZINE 122 (2021) — DOI 10.4171/MAG-54

we have to test. Then, in principle, it can be done, so what is the problem?

I think the notion of an efficient algorithm clarifies what the problem is. There may be an exponential number of things to try out, and you will never do it, right? If instead you have a fast algorithm for doing it, then it makes all the difference. The question whether such an algorithm exists becomes all important.

This understanding evolved. It caught up first with pioneers in the 70s in the area of combinatorics and in the area of discrete mathematics, because there it's most natural; at least it's easy to formulate problems, so that you can attach complexity to them. Gradually it spread to other parts of mathematics. Number theory is a great example, because there too there are discrete problems and discrete methods hiding behind a lot of famous number theoretical results. From there it gradually dispersed. I think by now it's pretty universal to understand the importance of discrete mathematics and theoretical computer science.



László Lovász
© Peter Badge/Abel Prize

Turing and Hilbert

This is admittedly a naïve question, but as non-experts we have few inhibitions, so here goes: Why is it that Turing's notion of what is today called a Turing machine captures the intuitive idea of an effective procedure, and, so to speak, sets the standard for what can be computed? How is this related to Hilbert's Entscheidungsproblem?

WIGDERSON. I think my first recommendation would be to read Turing's paper – in fact, to read *all* his papers. He writes so eloquently. If you read his paper on computing procedures and the Entscheidungsproblem, you will understand everything.

There are several reasons why the Turing machine is so fundamental and so basic. The first one is that it's simple – it's extremely simple. That was evident to Turing and to many others at that time. It's so simple that it could be directly implemented. And thereby he started the computer revolution. If you look at other notions of computability that people studied, Gödel and others – definitely Hilbert – with recursive functions and so on, they did not lend themselves to being able to make a machine out of them. So this was fundamental.

The second is that a few years later it was proved that all other notions of efficient computability were equivalent. So the Turing machine could simulate all of them. It encompassed all of them, but it was much simpler to describe.

Thirdly, one way Turing motivates his model is to look at what we humans do when we calculate to solve a problem, let's say multiply two long numbers. Look at what we do on a piece of paper, we abstract it and formalize it. And when we do that, we will automatically be lead to a model like the Turing machine.

The fourth reason is the universality, the fact that his model is a universal model. In a single machine you can have part of the data be a program you want to run, and it will just emulate this program. That is why we have laptops, computers and so on. There is just one machine. You don't need to have a different machine to multiply, a different machine to integrate, a different machine to test primality, etc. You just have one machine in which you can write a program. It was an amazing revolution and it encapsulates it in a really simple notion that everybody can understand and use, so that is the power of it.

Now, you asked about the relation to the Entscheidungsproblem. You know, Hilbert had a dream, and the dream had two parts: Everything that is true in mathematics is provable, and everything provable can be automatically computed. Well, Gödel shattered the first one – there are true facts, let's say, about integers, that can not be proven. And then Church and Turing shattered the second one. They showed that there are provable things that are not computable. Turing's proof is not only far simpler than Gödel's, with Turing's clever diagonal argument, it also implies the Gödel result if you think about it. This is usually the way most people teach Gödel's incompleteness theory today; well, I don't know if "most people" would agree to this, but it's using Turing's notions. So that's the connection. Turing was, of course, inspired by Gödel's work. The whole thing that led him into working on computability was Gödel's work.

LOVÁSZ. I have just one thing I would like to add. A Turing machine is really consisting of just two parts. It's a finite automaton and a memory. If you think about it, the memory is needed. Whatever computation you do you need to remember the partial results.

The memory in its simplest way is to just write it on a tape as a string. The finite automaton is sort of the simplest thing that you can define which will do some kind of, actually any kind of computation. If you combine the two you get the Turing machine. So it's also natural from this point of view.

P versus NP

Now we come to a really big topic, namely the P *versus* NP *problem, one of the Millennium Prize Problems. What is the* P *versus* NP *problem? Why is that problem the most important in theoretical computer science? What would the consequences be if* $\mathrm{P} = \mathrm{NP}$*? What do you envisage a proof of* $\mathrm{P} \neq \mathrm{NP}$ *would require of tools?*

LOVÁSZ. Well, let me again go back to when I was a student. I talked to Tibor Gallai, who was a distinguished graph theorist and my mentor. He said: Here are two very simple graph-theoretical problems. Does a graph have a perfect matching, that is, can the vertices be paired so that each pair is connected by an edge? The other one is whether the graph has a Hamiltonian cycle, i.e., does it have a cycle which contains all the nodes?

The first problem is essentially solved; there is a lot of literature about it. As for the other we only have superficial results, maybe nontrivial results, but still very superficial.

Gallai said, well, you should think about it, so that maybe you could come up with some explanation. Unfortunately, I could not come up with an explanation for that, but with my friend, Péter Gács, we were trying to explain it. And then we both went off – we got different scholarships: Gács went to Moscow for a year and I went to Nashville, Tennessee for a year. Then we came back and we both wanted to speak first, because we both had learned about the theory of P versus NP, which completely explains this. Peter Gács learned it from Leonid Levin in Moscow, and I learned it from listening in on discussions taking place at coffee tables at conferences.

The perfect matching problem is in P and the Hamilton cycle problem is NP-complete. This explained what really was a tough question. It was clear that this was going to be a central topic, and this was reinforced with the work of Karp proving the NP-completeness of lots of everyday problems. So, summing-up, the notions of P and NP they made order where there was such a chaos before. That was really overwhelming.

WIGDERSON. The fact that it puts an order on things in a world that looked pretty chaotic is the major reason why this problem is important. In fact, it's almost a dichotomy, almost all natural problems we want to solve are either in P, as far as we know, or are NP-complete. In the two examples Lovász gave, first the perfect matching, which is in P, we can solve it quickly, we can characterize it and do a lot of things, we understand it really well. The second example, the Hamiltonian cycle problem is a representative of an NP-complete problem.

The main point about NP-completeness is that every problem in this class is equivalent to every other. If you solve one, you have solved all of them. By now we know thousands of problems that we want to solve, in logic, in number theory, in combinatorics, in optimization and so on, that are equivalent.

So, we have these two classes that seem separate, and whether they are equal or not is the P versus NP question; and all we need to know is the answer to one of the NP-complete problems.

But I want to look at the importance of this problem from a higher point of view. Related to what I said about natural problems we want to compute, I often argue in popular lectures that problems in NP are really all the problems we humans, especially mathematicians, can ever hope to solve, because the most basic thing about problems we are trying to solve is that we will at least know if we have solved them, right? This is true not only for mathematicians. For example, physicists don't try to build a model for something for which, when they find it, they will not know if they have found it. And the same is true for engineers with designs, or detectives with solutions to their puzzles. In every undertaking that we seriously embark on, we assume that when we find what we were looking for we know that we have found it. But this is the very definition of NP: a problem is in NP exactly if you can check if the solution you got is correct.

So now we understand what NP is. If $\mathrm{P} = \mathrm{NP}$, this means that all these problems have an efficient algorithm, so they can be solved very quickly on a computer. In some sense, if $\mathrm{P} = \mathrm{NP}$ then everything we are trying to do can be done. Maybe find a cure for cancer or solve other serious problems, all these can be found quickly by an algorithm. That is why $\mathrm{P} = \mathrm{NP}$ is important and would be so consequential. However, I think most people believe that $\mathrm{P} \neq \mathrm{NP}$.

LOVÁSZ. Let me add another thought on how it can be proved that $\mathrm{P} \neq \mathrm{NP}$. There is a nice analogy here with constructions with ruler and compass. That is one of the oldest algorithms, but what can you construct by ruler and compass? The Greeks formulated the problems about trisecting the angle and doubling the cube by ruler and compass, and they probably believed, or conjectured, that these were not solvable by ruler and compass. But to prove this is not easy, even today. I mean, it can be taught in an undergraduate class, in an advanced undergraduate class, I would say. You have to deal with the theory of algebraic numbers and a little bit of Galois theory in order to be able to prove this. So to prove that these problems are not solvable by a specific algorithm took a huge development in a completely different area of mathematics.

I expect that $\mathrm{P} \neq \mathrm{NP}$ might be similar. Of course, we probably will not have to wait 2000 years for the solution, but it will take a substantial development in some area which we today may not even be aware of.

But we take it for granted that you both think that P *is different from* NP*, right?*

Wigderson. I do, but I must say that the reasons we have are not very strong. The main reason is that for mathematicians it seems obviously much easier to read proofs of theorems that are already discovered, than to discover these proofs. This suggests that P is different from NP. Many people have tried to find algorithms for many of the NP-complete problems for practical reasons, for example, various scheduling problems and optimization problems, graph theory problems, etc. And they have failed, and these failures may suggest that there are no such algorithms. This, however, is a weak argument.

In other words, I intuitively feel that $P \neq NP$, but I don't think it's a strong argument. I just believe it as a working hypothesis.

Problems versus theory

We often characterize mathematicians as theory builders or as problem solvers. Where would you place yourself on a scale ranging from theory builder to problem solver?

Wigderson. First of all, I love solving problems. But then I ask myself: Oh, this is how I solved it, but maybe this is a technique that can be applied other places? Then I try to apply it in other places, and then I write it up in its most general form, and that is how I present it. In this way I may also be called a theory builder. I don't know. I don't want to characterize myself in terms of theory builder or problem solver.

I enjoy doing both things, finding solutions to problems and trying to understand how they apply elsewhere. I love understanding connections between different problems, and even more between different areas. I think we are lucky in theoretical computer science that so many seemingly dispersed areas are so intimately connected, but not always obviously so, like with hardness and randomness. Theory is built out of such connections.

Lovász. I have similar feelings. I like to solve problems. I started out under the inspiration of Paul Erdős, who was really always breaking down questions into problems. I think that was a particular strength of his mathematics, that he could formulate simple problems that actually illustrated an underlying theory. I don't remember who said this about him: it would be nice to know the general theories that are in his head, which he breaks down into these problems that he feeds us so that we can solve them. And, indeed, based on his problems, whole new areas arose, extremal graph theory, random graph theory, probabilistic combinatorics in general, and various areas of number theory. So I started as a problem solver, but I always liked to make connections, and tried to build something more general out of a particular problem that I had solved.

László Lovász. © Hungarian Academy of Sciences/Institute for Advanced Study, Princeton, NJ, USA

Youth in Haifa

Professor Wigderson, you were born in 1956 in Haifa, Israel. Could you tell us when you got interested in mathematics and, in particular, in theoretical computer science?

Wigderson. I got interested in mathematics much earlier than in computer science. As a very young child, my father introduced me to mathematics. He liked to ask me questions and to look at puzzles, and I got interested. We found books that I could read, and in these there would be more problems. This was my main early interaction with mathematics. In high school we had a very good mathematics teacher who came from Ukraine, and he had a special class for interested kids. He taught us more exciting stuff, like college level stuff, and I got even more excited about mathematics. In college I got much more into it, but it's actually an accident that I got into computer science, and thereby to theoretical computer science.

After my army service, as I was applying to colleges in Israel, I thought that I wanted to do mathematics, but my parents suggested that it might be good to also have a profession when I graduated. So they said: "Why don't you study computer science, it will probably be a lot of math in it anyway, and you will enjoy it. Also, when you graduate you will have a computer science diploma." Nobody was thinking about academia at that time.

So I went to the computer science department at the Technion, and I think I was extremely lucky. I am sure that if I had gone to a math department I would have been interested in many other things, like analysis, combinatorics, geometry, and so on. Because I was in the computer science department, I took several theoretical courses. We had, in particular, a very inspiring teacher, Shimon Even, at the Technion. His courses on algorithms and complexity were extremely inspiring. When I applied to graduate school I applied for continuing to do this sort of stuff. This was how I was drawn into theoretical computer science.

But still, in an earlier interview you have described yourself as a beach bum and a soccer devotee. That contrasts rather starkly with what you have been telling us now, doesn't it?

Wigderson. I don't think there is a contrast. I mentioned that my dad was my main intellectual contact in mathematics. The schools in the neighbourhood of our home were not very good, it was pretty boring. The neighbourhood was situated by the beach, so everybody was at the beach. We were beach bums by definition. The weather in Israel is wonderful, so you can be altogether 300 days a year on the beach and in the water. So that was one pastime activity. The other thing you mentioned, soccer, is the easiest game to play. You need no facilities. And that was what we did, being involved in these two activities. When I was growing up I never saw myself as an intellectual. I loved math, but I also loved soccer, I loved swimming and I loved reading. And this is how I spent all my youth. There were no contrast and, if anything, it's probably good to do other things.

Youth in Budapest

Professor Lovázs, you were definitely not a beach bum.

Lovász. When I entered eighth grade, I did not have any special subject that I was particularly interested in. In the eighth grade I started to go to a math club, and I realized how many interesting problems there are. Then the teacher of the math club recommended that I should go to a particular high school which had just started and was specializing in teaching mathematically talented kids. We had ten classes of mathematics a week, which this group of students enjoyed very much, including myself. I appreciated very much the fact that I was among a fairly large group of students who had quite similar interests.

In elementary school, I was a little bit outside of the "cool" group of the class. I was not in the mainstream of the class, but in this new high school class I found myself much more at home. In fact, I felt so much at home that I married one of my classmates, Katalin Vesztergombi, and we are still together.

This high school was absolutely a great start for my life, that is how I feel about it. Before that … you just had to survive the school. I entered this new high school in the first half of the sixties. There were many good mathematicians in Budapest at that time, and they did not really have the chance to travel or to do anything outside Hungary. So they had more time, and they came to the high school and gave talks and they took a great interest in our group. We learned a lot from them. I should, of course, mention that Paul Erdős was often visiting the class and gave talks and gave problems. So it was all very inspiring.

Professor Lovász, to quote professor Wigderson: "In the land of prodigies and stars in Hungary, with its problem solving tradition, he (meaning you) stands out." We have a witness who recalls rushing home from school to watch the final of one of the competitions in which you participated on national television, where you solved a combinatorial problem in real time and won the competition. It's kind of hard to imagine doing such things now and in the West.

Lovász. You are right. That was one of the things that went on for a few years on Hungarian TV, but unfortunately it stopped. Unfortunately, because I thought it was a very good popularization of mathematics. You know, people like to watch competitions. The way it worked was that there were two glass cells, and the two students that were competing were sitting in separate cells. They got the same problem which they had to solve, and then verbally tell the solution; maybe there was a blackboard they could use as well. I think people like to watch youngsters sweating and doing their best to win. You know, most people cannot jump over two meters, but nevertheless we watch the Olympic Games. Even today I meet people, of course older people, who say: "Oh, yeah, I saw you on TV when you were in high school, and I was in the eighth grade of elementary school, and it was so nice to watch you." It was really something quite special.

A part of this story which is both funny and charming is that you told us that the solution to the final problem, with which you won the competition, you had previously learned from the other competitor. Isn't that correct?

LOVÁSZ. Yes, that's true. But we competitors were also good friends, and we still are very good friends. Especially the two people with whom I competed in the semi-final and the final, are very good friends of mine. One was Miklós Laczkovich. He came up with the proof of Tarski's conjecture about squaring the circle. And the other one was Lajos Posa. He is very well known in math education. He did a lot on developing methods to teach talented students.

Before we leave this subject, we should also mention that you won the gold medal three years in a row 1964–65–66 at the International Mathematical Olympiad, when you were 16–17–18 years old. These are impressive results! We don't know of anyone that has such record in that competition.

LOVÁSZ. Thank you, but there are others. Someone has won it five times. You can go to the website of the International Mathematical Olympiad, and there you find a list of achievements.

Lovász local lemma

Professor Lovász, you have published several papers – we think six papers altogether – with your mentor Paul Erdős. We think we know the answer to which one of these papers is your favourite, and you can correct us if we are wrong. A weak version of the important so-called Lovász local lemma was proven in 1975 in a joint paper with Erdős – that's the paper we have in mind. The lemma itself is very important as is attested to by Robin Moser and Gábor Tardos receiving the Gödel Prize in 2020 for their algorithmic version of the Lovász local lemma. Anyway, could you tell us what the Lovász local lemma is all about?

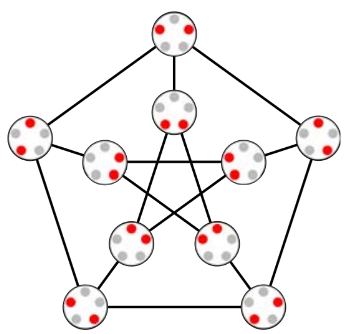
LOVÁSZ. Okay, I will try. Almost everything in mathematics, or at least in discrete mathematics, you can formulate like this: there are a number of bad events, and you want to avoid all of them. The question is whether you can give a condition so you can avoid all of these. The most basic thing is that if the probabilities of these events add up to something less than one, then with positive probability you will avoid all of them. That is a very basic trick in applications of probability in discrete mathematics. But suppose that the number is much larger, so that the probabilities add up to something very large, how do you handle that? Another special case is if they are independent events. If you can avoid each of them separately, then there is a positive probability that you avoid all of them, simply take the product of the probabilities for avoiding each one of them.

The local lemma is some kind of combination of these two ideas. If the events are not independent, but each of them is dependent only on a small number of others, and if the sum of the probabilities of those that it depends on is less than one – not the total sum, but just those it depends on – then you can still, with positive probability, avoid each of the bad events.

Maybe I should add one thing here. There was a problem of Erdős which I was thinking about, and I came up with this lemma. I was together with Erdős, actually at Ohio State for a summer school. We solved the problem, and we wrote a long paper on that problem and related problems, including this lemma. But Erdős realized that this lemma was more than just a lemma for this particular problem and made sure that it became known under my name. Normally it would be called the Erdős–Lovász local lemma, because it appeared in a joint paper of us, but he always promoted young people and always wanted to make sure that it became known if they had proved something important. I benefited from his generosity.

The Kneser conjecture

In 1955 Kneser conjectured how many colours you need in order to colour a type of naturally occurring graphs, now known as Kneser graphs. In 1978 you, professor Lovász, proved this conjecture by encoding the problem as a question of high dimensional spaces, which you answered by using a standard tool in homotopy theory, and so boosted the field of combinatorial topology. How did such a line of approach occur to you, and can you say something about the problem and your solution?



Kneser graph $K(5, 2)$

LOVÁSZ. It goes back to one of these difficult problems, the chromatic number problem: how many colours do you need to colour a graph properly, where properly means that neighbouring vertices must have different colours. That is a difficult problem in general, it's an NP-complete problem.

A first approach is looking at the local structure. If a graph has many vertices which are mutually adjacent, then, of course, you need many colours. The question is: is there always such a local reason? It was already known at that time that there are graphs which have absolutely no local structures, so that they don't have short cycles, but nevertheless you need many colours to colour them. It was an interesting question to construct such graphs. For example, just exclude triangles, or more generally, exclude odd cycles from the graph. There was a well known construction for such a graph by looking at the sphere, and then connecting two points if they are almost antipodal. Then Borsuk–Ulam's theorem says that you will need more colours than the dimension, so that the almost antipodal points have different colours. That was one construction, and the other one was the construction where the vertices would be a k-element subset of an n-element set, where $n > 2k$, and you can connect two of these if they are disjoint. Kneser conjectured what would be the chromatic number of such a graph.

It was an interesting problem going around in Budapest. Simonovits, a friend and colleague of mine, brought to my attention that these problems could actually be similar, or that these two constructions could be similar. So I came up with a reduction of one into the other, but then it turned out that the reduction was more general and gave a lower bound on the chromatic numbers of any graph in terms of some topological construction. So, that is how topology came in. It took actually quite some time to make it work. As I remember it, I spent about two years to make these ideas work, but eventually it worked.

Zero-knowledge proofs

Professor Wigderson, earlier in your career you made fundamental contributions to a new concept in cryptography, namely the zero-knowledge proof, which more than 30 years later is now being used for example in blockchain technology. Please tell us specifically what a zero-knowledge proof is, and why this concept is so useful in cryptography?

WIGDERSON. As a mathematician, suppose you found the proof of something important, like the Riemann hypothesis. And you want to convince your colleagues that you have found this proof, but you don't want them to publish it before you do. You want to convince them only of the fact that you have a proof of this theorem, and nothing else. It seems ridiculous, it seems absolutely ridiculous, and it's contrary to all our intuition that there is a way to convince someone of something they do not believe, without giving them any shred of new information.

This very idea was raised by Goldwasser, Micali and Rackoff in 1985, where they suggested this notion. They did not suggest it for paranoid mathematicians, but they suggested it for cryptography. They realized that in cryptography there are a lot of situations, in fact, almost all situations, of interactions between agents in a cryptographic protocol, in which no one trust the other ones. Nevertheless, each of them makes claims that they are doing something, or knowing something, which they don't want to share with you. For example, their private key in a public crypto system. You know, each one is supposed to compute their public key by multiplying two prime numbers, which they keep secret. I give you a number and I tell you: here is a number; I multiplied two secret prime numbers and this is the result. Why should you believe me? Maybe I did something else, and this is going to ruin the protocol. To fix this, it would be nice if there was a way for me to convince you that that's exactly what I did. Namely, there exist two prime numbers whose product is the number I gave you. That is a mathematical theorem, and I want to convince you of it, without giving you any idea what my prime numbers are, or anything else. Goldwasser, Micali and Rackoff suggested this extremely useful notion of a zero-knowledge proof.

They gave a couple of nontrivial examples, which was already related to existing crypto systems where this might be possible. And they asked the question: what kind of mathematical statements can you have a zero-knowledge proof for? A year later, with Goldreich and Micali, we proved that it was possible for any mathematical theorem. If you want it formally, it's true for any NP-statement.

So this is the content of the theorem. I am not going to tell you the proof of the theorem, though something can be said about it. The proof uses cryptography in an essential way. It's a theorem which assumes the ability to encrypt. Why it's useful in cryptography is exactly for the reasons I described, but, in fact, it's much more general, as we observed in a subsequent paper. Given a zero-knowledge proof, you can really automate the generation of protocols that are safe against bad players. The way to get a protocol that is resilient against bad players once you have a protocol that works if everybody is honest and doing exactly what they should, is just to intervene every step with a zero-knowledge proof, in which the potentially bad players will convince the others that they are doing the right thing. It's much more complex, zero-knowledge is not enough, you have to have a way to computing with secrets.

I want to stress that when we proved this theorem, it was a theoretical result. It was clear to us from the beginning that the protocol which enables zero-knowledge proof, is complex. We thought it unlikely to be of any use in cryptographic protocols that are running on machines.

The fact that they became practically useful is still an amazing thing to me, and I think that is a good point to make about many other theoretical results in theoretical computer science, in particular, about algorithms. People tend to complain sometimes about the notion of P being too liberal when describing efficient algorithms, because some algorithms, when they were first discovered, may have a running time that looks too large. It's polynomial time, but maybe it's n to the 10th power, and for n size a thousand, or size a million, which are problems that come up naturally in practice, it seemed useless, as useless as exponential time algorithms. But what you learn again and again, both in the field of cryptography and in the field of algorithms, is that once you have a theoretical solution with ideas in it that make it very efficient, then other people, especially if they are motivated enough, like in cryptography or in optimization, can make it much more efficient, and eventually practical. That's a general point I wanted to make.

Randomness versus efficiency

It's an amazing result, and we quote you saying: "This is probably the most surprising, the most paradoxical of the results that my colleagues and I have proved". Let us continue, professor Wigderson, with another topic to which you have made fundamental contributions. When you began your academic career in the late 1970s, the theory of computational complexity was in its infancy. Your contribution in enlarging and deepening the field is arguably greater than that of any other person. We want to focus here on your stunning advances in the role of randomness in aiding computation. You showed, together with co-workers Nisan and Impagliazzo, that for any fast algorithm that can solve a hard problem using coin flipping, there exists an almost as fast algorithm that does not use coin flipping, provided certain conditions are met. Could you please elaborate on this?

WIGDERSON. Randomness has always fascinated me. Specifically, the power of randomness in computation, but not only in computation. This is probably the area where I have invested most of my research time. I mean, *successful* research time! The rest would be trying to prove lower bounds or hardness results, like proving that P is different from NP, where I generally – like everybody else – failed.

So, back to randomness. Ever since the 1970s, people realized that randomness is an extremely powerful resource to have in algorithms. There were initial discoveries, like primality tests. Solovay/Strassen and Miller/Rabin discovered fast methods with randomness to test if a number is prime. Then in coding theory, in number theory, in graph theory, in optimization and so on, randomness was used all over the place. People just realized it's an extremely powerful tool to solve problems that we have no idea of how to solve efficiently without randomness. With randomness you can find the solution very fast. Another famous class of examples is Monte Carlo methods. So you explore a large chunk of problems using randomness. Without it, it seemed like it would take exponential time to solve them, and it was natural to believe that having randomness is much more powerful than not having it.

Nevertheless, mainly from motivations in cryptography, people started in computational complexity trying to understand pseudorandomness. You need randomness in cryptographic protocols for secrecy. On the other hand, sometimes random bits were not so available, and you wanted to test when random bits are good, as good as having independent coin flips – which you really assume when talking about probabilistic algorithms.

So, there was a quest to understand when a distribution of bits is as good as random. This started in cryptography with a very powerful work by Blum, Micali and Yao. Notions began to emerge which suggested that if you have computational hardness, if you somehow have a hard problem, then you can generate pseudorandom bits cheaply. So you can invest much less randomness in order to generate a lot, which is still useful, let's say for probabilistic algorithms.

This kind of understanding started in the early 1980s. It took about 20 years of work to really elucidate it and to be able to make the weakest assumptions on what hardness you need in order to have a pseudorandom outcome, which then corresponds to a full probabilistic algorithm. Parts of this were indeed developed in my papers with Nisan, and then with Babai and Fortnow, and then with Impagliazzo and Kabanets.

The upshot of this development is again a conditional result, right? You have to assume something, if you want the conclusion you stated. What you need to assume is that some problem is difficult. You can take it to be the problem of colouring graphs, you can take it to be any NP-complete problem you like, or even problems that are higher up, but you need a problem that is exponentially difficult. This is the assumption that the result is conditioned on. If you are willing to make this assumption, then the conclusion is exactly as you said, namely that every efficient probabilistic algorithm can be replaced by a deterministic algorithm which does the same thing. In fact, it does it without error and is roughly as efficient as the original one.

In other words, the power of probabilistic algorithms is just a figment of our imagination. It's only that we are unable to find deterministic algorithms that we can prove are as efficient. This result suggests that there is no such power and that randomness does not help to make efficient algorithms more efficient.

The hardness assumption that you need to make, were they something that you were expecting?

WIGDERSON. They are completely expected! First of all, they are expected in the sense that they were there from the beginning, specifically in the works of Blum, Micali and Yao that I mentioned, which do create pseudorandom generators that are good against efficient algorithms and which assume specific hardness assumptions like those used in cryptography. For example, that factoring

Avi Wigderson. © Andrea Kane, Institute for Advanced Study, Princeton, NJ, USA

is difficult, or that one way functions exist. These are very specific hardness assumptions, and these problems are unlikely to be NP-complete.

In my paper with Nisan, we realized that a much weaker assumption is enough. It was not enough to give the result stated at the end, because it's not efficient enough, but it already got us pretty close to the understanding that random algorithms are not as powerful as they seemingly are. It did not give the BPP $=$ P consequence, which is the final one. This was not surprising, the connection paradigm between hardness and randomness came from the very initial studies of computation of pseudorandomness, and, if I remember correctly, the paper of Blum and Micali, or perhaps even the Ph.D. thesis of Silvio Micali, is titled: *Hardness vs. Randomness*. There is an intimate connection there – it was there from the start – and the question is how tight the connection is.

I should probably mention that the consequence of what we just discussed is that hardness implies derandomization, and the question is whether the reverse hold also. If you have a good pseudorandom generator, or if you could derandomize all probabilistic algorithms, does it mean that you can prove something like P $\neq$ NP? The answer is that we have partial results like that. My paper with Impagliazzo and Kabanets is one, and there is another paper with just the two of them. So there are partial results for the converse, and we don't understand it fully. But it's a fascinating connection, because these two issues seem separate from each other. I think it's a very fundamental discovery of the field, this intimate connection between computational difficulty and the power of randomness.

The LLL-algorithm

Professor Lovász, we would like to talk about the LLL-*algorithm, an algorithm which has striking applications. For instance, it's claimed that the only crypto systems that can withstand an attack by a quantum computer use* LLL*. The algorithm appears in your paper together with the Lenstra brothers on factorization of polynomials, which more or less follows the expected path of reducing modulo primes, and then using Hensel's Lemma. But as far as we understand the breakthrough from you and the Lenstra brothers was that you were able to do the lift in polynomial time by an algorithm giving you an approximation to the shorter vector in a lattice. Tell us first how the collaboration with the Lenstras came about.*

Lovász. This is an interesting story about mathematics and the role of beauty, or at least elegance, in mathematics. With Martin Grötschel and Alexander Schrijver we were working on applications of the ellipsoid method in combinatorial optimization. We came up with some general theorem that stated some equivalence of separation and optimization. Actually, these were polynomial time equivalent problems under some mild additional conditions. But there was a case where the algorithm did not work, and that was when the convex body was lying in a lower dimensional linear subspace. One could always get around this, sometimes by mathematical methods, for example, by lifting everything into a higher dimensional space. But there was always some trick involved that we wanted to avoid.

At some point I realized that we can solve this if we can solve some really ancient mathematical problem algorithmically. That was Dirichlet's result that several real numbers can be simultaneously approximated by rational numbers with the same denominator, and the question was whether you could solve this algorithmically. Now one looks at the proof and one sees immediately that the proof is the opposite of being algorithmic; it's a pigeonhole principle proof, so it just shows the existence of such an approximation. After some trial and error, I came up with an algorithm which actually computed in polynomial time such an approximation with rational numbers with common denominator.

A little bit earlier I heard a talk of Hendrik Lenstra, where he talked about similar problems, but in terms of lattices and bases reduction in lattices. Now it's easy to reduce the Dirichlet problem to a shortest lattice vector problem. So I wrote to them, and it turned out that if I could solve the Dirichlet problem, then they could factor polynomials in polynomial time.

This was actually very surprising. One would think that factoring an integer should be easier than factoring a polynomial. But it turns out that it's the other way around, polynomials can be factored in polynomial time. So that is how this joint paper came about. Then a couple of years later Lagarias and Odlyzko discovered that this algorithm can be used to break the so-called knapsack crypto system. Since then this algorithm is used a lot in checking the security of various crypto systems.

As far as we understand it has applications way beyond anything that you imagined?

LOVÁSZ. Yes, definitely. For example, shortly after it was published it was used by Andrew Odlyzko and Herman te Riele in a very extended numerical computation to disprove the so-called Mertens conjecture about the ζ-function in prime number theory. But the point that I want to stress is that the whole thing started from something that was apparently not so important. Grötschel, Schriver and I just wanted to get the nicest possible theorem about equivalence of optimization and separation. This, however, was the motivation for proving something that turned out to be very important.

The ellipsoid method

Indeed, in 1981 you published a paper together with coauthors Grötschel and Schrijver entitled "The ellipsoid method and its consequence in combinatorial optimization", a paper which is widely cited, and which you touched upon in your previous answer. There is a prehistory to this, namely a paper by a Russian, Khachiyan, containing a result that was regarded as sensational. Could you comment on this, and how your joint paper is related to his?

LOVÁSZ. Khachiyan gave the first polynomial time algorithm for linear programming using what is called the ellipsoid method today. I should say that in the Soviet Union at that time there were several other people who worked on similar results, but he proved the necessary details. So it was Khachiyan who proved that linear programming can be solved in polynomial time.

Of course, everybody got interested. In the theory of algorithms before that there existed these mysterious problems that in practical terms could always be efficiently solved, but there was no polynomial time algorithm known to them. So we got interested in it, and we realized that to apply Khachyian's method you don't have to have an explicit description of the linear program. It's enough if the linear program is given in such a way that if you ask whether a point is a feasible point, then you should be able to tell this, and you should be able to find them if any constraints are violated. That observation was made by several people, including Karp and Papadimitriou, and I think Padberg and Rao. We realized that in combinatorial optimization there are many situations like this.

Then I met Martin Grötschel, and he came up with a way to apply these methods to another old problem, namely to find the chromatic number of a perfect graph in polynomial time, which was also an unsolved problem in those days. And for that it turned out that you have to apply this ellipsoid method, not only to linear programs, but to convex programs more generally. We worked on this together with Lex Schrijver, who visited the University of Szeged for a year where we shared an office, and started to see what happens in general in convex optimization and how to apply this. This is how we came up with this result that I mentioned, the equivalence of separation and optimization, it was sort of the main outcome of this study. Eventually we wrote a monograph about this subject.

The zig-zag product

Expander graphs have been a recurring theme for the Abel Prize. Last year we had Margulis, who constructed the first explicit expander graphs, after Pinsker had proven that they existed. Gromov, who won the Abel Prize in 2009, used expanders on Cayley graphs of fundamental groups, which were relevant for the study of the Baum–Connes conjecture. Also Szemerédi, who won the Abel Prize in 2012, made use of expander graphs. In 2000, you, professor Wigderson, together with Reingold and Vadhan, presented the zig-zag product of regular graphs, which is, as far as we understand, analogous to the semidirect product in group theory, by which you gave explicit constructions of very large and simple expanders. Could we just start by asking: what is the zig and what is the zag?

WIGDERSON. So, maybe I should start with what is an expander graph? You should think of networks, and you should think that one desirable property of networks is that there would be sort of fault tolerance. If some of the connections are severed you would still be able to communicate. It could be computer networks, or it could be networks of roads that you would like to be highly connected. Of course, you don't want to pay too much, so you would like these networks to be sparse, that is, you don't want to have too many connections. You want a large graph in which the degree of every vertex – that is, the number of connections to every vertex – is small, let's say constant, for example ten.

A random graph will have this property, and the whole question – this is what Pinsker realized – becomes: can you describe such graphs, and can you find them efficiently? Margulis gave the first construction using this deep algebraic concept, namely Kazhdan's property (T). They can also be built using results by Selberg and others.

Then people started to simplify the proofs. By the time I was teaching this material there were reasonably simple proofs, like the one given by Jimbo and Maruoka, and you could teach it in a class in an hour or two; it's just basically Fourier transform on finite groups. So you have everything you want, you have a very nice explicit construction, you can even prove it in a class to undergraduates, but to me it was, as with many proofs based on algebra, so mysterious. I mean, what is going on? What is really behind the fact that these are highly connected graphs? This was sort of an obsession of mine for years, and I did not know what to do with it.

In 2000, just after I moved to the IAS, I had two postdocs here, Salil Vadhan and Omar Reingold. We were working on a completely different project about pseudorandomness, where an important notion is the notion of an extractor, which has something to do with purifying randomness. I will not talk about that now, but we were trying to build better extractors. As we were doing this we realized that one of our constructions may be useful towards creating expanders. The constructions in the extractor business were often iterative, and they have a very different combinatorial nature than constructions, say, of the algebraic type. Once we realized this we understood that we had a completely different combinatorial construction of expanders, but more than that, one in which, for me, it was evident from the proof why these graphs are expanding.

This is the zig-zag result; the zig-zag name was actually suggested by Peter Winkler. The construction starts with a small graph which is expanding, and one uses it to keep boosting another graph to be an expander. So you plug this little graph in somehow, and you get a bigger expander, and then you repeat this to get a bigger one, and so on. So you can generate arbitrary large expanders. This local construction has some zig-zag picture in it if you look at it, but that is not the important thing.

There is another way of describing an expander which I think is more intuitive. An expander is a graph such that, no matter what distribution you have on the vertices, if you take a vertex from this distribution and go from this vertex to a random neighbour, the entropy of the distribution increases. That is another way to describe expanders, and this you see almost with your eyes in the zig-zag construction. You see how the entropy grows, and that is what I like about this way of looking at it.

To try to get a picture of what is going on: as far as we understand you have a graph and you place this other graph at all the vertices. Then you have to decide how to put the edges in. Then essentially what you are doing, just like in the semidirect product situation where you have the multiplication rule, you move a little bit in one of the vertices, then you jump all the way to the next vertex, and then you do the similar jump there. Is that correct, vaguely?

Wigderson. It's completely correct, and moreover the connection to semidirect products was something we realized two or three years later with Alexander Lubotzky and Noga Alon. It was sort of a challenge that I felt early on, namely that the graphs that we got were expanders, they were combinatorially generated, we understood them, and I was wondering whether our construction could be useful to construct Cayley graphs. And then with Noga Alon and Alexander Lubotzky we realized it's not just similar, but the zig-zag product is a combinatorial generalization of semidirect products of groups applied to Cayley graphs. It's more general and it specializes in the case of Cayley graphs to semidirect products. For example, because of this you can prove that Cayley graphs of groups that are not simple can be expanding with a constant number of generators. No algebraic method is known to give that.

This has been used extensively in many situations, and one of the things one perhaps should mention is that the symmetric logspace and the logspace are the same, as shown by Reingold in 2004. This seems to be an idea that really caught on. Are you still using it yourself, or have you let your "baby" grow up and run into the mathematical community?

Wigderson. I think it's great that we have a mathematical community. Many of our ideas have been taken to places beyond my imagination. There is something fundamental about this construction, and it was used like you said in this Reingold result, which can more simply be described as the logspace algorithm for connectivity in graphs. In fact, it goes back to a result of Lovász and his collaborators, and can be viewed as a randomization result.

Lovász with Karp, Aleliunas, Lipton and Rackoff showed in 1980 that if you want to test whether a large graph is connected, but you have no memory, you just need enough memory to remember where you are, then by tossing coins you can explore the whole graph. This is the random logspace algorithm for graph connectivity. Derandomizing this algorithm was another project of mine that I never got to do, but Reingold observed that if you take the zig-zag product and applied it very cleverly to their randomized algorithm, you get the deterministic logspace algorithm for the same problem. So it's a particular pseudorandom generator tailored to this. It was also used in the new PCP-theorem of Irit Dinur. So, yeah, there is something general with this zig-zag product that other people find extremely useful.

Mutual influence

Actually, this brings us to an interesting place in this interview, because here we are seeing connections between what the two of you were doing.

Wigderson. Let me mention one of the most influential things that happened to me in my postdoc years. It was in 1985. I was a postdoc in Berkeley, and there was a workshop going on in Oregon in which Lovász gave ten lectures. I don't remember exactly what it was called, but there were lectures on optimization, geometry of numbers, etc. It was a whole week of lectures and everybody wanted to hear Lovász' talk, and everybody appreciated how extremely clear his presentation was.

But the most important thing I got out of this is what Lovász described himself when you asked him the question about the LLL-algorithm, and its relation to the work on the ellipsoid and so on. He stressed how a high level point of view, rather than one focused on a specific problem, can connect lots and lots of areas of mathematics of great importance. Lovász described to you how a question that was a bit peculiar, namely about having a more elegant solution to a problem in optimization, led to solving the lattice basis reduction problem, and how it was connected to Diophantine approximation, as well as how it connects to cryptography, both to breaking crypto systems and creating crypto systems. And, you know, you get this panoramic view where everything fits in with everything. I was extremely influenced by this, it was an amazing memorable event in my early career.

Lovász. I think I have some similar memories. The zero-knowledge proof was such a shockingly exciting thing that I learned about, and it sort of showed me how powerful these new ideas of cryptography, and theoretical computer science in general, how very powerful they are. I was always very interested in Wigderson's work on randomness, even though I was sometimes trying to go the opposite direction, and find examples where randomness really helps.

One has to add that this is sometimes a matter of the model, of the computational model. I mentioned some results about convex optimization, convex geometry, algorithmic results in high dimensional convexity, and it's a basic problem there that if you have a convex body, how can you compute the volume? One of my Ph.D. students at the time, György Elekes, came up with a beautiful proof showing that you need exponential time to approximate this volume, even within a constant factor. That was in our model in which we formulated this equivalence of optimization and separation of convex bodies given by a separation oracle. A few years later, and that is actually another thing that Wigderson said, Dyer, Frieze and Kannan came up with a randomized algorithm to compute the volume, or to approximate the volume, in polynomial time with an arbitrary small relative error.

The interesting thing is the dependence on the dimension. If the dimension is n then their algorithm took n^{29} steps. Obviously this was very far from being practical, but that started their flow of research. I was also part of it and I really liked this result, and I was quite interested in making it more efficient and understanding why the exponent is so high. And then the exponent went down nicely from 29 to 17, to 10, to 7, to 5, to 4. It stood for a long time at 4, but a year ago it went down to 3. So now this is close to being practical. It's still not, n to the cube is still not enough to be a really fast algorithm, but it's definitely not ridiculously way off.

Two comments about this example. Firstly, because it's a different computational model, provably randomness helps. It's provable that without randomness it takes exponential time, and with randomness it's now down to a decent polynomial time. And the second is that polynomial time is an indicator that this problem has some deep structure. You explore this deep structure, and eventually you can improve the polynomial time to something decent.

Graphons

Here is a question to you, professor Lovász, on a subject where you have made major contributions: what is a limit theory for graphs, and what are graph limits good for? Also, explain what a graphon is.

Lovász. I will try to be not too technical. A graph is often given by an adjacency matrix, so you can imagine it as a zero-one matrix. And now, suppose that the graph is getting bigger and bigger, and you have this sequence of matrices. We always think of these as functions on the unit square, where we just cut into smaller squares, each square carrying a zero or a one. And now these functions in some sense tend to a function on the unit square, which may be continuous, or, at least not discrete any more, and that is a graphon. So, for example, if the graph is random, so each square is randomly one or zero, then it will tend to a grey square, that is, to an identical one half graphon. So a graphon is a function on the unit square, which is measurable and symmetric, and it turns out that you can exactly define what it means that a sequence of graphs converges to such a graphon.

Now, a lot of properties of the graphs are preserved, that is, if all the graphs in the sequence have a certain property, then the limit will also have this property. For example, if all these graphs have some good eigenvalue gap – a property that expanders have – then the limit will also have a good eigenvalue gap. Here we are considering dense graphs. So you look at this space of graphons, and then you have to prove – and there is a lot of technical details there – that the space of graphons in an appropriate metric is compact. This is very convenient to work with, because from then on you can, for example, take a graph parameter, let's say density of triangles. It can be defined in the limit graphon what the density of triangles is, and then in this limit graphon there will be a graphon which minimizes this under certain other conditions.

So you can play the usual game which you play in analysis, that studies the minimum, the minimizer, and then you try to determine whether it's a local minimum, or a global minimum. All these things that you can do in analysis, you can do in this setting, and this all has some translation back to the graph theory.

It's worthwhile mentioning that the Regularity Lemma of Szemerédi is closely related to the topology of graphons. In particular, compactness of the space of graphons implies a strong form of the Regularity Lemma.

The Shannon capacity

Professor Lovász, in 1979 you published a widely cited paper titled: "On the Shannon capacity of a graph". In this paper you determine the Shannon capacity of the pentagon by introducing deep mathematical methods. Moreover, you proved that there is a number, now called the Lovász number, which can be computed in polynomial time. The Lovász number is the upper bound of the Shannon capacity associated to a graph. Could you tell us a little more about that, and explain what the Shannon capacity is?

LOVÁSZ. I will not give a formal definition of what the Shannon capacity is, but you have an alphabet and you are sending messages composed of the letters of the alphabet. Now certain letters are confusable or confoundable, so they are not clearly distinguished by the recipient. You want to pick a largest subset of words which can be sent without danger of confusion. For any two words there should be at least one position where they are clearly distinguishable. So if the alphabet is described by the vertices of a graph, an edge between two letters means that those two letters are confusable. Shannon came up with this model, and he determined the capacity. If you are sending very long words, how many words can you send without causing confusion? That number grows exponentially, and the base of this exponential function is the Shannon capacity.

The pentagon graph was the first one for which the Shannon capacity was not known, and I introduced some technique called the orthogonal representation, which enabled me to answer this question.

This is an example of one of those things that occasionally happen, namely when you answer a question, then all of a sudden it begins to have its own life. For example, it was used to determine the chromatic number of perfect graphs. In a very different direction, recently a group of physicists found some quite interesting applications of it in quantum physics. So all of a sudden you hear that something you did has inspired other people to do something really interesting. That is very pleasing.

© Eirik Furu Baardsen/DNVA

The Erdős–Faber–Lovász conjecture

Our last mathematical question to you, professor Lovász, is about the so-called Erdős–Faber–Lovász conjecture, a conjecture that was posed in 1972. How did it come about, and what were your initial thoughts on how difficult it would be to prove it? Quite recently the conjecture has been proved by Kang, Kelly, Kühn, Methuku and Osthus. We should also add that apparently Erdős considered this to be one of his three most favourite combinatorial problems.

Lovász. The background for this problem was that there was a meeting in August 1972 at Ohio State University, where we discussed hypergraph theory, which was just beginning to emerge as an interesting topic. The idea is that instead of having a standard graph where an edge always has two endpoints, you can instead look at structures where an edge has three endpoints, or five endpoints, and so on. These are called hypergraphs, and the question was: given any particular question in graph theory, like chromatic number, connectivity, etc., how can this be generalized to hypergraphs?

One of these questions was what is called the edge chromatic number in graph theory. It's a well known variant of the chromatic number problem, in which case you colour the edges, not the vertices, and you want that edges incident with the same vertex should get different colours. And then you can ask the same question about hypergraphs and what upper bound you can give on the number of different colours needed. We came up with this observation that in all the known examples the number of vertices was an upper bound on the number of colours needed to edge-colour the hypergraph.

A few weeks after this meeting at Ohio State, I was visiting the University of Colorado, Boulder, and so was Erdős. Then Faber gave a party, and we began to discuss mathematics, that is what mathematicians usually do at parties, and so we came up with this question.

Erdős didn't really believe this to be true. I was more optimistic and thought maybe it is true. It certainly was a nice conjecture, stating that the number of vertices was an upper bound on how many colours is needed. Then we realized that the conjecture had some nontrivial special cases, like something called the Fisher inequality in the theory of block designs. And that is where we got stuck. The conjecture became more and more famous, it's a very elementary question, very simple to ask. Nobody could actually get a good grip on it. Eventually Jeff Kahn was able, maybe 10 years ago or so, to prove it with a factor of $1 + \varepsilon$, for every positive ε.

A year ago, Daniela Kühn and her students were able to prove it, at least for every large enough n. One peculiar feature of this conjecture is that you make a conjecture based on small n, and then you can prove it for very large n. And what is in between often remains a question mark. She gave a talk about it at the European Congress a couple of months ago, and it was very convincing, so I think it's now proved.

Quantum interactive proofs

In January 2020, five people, Ji, Natarajan, Vidick, Wright and Yuen announced that they had proved a result in quantum complexity theory that implied a negative answer to Connes' embedding problem in operator algebra theory. This came as a total surprise to a lot of people, included the two of us, as we are somewhat familiar with the Connes problem, a problem whose proof has withstood all attacks over the last forty plus years. That a problem which seems to have nothing to do with quantum complexity theory should find its solution within the latter field is astonishing to us. Professor Wigderson, do you have any comments?

Wigderson. Ever since this result came out I have tried to give popular lectures about the evolution of the particular field that is relevant to this result, namely interactive proofs, specifically the study of quantum interactive proofs and how it connects to the $\mathrm{MIP}^* = \mathrm{RE}$ result, as well as to particular questions, like the Connes embedding problem and the Tsirelson problem in quantum information theory. Of course, every particular result might be surprising, but I am not at all surprised by this connection. By now we have lots and lots of places all over mathematics where ideas from theoretical computer science, algorithms and, of course, discrete mathematics, are present and reveal their power.

As for the connection to operator algebras, and specifically to von Neumann algebras, it's not so mysterious as it may seem, because of quantum measurements involving applications of operators. The question of whether these operators commute is fundamental in the understanding both of quantum information theory and in the power of quantum interactive proofs. I was more focused on the reason that possibly a proof could be obtained in the realm of quantum interactive proofs, and not in classical quantum information theory.

If you look at the formulation of quantum interactive proofs – particularly the MIP^* ones of multiple provers – and you compare them to the EPR paper, the famous Einstein–Podolsky–Rosen Gedankenexperiment testing quantum mechanics, you see the same picture. You see there a two-prover interactive proof like you see in the more recent complexity theoretic quantum interactive proofs. If you look at the history of studying such experiments or proofs, in the physics world the focus was on particular different types of problems. There are several famous ones, like the Bell inequalities. Whereas it's very natural for people studying quantum interactive proofs to study them as a collection. There is a collection of games, some games reducible to each other, and the proof that $\mathrm{MIP}^* = \mathrm{RE}$ is a sequence of amazing reductions and amplification results using various quantum coding theory techniques and so on, even PCP techniques. This complexity-theoretic way of looking at things builds a better understanding of how they behave as a whole, and I think that is the source of the power of this approach, and the applications come from the final result just because the objects of study are operators on a Hilbert space.

Non-commutative optimization

Professor Wigderson, you are currently working on something that appears to us to be quite different from what you have been working on earlier. You call it noncommutative optimization, and it seems to us that you are doing optimization in the presence of symmetries of certain noncommutative groups, general linear groups and stuff like that. It seems like a truly fascinating project with connections to many areas. Would you care to comment a little bit on what you are doing here?

Wigderson. First of all, it's completely true that it's very different from anything that I have done before, because it's more about algorithms than about complexity. Even more, it's using far more mathematics that I did not know about beforehand. So I had to learn, and I still have to learn much more mathematics, especially invariant theory, representation theory and some algebraic geometry that I certainly was not aware of and never needed before.

This again shows the interconnectivity in mathematics, in particular, what is used from different areas of mathematics in order to obtain efficient algorithms and for obtaining other results in discrete mathematics. This connection, of course, goes in the other direction as well and enriches these mathematical areas.

This project started from something that is very dear to me, namely the derandomization project that I have been thinking about for thirty years. One of the simplest problems which we know has a probabilistic algorithm, but that we don't know have a deterministic counterpart – I mean without assumptions – is the testing of algebraic identities. You can think of the Newton identities between symmetric polynomials, you can think of the Vandermonde identity, there are lots and lots of algebraic identities in mathematics.

If anybody conjectures an algebraic identity, what do you do, how do you check it? You may think about these as polynomials with many variables. Of course, you can not expand them and compare coefficients, because this would take exponential time since there are exponentially many coefficients. Well, there is a sure probabilistic way. What we do is just to plug random numbers into the variables and evaluate the polynomials in question, and compare the results. This will be correct with high probability. So there is a fast probabilistic algorithm for this problem of polynomial identity testing, and we don't know if a fast deterministic one exists.

About twenty years ago Kabanets and Impagliazzo realized something absolutely fundamental, namely that if you find a deterministic polynomial time algorithm for this problem, you would have proved something like P different from NP. The analogue in algebraic complexity theory is that you would have proven that the permanent is exponentially harder to compute than the determinant. In short, a hardness result which will be a breakthrough in computer science and mathematics!

First of all, I would like to say that this statement should be shocking, because a fast algorithm implies hardness of a different problem. It implies a computational hardness result, which is amazing. Even before this result it was a basic problem to try to derandomize, and there were various attempts in many special cases that I worked on and others worked on. And, of course, this result made these attempts far more important.

Some years ago the issue of what happens with polynomials or rational functions that you are trying to prove are equivalent, are not with commuting variables, but are rather with noncommutative variables. It became evident that we needed it in a project here with two postdocs, Pavel Hrubes and Amir Yehudayoff. We started working on the noncommutative version of testing algebraic identities; it's basically the word problem for skew fields, so it's a very basic problem. It became apparent from our attempts that invariant theory was absolutely crucial for this problem. So understanding the invariants of certain group actions on a set of matrices, as well as understanding the degree of the generating invariants of such actions, became essential.

So I started learning about this and kept asking people in this area, and then I started collaborating with two students in Princeton, Ankit Garg and Rafael Oliveira. Eventually, cutting a long story short, together with Leonid Gurvits we found a deterministic polynomial time algorithm for solving this problem in a noncommutative domain, for noncommutative variables. Nothing like this was known, even a randomized algorithm was not known, and it uses essentially results in invariant theory.

And then we were trying to understand what we did. For the last five years I have repeatedly attempted to better understand what we did, to understand the extent of the power of these types of algorithms. What are the problems they are related to or can solve, and what these techniques can do, and what, in general, is the meaning of this result?

I should say something about applications of this. It turns out that it captures a lot of things that seemed to be unrelated. It's useful not just for testing identities, but also for testing inequalities, like the Brascamp–Lieb inequalities. It's good for problems in quantum information theory, it's good for problems in statistics, for problems in operator theory. It seems to be very broad.

Now all these algorithms just evolve along the orbit of a group action on some linear space. That is the nature of all of them. Many of these problems we are looking at are not convex, so standard convex optimization methods don't work for them. But these algorithms work. And what we understood was that these algorithms can be viewed as doing convex optimizations, standard first order, second order methods, that are used in convex optimization, except, instead of taking place in Euclidean space, they take place in some Riemannian manifold, and the convexity that you need is the geodesic convexity of that space.

By now we have a theory of these algorithms, but, of course, there is plenty that we don't understand. The growing number of application areas of this I find very fascinating. Of course, I am hoping that eventually it will help us to solve the commutative case and understand what works and what does not work there.

LL and AW are super heros

To our delight also some young Koreans have discovered that you are mathematical super heroes. Your two sons have a common Ph.D. advisor at Stanford, Jacob Fox, and this was seized upon by a South Korean popular science journal aimed at a younger audience, where you and your sons are depicted as various characters from Star Wars. As high profile scientists, do you feel comfortable being actual heroes with lightsabers and what not?

LOVÁSZ. I always like a good joke, so I think this was a great cartoon.

WIGDERSON. I also loved it, and I think that it just shows that one can always be more creative in getting younger audiences excited about mathematics in ways that you did not expect before.

© dongascience <mathdonga>

Is science under pressure?

There is a question we would like to ask that has nothing to do as such with mathematics, and that is: do you feel that science is under pressure and is this something that mathematicians can and should engage in?

LOVÁSZ. I think that is true, science is under pressure. The basic reason for that, as far as I see it, is that it has grown very fast, and people understand less and less of what is going on in each particular science, and that makes it frightening, that makes it alien. Furthermore, that also makes it more difficult to distinguish between what to believe and what not, to distinguish between science and pseudoscience. This is a real problem. I think we have to very carefully rethink how we teach students in high school. Now, mathematics is one of the areas where the teaching of it is really not up to what it could be. I guess about 90 % of the people I meet say: I have always hated mathematics.

I think we are not doing our job of teaching well. I am saying this in spite of that some of my best friends are working on trying to improve mathematics education. Many people realize that there is a problem there, but it's very difficult to move ahead. I have less experience with other areas, but just looking from the outside I can see how biology today is different from what I studied in biology in high school. It's clear that it's a huge task there in front of the scientific community.

Mathematics should play central role because a lot of sciences are using more and more mathematics, not only statistics, which is sort of standard. For example, network theory or, of course, analysis and differential equations, and quantum physics, which is really also mathematics; it's a complicated area of multilinear algebra, so to say. I think the problem is there and that we should do something about it.

On behalf of the Norwegian Mathematical Society and the European Mathematical Society, and the two of us, we would like to thank you for this very interesting interview, and again, congratulations with being awarded the Abel Prize!

WIGDERSON. Thank you!

LOVÁSZ. Thank you!

Mudumbai Seshachalu Narasimhan (1932–2021)

Oscar García-Prada

On May 15, 2021, the eminent Indian mathematician Mudumbai Seshachalu Narasimhan passed away at his home in Bangalore. His work in the field of geometry is internationally recognised, having deep connections with different branches of mathematics and theoretical physics. Narasimhan spent much of his career at the Tata Institute of Fundamental Research (TIFR) in Mumbai, where he was a key figure in the creation and development of the internationally acclaimed modern Indian school of algebraic geometry. After retiring from TIFR, from 1993 to 1999, Narasimhan was Head of the Mathematics Section of the International Centre for Theoretical Physics (ICTP) in Trieste, an institution created in 1964 by the Pakistani 1979 Nobel Laureate in Physics Abdus Salam.

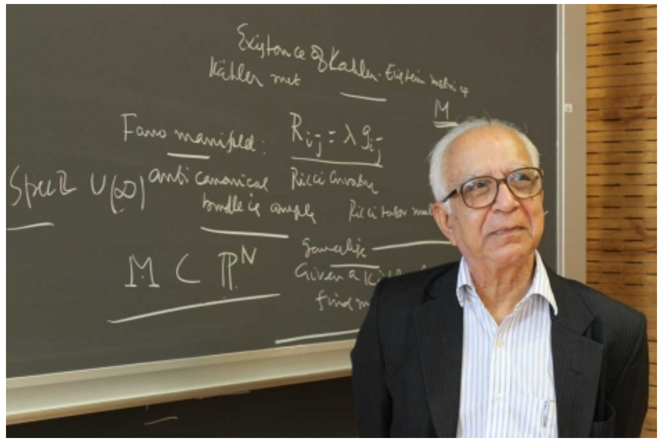


Figure 1. M. S. Narasimhan, ICMAT, Madrid, 2017

1 Life and career

Narasimhan was born on 7 June 1932 in Thandarai, a small town in Tamil Nadu (India), to a prosperous farming family. Although their circumstances were somewhat reduced after his father passed away when he was only thirteen, his family encouraged him to do what he wanted. From a young age he showed a great interest in mathematics and already in school he decided to become a researcher, even before really knowing what that meant. He completed his first university studies at Loyola College in Madras, in the heart of British India. There, he had as a teacher the French Jesuit Father Charles Racine, who was in contact with legendary figures of mathematics such as Elie Cartan, Jacques Hadamard, André Weil and Henri Cartan. Racine introduced him to modern mathematics, unknown in India, and, in particular, to the great French school. At Loyola College Narasimhan met C. S. Seshadri – also deceased in 2020 – who would later become one of his main collaborators.

Following his studies at Loyola College and on the advice of Father Racine, Narasimhan moved in 1953 to the newly created TIFR in Bombay to do his doctorate under the direction of K. Chandrasekharan, one of the founders of the centre's School of Mathematics. There he was able to interact with first-rate mathematicians who came as visitors to teach courses of two or three months. Among them was Laurent Schwartz – Fields medallist in 1950 – who would have a great influence on Narasimhan and would be his mentor during his three-year stay in Paris in the late 1950s, where he would also coincide with Seshadri. In the initial period of his stay in Paris he could not completely concentrate on mathematics as he was hospitalised due to a sickness. However, he used that time to read the paper of Kodaira and Spencer on deformations of complex structures which eventually played a great role in his future work. During his time in France he also collaborated with Japanese mathematician Takeshi Kotake, who was also in Paris to work with Schwartz.

When he returned to TIFR in 1960, Narasimhan and Seshadri started an intense collaboration that resulted in the famous Narasimhan–Seshadri theorem, published in 1965. A bit later, he began his long and fruitful collaboration with S. Ramanan. Along with Ramanan, who was his first student, Narasimhan's student roster includes other such illustrious names as N. Nitsure, R. Parthasarathy, V. K. Patodi, M. S. Raghunathan, T. R. Ramadas and R. R. Simha, who have made essential contributions to various areas of mathematics. Narasimhan's presence at TIFR was indeed a source of inspiration to several generations of young mathematicians.

During his time at TIFR, Narasimhan had also important administrative activity. In particular, he was the first Chairman of the

EMS MAGAZINE 122 (2021) — DOI 10.4171/MAG-60

National Board for Higher Mathematics, which was set up in 1983 by the Government of India, under the Department of Atomic Energy, to foster the development of higher mathematics in the country. Together with S. Ramanan, who acted as Secretary, Narasimhan undertook the task of setting it up in the initial years. He was also a member of the Executive Committee of the International Mathematical Union (IMU) during the period 1983–1986, as well as President of IMU's Commission on Development of Exchange.

After retiring from TIFR, Narasimhan was the Head of the ICTP Mathematics Section from 1993 to 1999. In this position, he carried out in particular important work in supporting young mathematicians from developing countries. When he retired from ICTP, he continued to be an adviser of ICTP and served as a member of its Scientific Council. In 2020, he was awarded the Spirit of Abdus Salam Award by the family of the ICTP founder at a ceremony where numerous mathematicians from around the world showed him their great admiration, respect and affection.

After his stay at the ICTP, Narasimhan spent three years at SISSA (Trieste), before returning to India, where he continued his mathematical activity at the Indian Institute of Science in Bangalore.

Narasimhan's work earned him many prestigious awards, including the Shanti Swarup Bhatnagar Prize (1975), Third World Academy of Sciences Prize for Mathematics (1987), the Srinivasa Ramanujan Medal (1989), the French Ordre National du Mérite (1990), the Padma Bhushan Award by the President of India (1990), the C. V. Raman Birth Centenary Award of the Indian Science Congress (1994), and the 2006 King Faisal International Prize in Science that he shared with Sir Simon Donaldson. He was also a Fellow of the Indian National Sciences Academy, Indian Academy of Sciences, the Royal Society of London and the Third World Academy of Sciences.

Narasimhan was a great fan of detective novels, and literature in general, in Tamil, English and French. He also liked Indian classical music, as well as Western classical music.

Narasimhan was married to Sakuntala Narasimhan, a renowned Indian classical music singer and journalist. The couple had a daughter, Shobhana Narasimhan, a physics researcher and professor at the Jawaharlal Nehru Center for Advanced Scientific Research, and a son, Mohan Narasimhan, who, after obtaining an MBA and having worked in the US for several years, returned to India, where he teaches martial arts.

2 Work

Narasimhan made important contributions in several areas of mathematics, including algebraic geometry, differential geometry, representation theory of Lie groups and analysis. Here, we will focus mostly on his work in algebraic geometry, and specially in the theory of moduli spaces of vector bundles on Riemann surfaces, with particular reference to works that are more familiar to the author. For details, one can consult the Collected Papers of M. S. Narasimhan [10].

The theorem of Narasimhan and Seshadri

Upon his return to TIFR in 1960, Narasimhan embarked on an intense collaboration with Seshadri that resulted in the famous Narasimhan–Seshadri theorem, published in 1965. This theorem captures the interconnection between various branches of geometry, topology and theoretical physics, and was the basis for later fundamental works by some of the greatest mathematicians of our time such as Michael Atiyah, Raoul Bott, Simon Donaldson, Karen Uhlenbeck, Shing-Tung Yau, Nigel Hitchin and Carlos Simpson, among others.

The problem of classifying holomorphic vector bundles over a compact Riemann surface X of genus g is a central one in algebraic geometry. The set of equivalence classes of holomorphic line bundles on X is given classically by the Picard group of X. For genus $g = 0$ higher rank holomorphic vector bundles were classified by Grothendieck (1957), and in a different fashion by earlier work of Birkhoff (1909). The case of elliptic curves ($g = 1$) was solved by Atiyah (1957).

For genus $g \geq 2$ the problem is much harder. Inspired by some remarks in the 1938 paper of A. Weil on "Généralisation des fonctions abéliennes", Narasimhan and Seshadri started looking in 1961–62 at unitary vector bundles. A unitary representation ρ of dimension n of the fundamental group of X defines a holomorphic vector bundle E_ρ of rank n and degree 0, which is referred to as a *unitary vector bundle*. This is called an *irreducible unitary vector bundle* if ρ is irreducible. They showed that the infinitesimal deformations of a unitary vector bundle E_ρ as a holomorphic bundle can be identified with the infinitesimal deformations of the representation ρ. From this, they deduced that the set of equivalence classes of unitary vector bundles had a natural structure of a complex manifold, and were able to compute the expected dimension.

A breakthrough came with the work of Mumford on Geometric Invariant Theory. In the 1962 International Congress in Stockholm, he introduced the notion of stability of a vector bundle on a compact Riemann surface, and proved that the set of equivalence classes of stable bundles of fixed rank and degree has a natural structure of a non-singular quasi-projective algebraic variety, projective if the rank and degree are coprime. Let E be a holomorphic vector bundle over X. Define the *slope* of E as

$$\mu(E) = \frac{\deg(E)}{\operatorname{rank}(E)}.$$

The holomorphic vector bundle E is said to be *stable* if $\mu(F) < \mu(E)$ for every proper holomorphic subbundle $F \subset E$. One can similarly define *semistability* replacing the strict inequality by $\leq$ for every subbundle $F \subseteq E$.

After they became aware of Mumford's work, the relation with unitary bundles was clear to them. Narasimhan and Seshadri proved that an irreducible unitary bundle is stable. For arbitrary degree they showed that the stable vector bundles on X are precisely the vector bundles on X which arise from certain irreducible unitary representations of suitably defined Fuchsian groups acting on the unit disc and having X as quotient. The result that they proved in [8] can be easily reformulated as saying that a holomorphic vector bundle over X is stable if and only if it arises from an irreducible projective unitary representation of the fundamental group of X. From this, one deduces that a reducible projective unitary representation of the fundamental groups corresponds to a direct sum of stable holomorphic vector bundles of the same slope (what is nowadays referred as a *polystable* vector bundle). One can observe that the projective unitary representations lift to unitary representations of a certain *central extension* of the fundamental group of X.

The Narasimhan–Seshadri theorem has been a paradigm and an inspiration for almost 60 years now for many important developments. The theorem was generalised by Ramanathan (1975) to representations into any compact Lie group. The gauge-theoretic point of view of Atiyah and Bott (1982), using the differential geometry of connections on holomorphic bundles, and the new proof of the Narasimhan–Seshadri theorem given by Donaldson (1983) following this approach, brought new insight and new analytic tools into the problem. In this approach a projective unitary representation of the fundamental group is the holonomy representation of a unitary projectively flat connection.

The case of representations into a non-compact reductive Lie group G required the introduction of new holomorphic objects on the Riemann surface X called *G-Higgs bundles*. These were introduced by Hitchin (1987), who established a homeomorphism between the moduli space of reductive representation in $\mathrm{SL}_2(\mathbb{C})$ and polystable $\mathrm{SL}_2(\mathbb{C})$-Higgs bundles. This correspondence was generalised by Simpson (1988) to any complex reductive Lie group (and in fact, to higher dimensional Kähler manifolds). The correspondence in the case of non-compact G needed an extra ingredient – not present in the compact case – having to do with the existence of twisted harmonic maps into the symmetric space defined by G. This theorem was provided by Donaldson (1987) for $G = \mathrm{SL}_2(\mathbb{C})$ and by Corlette (1988) for arbitrary G. It is perhaps worth pointing out that this theorem is a twisted version of an existence theorem of harmonic maps of Riemannian manifolds proved by Eells–Sampson (1964) pretty much around the same time as the theorem of Narasimhan and Seshadri. Corlette's theorem, which holds for any reductive real Lie group, can be combined with an existence theorem for solutions to the Hitchin's equations for a G-Higgs bundle, given by the author in collaboration with Bradlow, Gothen and Mundet i Riera (2003, 2009) to prove the correspondence for any real reductive Lie group G. Earlier, Simpson (1992) gave an indirect proof of this by embedding G in its complexification.

There is another direction in which the Narasimhan–Seshadri theorem has been generalised. This is by allowing punctures in the Riemann surface. Here one is interested in studying representations of the fundamental group of the punctured surface with fixed holonomy around the punctures. These representations now relate to the parabolic vector bundles introduced by Seshadri (1977). The correspondence in this case for $G = \mathrm{U}_n$ was carried out by Mehta and Seshadri (1980). A differential geometric proof modelled on that of Donaldson for the parabolic case was given by Biquard (1991). The case of a general compact Lie group has been studied by Bhosle–Ramanathan (1989), Teleman–Woodward (2003), Balaji–Seshadri (2015), Balaji–Biswas–Pandey (2017) and others, under suitable conditions on the holonomy around the punctures.

The non-compactness in the group and in the surface can be combined to study representations of the fundamental group of a punctured surface into a non-compact reductive Lie group G. Simpson (1990) considered this situation when $G = \mathrm{GL}_n\,\mathbb{C}$. Biquard and Mundet i Riera in collaboration with the author (2020) extended this correspondence to the case of an arbitrary real reductive Lie group G (including the case in which G is complex), establishing a one-to-one correspondence between reductive representations of the fundamental group of a punctured surface X with fixed arbitrary holonomy around the punctures and polystable parabolic G-Higgs bundles on X.

In 1972 Takemoto generalised Mumford's stability to holomorphic vector bundles on a higher dimensional complex projective variety. This was easily extended to any compact Kähler manifold and, in this setup, the projectively flat condition of the theorem of Narasimhan and Seshadri generalises to the Hermitian–Yang–Mills equation, whose existence of irreducible solutions is equivalent to Mumford–Takemoto stability of the bundle, as proved by Donaldson (1986, 1987) in the algebraic case, and by Uhlenbeck and Yau (1986) in the general Kähler situation.

In a very different direction, partial p-adic analogues of the Narasimhan–Seshadri theorem and the Hitchin–Simpson correspondence have been studied by Deninger–Werner (2005, 2010), Faltings (2005, 2011), Ogus–Vologodsky (2007), as well as Abbes–Gros (2016) and Xu (2017).

Collaboration with S. Ramanan

After his return to TIFR in 1960, Narasimhan also began his long and fruitful collaboration with S. Ramanan. Together, they developed over more than two decades the theory of moduli spaces of vector bundles on Riemann surfaces.

Their first collaboration, however, was in the area of differential geometry, proving the existence of universal connections. In a first paper (1961) they proved that for the unitary group, namely the Stiefel bundle over the Grassmannian, there was a natural homogeneous connection which could serve as a universal connection.

They later generalised this result to all compact Lie groups and in fact to all Lie groups (1963). This result has been extensively used by physicists and geometers, for instance in Chern–Simons theory and in the work of Quillen on superconnections.

After the work of Narasimhan and Seshadri, using Mumford's theory, Seshadri (1967) showed that on the set $M(n,d)$ of semistable vector bundles of rank n and degree d on X of genus $g \geq 2$, under a certain notion of equivalence introduced by Seshadri – what later was called S-equivalence –, there is a natural structure of a normal projective variety. In [6] Narasimhan and Ramanan showed that the smooth points of $M(n,d)$ correspond precisely to the stable vector bundles, except for the case $n = 2$, $g = 2$ in which case $M(2,0)$ is smooth. They also gave an explicit description of $M(2,0)$ and $M(2,1)$ when $g = 2$. The explicit description of $M(2,1)$ had also been given independently by Newstead (1968) using different methods, and was later extended by Desale–Ramanan (1976) to hyperelliptic curves. Later Narasimhan and Ramanan began studying the case of genus $g = 3$ for which an earlier purely geometric study by Coble was very helpful.

Their next joint endeavour was to study the geometry of the moduli spaces $M(n,d)$, in general, using the geometry of X. Narasimhan and Ramanan [7] proved an analogue for the moduli spaces of vector bundles of the Torelli theorem regarding the Jacobian of X. A significant difference is that, unlike the Jacobian, which can be deformed into abelian varieties which are not necessarily Jacobians, the deformations of the moduli spaces of fixed determinant are obtained only from deformations of the Riemann surface. In [7] they introduced and exploited the notion of *Hecke correspondence*. In particular, when the genus is 2, this is a correspondence between the moduli spaces $M(2,0)$ and $M(2,1)$ with fixed determinants that they had explicitly described. The Hecke correspondence has been extensively used in the study of moduli spaces and plays a central role in the Geometric Langlands Programme.

Figure 2. From left to right: M. S. Narasimhan, the author, C. S. Seshadri, S. Ramanan and M. S. Raghunathan, Indian Institute of Science, Bangalore, 2012

They later looked at direct images of line bundles on etale coverings of the Riemann surface, and described them as fixed-point subvarieties of the moduli space of vector bundles under a natural action given by tensoring by a line bundle of finite order. Using the fixed point theorems, they were able to compute some topological invariants of the moduli space. This provided a higher rank generalisation of the Prym construction that has been recently generalised to moduli spaces of principal bundles and Higgs bundles in joint work of the author with Ramanan (2019), and with Barajas (2021).

Jointly with A. Beauville, Narasimhan and Ramanan (1989) generalised the Hitchin integrable system, given by the moduli space of Higgs bundles, to the situation in which the Higgs field is twisted by an arbitrary line bundle. This was extensively used by Ngô (2010) in his proof of the fundamental lemma of the Langlands Programme. A generalisation of this system twisting by a higher rank vector bundle was given recently by Narasimhan in collaboration with G. Gallego and the author [2]. This generalisation was motivated by a problem in supersymmetric gauge theory, and made use of ideas of Chen and Ngô (2020) in their study of the Hitchin fibration for higher dimensional varieties. A generalisation of the results by Beauville–Narasimhan–Ramanan for higher dimensional varieties was given by Narasimhan and Hirshowitz (1994).

The Harder–Narasimhan filtration

Another seminal contribution of Narasimhan is his joint work with G. Harder [3] on the computation of the cohomology of the moduli space of vector bundles $M(n,d)$ with n and d coprime. Their number theoretical approach, counting points over finite fields, was based on the Weil conjectures that had just then been proved by Deligne (1974), and Siegel's formula. Earlier, Harder (1970), using the work by Newstead (1968) on the computation of the Betti numbers for $M(2,1)$, had established a connection between the cohomology groups of the rank 2 moduli space and the Tamagawa number of $SL_2(\mathbb{C})$. This method, pursued by Desale–Ramanan (1975), led to an explicit inductive formula for the Betti numbers of the moduli spaces $M(n,d)$ in the coprime situation. Later, Atiyah and Bott (1983) used Yang-Mills theory to give an alternative computation of the Betti numbers.

An important concept introduced in [3] is that of *Harder–Narasimhan filtration*. Harder and Narasimhan proved that given any vector bundle E there is a canonical filtration

$$0 = E_0 \subset E_1 \subset \cdots \subset E_k = E$$

such that E_i/E_{i-1} is semistable for $i = 1, \ldots, k$, and

$$\mu(F_i/F_{i-1}) > \mu(F_{i+1}/F_i) \quad \text{for } i = 1, \ldots, k-1.$$

The Harder–Narasimhan filtration also played a central role in the approach of Atiyah and Bott, using the differential geometry of connections and holomorphic structures on vector bundles.

The notion of Harder–Narasimhan filtration has been extended to principal bundles, Higgs bundles and other similar objects with important applications. Analogues of this filtration have been used extensively in numerous other contexts in algebraic geometry and number theory.

Other contributions

There are many other important contributions of Narasimhan, some of which would deserve a section of their own, but for lack of space we will just briefly describe some of them here. For a more complete account we refer to [10].

It was in the mid 1950s that the first papers of Narasimhan appeared. They were devoted to the study of the Laplace operator on Riemannian manifolds (1956) and certain extensions of elliptic operators (1957). After these, he wrote a paper giving a new approach to the construction of Green's function of an open Riemann surface (1960), and another paper studying the local properties of variations of complex structures on a relatively compact subdomain of an open Riemann surface (1961).

Together with T. Kotake, Narasimhan proved a theorem characterising real analytic functions by Cauchy-type inequalities satisfied with respect to powers of a linear elliptic operator with analytic coefficients (1962). This result was used in the original proof of the Atiyah–Bott fixed point theorem, and has been generalised in several directions by many authors, including Lions–Magenes, Bouendi–Goulaouic, Bouendi–Metvier and Bolly–Camus–Mattera.

Narasimhan and R. R. Simha [9] proved, using differential geometric methods, that the set of isomorphism classes of complex structures with ample canonical line bundle on a compact connected real analytic manifold has a natural structure of a Hausdorff complex space.

Jointly with K. Okamoto [4], Narasimhan made an important contribution to the theory of representation of Lie groups. It had been suggested by Langlands (1966) that, in analogy to the Borel–Weil–Bott theorem for compact groups, the Harish-Chandra discrete series of a real semisimple non-compact Lie group defining a symmetric space of Hermitian type could be realised as square-integrable harmonic forms in certain holomorphic vector bundles. The work of Narasimhan and Okamoto was the first breakthrough in the proof of this conjecture. Although Narasimhan did not pursue this any further, his student Parthasarathy has contributed in an important way to this field.

Narasimhan wrote two joint papers with H. Lange: the first one (1983) on the study of maximal subbundles of rank two vector bundles on curves, and a second one (1989) on squares of ample line bundles on abelian varieties.

J.-M. Drezet and Narasimhan [1] proved that the moduli space of vector bundles on a curve is locally factorial and determined the Picard group, showing that this is isomorphic to the integers. Their results enable one to define a generalisation of the Riemann theta divisor of the Jacobian. The famous Verlinde formula gives the dimension of the space of sections of powers of the theta line bundle (generalised theta functions) on the moduli space. Tsuchiya–Ueno–Yamada (1989) had proved factorisation theorem and the Verlinde formula in the context of Conformal Field Theory. Narasimhan and Ramadas [5] gave an algebro-geometric proof of this in the rank 2 case, which they extended also to parabolic bundles. In a previous collaboration, Narasimhan and Ramadas (1979) studied Yang–Mills theory on the product of the 3-sphere with the real line, using topological and differential geometric techniques to identify the configuration space as the base space of a principal bundle with the gauge group as structure group.

In joint work with S. Kumar (1997), Narasimhan extended his result with Drezet to the moduli space of principal bundles over a compact Riemann surface with a simple, simply-connected connected complex affine algebraic structure group. And with S. Kumar and A. Ramanathan (1997), using the relation between principal bundles and infinite Grassmanians, they elucidated the relation between conformal blocks and generalised theta functions. This enables one to compute the dimension of the space of generalised theta functions using the Verlinde formula. This was also proved by Beauville–Lazlo (1994) in the vector bundle case.

Narasimhan and M. Nori (1981) proved that there are only finitely many smooth curves having a given abelian variety as the Jacobian. I. Biswas and Narasimhan (1997) studied Hodge classes of moduli spaces of parabolic bundles on general curves. With Y. I. Holla (2001), Narasimhan proved a generalisation of a theorem of Nagata on a ruled surface to the case of a bundle of flag varieties associated to a principal bundle.

Narasimhan also worked on vector bundles on higher dimensional varieties. He studied the moduli space M of stable vector bundles of rank 2, vanishing first Chern class and second Chern class $c_2 = 2$ on complex projective 3-space. With A. Hirschowitz (1982) he proved that M is rational. A compactification of M was given by Narasimhan and G. Trautmann (1990) as the closure in the moduli space of sheaves constructed by Maruyama (1978). Later, Narasimhan and Trautmann (1991) computed the Picard group of the compactification. With W. Decker and F.-O. Schreyer (1990), he studied rank 2 vector bundles on projective 4-space, developing a construction by Barth (1980) of irreducible rank 2 bundles with first Chern class $c_1 = -1$. G. Elencwajg and Narasimhan (1983) wrote a paper on projective bundles on complex tori.

Jiayu Li and Narasimhan (1999) proved a correspondence relating the existence of a Hermitian–Einstein metric on a rank 2 parabolic bundle over a Kähler surface to the stability of the parabolic bundle. This was related to work by Munari (1993) and Biquard (1997).

3 Some personal reminiscences

I first met Narasimhan quite soon after having completed my doctoral thesis in 1991. From the very beginning, he was very kind to me, and extremely generous in the exchange of ideas. In those years, we mostly met in conferences in Europe, but thanks to my collaboration with S. Ramanan, whom I had met soon after Narasimhan, I started travelling regularly to India, where we also met.

We were very lucky to have Narasimhan in Madrid on several memorable occasions. In 2006 he participated in a panel, jointly with Sir Michael Atiyah, Jean-Pierre Bourguignon, Philip Candelas, José Manuel Fernández de Labastida, and Shing-Tung Yau on "New Interactions between Geometry and Physics", organised in the context of a conference in honour of Nigel Hitchin for his 60th birthday, that took place in Madrid soon after the International Congress. In 2012, the Instituto de Ciencias Matemáticas (ICMAT) in Madrid organised a conference in his honour for his 80th birthday, and later in 2017 he was invited as a special guest for a conference that ICMAT organised celebrating Ramanan's 80th birthday. On that occasion he participated in a special panel jointly with Antonio Córdoba, Nigel Hitchin and S. Ramanan on "Mathematics in India and Europe". A photographic exhibition on "Kolam, an Ephemeral Women's Art of South India" by photographer and anthropologist Claudia Silva was opened after the panel.

Over the years, we had many discussions on the possibility of establishing a scheme for mathematical collaboration between India and Europe in our research field. There had been some bilateral programmes between France and India, and we were contemplating the idea of bringing that to a larger context. It took a long time, but eventually we established a collaboration programme involving four nodes in Europe (Aarhus, Madrid, Oxford and Paris) and four in India (Bangalore, Mumbai and two in Chennai). This was the Indo-European Project on Moduli Spaces that was operating during the period 2013–2017, involving more than eighty mathematicians, funded under the Marie Curie Programme by the European Commission, and coordinated by ICMAT in Madrid. Narasimhan played an important role in the gestation of this project.

Figure 3. From left to right: M. S. Narasimhan, S. Ramanan, N. Hitchin and A. Córdoba, ICMAT, Madrid, 2017

Figure 4. From left to right: Guillermo Barajas, Guillermo Gallego, Gadadhar Misra and M. S. Narasimhan, Bangalore, 2020

In addition to discussing mathematics and scientific collaboration, Narasimhan and I very much liked to enjoy a glass (or two!) of good red wine, very often in company of our common friend and collaborator Ramanan, and other good friends. My wife and I were very fortunate to enjoy his great hospitality and that of his wife Sakuntala and daughter Shobhana, at his home during our many visits to Bangalore over the last few years.

I last saw Narasimhan in person in Bangalore in February 2020, during an activity on Moduli Spaces organised at the International Centre for Theoretical Sciences (ICTS). On that occasion we also had the opportunity to have a very nice dinner, accompanied as usual by good red wine, with our friend Gadadhar Misra and other friends. After the ICTS meeting, I went to Chennai for a few days, for a visit to the Chennai Mathematical Institute (CMI), where as a matter of fact I also saw C. S. Seshadri for the last time. I had actually met Seshadri in the late 1980s when I was still a graduate student at Oxford, where he gave a talk on parabolic bundles, a subject of great interest at the time in relation to Jones–Witten theory and the Atiyah–Segal approach to topological quantum field theory.

My students Guillermo Barajas and Guillermo Gallego also came to the workshop at ICTS in February 2020, and after that, while I was visiting the CMI, they went to the Indian Institute of Science to discuss with Narasimhan for a week. As always Narasimhan was extremely generous, spending a lot of time talking with them and, together with Gadadhar Misra, entertaining them (Figure 4). The last paper of Narasimhan, written jointly with Gallego and the author [2] appeared just a few days before his passing. The discussions with Barajas were very useful in connection with a joint paper of Barajas and the author (2021), which generalises to principal bundles and Higgs bundles the Prym-type construction given by Narasimhan and Ramanan (1975).

In addition to being a great mathematician, Narasimhan was a wonderful human being. He was kind, generous and sympathetic, and is very much missed by many people who loved him.

References

[1] J.-M. Drezet and M. S. Narasimhan, Groupe de Picard des variétés de modules de fibrés semi-stables sur les courbes algébriques. *Invent. Math.* **97**, 53–94 (1989)

[2] G. Gallego, O. García-Prada and M. S. Narasimhan, Higgs bundles twisted by a vector bundle. arXiv:2105.05543 (2021)

[3] G. Harder and M. S. Narasimhan, On the cohomology groups of moduli spaces of vector bundles on curves. *Math. Ann.* **212**, 215–248 (1975)

[4] M. S. Narasimhan and K. Okamoto, An analogue of the Borel–Weil–Bott theorem for hermitian symmetric pairs of non-compact type. *Ann. of Math. (2)* **91**, 486–511 (1970)

[5] M. S. Narasimhan and T. R. Ramadas, Factorisation of generalised theta functions. I. *Invent. Math.* **114**, 565–623 (1993)

[6] M. S. Narasimhan and S. Ramanan, Moduli of vector bundles on a compact Riemann surface. *Ann. of Math. (2)* **89**, 14–51 (1969)

[7] M. S. Narasimhan and S. Ramanan, Deformations of the moduli space of vector bundles over an algebraic curve. *Ann. of Math. (2)* **101**, 391–417 (1975)

[8] M. S. Narasimhan and C. S. Seshadri, Stable and unitary vector bundles on a compact Riemann surface. *Ann. of Math. (2)* **82**, 540–567 (1965)

[9] M. S. Narasimhan and R. R. Simha, Manifolds with ample canonical class. *Invent. Math.* **5**, 120–128 (1968)

[10] M. S. Narasimhan, *The Collected Papers of M. S. Narasimhan*, two volumes, edited by N. Nitsure, Hindustan Book Agency (2007)

Oscar García-Prada is a CSIC research professor at the Instituto de Ciencias Matemáticas – ICMAT in Madrid. He obtained a DPhil in mathematics at the University of Oxford in 1991 and had postdoctoral appointments at the Institut des Hautes Études Scientific (Paris), the University of California at Berkeley and the University of Paris-Sud, before holding positions at the University Autónoma of Madrid and the École Polytéchnique (Paris). In 2002, he joined the Spanish National Research Council (CSIC). His research interests lie in the interplay of differential and algebraic geometry with differential equations of theoretical physics, specifically in the study of moduli spaces and geometric structures.

oscar.garcia-prada@icmat.es

Teaching *school mathematics* to prospective teachers

Hung-Hsi Wu

What kind of mathematics should be taught to prospective mathematics teachers has been a longstanding open problem in mathematics education. We contend that we should teach them exactly what they need for their work: *school mathematics.*

1 Introduction

Good school mathematics education requires teachers who are mathematically knowledgeable. After all, one can't teach what one doesn't know. However, at least in America, we are still none too sure about what kind of mathematics we should teach prospective teachers to make them knowledgeable (cf. [12]). In a well-known article back in 1990 [1], Deborah Ball reported on her study of the subject matter knowledge of 252 prospective mathematics teacher candidates (217 elementary school teachers and 35 high school teachers) in five universities. The study zeroed in on one topic: division of fractions. When presented with the division of $1\frac{3}{4} \div \frac{1}{2}$ and four story problems, only 30 % of them were able to select the problem that correctly represented this division. In a smaller study, 35 of the 217 teachers (25 elementary and 10 high school) were asked to *create* a word problem of their own to correctly represent this division. Only 4 out of the 35 teachers (thus 11 %) could give a satisfactory answer and all 4 were high school teachers. Ball's (separate) interviews—on the same topic of fraction division—with mathematics majors in college who did not plan to go into teaching did not produce better results. Her conclusion was that the subject matter preparation of prospective teachers was in dire need of our serious reappraisal.

The inquiry into how best to help prospective teachers acquire the needed understanding of mathematics for teaching naturally predated Ball's study and went back to at least the beginning of the 20th century. In the waning days of the New Math phase of the 1960s, E. G. Begle also pondered over the possible correlation between teachers' knowledge of the subject matter and their students' achievements. In his 1972 study of 308 teachers of high school algebra [2], he found no evidence that the amount of teacher training in mathematics led to increased student achievement. This finding was further confirmed in 1979 [3].

The decades since the works of Begle and Ball have lent clarity to the phenomenon they uncovered. We will first analyze Ball's data about $1\frac{3}{4} \div \frac{1}{2}$, and then put the data in the proper perspective by coming to terms with the fact that *school mathematics* is a separate discipline distinct from the mathematics we teach in universities.

2 The division of fractions: two views

We will approach the topic of fraction division from two perspectives. First, we describe what elementary students need to know to answer Deborah Ball's questions and, second, what university students in a course on algebra can learn about fraction division. Due to length limitations, we will focus only on the critical *mathematical* differences between the two without addressing the pedagogical ramifications.

When the topic of fraction division is brought up in upper elementary school, students face a real conceptual challenge: the concept of a fraction is a higher level of abstraction than anything they have ever faced, and the concept of division is the most elusive of the four arithmetic operations on fractions. Students cannot overcome either obstacle if they are not told *exactly* what these concepts mean. As an Arizona elementary school teacher Kyle Kirkman said:

> I have learned that precise mathematical definitions are critical. If precision is lacking, students will fill in any missing or vague elements of the definition with whatever happens to be present in their paradigm that seems to fit the idea. Not all of mathematics is intuitive in nature, so this can definitely lead to erroneous conclusions. [12, Section 4.2.4]

Unfortunately, it is the case that school mathematics usually explains fractions to students in terms of vague metaphors without giving a precise definition, at least not one that students can use for reasoning about the four operations on fractions. We have to first describe a remedy for this deplorable situation. We will define a fraction in terms of something that feels "real" and "tangible"

Figure 1

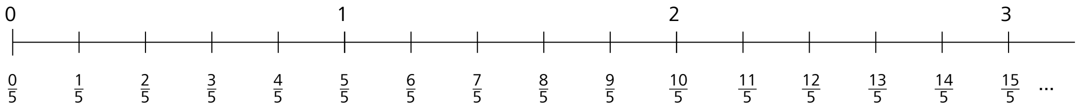
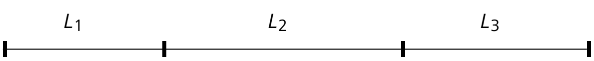


Figure 2

to elementary students, and the commonly accepted definition nowadays is as a point on the so-called *number line* (see [9, Sections 12.1 and 12.2] or [11, pp. 1–18]), as follows. We assume that we can tell whether two segments (i.e., closed intervals) on a line have *equal length* or not. A *number line* is a horizontal line on which the whole numbers have been identified as points so that the numbers $1, 2, 3, \dots$ are placed successively to the right of 0 and the segments $[0, 1], [1, 2], [2, 3], \dots$ *all have equal length* (Figure 1). The fractions with denominators equal to 5 (for example) consist of the whole numbers together with the division points when each of the segments $[0, 1], [1, 2], [2, 3], \dots$ is divided into 5 *equal parts*, i.e., 5 segments of equal length (Figure 2). We call this sequence the *sequence of fifths*. We can likewise introduce the *sequence of n-ths* for each nonzero whole number n. (Observe the resemblance of the sequence of n-ths for each n to the sequence of whole numbers.) *Fractions* are by definition the totality of all the points in the sequence of n-ths for all nonzero whole numbers n.

Next, we introduce the concept of *length* for certain segments. By definition, the *length* of the segment $[0, \frac{a}{b}]$ ($\frac{a}{b}$ a fraction) is $\frac{a}{b}$. Thus a segment with the same length as $[0, \frac{a}{b}]$ now also has length $\frac{a}{b}$. To put this definition to use, we introduce the concept of the *concatenation* of a collection of segments—say L_1, L_2, and L_3—to be the segment formed by putting these segments together end-to-end:



It follows that the length of the concatenation of 3 of the parts when $[0, 1]$ is divided into (let us say) 7 equal parts is $\frac{3}{7}$ because this segment has the same length as $[0, \frac{3}{7}]$.

Since division is based on multiplication, we will come straight to *fraction multiplication* without discussing *equivalent fractions* or *fraction addition*. By definition, $\frac{2}{5} \times \frac{3}{4}$ is the length of the concatenation of 2 of the parts when the segment $[0, \frac{3}{4}]$ is divided into 5 equal parts. The multiplication of two fractions in general is defined similarly (see, e.g., [11, Section 1.5] or [15, Section 1.4]). It becomes a nontrivial fact (for elementary students) to prove the following *product formula*:

$$\frac{2}{5} \times \frac{3}{4} = \frac{2 \times 3}{5 \times 4}. \tag{1}$$

See, e.g., [11, Theorem 1.5, p. 60].

This definition of fraction multiplication did not come out of the blue. If, in the definition of $\frac{2}{5} \times \frac{3}{4}$, we replace the fraction $\frac{3}{4}$ by 1 ($= \frac{1}{1}$), then the definition of $\frac{2}{5} \times 1$ ("the total length of 2 of the parts when $[0, 1]$ is divided into 5 equal parts") becomes exactly the above definition of $\frac{2}{5}$, so that (not surprisingly) $\frac{2}{5} \times 1 = \frac{2}{5}$. Furthermore, if we consider the product of whole numbers, say 2×3, we can also regard it as the multiplication of the fractions $\frac{2}{1}$ and $\frac{3}{1}$. Then the definition of fraction multiplication says that this product is the total length of 2 of the parts when the segment $[0, 3]$ is divided into 1 equal part, i.e., when each part is the segment $[0, 3]$ itself. In other words, the product 2×3, whether considered as the product of two *whole numbers* or the product of two *fractions*, is just $3 + 3$. In this light, we see that this definition of fraction multiplication is a very natural outgrowth of familiar concepts.

How is this concept of multiplication related to the real world? To elementary students, this is an important concern, as the following problem shows.

Example 1. If $4\frac{2}{3}$ buckets of water fill a water container exactly, what is the volume of the container if the volume of the bucket is 5.5 liters?

Solution. The important thing is to understand the given data. Since $4\frac{2}{3} = 4 + \frac{2}{3}$ by definition, the container contains 4 buckets and $\frac{2}{3}$ of a bucket of water. The total volume of 4 buckets is clear: $4 \times 5\frac{1}{2}$ liters. Now, students have to understand (and a teacher should explain) that "$\frac{2}{3}$ of $5\frac{1}{2}$ liters" means that it is the total volume of "2 of the parts when $5\frac{1}{2}$ liters is divided into 3 parts of equal volume". *By our definition of fraction multiplication,* this is precisely $\frac{2}{3} \times 5\frac{1}{2}$ liters on the number line whose "1" is interpreted as "1 liter". By the distributive law, the volume of the container is

$$\left(4 \times 5\frac{1}{2}\right) + \left(\frac{2}{3} \times 5\frac{1}{2}\right) = 4\frac{2}{3} \times 5.5 \text{ liters.}$$

Thus, "$4\frac{2}{3}$ *of* 5.5 liters" is equal to "$(4\frac{2}{3} \times 5.5)$ liters".

Incidentally, this *explains* why, when textbooks do not define fraction multiplication, they give the rote instruction that the word "of" means "multiply".

Next, division. We must first review the concept of division *among whole numbers* (see [9, pp. 97–100]). Observe that, whereas we can add or multiply *any* two whole numbers, we are not free to subtract or divide any two whole numbers. For example, in the context of whole numbers, the subtraction $3 - 7$ is not allowed, nor is $21 \div 5$. Let us explain the latter: within whole numbers, we can write $21 \div 7$ (respectively, $15 \div 3$) only because we know ahead of time that 21 (resp. 15) is a *whole number* multiple of 7 (resp. 3). For example, the definition of division $21 \div 7$ is:

$$21 \div 7 = \{\text{the whole number } k \text{ so that } k \times 7 = 21\}. \qquad (2)$$

This is why $21 \div 7 = 3$. The definition makes it perfectly clear that, without a prior guarantee that 21 is a multiple of 7, *the whole number* $21 \div 7$ would be impossible to define. Equivalently, if we do not know that 21 objects can be partitioned into 3 equal groups of 7, then we cannot talk about $21 \div 7$. If students find equation (2) to be confusing, remind them that (2) is no different from the definition of subtraction:

$$21 - 7 = \{\text{the whole number } \ell \text{ so that } \ell + 7 = 21\}.$$

Why this review is important is that the division among whole numbers serves as a model for the division among fractions, because whole numbers are also fractions (see [9, pp. 284–289]). So, according to (2), the division $1\frac{3}{4} \div \frac{1}{2}$ ($1\frac{3}{4}$ is just $\frac{7}{4}$) would make no sense unless $1\frac{3}{4}$ is a *fractional multiple* of $\frac{1}{2}$ in the sense that $1\frac{3}{4} = \frac{m}{n} \times \frac{1}{2}$ for some fraction $\frac{m}{n}$. (This $\frac{m}{n}$ is unique; see [9, Lemma, p. 286] or [11, Lemma 1.7, p. 75].) *Assuming there is such an* $\frac{m}{n}$, then we can define $1\frac{3}{4} \div \frac{1}{2}$ in exactly the same way as in (2):

$$1\frac{3}{4} \div \frac{1}{2} = \Big\{\text{the fraction } \frac{m}{n} \text{ so that } \frac{m}{n} \times \frac{1}{2} = 1\frac{3}{4}\Big\}. \qquad (3)$$

See [9, p. 289] or [11, p. 75].

Surprisingly, in contrast with the case of whole numbers, it turns out that such a fraction $\frac{m}{n}$ on the right side of (3) can always be found as follows:

$$\begin{aligned} 1\frac{3}{4} &= 1 \times 1\frac{3}{4} = \Big(\frac{1}{2} \times \frac{2}{1}\Big) \times 1\frac{3}{4} \\ &= \frac{1}{2} \times \Big(\frac{2}{1} \times 1\frac{3}{4}\Big) \quad \text{(associative law of mult.)} \\ &= \Big(\frac{2}{1} \times 1\frac{3}{4}\Big) \times \frac{1}{2} \quad \text{(commutative law of mult.).} \qquad (4) \end{aligned}$$

From (4), we see that if we let $\frac{m}{n} = \frac{2}{1} \times 1\frac{3}{4}$, then $1\frac{3}{4} = \frac{m}{n} \times \frac{1}{2}$ and (3) would allow us to conclude that

$$1\frac{3}{4} \div \frac{1}{2} = \frac{2}{1} \times 1\frac{3}{4}.$$

This is of course the *invert and multiply rule* for fraction division. This reasoning is seen to be perfectly general.

We now give a word problem whose solution requires the use of the division $1\frac{3}{4} \div \frac{1}{2}$ in Ball's article [1] *and we will also explain how this comes about*.

Example 2. How many cups of water will fill a jar with a volume of $1\frac{3}{4}$ liters if the cup holds $\frac{1}{2}$ liters?

Solution. Let $\frac{m}{n}$ cups of water fill the jar. Using the reasoning in Example 1 about the volume of a water container, we see that

$$\frac{m}{n} \times \frac{1}{2} = 1\frac{3}{4}.$$

By the definition of fraction division, this means

$$\frac{m}{n} = 1\frac{3}{4} \div \frac{1}{2} = \frac{2}{1} \times 1\frac{3}{4} = 3\frac{1}{2},$$

where the last equality is a routine calculation.

We have now done enough to show the minimal mathematical knowledge a school teacher needs to teach fraction division correctly to elementary students. We point out once again that this minimal knowledge is *not* typically what elementary students are taught in schools. Be that as it may, it is time to take up the other concern in Ball's 1990 article about why university mathematics majors may not possess such knowledge either. We will only be able to provide the barest outline in the following discussion.

A university course on abstract algebra that includes the mathematically correct way to define fractions is essentially students' first introduction to abstract mathematics. The main purpose of such a course is to guide students' first steps in the new environment of what is called *abstract mathematics*. Hence the relentless emphasis in such courses is on correct definitions and proofs, and on reducing all complex mathematical phenomena—by the use of logic—down to the bare essentials. For the case at hand, let us put ourselves at the juncture where students are already in possession of the integers, to be denoted by $\mathbb{Z}$, and are made aware that the main defect of $\mathbb{Z}$ from an abstract point of view is that no nonzero integer other than 1 and -1 has a *multiplicative inverse*, i.e., given an integer z, $z \neq 1$ or -1, there is no integer z' so that $zz' = z'z = 1$. The way to eliminate this defect is to expand $\mathbb{Z}$ by including the desired multiplicative inverses to form the *field of quotients* $\mathbb{Q}$. This $\mathbb{Q}$ is of course what we call the *rational numbers* (the fractions and negative fractions), but in the abstract setting, we cannot just adjoin the new numbers $\pm\frac{1}{2}$, $\pm\frac{1}{3}$, etc., to $\mathbb{Z}$ and declare, "There you are!". After all, *what are these new numbers and how do we add and multiply them*? We want students to learn how to use a similar reasoning to expand any *integral domain* into a *field* so that every nonzero element of the integral domain will have a multiplicative inverse in the field. The way to do this is to form the set of all ordered pairs of integers $\{\langle u, v\rangle\}$ (where u and v are integers with $v \neq 0$) and introduce into this set an *equivalence* *relation* (which essentially declares that the cross-multiplication algorithm is valid), and then $\mathbb{Q}$ is by definition the set of *equivalence classes*. After that, we can show that each integer u in $\mathbb{Z}$ can be identified with the equivalence class containing $\langle u, 1\rangle$, and we also write $\frac{u}{v}$ for the equivalence class containing $\langle u, v\rangle$ so as to align the new notation with the old. In particular, this means $\frac{u}{1}$ is identified with the integer u for each u.

For beginners, just getting used to this general construction and being at ease with the idea that each "number" in $\mathbb{Q}$ is now an equivalence class (each containing an infinite number of elements) is already a full-time job. But more is yet to come. So far, we only have a larger set $\mathbb{Q}$ containing $\mathbb{Z}$ but we do not yet know how to do arithmetic in $\mathbb{Q}$, i.e., given two arbitrary elements of $\mathbb{Q}$, we do not know as yet how to add them or multiply them. The next step is therefore to *define* the rules for adding and multiplying elements in $\mathbb{Q}$ (which are equivalence classes) with the goal of showing that $\mathbb{Q}$ in fact forms a *field*, which means in particular that each nonzero element z of $\mathbb{Q}$ will have a multiplicative inverse z^{-1}, i.e., so that $zz^{-1} = z^{-1}z = 1$. Here are the relevant definitions: for u, v, s, t in $\mathbb{Z}$ with $v \neq 0$ and $t \neq 0$,

$$\frac{u}{v} + \frac{s}{t} \overset{\text{def}}{=} \frac{ut + sv}{vt},$$
$$\frac{u}{v} \times \frac{s}{t} \overset{\text{def}}{=} \frac{us}{vt}. \tag{5}$$

We underscore the momentous shift in perspective that has just taken place here. In school mathematics, fractions are considered to be a part of nature that students should get to know; the idea that two fractions can be multiplied is taken for granted. What needs to be to explained is how the product of two fractions is related to the daily phenomena around us and why the product formula (1) is correct. By contrast, abstract mathematics progresses from $\mathbb{Z}$ to $\mathbb{Q}$ by regarding only the integers as known so that how to add or multiply the *unknown* non-integer rational numbers is a total blank that is waiting to be filled in; this is done by judiciously defining what the sum and product of two rational numbers must be. The internal structure of $\mathbb{Q}$ is the sole concern here, not how $\frac{u}{v} \times \frac{s}{t}$ is related to daily phenomena. In particular, whereas equation (1) is a theorem in school mathematics, the same statement (5) is merely a *definition* in university mathematics.

We can now explain why university mathematics majors are generally not capable of explaining to elementary students how to multiply two fractions. First of all, most if not all of these math majors were not provided with this kind of knowledge when they were in elementary school themselves (see, e.g., [16]). More to the point, what they learn about fractions in college mathematics courses is about the abstract structure of the rational numbers as a field, not about how fractions are related to daily experiences. Therefore, it is not that university mathematics majors are ignorant about fractions, but that their understanding of fractions is divorced from the concerns of elementary students. To the extent that multiplication is the foundation of division, the same comment will apply to the school mathematics of fraction division, as we now show.

As part of the mission of university mathematics to reduce all phenomena to bare essentials, the four arithmetic operations in school mathematics are reduced to only two, namely, addition and multiplication. In a field, subtraction $a - b$ is *by definition* the addition $a + (-b)$, where $-b$ is the *additive inverse* of b, and division $a \div b$ ($b \neq 0$) is *by definition* just the multiplication $a \times b^{-1}$, where b^{-1} is the multiplicative inverse of b. Since the multiplicative inverse of a nonzero rational number $\frac{s}{t}$ is clearly just the reciprocal $\frac{t}{s}$, the invert-and-multiply rule is now—like the product formula (5)—a matter of definition:

$$\frac{u}{v} \div \frac{s}{t} \overset{\text{def}}{=} \frac{u}{v} \times \left(\frac{s}{t}\right)^{-1} = \frac{u}{v} \times \frac{t}{s}. \tag{6}$$

From the point of view of abstract mathematics, "division" is just an afterthought once multiplication is in place. Mathematics majors would usually be busy with exploring the new algebraic structures (groups, fields, rings, etc.) at this point and any puzzlement over division or its ramifications in real life simply does not enter the picture. If they cannot help elementary students overcome the fear of "Ours is not to reason why, just invert and multiply", it is—again—not because they know less than school teachers but because they know something *different* from the concerns of elementary students.

3 What is school mathematics?

Through one small topic—fraction division—we get to see the critical difference between what may be called *university mathematics* (the mathematics taught in universities to prepare students for mathematical research) and *school mathematics* (the mathematics taught in K-12 schools). A main goal of the former is to introduce students to abstract mathematics, and the main emphasis is on logical completeness and the use of abstractions to achieve this goal. No matter how gently this is done, it is too austere and too sophisticated to be suitable for use in schools. School students who come mostly from the world of tactile experiences need a bridge to help them transition to the world of abstractions. School mathematics is that bridge, and it should be recognized as an independent discipline devoted to the *customization* of university mathematics to meet the needs of school students, in the same way that chemical engineering is the discipline that customizes abstract chemical principles to meet human needs. In this sense, *school mathematics* is *mathematical engineering* (see [7]).

Now, there is good engineering and there is also bad engineering. Good engineering always observes the basic principles of its associated scientific discipline—for example, mechanical engineering does not engage in designing perpetual motion machines—but bad engineering can do just the opposite. In the case of mathematics, bad mathematical engineering has been at work for a long time at least in America; it has produced school mathematics that seems to make a mockery of the fundamental principles of mathematics (see, e.g., [16]). But before proceeding any further, let us state one version of the *fundamental principles of mathematics* [8]:

(i) *Clear definitions.* Each concept is precisely defined so as to be usable for reasoning.
(ii) *Logical reasoning.* Every claim is supported by reasoning that explains *why* it is true. (It is understood that in a few special cases, such as the fundamental theorem of algebra, the reasoning can be deferred.)
(iii) *Precise language.* There is no place for ambiguity in a discipline where the difference between true and false is absolute.
(iv) *Coherence.* The concepts and skills are not fragmented bits and pieces but are part of a coherent whole.
(v) *Purposefulness.* Each concept or skill is there for a purpose.

We have seen all of them in action in the preceding discussion of fraction division. Thus, fraction, fraction multiplication, and fraction division were all precisely defined to make possible the use of reasoning to explain formulas (5) and (6). An example of the precision that is in school mathematics is the definition of *division among whole numbers* that shows why "$m \div n$" does not always make sense for two arbitrary whole numbers m and n. As for "coherence", we took pains to explain how the definition of fraction multiplication evolves from the definition of a fraction as well as from the definition of whole-number multiplication. We also showed that the definition of fraction division is modeled on the definition of whole-number division. Finally, although the *purpose* of the concepts of fraction multiplication and division is all too obvious, there are many other concepts or skills whose presence in the school curriculum is not well explained, e.g., why learn how to round to the nearest ten or nearest thousand (see [9, Chapter 10]), why take the absolute value of a real number (see [15, pp. 130–131] and [14, pp. 120, 123]), etc. Also see the discussion of *slope* below.

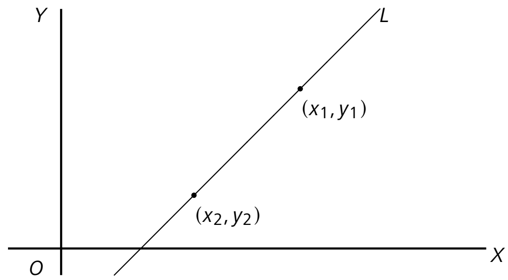
We will refer to school mathematics that observes the fundamental principles of mathematics as *PBSM* (Principles-Based School Mathematics; see [5]).

We now have the necessary tools to revisit the problem concerning the mathematical education of teachers that Begle, Ball, and others uncovered but did not clearly articulate. In our language, their message is that to get mathematically knowledgeable teachers, we have to teach teachers *PBSM* instead of university mathematics. This is because school mathematics and university mathematics are related but essentially distinct disciplines, so that knowing university mathematics does not imply knowing PBSM. We have underscored their differences using a small topic—that of fraction division—but there are many other such examples. Let us briefly look at two additional ones to further plead our case: the concept of the *slope* of a line, and the broad issue of the school geometry curriculum. Similar examples are pointed out throughout the six volumes [9–11, 13–15].

First, consider how school mathematics handles "slope". The typical starting point is to let students retain their naive conception of a line as in Euclidean geometry and define *slope* in terms of this naive conception. Thus, let a line L in the coordinate plane $\mathbb{R}^2$ be given. Suppose L is not *vertical* (i.e., not parallel to the y-axis). Then school mathematics defines the *slope* of L as the quotient

$$\text{slope of } L \stackrel{\text{def}}{=} \frac{y_1 - y_2}{x_1 - x_2}, \tag{7}$$

where (x_1, y_1) and (x_2, y_2) are any two distinct points on L.



We can explain to students that the slope of a (nonvertical) line is a measurement of its "slant" relative to the y-axis (see [15, pp. 338–346]). Incidentally, this explanation is an example of the *purposefulness* of a concept. In any case, the central fact concerning slope is the following theorem [14, Theorem 6.11, p. 354].

Theorem 1. *The graph of a linear equation $y = mx + b$ (m and b are constants) is a line with slope m, and conversely a line with slope m is the graph of an equation $y = mx + b$.*

There is a subtlety hidden in the definition of slope: how do we know that the right side of (7) does not change no matter which two points (x_1, y_1) and (x_2, y_2) are chosen on L? Most school textbooks evade this question, leading to much confusion in students' understanding of slope. The fact is that to answer this question, we need the theorem that two triangles are similar if they have a pair of equal angles. Rare is the school curriculum that has covered similar triangles by the time it takes up the topic of slope. Consequently, slope is rarely defined correctly. If there is no correct definition for a concept, then there can be no theorem involving the concept. Consequently, Theorem 1 is almost never proved in school mathematics.

Not surprisingly, university mathematics approaches slope by ignoring any reference to students' naive knowledge and simply *defining* a line in the plane as the graph of an equation $y = mx + b$ (m and b being constants) or $x = b$ (a vertical line). Then the *slope* of the graph of $y = mx + b$ is by definition m. Very simple! Therefore, brevity and total clarity are achieved at the expense of students' intuition. (Unfortunately, there are mathematics textbooks for teachers that ignore the need for mathematical engineering and also define a line the same way.) Clearly, such an understanding of the slope of a line, while mathematically correct, will not help secondary school students to come to terms with the concept of slope.

Finally, a few passing remarks about the school geometry curriculum. There are obvious defects in this curriculum that cry out for correction. We have already brought up the need for coordinating the teaching of similar triangles with the teaching of slope; this need is generally not met. There is also the need to explain the concepts of *congruence* and *similarity* because they come up naturally in daily life. However, the school curriculum usually teaches only *triangle* congruence and similarity in the course on Euclidean geometry but never the congruence and similarity of general geometric figures. This is not only defective as general education but also detrimental to the school mathematics curriculum itself as a general knowledge of the congruence and similarity of parabolas would greatly clarify the subject of quadratic equations and functions (see [13, Sections 2.1 and 2.2]). Last but not least, the course on Euclidean geometry is usually flaunted as the crown jewel of school education on teaching students how to use logic to prove *everything* strictly on the basis of axioms. The sooner we can disabuse school students of this illusion the better! Indeed, we have known since the work of Hilbert (1862–1943) that the axiomatic system of Euclidean geometry is extraordinarily subtle and its inner workings are not suitable for the education of school students (see the early chapters in Hartshorne's book [4]; they will tax the dedication of even university mathematics majors). School mathematics education should steer away from this make-believe about axiomatic systems of Euclidean geometry and, instead, try to introduce a reasonably large number of *redundant* assumptions into Euclidean geometry to minimize students' need to prove many boring, obvious, and difficult-to-prove theorems at the beginning. Compare [15, Chapters 4–5] and [13, Chapters 6–8].

Needless to say, no part of university mathematics will ever address any of these issues in the presentation of high school geometry. Serious mathematical engineering is called for here to make plane geometry truly consumable by high school students.

4 An existence proof

Thus far, we have advocated for the need to teach prospective teachers PBSM. The implicit assumption is that PBSM has always been around and is ready for the taking. This is a pleasant assumption to make and an even more pleasant assumption to believe. However, it is sobering to realize that, with all kinds of defective school mathematics out in the world, there is a distinct possibility that university mathematics can never be customized for the consumption of school students without violating one or more of the fundamental principles of mathematics. Alan Schoenfeld seems to be the first among educators to acknowledge in 1994 that, although he believed that something like PBSM should exist, there was as yet no documented proof that such was the case [6]. What we can report in 2021 is that there is now at least one systematic exposition of PBSM from kindergarten to grade 12 in the form of six volumes: [9] for teachers of grades K-5, [10, 11] for teachers of grades 6–8, and [13–15] for teachers of grades 9–12.

We can explain the need for such a *complete* exposition of thirteen years of PBSM. There have been articles and books that demonstrate the possibility of introducing reasoning to a specific topic or two in school mathematics, but discussions on such a small scale cannot bring out the essence of the fundamental principles of mathematics. For example, to expose teachers to the need for precise definitions, we cannot show them PBSM on just a few key topics because teachers need to experience this need in *every* aspect of school mathematics, including the definitions of the most mundane of concepts such as percent, ratio, speed, equation, variable, angle, graph of an inequality, etc. Or, consider the issue of coherence: it is usually invisible when school mathematics is viewed through a microscope, such as a focus on fraction addition or fraction division. But when the subject of fractions is taken as a whole, then the way the theorem of equivalent fractions pulls all the diverse parts of fractions together becomes somewhat breathtaking (see, e.g., [11, pp. 28–86]). On a slightly larger scale, one also gets to witness coherence at work when the concept of division is shown to be qualitatively the same for whole numbers, fractions, rational numbers, and real numbers (cf. [9]). We should add that, without such a longitudinal overview of school mathematics, the defects of the school geometry curriculum might not have been detected.

The 6-volume exposition of PBSM, beyond providing a foundation for student textbooks in school mathematics, shows in detail how we can achieve a better mathematical education for teachers. In America, teachers are taught in three grade-bands: elementary (K-grade 5), middle school (grades 6–8), and high school (grades 9–12). As noted above, the six volumes in question have been written with these grade-bands in mind so that, collectively, they now provide one answer to the original question implicitly raised by Begle, Ball, et al., namely, what kind of mathematics should we teach teachers? (A more detailed answer to this question is given in [15, p. xxi].) It goes without saying that school mathematics curricula are not now—and won't ever be—all alike, but we hope such a complete exposition of PBSM will nevertheless contribute to better school mathematics education by freeing educators from the need to perform the necessary mathematical engineering. It should now be relatively easy to freely modify this existing model [9–11, 13–15] to meet diverse needs.

Acknowledgements. I am very indebted to Larry Francis for his important suggestions for improvement.

References

[1] D. L. Ball, The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal* **90**, 449–466 (1990)

[2] E. G. Begle, Teacher knowledge and student achievement in algebra. SMSG Reports, No. 9, https://files.eric.ed.gov/fulltext/ED064175.pdf (1972)

[3] E. G. Begle, *Critical Variables in Mathematics Education: Findings from a Survey of the Empirical Literature*. Mathematical Association of America (1979)

[4] R. Hartshorne, *Geometry: Euclid and Beyond*. Springer (1997)

[5] R. C. Poon, Principle-based mathematics: an exploratory study. Dissertation at University of California, Berkeley, http://escholarship.org/uc/item/4vk017nt (2014)

[6] A. Schoenfeld, What do we know about Mathematics Curricula? *J. Math. Behavior* **13**, 55–80 (1994)

[7] H. Wu, How mathematicians can contribute to K-12 mathematics education. In *Proceedings of International Congress of Mathematicians, Vol. III (Madrid, 2006)*, European Mathematical Society, 1676–1688, http://math.berkeley.edu/~wu/ICMtalk.pdf (2006)

[8] H. Wu, Phoenix rising. Bringing the Common Core State Mathematics Standards to life. *Amer. Educator* **35**, 3–13, www.aft.org/pdfs/americaneducator/fall2011/Wu.pdf (2011)

[9] H. Wu, *Understanding Numbers in Elementary School Mathematics*. American Mathematical Society, Providence (2011)

[10] H. Wu, *Teaching School Mathematics: Algebra*. American Mathematical Society, Providence (2016)

[11] H. Wu, *Teaching School Mathematics: Pre-Algebra*. American Mathematical Society, Providence (2016)

[12] H. Wu, The content knowledge mathematics teachers need. In *Mathematics Matters in Education*, edited by Y. Li, W. J. Lewis and J. Madden, Springer, Cham, 43–91, https://math.berkeley.edu/~wu/Contentknowledge1A.pdf (2018)

[13] H. Wu, *Algebra and Geometry*. American Mathematical Society, Providence (2020)

[14] H. Wu, *Pre-Calculus, Calculus, and Beyond*. American Mathematical Society, Providence (2020)

[15] H. Wu, *Rational Numbers to Linear Equations*. American Mathematical Society, Providence (2020)

[16] H. Wu, Learnable and unlearnable school mathematics. https://math.berkeley.edu/~wu/AE2020A.pdf (2021)

Hung-Hsi Wu is Professor Emeritus of Mathematics at the University of California at Berkeley, U.S.A. He is a geometer but has spent the last twenty-five years educating school teachers.

wu@berkeley.edu

RICAM, the Johann Radon Institute for Computational and Applied Mathematics

Philipp Grohs, Peter Kritzer, Karl Kunisch, Ronny Ramlau and Otmar Scherzer

The Johann Radon Institute for Computational and Applied Mathematics (RICAM) of the Austrian Academy of Sciences focuses on basic research in applied mathematics. The institute is based in Austria's third-largest city and industrial hub, Linz. Researchers from all around the globe collaborate on common core areas in mathematical modeling, simulation, inverse problems, and optimization. RICAM stands for excellence in research, as can be seen from a high level of publications and the popularity of the Institute's Special Semesters within the academic community. The work groups at RICAM provide a broad field of expertise over a whole range of different subjects, and together they create an exciting atmosphere to carry out research in applied mathematics.

1 Introduction: The structure of RICAM

The *Johann Radon Institute for Computational and Applied Mathematics* (RICAM) was founded in 2003 by Heinz Engl, who is now acting as the Rector of the University of Vienna, with the goal of establishing an internationally visible and successful research institute in the field of applied mathematics. Since then, RICAM has carried out basic research in computational and applied mathematics according to highest international standards and has emphasized interdisciplinary cooperation between its workgroups and with institutions with similar scope and universities around the globe. The researchers also cooperate with other disciplines, in particular within the framework of Special Semesters on topics of major current interest. One of the institute's goals is to support young scientists. Indeed, the positions at the institute are usually for PhD students and PostDocs, and most of the members are young scientists who are in a stage of their career between the very beginning of their doctorate and the final step of obtaining a permanent position at a university or another research institution. The leaders of the work groups, which usually comprise 2–10 members, are typically university professors in Austrian universities. Through its position as the biggest mathematical non-university institute in Austria, through its work, and by its efforts in public outreach, RICAM promotes the role of mathematics in science, industry, and society.

RICAM currently consists of the following work groups with their corresponding leaders:

- *Computational Methods for PDEs* led by Ulrich Langer/Herbert Egger (Johannes Kepler University Linz),
- *Geometry in Simulations* led by Bert Jüttler (Johannes Kepler University Linz),
- *Inverse Problems and Mathematical Imaging* led by Otmar Scherzer (University of Vienna),
- *Mathematical Data Science* led by Philipp Grohs (University of Vienna),
- *Mathematical Methods in Medicine and Life Sciences* led by Luca Gerardo-Giorda (Johannes Kepler University Linz),
- *Optimization and Optimal Control* led by Karl Kunisch (Karl Franzens University Graz),
- *Symbolic Computation* led by Josef Schicho (Johannes Kepler University Linz),
- *Transfer Group* led by Ronny Ramlau (Johannes Kepler University Linz),
- *Applied Discrete Mathematics and Cryptography* led by Arne Winterhof (RICAM),
- *Multiscale Modeling and Simulation of Crowded Transport in the Life and Social Sciences* led by Marie-Therese Wolfram (University of Warwick), being phased out in 2021,
- *Multivariate Algorithms and Quasi-Monte Carlo Methods* led by Peter Kritzer (RICAM).

Prof. Heinz Engl was the first Director of the Institute until the end of 2011, when he became Rector of the University of Vienna.

Currently, the institute is led by Managing Director Prof. Ronny Ramlau and Deputy Director Prof. Karl Kunisch.

In the following sections, we would like to present selected recent research activities at RICAM.

EMS MAGAZINE 122 (2021) — DOI 10.4171/MAG-37

Figure 1. Left to right: Managing Director Ronny Ramlau, Deputy Director Karl Kunisch, and the founder of RICAM, Heinz Engl. Photos by Claudia Börner.

Figure 2. The institute is housed in one of the Science Park buildings on the campus of JKU Linz. Photo by Claudia Börner.

2 Scientific machine learning

The usage of (data-driven) machine learning methods to tackle model-based problems, for example from the natural sciences, has recently evolved into a promising research area often termed "scientific machine learning". In this section we describe partly related efforts in this direction by the two research groups "Mathematical Data Science" and "Optimization and Optimal Control".

The group "Mathematical Data Science" represents the area of data science, particularly in relation to different aspects of applied mathematics. Its research is motivated by interdisciplinary applications and ranges from theory over algorithm development to the solution of real-world problems. A particular focus is currently put on the use and analysis of deep learning methods for the numerical solution of mostly high dimensional partial differential equations (PDEs) that are burdened by the curse of dimensionality in the sense that the computational complexity of most known algorithms scales at least exponentially in the underlying problem dimension. Prominent examples include the Black–Scholes (BS) PDE from computational finance, the Hamilton–Jacobi–Bellmann (HJB) PDEs arising from stochastic optimal control problems, or the many-electron Schrödinger equation from computational chemistry. For BS PDEs and certain nonlinear HJB PDEs we were recently able to prove that neural networks are capable of representing their solutions without incurring the curse of dimensionality [9, 10], and that such solutions can be numerically found by solving an empirical risk minimization (ERM) problem of a size scaling only polynomially in the problem dimension [3]. While the analysis of the computational complexity of the ERM problem remains wide open, there are some empirical results suggesting that its scaling does not suffer from the curse of dimensionality either, see [2] and Figure 3. In another direction, we investigate the numerical solution of the many-electron Schrödinger equation using a deep

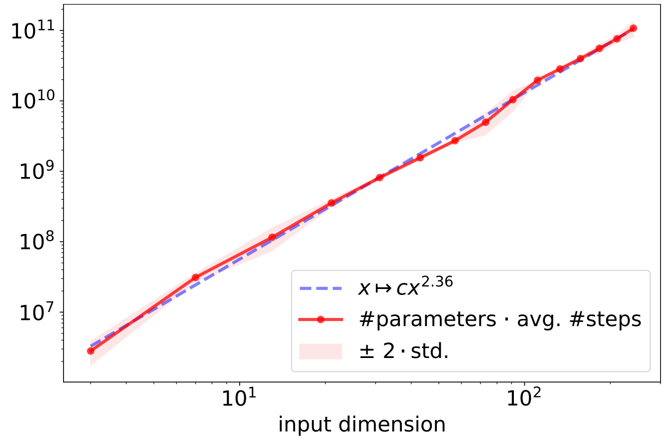


Figure 3. Empirical evidence shows that the total computational cost for solving parametric Black–Scholes PDEs by the deep learning based algorithm of [2] does not suffer from the curse of dimensionality.

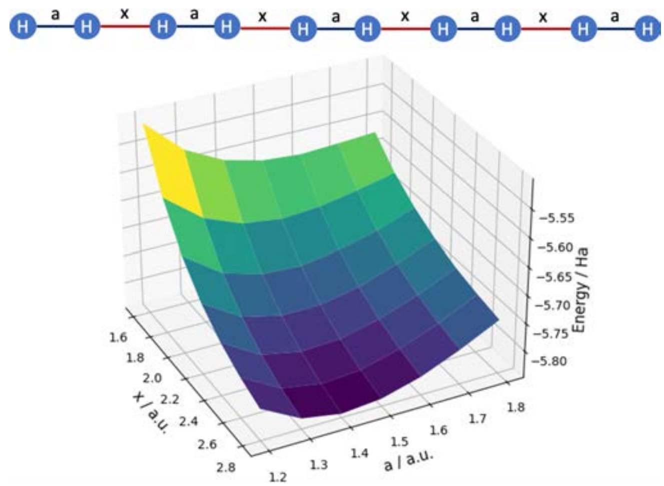


Figure 4. Potential energy surface of H10 chain computed by the deep learning based algorithm of [15].

learning based Quantum Monte Carlo ansatz. We show that a judicious weight sharing strategy enables the efficient computation of potential energy surfaces with chemical accuracy, see [15] and Figure 4. We see that the use of modern machine learning methods for problems in scientific computing has the potential to overcome the curse of dimensionality for several important problems arising for example in computational finance, computational chemistry or optimal control. It therefore presents exciting opportunities to render previously intractable problems in these fields feasible. The extent to which this potential can be fully realized will constitute a central topic in future work.

The "Optimization and Optimal Control" group combines techniques from scientific machine learning with efforts to efficiently solve deterministic closed loop optimal control problems governed by partial differential equations, and problems in computer vision and medical imaging.

In [6] a tensor decomposition approach for the solution of high-dimensional, fully nonlinear HJB equations arising in deterministic optimal feedback control of nonlinear dynamics is presented. It combines a tensor train approximation for the value function with a Newton-like iterative method for the solution of the resulting nonlinear system. In numerical tests the tensor approximation leads to a polynomial scaling with respect to the dimension, thus partially circumventing the curse of dimensionality. In an alternative approach [13], rather than obtaining the feedback from the HJB equation directly, the feedback gains are approximated by neural networks, which are trained by open loop optimal controls. A third promising approach for the computation of high-dimensional optimal feedback laws is based on sparse regression exploiting the control-theoretical link between HJB equations and first-order optimality conditions via Pontryagin's Maximum Principle [1]. Combined with model reduction techniques these methods have the potential of solving closed loop optimal control for complex partial differential equations. In another line of research [7], variational formulations of mathematical imaging problems are combined with deep learning techniques by introducing a data-driven total deep variation regularizer. We take advantage of the well-known phenomenon that typically the best image quality is achieved when the gradient flow process is stopped before converging to a stationary point. This paradox originates from a trade-off between optimization and modeling errors of the underlying variational model. An optimal stopping time is introduced, which is learned from data by means of an optimal control approach.

3 Inverse problems in science and industry

For many years, Linz has been a center of research in the area of "Inverse Problems". Given – possibly noisy – measured data y, the goal of this research area is the development and analysis of methods that allow for a reconstruction of the underlying quantity x using a suitable model F that connects y and x. This results in the task of solving a (possible) nonlinear operator equation $F(x) = y$. In many applications of interest the solution x of the operator equation does not depend continuously on the data y, and requires the use of *regularization techniques* for a stable reconstruction [8]. A well known example is Computerized Tomography (CT), where the measured damping of x-rays passing through a body is used to reconstruct the density distribution of the body. In this case, the connection between the density and the data is given by the Radon transform R [14].

An active research area of the "Inverse Problems and Imaging Group" concerns the analysis and implementation of regularization methods, in particular total variation (TV) regularization (Figure 5) and tomographic inversions with uncertainties in the model (cf., e.g., [12]). One particular application of the latter is 3D visualization of optically and acoustically trapped particles from recorded diffraction images.

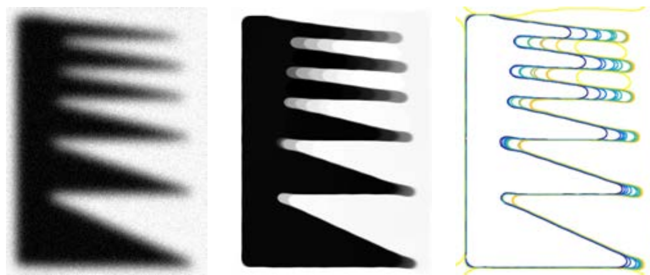


Figure 5. Total variation deblurring with decreasing noise level and regularization parameter. Left to right: input image blurred and with additive noise, numerical deconvolution results, some of the level lines approaching those of the true data (geometrical convergence).

The "Transfer Group" is involved in several research projects with industry as well as other branches of science, with the goal of developing state of the art mathematical methods for applications. The group has a very tight collaboration with the "Inverse Problems and Imaging Group" within the special research program (SFB) *Tomography across the Scales*, which is formed by a group of researchers from mathematics, physics, astronomy and medical engineering. The Transfer Group has been particularly active in the field of Adaptive Optics (AO). AO was initially invented for the correction of degraded images from astronomical images, caused by turbulences in the atmosphere above earth-bound telescopes. To this end, incoming wavefront data from reference guide stars is measured by a wavefront sensor, and the reconstructed wavefront is used to compute an appropriate shape of deformable mirrors that are employed to correct the science images of the telescope (Figure 6). The whole correction process involves a series of mathematical tasks, e.g., modeling, analyzing and simulation

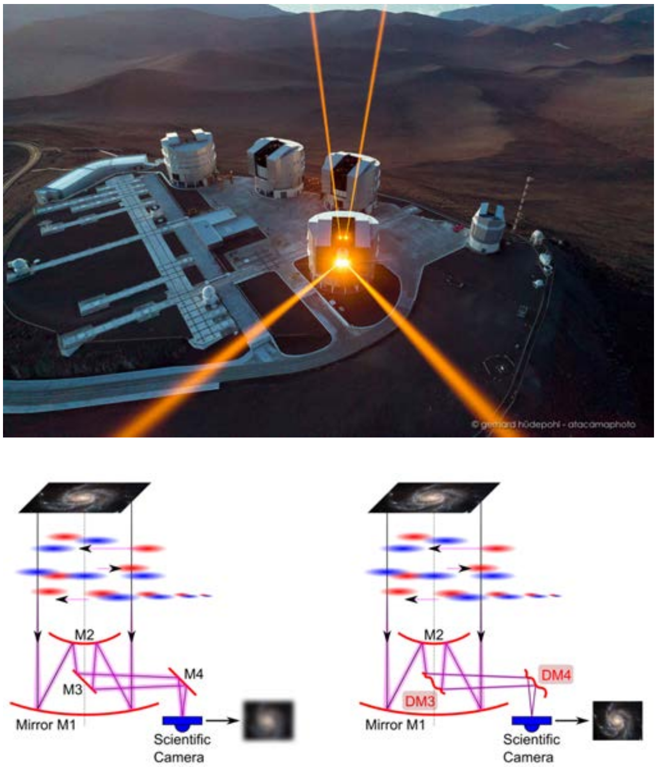


Figure 6. Top: ESO's Very Large Telescope in the Atacama desert (Chile) using Laser guide stars for a tomography of the atmosphere (Source: ESO). Bottom: Sketch of an AO system. The light from the scientific object is collected by the main mirror M1 and then corrected by the deformable mirrors DM1 and DM2.

of wavefront sensor devices, stable reconstruction of wavefronts from measured data [11], reconstruction of the turbulence distribution above the telescope, optimal control of the correction of the deformable mirrors and computation of the overall Point Spread function of the observation. All the computations involved in this process must be done in real time, as the atmosphere changes every few milliseconds. RICAM is a partner of the European consortia developing the instruments METIS and MICADO [5] for the Extremely Large Telescope of the European Southern Observatory (ESO), focusing on developing methods and software for its AO systems.

More recently, AO systems have also been used for image improvement in Ophthalmology and Microscopy. Within the aforementioned SFB, the Transfer Group works, in cooperation with the Medical University of Vienna, on the development of methods for AO systems in ophthalmic OCT systems that achieve an improved imaging quality of, e.g., the human retina, see Figure 7.

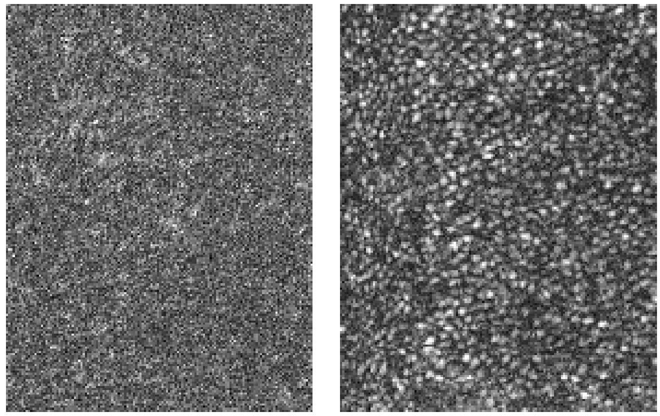


Figure 7. Image of the retina of the human eye: without AO (left) and with AO based on a pyramid wavefront sensor (right), image taken from [4].

4 Cooperations between groups

As visible from the research work presented above, RICAM is more than just the sum of its individual work groups; cooperations between the groups and an atmosphere of creativity and innovation are among the strengths of the institute. Another is the special mix of mathematical disciplines present at RICAM. Apart from the groups highlighted in the previous sections, there is a group working on computational methods for PDEs, which has been led by Ulrich Langer until this year, and which is now in a transition phase after which it will be taken over by Langer's successor, Herbert Egger. The group will focus on the development, analysis, and efficient realization of numerical methods for the simulation of coupled and multiscale phenomena in physics, engineering, and material and life sciences. Tight collaborations and synergies with several existing RICAM groups can be expected. In particular, this group is strongly connected to the "Geometry in Simulations" group, led by Bert Jüttler.

What is more, in contrast to several other large mathematics institutes, RICAM has a very strong and productive group working on symbolic computation, algebra, and combinatorics. A recent addition to the work groups is a team led by Luca Gerardo-Giorda focusing on applications to the life sciences. Further teams working on cryptography, simulation of crowded transport, and quasi-Monte Carlo methods round off the picture.

5 Special Semesters

RICAM *Special Semesters* attract world leaders in their respective research fields, and are the starting point for many international cooperations with RICAM scientists. Furthermore, they contribute significantly to the international visibility of the institute and to the

Figure 8. Cooperations among the different work groups is one of the strengths of the institute. Photo by ÖAW, Daniel Hinteramskogler.

exchange of the newest trends and developments. In general, at least one group leader, one other member of a group, and one member of the administrative staff are responsible for the organization of each Special Semester, frequently in cooperation with leading scientists from international research institutions. RICAM Special Semesters regularly attract around 200–300 guests. These may be either short term guests, who usually attend selected workshops, or long-term guests staying for longer periods during the Special Semester. The next Special Semester, "Tomography Across the Scales", will be held in the fall of 2022, subject to positive developments with respect to the current global pandemic.

Figure 9. Special Semesters are the most important events organized by RICAM. Photo by ÖAW, Daniel Hinteramskogler.

References

[1] B. Azmi, D. Kalise and K. Kunisch, Optimal feedback law recovery by gradient-augmented sparse polynomial regression. *Journal of Machine Learning Research*, **22**, 1–32 (2021)

[2] J. Berner, M. Dablander and P. Grohs, Numerically solving parametric families of high-dimensional Kolmogorov partial differential equations via deep learning. In *Advances in Neural Information Processing Systems*, volume 33, pages 16615–16627, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan and H. Lin, Curran Associates, Inc. (2020)

[3] J. Berner, P. Grohs and A. Jentzen, Analysis of the generalization error: empirical risk minimization over deep artificial neural networks overcomes the curse of dimensionality in the numerical approximation of Black–Scholes partial differential equations. *SIAM J. Math. Data Sci.* **2**, 631–657 (2020)

[4] E. F. Brunner, I. Shatokhina, M. F. Shirazi, W. Drexler, C. K. F. Hitzenberger, R. A. Leitgeb, R. Ramlau and M. Pircher, In-vivo demonstration of AO-OCT with a 3-sided pyramid wavefront sensor (Conference Presentation). In *Ophthalmic Technologies XXX*, volume 11218, edited by F. Manns, A. Ho and P. G. Söderberg, International Society for Optics and Photonics, SPIE (2020)

[5] R. Davies, J. Alves, R. Ramlau, R. Wagner, V. Hutterer, J. Niebsch et al., The MICADO first light imager for the ELT: overview, operation, simulation. In *Proceedings of the SPIE Astronomical Telescopes and Instrumentation Conference, Vol. 10702 (Austin, Texas, June 2018): Ground-based and Airborne Instrumentation for Astronomy VII*, edited by C. J. Evans, L. Simard and H. Takami, SPIE, 570–581 (2018)

[6] S. Dolgov, D. Kalise and K. K. Kunisch, Tensor decomposition methods for high-dimensional Hamilton–Jacobi–Bellman equations. *SIAM J. Sci. Comput.* **43**, A1625–A1650 (2021)

[7] A. Effland, E. Kobler, K. Kunisch and T. Pock, Variational networks: an optimal control approach to early stopping variational methods for image restoration. *J. Math. Imaging Vision* **62**, 396–416 (2020)

[8] H. W. Engl and R. Ramlau. Regularization of inverse problems. In *Encyclopedia of Applied and Computational Mathematics*, edited by B. Engquist, Springer, Heidelberg (2015)

[9] P. Grohs and L. Herrmann, Deep neural network approximation for high-dimensional parabolic Hamilton–Jacobi–Bellman equations. arXiv:2103.05744 (2021)

[10] P. Grohs, F. Hornung, A. Jentzen and P. Von Wurstemberger. A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black–Scholes partial differential equations. *Memoirs of the AMS*, to appear (2020)

[11] V. Hutterer, R. Ramlau and I. Shatokhina, Real-time adaptive optics with pyramid wavefront sensors: part I. A theoretical analysis of the pyramid sensor model. *Inverse Problems* **35**, 045007 (2019)

[12] J. A. Iglesias and G. Mercier, Influence of dimension on the convergence of level-sets in total variation regularization. *ESAIM Control Optim. Calc. Var.* **26**, Paper No. 52 (2020)

[13] K. Kunisch and D. Walter, Semiglobal optimal feedback stabilization of autonomous systems via deep neural network approximation. *ESAIM Control Optim. Calc. Var.* **27**, Paper No. 16 (2021)

[14] R. Ramlau and O. Scherzer (eds.), *The Radon transform. The First 100 Years and Beyond*. Radon Ser. Comput. Appl. Math. 22, De Gruyter, Berlin (2019)

[15] M. Scherbela, R. Reisenhofer, L. Gerard, P. Marquetand and P. Grohs, Solving the electronic Schrödinger equation with weight-sharing deep neural networks for multiple nuclear geometries. arXiv:2105.08351 (2021)

Philipp Grohs is professor at the University of Vienna and group leader at RICAM. His main research interests are in numerical analysis, signal processing, and mathematics of machine learning. He was awarded the ETH Latsis Prize and selected for an Alexander-von-Humboldt Professorship award.

philipp.grohs@ricam.oeaw.ac.at

Peter Kritzer is leading the RICAM work group “Multivariate Algorithms and Quasi-Monte Carlo Methods”. His research interests include high-dimensional numerical integration, function approximation, and information-based complexity. He received the Prize for Achievement in IBC in 2015.

peter.kritzer@ricam.oeaw.ac.at

Karl Kunisch is professor emeritus at the Karl-Franzens University Graz and group leader at RICAM. He is an applied mathematician whose main research interests lie in control and optimization in the context of partial differential equations. He was awarded an ERC advanced grant, and received the W. T. and Adalia Reid Prize.

karl.kunisch@ricam.oeaw.ac.at

Ronny Ramlau is professor for Industrial Mathematics at the Johannes Kepler University Linz and director of RICAM. His main research is on inverse problems and imaging, with a focus on astronomical and medical applications. Additionally, he is involved in several cooperations with industrial partners.

ronny.ramlau@ricam.oeaw.ac.at

Otmar Scherzer is professor at the University of Vienna and group leader at RICAM. He is working in applied mathematics, with a special focus on inverse problems and computational microscopic imaging. Furthermore, he is the editor-in-chief of the flagship journal “Inverse Problems”.

otmar.scherzer@univie.ac.at

The Bosnian Mathematical Society and the mathematical life in Bosnia and Herzegovina

Muharem Avdispahić

The Bosnian Mathematical Society grew out of the Mathematics Section of the Society of Mathematicians, Physicists and Astronomers of Bosnia and Herzegovina. The latter was founded in 1949 during the preparation process for the First Congress of Mathematicians and Physicists of the Federal People's Republic of Yugoslavia held in Slovenia that year.

As far as the earlier periods are concerned, one could say that the presence of university level educated teachers of mathematics in Bosnia and Herzegovina was closely related to the emergence of secular grammar schools in the period of Austro-Hungarian Monarchy, mainly between 1879 and 1899. (For a broader audience it might be of interest to learn that two Nobel Prize winners of Bosnian origin, Ivo Andrić (Literature, 1961) and Vladimir Prelog (Chemistry, 1975) attended the First Grammar School Sarajevo. Indeed, Ivo Andrić then enrolled in the mathematics and sciences programme at the University of Zagreb, but soon switched to humanities.)

A rather high level of school mathematics in Bosnia and Herzegovina was upheld during the period of the Kingdom of Yugoslavia; however, the number of such schools was still very limited; their teachers were mostly receiving their degrees from abroad.

The wide spread of the upper school system after WWII and the increased demand for qualified schoolteaching made it necessary to establish a first Teacher Training College in Sarajevo in 1946, followed by the creation of analogous institutions in Banja Luka, Mostar and Tuzla. Four-year university study programmes in mathematics and physics opened in 1950 at the Philosophical Faculty of the University of Sarajevo. Ten years later, the Faculty of Natural Sciences and Mathematics was established, with a Mathematics as one of its five departments. Soon after this, the non-educational sector began to express the need for qualified mathematicians as well.

As expected, the Society of Mathematicians and Physicists, its Mathematics Section in particular, concentrated its efforts during this phase on mathematics teaching, curricula design, organization of seminars for teachers and popularization of mathematics. In this period, it published a regular Bulletin containing papers related in principle to various topics in school mathematics.

One particular activity of the Society was to search for mathematically gifted students. The first competition for high school students at the level of Bosnia and Herzegovina was organized in 1959, as a part of the rich process through which participants were selected for the BMO (Balkan Mathematical Olympiad) and the IMO. This tradition has remained alive ever since, becoming enriched over time with regular participation at the JBMO, Mediterranean Mathematical Competition, European Girls' Mathematical Olympiad, etc.

According to the report on the history of research mathematics in Bosnia and Herzegovina presented by Academician Mahmud Bajraktarević at the meeting in the Academy of Sciences and Arts of Bosnia and Herzegovina in 1978, the first modern research paper ever published by a Bosnian mathematician was the research note by Vera Šnajder in *Comptes Rendus de l'Académie des Sciences*, Paris, T. 192 (1931), 1703–1706. Mahmud Bajraktarević himself was the first Bosnian to obtain a doctoral degree in mathematics, with the thesis *Sur certaines suites itérées* which he wrote and defended in Paris in 1953.

The opening of a postgraduate programme at the Department of Mathematics in Sarajevo in 1966 marked a new phase in the development of mathematics in the country.

By the time of the above-mentioned report, mathematicians from Bosnia and Herzegovina had published around 300 papers total, in journals covered by *Mathematical Reviews*, *Zentralblatt für Mathematik* or *Referativnij Zhurnal*. The most visible results concerned summability, Fourier analysis and functional equations, accompanied by emerging interests in roughly twenty five other disciplines of pure and applied mathematics and early computer science. Until 1973, Sarajevo was the only centre of mathematical research in Bosnia and Herzegovina. The situation gradually changed with the opening of universities in Banja Luka, Tuzla and Mostar (between 1975 and 1977) and with some research activities also occurring in other industrial centres such as Zenica.

The Academy of Sciences and Arts, the University of Sarajevo Department of Mathematics and the Mathematics Section of the Society of Mathematicians, Physicists and Astronomers joined efforts to cofound the scientific journal *Radovi Matematički*, which in 1985 replaced an earlier publication of the Academy, *Radovi Odjeljenja* *prirodnih i matematičkih nauka Bosne i Hercegovine*. The decision to found this journal reflected the growing need for an independent mathematical journal with an international editorial policy.

The cover page of the journal was recognizable by reproducing interesting details from the old mathematical manuscripts kept in local libraries and museums. Multiplication table on the cover of the very first issue stem from a unique fourteenth century transcript dealing with elements of algebra and geometry and being kept in the Ghazy Husrev Bey Library established in Sarajevo in 1537.

In addition to research papers, the second issue of each volume of *Radovi matematički* contained a Chronicle, which served as a useful source of information about the period that brought a substantial new quality to mathematical life in Bosnia and Herzegovina. The Chronicle reported annually on the research publications and research communications of Bosnian mathematicians, colloquia and seminars at mathematical departments in the country, research fellowships and visits, published books and lecture notes, and other academic news such as data on new Ph.D.s in mathematics and new job appointments. An increasing number of graduates also started undertaking doctoral studies at prestigious universities abroad, mostly in the U.S. The Society contributed a section on its own activities, often concerning competitions and summer schools for young mathematicians.

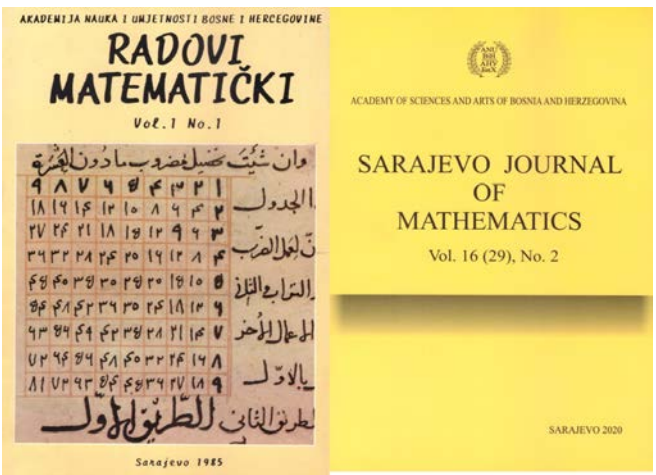


Figure 1. The co-edited journal and its successor

The year 1992 was a turning point with far-reaching implications. In the process of dissolution of former Yugoslavia, Bosnia and Herzegovina gained its independence in April of that year. War broke out on the very day the country achieved international recognition. Sarajevo, its capital, was put under a siege that lasted for the next 1426 days.

By keeping the academic activities alive and fighting to preserve civilized standards under the most difficult conditions encountered

Figure 2. 1995 photo of war-torn premises at the Faculty of Natural Sciences and Mathematics in Sarajevo

by any higher education institution in Europe in the second half of the 20th century, the University of Sarajevo received broad international respect. On the initiative of Professor Friedrich Hirzebruch, the Bosnian Mathematical Society was accepted into the European Mathematical Society at the Council meeting in Zurich in 1994.

Soon after the Dayton-Paris Peace Accord was signed in December 1995, the TEMPUS projects CME *Information Technology Development* and JEP *Developing the Faculty of Science Activities* coordinated by University of Sarajevo contributed to the academic reconstruction of studies in mathematics and natural sciences at all public universities in Bosnia and Herzegovina.

The graduate school at the Department of Mathematics in Sarajevo enrolled a new generation of students in 1997. Volume 8 of *Radovi matematički* received the timestamp 1992–1996. The BMS continued to co-edit this journal until 2005, when the role of *Radovi matematički* was taken over by its successor, the *Sarajevo Journal of Mathematics* published by the Academy.

Bosnia and Herzegovina became a full member of the IMU in 2002, represented through the BMS as an adhering organization.

BMS was also a founding member of the MASSEE (Mathematical Society of South Eastern Europe) in 2003. An initiative of a BMS representative at the annual MASSEE meeting in 2008 resulted in a three years TEMPUS JEP *Doctoral Studies in Mathematical Sciences in South East Europe* coordinated by University of Sarajevo and involving a consortium of 11 institutions from five West Balkan and three EU countries. The efforts invested towards harmonized networking programmes up to EHEA-ERA standards are well represented by six joint Ph.D. courses held during 2011 – the Year of Mathematics in South East Europe (cf. also two special issues of *Mathematica Balkanica* in 2010 and 2011 consisting of contributions from SEE Young Researchers Workshops).

Young aspiring mathematicians were given numerous opportunities to upgrade their capacities through summer schools within a long term DAAD project under the Stability Pact for South East Europe's *Center of Excellence for Applications of Mathematics*.

Other valuable contributors to mathematical life in Bosnia and Herzegovina are also the *Bulletin* and *Journal of the International Mathematical Virtual Institute*, successors to the *Bulletin of Society of Mathematicians Banja Luka*.

The country's constitution, adopted as Annex IV to the Dayton Peace Agreement, gives full responsibility in the area of education and research to lower level administrative units. There exist cantonal associations of mathematicians with seats in Sarajevo, Tuzla, Bihać, Travnik Zenica, two associations with seats in Mostar, on the territory of the entity Federation of Bosnia and Herzegovina, and two associations with seats in East Sarajevo and Banja Luka, on the territory of the entity Republika Srpska. The challenges facing any state level organization (with the exception of sports) are amply illustrated by the fact that the state level Ministry of Justice refuses to implement the decision of the Constitutional Court reached in the matter of the status of the Academy of Sciences and Arts of Bosnia and Herzegovina.

This being said, events such as the 3rd EU/US Summer School and Workshop on Automorphic Forms and Related Topics 2016, the BMS Mathematical Conference in 2018, the Sarajevo Stochastic Analysis Winter School in 2019, the International Conference on Fibonacci Numbers and their Applications in 2020 and the 26th International Conference on Difference Equations and Applications in 2021 demonstrate the level of research activities attained nowadays.

MathSciNet reports 85 papers published by Bosnian mathematicians in 2020 alone. The new areas of strength in the country appear to lie in number theory and difference equations, with noticeable advances in mathematical logic, associative rings and algebras, measure and integration theory, operator theory and general topology.

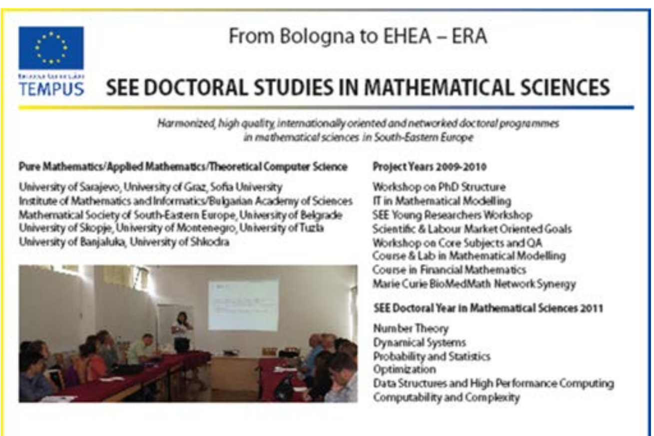


Figure 3. Flyer for SEE networking Ph.D. programme(s)

Figure 4. Major international conferences hosted over the last four years

Muharem Avdispahić is a professor at the Department of Mathematics of University of Sarajevo, with research interests in number theory and harmonic analysis. He was a Fulbright fellow at Princeton University (1989–1990) and DAAD scholar at Max Planck in Bonn (summer 1995). He served as a Dean of Science (1992–2000), vice-rector for academic affairs (1995–2000) and rector of the University of Sarajevo (2012–2016). Under his leadership, the BMS became a member of the EMS and an adhering organization of the IMU.

mavdispa@pmf.unsa.ba

Her Maths Story – Sharing stories of women in mathematics in all walks of life

Joana Grah, Tamara Grossmann and Julia Kroos

Her Maths Story (hermathsstory.eu) is a platform portraying stories of women in mathematics in all walks of life. Motivated by the lack of women role models, Her Maths Story strives to bring light to the variety of careers, non-linear paths and individual decision-making processes of women mathematicians in today's society – mathematicians like you and me. The purpose of this platform is to encourage and empower women to pursue studies and careers in maths, spark their curiosity and transmit enthusiasm for technical subjects.

Historically, representation of women in STEM and particularly mathematics has been quite poor. While women were barred from studying in most European countries as recently as 100 years ago, the continuing underrepresentation of women in mathematics in the past decades can also be partly attributed to a scarcity of role models. The aim of Her Maths Story is to bring the existing role models into focus and present them as a collection of different careers and life paths – because they do exist, these women in mathematics.

Being a young woman entering the world of academia or industry can be quite intimidating in the beginning, and all the more so if there is no other woman in the same professional circle to connect with, talk to or look up to. But even well-established women mathematicians in academia and industry share common challenges, be it the subtleties of regularly not being listened to, not being taken seriously, or juggling societal expectations with a successful career.

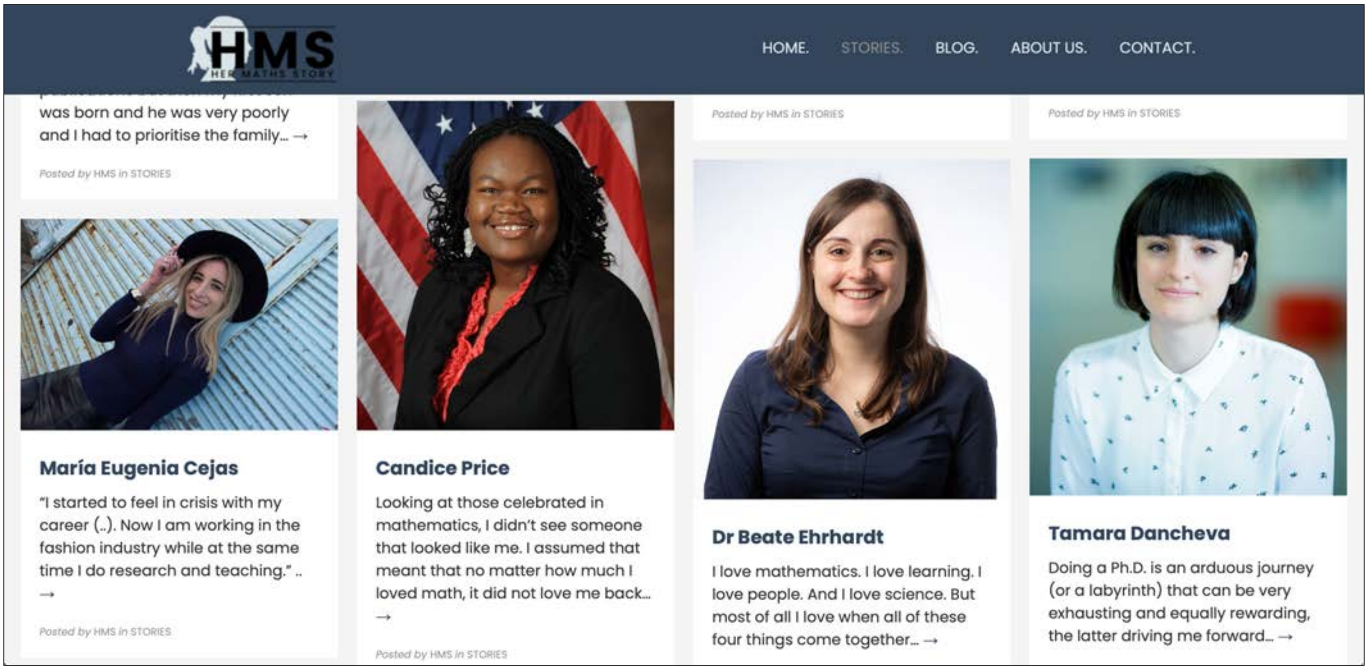
An increasing number of organisations, universities and companies are beginning to recognise that the number of women in their institutions is low, although there is a proven benefit in having gender-balanced teams[1]. Awareness of the importance of equality and diversity is increasing, and the path to change in the mathematical workforce can in fact benefit from the experience of women.

Therefore, the motivation for Her Maths Story is twofold: Giving women the space to share their experiences and learn from others, and offering insights into the life of a woman mathematician to the wider public.

The platform Her Maths Story includes both a website and various social media channels. Each week sees the publication of a blog entry or the personal story of a mathematician, both made accessible to a general audience. Stories consist of a text that can be read in 5–10 minutes describing the individual journey of a mathematician, and include a short CV with a photo. Blog entries, on the other hand, are longer texts that offer the opportunity to go into more detail on specific topics. We promote these stories and articles on our social media channels on Facebook, Instagram, LinkedIn and Twitter to reach as wide and as diverse a target group as possible.

We have published 33 stories since the beginning of this year. Of these, more than half are about women from Europe. The stories are as multifaceted as the career paths themselves. We have contributors from academia and industry, from undergraduates to PhD students to professors, from biostatisticians at AstraZeneca to data scientists at the energy company Enel to software developers in Airline IT. However, in our globalised world, cooperation with mathematicians from all countries is important, so we have made an effort not to focus solely on Europe, but to reach out to portray women from many different cultures and with highly varied life paths. For example, we featured the founder of the American organisation "Mathematically Gifted and Black", who is a professor of mathematics, and an Argentinian professor of mathematics who is also a fashion consultant. These stories deal not only with the common passion for mathematics, but also with personal motivations

[1] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi and T. W. Malone, Evidence for a collective intelligence factor in the performance of human groups. *Science* **330**, 686–688 (2010)



A screenshot of the stories section of Her Maths Story.

and societal and social hurdles, as well as the various influences in the individual careers.

The platform thrives on each and everyone's personal stories, experiences and advice. It is an ever-changing collection and there are still many more journeys to share and insights to gain.

The Her Maths Story team is always looking for new contributions, so if you want to share your journey or know someone with an inspiring, encouraging or unique maths story, contact them via hermathsstory@gmail.com.

Joana Grah is a scientific associate at the Heine Center for Artificial Intelligence and Data Science at the Heinrich Heine University in Düsseldorf, Germany, where she is responsible for education on the subjects of AI and data science as well as science communication and public outreach. After studying mathematics at the University of Münster, she received her PhD in applied mathematics from the University of Cambridge, UK, and worked as a postdoc at The Alan Turing Institute in London, UK, and the Graz University of Technology, Austria. She is passionate about encouraging women, equal opportunities and smashing the patriarchy. In her free time, she loves to dance.

joana.s.grah@gmail.com

Tamara Grossmann is a PhD student in applied mathematics at the University of Cambridge, UK. She is interested in image analysis and combining classical mathematical approaches with machine learning. Starting Her Maths Story has introduced her to many encouraging, fascinating and simply inspiring women mathematicians – a source of motivation for which she is very thankful.

t.g.grossmann@gmail.com

Julia Kroos is an applied mathematician at Bayer AG in Leverkusen, Germany, where she develops complex mathematical models for different questions in the pharmaceutical industry. After studying at the University of Münster, Germany, she received her PhD in applied mathematics from the University of the Basque Country in Bilbao, Spain and worked as a postdoc at the Basque Institute for Applied Mathematics in Bilbao, Spain and at the Johannes Kepler University in Linz, Austria. She is enthusiastic about the empowerment of women and about outdoor sports.

julia.m.kroos@gmail.com

Survey on Research in University Mathematics Education at ICME 14

Marianna Bosch, Reinhard Hochmuth, Oh Nam Kwon, Birgit Loch, Chris Rasmussen, Mike Thomas and María Trigueros

At the 14th International Congress on Mathematics Education, which took place in a hybrid mode in Shanghai from July, 11 to 18, 2021, the survey team on Research in University Mathematics Education (RUME) presented an overview of their work. As noted in the presentation, it is an exciting time for RUME. There are now several major conferences every year across the globe, as well as the fairly new *International Journal of Research in Undergraduate Mathematics Education*, now in its seventh year. The significant growth in the number of researchers focused on university mathematics education has led to the development of research groups and the consolidation of a diverse academic community; RUME is coming to age as a field of research that is beginning to coalesce and develop an identity.

To explore this identity, we surveyed 218 RUME scholars across the world, both well-established scholars and rising stars. We invited these scholars to respond to the following prompt:

> What do you see as the most significant advances, changes, and/or gaps in the field of research in university mathematics education? These advances, changes, or gaps might relate to theory, methodology, classroom practices, curricular changes, digital environments, purposes and roles of universities, social policies, preparation of university teachers, etc. Please elaborate on just *one or two advances, changes, or gaps* most relevant to your experience and expertise.

We received 119 responses. Our next step was to conduct a thematic analysis[1], which led to the identification of five areas in which there has been considerable progress (Theoretical Perspectives, Instructional Practices, Professional Development of University Teachers, Digital Technology, and Service-Courses in University Mathematics Education) and seven, non-disjoint areas in need of further research (Theories and Methods, Linking Research and Practice, Professional Development of University Teachers, Digital Technology, Curriculum, Higher Years, and Interdisciplinarity). We then conducted a literature review, guided by the identified themes.

We hope that this brief report offers those less familiar with RUME an overview of the progress to date and spurs interest in areas in which the reader might want to contribute to the knowledge base.

One of the field's major advances is that we now have a plethora of theoretical perspectives, and hence tensions among them can sharpen their constructs and methodologies and open the possibility of finding commonalities. This diversification has contributed to the development of new methods, research topics, and the development and research on theory-based teaching experiences. Recent years have seen the emergence of an interdisciplinary group of scholars interested in using a variety of approaches (logical, cognitive, historical, philosophical, etc.) to address questions which have always been of interest to RUME. Another theoretical advance that is of growing interest is the use of theories that enable insights into the interrelatedness of knowledge, identity, power, and social discourses [1]. While there is still much research that is needed here, we see this new direction as an important advance for the field of university mathematics education research.

The research of instructional practices at university level is another rapidly developing area of research. Much of the research on this topic relates to active or inquiry-based mathematics education [2, 6]. Given the myriad calls for instructional reform in university mathematics classrooms, researchers and educators have challenged conventional lecture-based instruction by conducting studies that have provided evidence for the positive effects of innovative student-centered instructions on students' cognitive and affective development. Active learning, broadly defined as classroom practices that engage students in activities such as reading, writing, discussing, or problem solving, that promote higher-order thinking, has repeatedly been shown to improve student success and to reduce the equity gap for women and underrepresented students [3, 7]. For example, a meta-analysis of 225 studies that compared student success in traditional lecture versus active learning in postsecondary science, engineering, and mathematics courses and found that average examination scores improved by about 6 % in active learning sections, and that students in classes with traditional lecturing were 1.5 times more likely to fail than were students in classes with active learning; further, the effectiveness of active learning was found across all

[1] Special thanks to Antonio Martinez and Talia LaTona-Tequida, graduate students at San Diego State University, for their help in this analysis.

class sizes [3]. On the other hand, RUME has only just begun to deeply explore the culture, experiences, and gendered/racialized interactions in these classes, and how those social factors may be obstructing the students' opportunities to learn [4].

Another area in need of research is on the learning and teaching of advanced mathematics. Historically, the work of Felix Klein is most relevant here. Core parts of his "Elementary Mathematics from a Higher Standpoint" [5] actually refer to mathematics that many of today's future teachers do not even get to know in the course of their academic studies. This applies, for example, to knowledge of Fourier analysis that goes beyond the basics, but especially also to knowledge of function theory, e.g. Riemann surfaces and value assignment theorems. Even when students hear about function theory, for example, they usually do not get as far as understanding what Felix Klein considered, more than a century ago, appropriate knowledge for prospective teachers. Klein considered this knowledge appropriate because it explains why, for example, certain elementary operations have to be restricted in certain ways for mathematical reasons (and not just for didactic reasons of reduction!), and related curricular decisions.

Also ripe for further investigation is the cooperation with mathematicians, engineers, economists, psychologists, etc. For many years, there have been many different kinds of cooperation, for example, agreements between faculties with regard to teaching. What does not seem to exist so far is, among other things, systematic research on these cooperations. What are the benefits of these? How do they take shape? How do they function? Possibilities, limits, etc.? Related to these cooperations is the relationship of mathematics to other sciences or the use of mathematics in other sciences. There are several places, such as philosophy or the history of science, in which such connections are examined and the question of what distinguishes mathematics itself and its respective role in other sciences is explored. Research on this is dependent on the respective ideological assumptions, and accordingly there are no unambiguous and generally accepted answers here. From the point of view of didactics, however, clarifications in this regard could certainly be regarded as desirable, since they would be of great help in answering the question of which goals, and how mathematicians and even more engineers, economists, psychologists, etc., are to be taught.

Last but not least are questions concerning mathematics itself. Mathematics, too, changes its inherent orientation, and to some extent its character, over time. New fields, such as Big Data and Data Science, are continually emerging. Correspondingly, there are new fields of application in other sciences, such as discrete mathematics in electrical engineering, numerical methods in psychology, etc. This leads directly to questions of what should be taught in service courses.

The video of the oral ICME presentation on the survey can be found here: https://drive.google.com/file/d/1LTBDl_KNZ371SL5ahvN2x_TA09PQQRlB/view?usp=sharing.

As we noted at the start of this brief overview, there is now much research-based wisdom, while at the same time there are exciting opportunities for new research. In particular, research mathematicians are welcome to join the systematic reflection and empirical investigation of university mathematics teaching.

References

[1] A. P. Adiredja and C. Andrews-Larson, Taking the sociopolitical turn in postsecondary mathematics education research. *Int. J. Res. Undergrad. Math. Educ.* **3**, 444–465 (2017)

[2] M. Artigue and M. Blomhøj, Conceptualizing inquiry-based education in mathematics. *ZDM–Math. Educ.* **45**, 797–810 (2013)

[3] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **111**, 8410–8415 (2014)

[4] E. Johnson, C. Andrews-Larson, K. Keene, K. Melhuish, R. Keller and N. Fortune, Inquiry and gender inequity in the undergraduate mathematics classroom. *J. Res. Math. Educ.* **51**, 504–516 (2020)

[5] F. Klein, *Elementary mathematics from an advanced standpoint: Arithmetic, algebra, analysis* (Vol. 1). Courier Corporation (2004)

[6] S. L. Laursen and C. Rasmussen, I on the prize: Inquiry approaches in undergraduate mathematics. *Int. J. Res. Undergrad. Math. Educ.* **5**, 129–146 (2019)

[7] E. J. Theobald et al., Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proc. Natl. Acad. Sci. USA* **117**, 6476–6483 (2020)

Marianna Bosch is an associate professor of didactic of mathematics at the University of Barcelona. She works in the development of the Anthropological Theory of the Didactic (ATD) created by the French researcher Yves Chevallard (Marseilles) and its contributions to the development of curricula and professional teachers. In the case of university mathematics education, she investigates the conditions needed to link study-based and inquiry-based proposals through a specific instructional format called *study and research paths*. She recently co-chaired the 3rd conference of the International Network for Didactics Research in University Mathematics (INDRUM).

marianna.bosch@ub.edu

Reinhard Hochmuth is professor of didactic of mathematics in the Department for Mathematics and Physics at the Leibniz University Hannover, Germany. His research focuses on mathematics for teacher education, the use of mathematics in empirical sciences and the development of inquiry-oriented approaches to the teaching and learning of advanced mathematics. The theoretical background of the research is the anthropological theory of didactics and a societally oriented subject-science. He is director of the Competence Centre for Higher Education Didactics in Mathematics (khdm) and founding member of the INDRUM Network coordination group.

hochmuth@idmp.uni-hannover.de

Oh Nam Kwon is professor in the Department of Mathematics Education at Seoul National University, Korea. Her research interests include curriculum development, teaching and learning of collegiate mathematics, and professional development of mathematics teachers. She was a member of 2015 OECD/PISA Mathematics Expert Group. She received a number of awards such as the Best Teaching Awards, Seoul National University in 2009, 2020 President's Commendation, and the 2021 Svend Pedersen Lecture Award. She serves currently as the co-chair of the 9th ICMI-East Asian Regional Conference on Mathematics Education.

onkwon@snu.ac.kr

Birgit Loch is Dean of Science, Agriculture, Business and Law at the University of New England, Australia. Her research interests include open educational resources, blended and online learning, students as partners and innovative use of educational technologies. She is a member of the International Steering Committee for Delta, the Southern Hemisphere Conference on Undergraduate Mathematics and Statistics Education, a national teaching award winner, Fellow of the Australian Mathematical Society and Principal Fellow of AdvanceHE (PFHEA).

Birgit.Loch@une.edu.au

Chris Rasmussen is professor in the Department of Mathematics and Statistics at San Diego State University. His research investigates inquiry-oriented approaches to the learning and teaching of differential equations as well as the process of departmental and institutional change to improve introductory mathematics courses. He is a founding editor for the *International Journal of Research in Undergraduate Mathematics Education* and was co-chair for the 13th International Congress on Mathematical Education Topic Study Group on Mathematics Education at the Tertiary Level.

crasmussen@sdsu.edu

Mike Thomas is an emeritus professor in the Mathematics Department, Auckland University, New Zealand. His research encompasses advanced mathematical thinking, semiotic analysis of gestures and representations, versatile thinking and the role of digital technology in the learning and teaching of mathematics at school and university. He was a founding editor of the *International Journal of Research in Undergraduate Mathematics Education* and led the 12th International Congress on Mathematical Education survey team on mathematical difficulties in the transition from school to university.

moj.thomas@auckland.ac.nz

María Trigueros is an associate professor in the Department of Mathematics at Instituto Tecnológico Autónomo de México. She works on the development and application of APOS theory. Through her research she has contributed to the study of the Schema notion in APOS Theory and to its use in investigating the learning and teaching of two-variable calculus and linear algebra. She is also interested in linking APOS to other theories and in developing and testing modeling tasks in the teaching of linear algebra. She now serves as chair of the 4th conference of the INDRUM-2022.

mtriguerosg@gmail.com

ERME column

regularly presented by Jason Cooper and Frode Rønning

In this issue, with a contribution by
Michèle Artigue

The Covid-19 pandemic: Challenging times for the mathematics education community

The Covid-19 pandemic started almost two years ago, turning our personal and professional lives upside down. As I wrote together with Ingrid Daubechies, former president of the International Mathematical Union, in the presentation of the plenary panel we coordinated at ICME-14 last July[1], this pandemic has brought mathematics to the forefront, in particular through the mathematical models used for understanding the course of the pandemic, and anticipating and weighting the possible consequences of different policy decisions. This has created a special responsibility for all those, mathematicians, mathematics educators and teachers, who play a role in helping students and the general public make sense of these models and their uses, and to understand more broadly what mathematical modelling is about, with all of its potential and its limitations. In his contribution to the panel, the Fields medallist Timothy Gowers stressed the importance of this mediating role accepted by many mathematicians beyond those directly involved in epidemiologic research, such as the South African mathematician Jean Lubuma, who also took part in the panel session.

Mathematics teachers and didacticians have endorsed this role, and they were certainly helped by the increasing importance given to modelling in mathematics curricula and educational research. However, for them, with the sudden transition to distance and online teaching, the pandemic represented first and foremost a disruption to the usual forms of teaching and interaction with their pupils and students. This disruption has dramatically exacerbated already existing educational inequalities, and Nelly León, a Venezuelan mathematics educator, especially developed this point during the panel session. But she also emphasized some positive dynamics generated by the pandemic: a radical evolution in the relationship with digital technologies; the production of a multitude of resources; new relationships between parents, students and teachers, and new solidarities. The responses to the call for papers launched by the journal *Educational Studies in Mathematics* in March 2020 also testify to the exceptional mobilization of teachers and researchers. The fourth contributor to the panel, David Wagner, who coordinated this editorial project with Man Ching Esther Chan and Christina Sabena, spoke of receiving 161 contributions from 36 countries; two special issues of the journal will soon be published (see [2] for a global view).

Teachers' testimonies and research studies show just how little we were prepared for this massive disruption. They allow us to identify research needs, for example the imperative need for research on online or hybrid teaching at all educational levels, not only at the levels of university and teacher education on which it has focused until now. Beyond that, there is a need for conceptual constructs allowing us to better approach the increasing complexity of human-digital artefacts relationships, as proposed by Borba [1]. This pandemic experience also prompts us to question our research priorities, to ask ourselves what place we desire to give to the question of educational inequalities and to the study of the new or underestimated expressions of these inequalities that the pandemic has brought to light, and to question how, beyond simply bringing understanding, didactic research can actually support action more effectively than it has done so far.

As has been repeatedly emphasized, the pandemic is not an isolated crisis. At a time when, thanks to vaccination, we seem more able to control the pandemic evolution, schools and universities have reopened, and people plan to meet again face-to-face at conferences such as the forthcoming CERME next February, we must not forget this. Other crises are looming, undoubtedly much more complex, dangerous and lasting ones. To understand this complexity and contribute to the public debate on the political decisions envisaged or taken, quality mathematics education for all is more necessary than ever. Beyond the epidemiological modelling highlighted by the pandemic, there is an increasing need for probabilistic and statistical education including risk issues [3]. This education should enable students to question the "formatting" role played by mathematics in our societies, as called for by critical mathematics education [4]. At a time when mathematics curricula are increasingly opening up to algorithmic and computational thinking, we cannot avoid thinking about how mathematical

[1] Plenary panel 3 entitled *Pandemic times: Challenges, responsibilities and roles for mathematics and mathematics education communities* (www.icme14.org)

algorithms shape our world, for better or for worse. To face these challenges, collaboration between the mathematics and didactic communities is more important than ever, and in the panel session, Ingrid and I expressed our hope that this crisis situation would strengthen synergies between communities.

References

[1] M. C. Borba, The future of mathematics education since COVID-19: Humans-with-media or humans-with-non-living-things. *Educ. Stud. Math.*, DOI 10.1007/s10649-021-10043-2 (2021)

[2] M. C. E. Chan, C. Sabena and D. Wagner, Mathematics education in a time of crisis—a viral pandemic. *Educ. Stud. Math.*, DOI 10.1007/s10649-021-10113-5 (2021)

[3] N. Radakovic and E. Chernoff, Risk education. In *Encyclopedia of Mathematics Education*, edited by S. Lerman. 2nd ed., Springer, Cham, 754–756 (2020)

[4] O. Skovsmose, Mathematics and crises. *Educ. Stud. Math.*, DOI 10.1007/s10649-021-10037-0 (2021)

Michèle Artigue is emeritus professor at the Université de Paris, and researcher in mathematics education. She has been vice-president then president of ICMI, the international Commission on Mathematical Instruction.

michele.artigue@univ-paris-diderot.fr

The “Jahrbuch über die Fortschritte der Mathematik” as a part of zbMATH Open

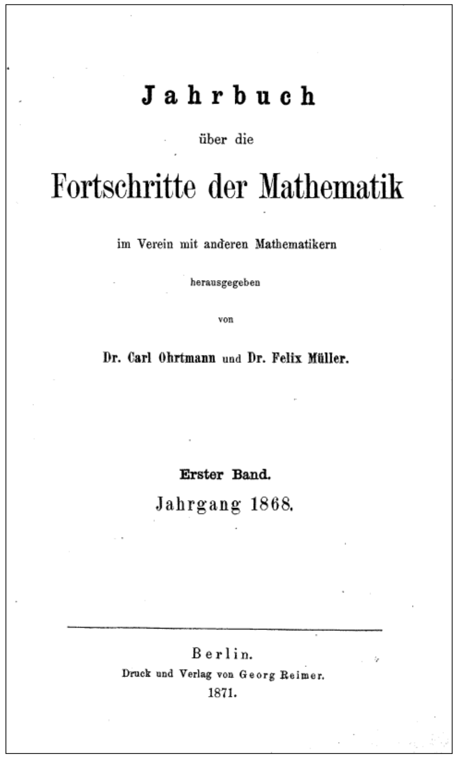
Olaf Teschke

2021 marks the 150th anniversary of the publication of the first volume of the “Jahrbuch über die Fortschritte der Mathematik” (JFM), as well as the transition of zbMATH to an open service 90 years after its founding. Initially digitised in the ERAM Project 1998–2004, JFM data have benefitted significantly from the subsequent integration into zbMATH, and are now available in a much enhanced form. We describe the improvements of the digital JFM version during the last decade, which are now available both as Open Access Database and Open Data.

1 Pieces of JFM history

According to most sources, Carl Ohrtmann and Felix Müller founded JFM at the end of 1868 with the aim of collecting, indexing, and reviewing the global mathematical literature in annual volumes. It followed the ideas of already established review journals like *Pharmaceutisches Central-Blatt* (later *Chemisches Zentralblatt*, founded in 1830), or *Fortschritte der Physik* (founded in 1847). The first JFM volume, covering 838 publications from 1868, was published in February 1871 – even in the beginning, the ideal of a complete and classified collection could only be achieved by a considerable delay.

Similar initiatives were at this time Boncompagni’s *Bullettino di bibliografia e di storia delle scienze matematiche e fisiche* (1868), Darboux’s *Bulletin des sciences mathématiques et astronomique* (1870) and the Dutch *Revue semestrielle des publications mathématiques* (1896), but JFM prevailed due to its broad community of expert reviewers, supported by the leading role of German mathematicians at that time. Until WWI, JFM defined the standard to judge contemporaneous research, and shaped both scope and classification of mathematics. The setback of the war, however, could never be regained: the admired comprehensiveness and the detailed knowledge of the organisation became a liability, leading to growing delays of up to seven years. Moreover, German was no longer the primary language of math publications, calling into question the existence of German-language-only reviews. In this critical situation, additional resources for JFM provided by the Prussian Academy since 1930 turned out to be double-edged: though they helped with the catching up, they forced JFM to follow its rather conservative politics, pursuing the ultimately infeasible aim of restoring JFM’s pre-WWI status in a changed environment. This inflexibility intensified the decline of JFM, which ultimately deteriorated due to the outcomes of Nazi politics (the details of decline and fall of JFM are investigated in [7]). The *Zentralblatt für Mathematik und ihre Grenzgebiete* (ZfM, now zbMATH Open), founded in 1931 by Otto Neugebauer with the focus on timeliness instead of systematic comprehensive annual volumes, which assembled a new generation of a global community providing multilingual reviews, soon replaced JFM as the primary source of current research information, with the average review becoming available only after an impressive 0.58 years. After Neugebauer’s emigration to the


Jahrbuch

über die

Fortschritte der Mathematik

im Verein mit anderen Mathematikern

herausgegeben

von

Dr. Carl Ohrtmann und Dr. Felix Müller.

Erster Band.

Jahrgang 1868.

Berlin.

Druck und Verlag von Georg Reimer.

1871.

Figure 1. The first page of the first JFM volume

U.S., the second journal he founded, *Mathematical Reviews* (MR) took the lead following similar principles, leaving JFM only the role of a well-organised collection of historical interest. Consequently, attempts to revive JFM after WWII failed, resulting in volume 67 of 1942 being the last one ever published – though former JFM staff members played a crucial role in the successful resurrection of *Zentralblatt* after the war.

2 JFM digitisation

Still after its end, JFM was used as a standard reference by MR and ZfM for earlier literature. Hence, at the dawn of math digitisation, it was natural to include information from JFM covering publications from 1868 through 1942. Initially suggested by then-EiCs of MR and ZfM Keith Dennis and Bernd Wegner as a joint National Science Foundation (NSF) and Deutsche Forschungsgemeinschaft (DFG) project in 1997, JFM was digitised as a part of the Electronic Research Archive for Mathematics (ERAM) project (1998–2002) funded by DFG and conducted by TU Berlin, Göttingen State and University Library (SUB) Göttingen and FIZ Karlsruhe. Along with the digitisation of important historical mathematical sources at SUB, JFM data were seen as a building block of the World Digital Mathematics Library initiative championed by the International Mathematics Union (IMU). Although clearly German funds alone could not be sufficient to achieve all these desirable objectives, significant results were obtained: all volumes of JFM were transformed into LaTeX and made freely available in a database, allowing for a search distinguishing author and reviewer names, titles, and review texts. At that time, it was likely the largest free LaTeX transcription project to have ever existed, and beyond its actual output, it provided some insights into feasible procedures. The combination of OCR techniques and manual transcription of formulae proved to be a manageable approach, although there was a relatively significant spread in the error rates of the various companies given the initial samples. Fortunately, DFG funds made it possible to choose the companies with the lowest error rates for the remaining parts; though some of the numerous errors coming from the initial low-cost alternatives can still be found in the data, and represent challenges both to the reader and for derived information like author disambiguation. As a collection, the LaTeX transcription of the Jahrbuch also provides a good gold data for next-level digitisation approaches like those outlined in [2].

3 Integration of JFM into zbMATH

The results of the digitisation project were made freely available after its ending in the JFM database, but a lot of work remained to be done (apart from a good compilation of earlier publications on JFM, one can find in [4] an account of the missed objectives within the ERAM project). Among the desirable features which were goals of the project that were not attained were: the standardisation of journals, author disambiguation, and interlinking with full-texts. Since these projects were being undertaken at the same time on a broader scale in the zbMATH database, JFM stakeholders allowed for the integration of the project data into zbMATH under the condition of providing resources for their enhancement. About a decade ago, this column [3] described the status at the start of the integration process. Since then, vast improvements have been achieved. The bibliographical sources, initially just a string, which could vary for a single journal from, say, Clebsch Annalen to Klein Annalen to Mathematische Annalen, with different abbreviations and mixed Arabic and Roman volume numbers. They have now not only been standardised, but have evolved into a full-scale journal database facilitating faceted searches including granular information such as titles, publishers, ISSN, main subjects, time periods, countries, languages or Open Access information and issue-level browsing. Thanks to these assignments, automated generation of full-text links is now possible. While the first JFM database did not contain a single DOI, these are now available for more than 20 % of the 223,276 JFM documents, along with 19,015 links to free EuDML entries, 8,587 to Gallica, and many to a number of other free digital libraries.

Author disambiguation has been particularly challenging for JFM entries: first names were usually abbreviated or completely missing, and typos from the chunks digitised with lower quality complicated the situation further. Approaches which work well for modern publications such as analysis of coauthor and reference networks fail due to the lack of reference data and the fact that at that time most publications were single-authored. Thus, progress depended on purely human checking of authorship assignments; fortunately this has now been done for a large part of the JFM data. As a result, JFM authorship data now contribute to comprehensive author profiles for mathematicians of three centuries. The situation is, however, less ideal for reviewer information: reviewer signatures mostly lack first names, which makes their identification very complicated[1]. A precise disambiguation and integration into person profiles remains to be done.

Likewise, the integration of JFM into zbMATH has led to its extension to a citation database. Not only could citation data be added for more than 6,000 JFM documents, but references from later publications to JFM could be matched within the integrated corpus. The resulting dataset provided the opportunity for unique analysis of long-term citation behaviour in mathematics, part of which was reported in this column [1].

[1] This is also likely the reason why the number of reviews for prominent mathematicians in earlier publications turn out to be frequently incorrect.

4 JFM as part of zbMATH Open

One major drawback, however, was that these improvements – facilitated by zbMATH resources – were only available within the commercial zbMATH database, and hence only partially accessible. Although limited to subscribers (except for the reduced free results), the improved functions led to a gradual shift in usage from the free project version of JFM to the zbMATH subset. In 2020, JFM documents were > 20 times more often accessed in their zbMATH version compared to the old JFM database. The transition of zbMATH to the zbMATH Open service at the beginning of 2021 [5] resolved this dilemma: the enhanced JFM data within zbMATH now became completely free, and also provides all the information of the old JFM database as a subset[2]. In fact, this transition achieved even more: the EMS, named by the project partners as the holder of the JFM data, agreed to make it available under a CC-BY-SA 4.0 data. In particular, this dataset is completely available via the zbMATH Open API introduced earlier in this column [6]. However, there still remains a great deal to do. As mentioned above, reviewer disambiguation is lacking. There is a small overlap for the years of 1931–1945 when JFM and zbMATH were published in parallel; corresponding items should ideally be merged. Furthermore, many comments from the mathematics history community have been collected during the project; they are not in publishable condition, however, due to their heterogeneous nature. Perhaps most importantly, the digitisation of historical full-texts may benefit greatly from recent technological developments, and their integration would facilitate additional functions like full-text or formula search for mathematical content over a period of 150 years.

[2] This also allowed to discontinue the technically outdated old JFM interface. Note that search results in zbMATH Open can be filtered to JFM documents by adding "dt:JFM" to a query.

References

[1] A. Bannister and O. Teschke, An update on time lag in mathematical references, preprint relevance, and subject specifics. *Eur. Math. Soc. Newsl.* **106**, 41–43 (2017)

[2] M. Beck, I. Beckenbach, T. Hartmann, M. Schubotz and O. Teschke, Transforming scanned zbMATH volumes to LaTeX: planning the next level digitisation. *Eur. Math. Soc. Newsl.* **117**, 49–52 (2020)

[3] S. Göbel, Jahrbuch über die Fortschritte der Mathematik and Zentralblatt MATH – reporting on more than 140 years of mathematics. *Eur. Math. Soc. Newsl.* **73**, 45–46 (2009)

[4] S. Göbel, W. Sperber and B. Wegner, 150 Jahre—Ein Rückblick auf das Jahrbuch über die Fortschritte der Mathematik. *Mitt. Dtsch. Math.-Ver.* **28**, 95–107 (2020)

[5] K. Hulek and O. Teschke, The transition of zbMATH towards an open information platform for mathematics. *Eur. Math. Soc. Newsl.* **116** 44–47 (2020)

[6] M. Schubotz and O. Teschke, zbMATH Open: Towards standardized machine interfaces to expose bibliographic metadata. *EMS Magazine* **119**, 50–53 (2021)

[7] R. Siegmund-Schultze, *Mathematische Berichterstattung in Hitlerdeutschland*. Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte der Mathematik [Studies on the Scientific, Social and Educational History of Mathematics] 9, Vandenhoeck & Ruprecht, Göttingen (1993)

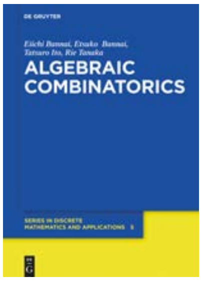
Olaf Teschke is managing editor of zbMATH Open and vice-chair of the EMS Committee on publications and electronic dissemination.

olaf.teschke@fiz-karlsruhe.de

Book reviews

Algebraic Combinatorics by Eiichi Bannai, Etsuko Bannai, Tatsuro Ito and Rie Tanaka

Reviewed by Tullio Ceccherini-Silberstein



The book under review is a most welcome, completely revised, widely expanded, and updated version of the celebrated and most influential book *Algebraic Combinatorics I; Association Schemes* [2] by Eiichi Bannai and Tatsuro Ito, published in 1984.

According to these authors, Algebraic Combinatorics is "the approach to combinatorics which was formulated in P. Delsarte's monumental thesis [3] in 1973, enabling us to look at a wide range of combinatorial problems from a unified viewpoint". As stated immediately thereafter, "the origins of this approach can be found in the previous work on character theory of finite groups and permutation group theory by I. Schur, G. F. Frobenius, and W. Burnside as well as in that on experimental designs and association schemes by R. C. Bose". This unified approach intertwines algebraic aspects of graph theory, coding theory, design theory, and finite geometries, with methods of Schur rings and of intersection matrices in permutation group theory. All this said, one may define Algebraic Combinatorics as "a group theory without groups"! To justify this statement and, possibly, give the reader a taste of the mathematics involved, we limit ourselves to present the definition of the central and unifying concept of the theory, namely of an association scheme, together with a couple of examples.

An *association scheme* is a pair $\mathfrak{X} = (X, \mathcal{R})$ where X is a finite set and $\mathcal{R} = (R_j)_{j=0}^N$ is a partition of $X \times X$, where the sets R_j, called the *associate classes*, satisfy the following properties:

(1) $R_0 = \{(x,x) : x \in X\}$ is the diagonal;
(2) for each $j = 1, 2, \ldots, N$, there exists $1 \leq j^* \leq N$ such that $R_{j^*} = \{(y,x) \in X \times X : (x,y) \in R_j\}$;
(3) there exist nonnegative integers $p_{i,j}^k$, $i, j, k = 0, 1, \ldots, N$, called the *parameters*, such that $|\{z \in X : (x,z) \in R_i,\ (z,y) \in R_j\}| = p_{i,j}^k$ for all $(x,y) \in R_k$.

An association scheme $\mathfrak{X}$ is said to be *commutative* (resp. *symmetric*) provided that $p_{i,j}^k = p_{j,i}^k$ (resp. $R_j = R_{j^*}$) for all $0 \leq i, j, k \leq N$. Note that symmetry implies commutativity. The matrices $A_j = (A_j(x,y))_{x,y \in X}$, $j = 0, 1, \ldots, N$, defined by setting

$$A_j(x,y) \coloneqq \begin{cases} 1 & \text{if } (x,y) \in R_j, \\ 0 & \text{otherwise,} \end{cases}$$

generate a subalgebra $\mathcal{A} \subseteq \mathrm{End}(\mathbb{C}[X])$, called the *adjacency algebra* (or *Bose–Mesner algebra* when it is commutative) associated with $\mathfrak{X}$.

Let us give a few examples. Let G be a finite group. Then setting $R_g \coloneqq \{(x,y) \in G \times G : x^{-1}y = g\}$ for all $g \in G$ yields an association scheme whose associated adjacency algebra is isomorphic to the group algebra $\mathbb{C}[G]$ of G. Also, if $K \leq G$ is a subgroup, then G acts transitively on the coset space $X = G/K$ yielding an association scheme $\mathfrak{X}$ whose associated classes are the G-orbits on $X \times X$ and whose adjacency algebra is isomorphic to the subalgebra of bi-K-invariant functions in $\mathbb{C}[G]$; moreover, $\mathfrak{X}$ is commutative exactly if (G,K) is a *Gelfand pair* (this is an important notion in Harmonic Analysis: it was used by Diaconis in his applications of Representation Theory to Probability and Statistics [4]). Finally, a finite regular undirected graph $\mathcal{G} = (X, E)$ with no loops is called *distance-regular* if there exist two sequences of constants, called the *parameters*, $(b_0, b_1, \ldots, b_N)$ and $(c_0, c_1, \ldots, c_N)$, where N is the diameter of $\mathcal{G}$, such that, for any pair of vertices $x, y \in X$ with graph distance $d(x,y) = i$ one has

$$|\{z \in X : d(x,z) = 1,\ d(y,z) = i+1\}| = b_i,$$
$$|\{z \in X : d(x,z) = 1,\ d(y,z) = i-1\}| = c_i$$

for all $i = 0, 1, \ldots, N$. Setting $R_i \coloneqq \{(x,y) \in X \times X : d(x,y) = i\}$, for $i = 0, 1, \ldots, N$, yields a symmetric association scheme. The corresponding Bose–Mesner algebra is singly-generated by A_1; in fact, for every $i = 0, 1, \ldots, N$ there exists a polynomial $p_i \in \mathbb{R}[t]$ of degree i such that $A_i = p_i(A_1)$. This is a prototype of a so-called *P-polynomial scheme*.

The history of this monograph is, briefly, as follows. The original project of "Algebraic Combinatorics" as the first comprehensive and foundational treatment of the theory, consisted of two parts:

the first one [2], on "Association Schemes" – lecture notes based on graduate courses given by Eiichi Bannai during his professorship at The Ohio State University (1978–1982) and arranged in collaboration with Tatsuro Ito – included: (1) Representations of finite groups, (2) Association schemes, and (3) Distance-regular graphs, and P- and Q-polynomial association schemes. A second part, on "Delsarte Theory, Codes and Designs", should have been published a couple of years later, but – "as the developments seemed not sufficient to complete a book and, at the same time, the range of mathematical objects they were interested in had expanded too widely to be handled" – it has never come to light. However, in 1999 the first two named authors published (in Japanese) *Algebraic Combinatorics on Spheres* [1] which was not translated into English, as the original plan to write the sequel to [2] was still alive.

The present book under review is the English translation, with the cooperation of Rie Tanaka – herself a mathematician working on association schemes – of the book *Introduction to Algebraic Combinatorics* by the first three named authors, published in Japanese in 2016. As we have mentioned at the beginning, it is not a sequel to [2] but, rather, a completely revised, widely expanded, and updated version of it, serving – this is the clear intent of the authors – as a "preparation for a second part to be hopefully accomplished by the younger generations of algebraic combinatorialists" coming from the flourishing school the authors have initiated in the US, in Japan, and in China during their life carriers (according to the Mathematics Genealogy Project, just Eiichi Bannai has 31 students and 60 descendants).

The contents of the book, according to its subdivision into chapters, are as follows:

(1) Classical design theory and classical coding theory (including an introduction to graph theory);
(2) Association schemes, Bose–Mesner algebras, and Terwilliger algebras;
(3) Codes and designs in association schemes (Delsarte theory on association schemes);
(4) Codes and designs in association schemes (continued);
(5) Algebraic combinatorics on spheres and general remarks on algebraic combinatorics;
(6) P- and Q-polynomial schemes.

The writing is very elegant, both in the style of the authors and in the graphic design of De Gruyter, the exposition is crystal clear – the proofs are carefully detailed, and plenty of worked-out examples, often described by beautiful pictures, friendly turn the reader familiar with the concepts gradually introduced – reflecting at the same time the deepest and masterful knowledge of the subject by the authors as well as their long didactic experience. For this reason, the book may be used as an excellent text for advanced undergraduate and graduate courses on Algebraic Combinatorics, and, at the same time, may also serve as a most precious reference for the more advanced and mature mathematician. The authors clearly enjoyed writing it: I am sure that all readers will enjoy reading it as well.

References

[1] E. Bannai and E. Bannai, *Algebraic Combinatorics on Spheres* (Japanese). Springer, Tokyo (1999)

[2] E. Bannai and T. Ito, *Algebraic combinatorics. I*. The Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA (1984)

[3] P. Delsarte, An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.* 10, Ann Arbor (1973)

[4] P. Diaconis, *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series 11, Institute of Mathematical Statistics, Hayward, CA (1988)

Eiichi Bannai, Etsuko Bannai, Tatsuro Ito and Rie Tanaka, *Algebraic Combinatorics.* De Gruyter Series in Discrete Mathematics and Applications 5, De Gruyter, 2021, 444 pages, Hardback ISBN 978-3-1106-2763-3, eBook ISBN 978-3-1106-3025-1.

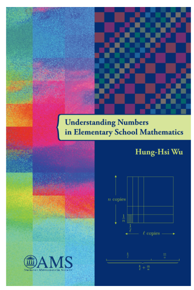
Tullio Ceccherini-Silberstein is a professor of mathematical analysis at the University of Sannio (Italy). His interests lie within harmonic and functional analysis, geometric and combinatorial group theory, dynamical systems and ergodic theory, and theoretical computer science.

tullio.cs@sbai.uniroma1.it

DOI 10.4171/MAG-59

Understanding Numbers in Elementary School Mathematics
by Hung-Hsi Wu

Reviewed by António de Bivar Weinholtz



For over half a century, the teaching of Mathematics in elementary, middle and high school has been the subject of intense debate among educators, mathematicians and politicians around the world. Dramatic changes have been introduced throughout the past decades in maths curricula and teaching methods in many countries and, taking into account the consequences of these policies, it has become clearer and clearer how some of their main features may have had a strongly negative impact on the ability of students to learn real mathematics, exactly as predicted by certain critics. The fact that this is still non-consensual, mainly among math educators and politicians, and that in some countries we can still witness successive contradictory movements in the area of math education, shows just how important it is to have at our disposal works of unquestionable quality devoted to that beautiful part of Mathematics that can and should be taught to students in the pre-university grades.

Prof. Hung-Hsi Wu has been one of the main members of the mathematical community to devote years of professional activity to the improvement of the teaching of pre-university mathematics. He was the coordinator and one of the main authors of the Common Core Mathematics Standard, a milestone in the rebirth of a sound K-12 curriculum for American schools, that has served as an inspiration to similar movements in many other countries. For the past decade, he has committed himself to write a series of volumes covering this curriculum, starting with the present one, which covers the most substantial part of the K-6 curriculum, namely numbers and operations.

The book is written for elementary school teachers, as a fundamental instrument for their mathematical education both during pre-service years and for their professional development. It also aims to provide a much needed resource for authors of textbooks. As the author points out, it was written after more than ten years of experimentation, in an effort to teach mathematics to elementary and middle school teachers. As it is clear that teachers should be able to go beyond what they have to teach in their understanding of school mathematics, the book also contains some topics that could be considered more appropriate for grades 7 or 8.

The book is written with the assumption, whose validity the author does an excellent job of explaining to the reader, that school mathematics is not a set of trivial topics that could be served to students with some degree of carelessness regarding the systematic and comprehensive approach that any mathematical theory requires, under the illusion that it is enough to comply with some more or less widely accepted, although rather arguable, pedagogical principles. On the contrary, to cite the author:

> "If we want a coherent curriculum and a coherent progression of mathematics learning, we must have at least one default model of a logical, coherent presentation of school mathematics which respects students' learning trajectory. It is unfortunately the case that, for a long time, such a presentation has not been readily available. The mathematical community has been derelict in meeting this particular social obligation."

With these principles in mind, it is not difficult to guess that the reading of this volume can be a delight to anyone with the ability to appreciate the beauty of the use of human reasoning in our quest to understand the world around us. The set of its potential readers should thus surely not be restricted to those for whom it was primarily intended, but should include anyone with the basic capacity and will to make the necessary effort required here, as for any other really worthwhile enterprise.

In principle, reading this book requires no previous mathematical knowledge, as should be clear by the fact that the first section of Chapter 1 of Part 1 is entitled "How to Count". From there on, the author uses precise definitions and logical reasoning to treat all subsequent topics; but this means that while one can find everything one needs to follow these developments inside the book, the successive steps do require a serious effort. In practice, as the author points out, "it can be too much of a challenge if you are unfamiliar with the procedural aspect of elementary school mathematics".

The precise content of the book has been carefully chosen to cover all the required topics while simultaneously allowing for a systematic mathematical development suitable as a background to organize the teaching of numbers and operations to K-6 students. Although one could devise approaches that differ in some details, this is not a field where too much can be left to the imagination of too many members of the educational community. A deep knowledge of mathematics is required, but also a deep respect for what past generations can teach us on these subjects, since in the case of basic mathematics one can benefit from literally hundreds, and in some cases thousands of years of a successful chain of transmission of knowledge from generation to generation. Thus, not only does this work fill a long standing gap in the school mathematics literature, but it does so in a basically unavoidable way.

After a first part on the introduction of whole numbers, including operations and algorithms, one finds perhaps the core of the book and of the whole of the K-6 curriculum, which is the second part, devoted to fractions. Generations of students have been deprived of the possibility of attaining a minimal understanding of this topic by the mistreatment of the subject in schools around the world; it is sad to verify that even where some efforts have already been made to correct this situation, there have been recent educational policies reversing these corrections. The author not only explains what lies behind these mistreatments, but sets out to establish a sound and detailed basis for a correct way to teach this central subject throughout the years of elementary school. Three more parts follow, on rational (relative) numbers, some elementary beautiful and basic topics of number theory and decimal expansions. On each topic the author provides the reader with numerous illuminating activities whose solutions can be found online, and an excellent choice of a wide range of exercises.

Hung-Hsi Wu, *Understanding Numbers in Elementary School Mathematics.* American Mathematical Society, 2011, 551 pages, Hardback ISBN 978-0-8218-5260-6, eBook ISBN 978-1-4704-1210-4.

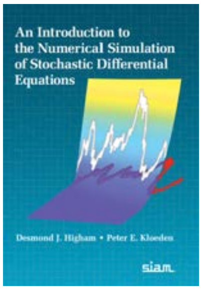
António de Bivar Weinholtz is a retired associate professor of Mathematics of the University of Lisbon Faculty of Science, where he taught from 1975 to 2009. He was a member of the scientific coordination committee of the new curricula of Mathematics for all the Portuguese pre-university grades (2012–2014).

abivar@sapo.pt

DOI 10.4171/MAG-56

An Introduction to the Numerical Simulation of Stochastic Differential Equations by Desmond J. Higham and Peter E. Kloeden

Reviewed by David Cohen



The aim of the book under review is to "provide a lively, accessible introduction to the numerical solution of stochastic differential equations (SDEs)" (taken from the first sentence of the preface). Knowing the scientific expertise of the two authors of this book as well as their communication skills, it was clear to me from the start that this aim would be reached. I hope that my review will illustrate this fact.

SDEs and their numerical simulations are indispensable tools in a multitude of disciplines. They have been used in various application areas such as chemical kinetics, engineering, epidemic modeling, financial mathematics, neuroscience, physics, social sciences, etc. To understand these models and simulate solutions to SDEs efficiently requires advanced knowledge in probability theory, stochastic analysis, as well as numerical analysis. However, to understand the main ideas and concepts of SDEs and their numerical simulations using this book, it is enough to start with a good knowledge of algebra, calculus and some familiarity with numerical analysis and probability theory.

By reading this book (and doing the exercises!), the reader will develop an intuitive feeling for the main ideas around SDEs, learn to implement first numerical methods to efficiently simulate SDEs in various applications, and get ideas of the proofs of some of the fundamental theorems in numerics for SDEs. All this is accomplished thanks to a constant focus on the essential (avoiding technicalities), an engaging and pedagogical way of presenting the necessary material, and appropriate use of well-chosen (computational) examples and inviting (theoretical and computational) exercises.

The book under review is not a rigorous text in mathematics, and the authors acknowledge this in many places. The interested reader is often referred to the more advanced and technical literature on the subject. For more rigorous and advanced books on the numerical treatment of SDEs, one could consult for instance: *Numerical Solution of Stochastic Differential Equations* by Kloeden, Platen, or *Stochastic Numerics for Mathematical Physics* by Milstein, Tretyakov, or *Numerical solution of SDE through computer experiments* by Kloeden, Platen, Schurz.

The present book contains 20 chapters, ordered in a very didactic manner in increasing order of difficulty (at least according to me). Each chapter starts with a compact outline in bullet points, a small picture reflecting the content of the chapter, and a clear and concise statement of motivation, and each one is illustrated with several accessible (computational) examples. Every chapter ends with notes and references (e.g. for further advanced reading), a list of both theoretical and computational exercises (with solutions available at the publisher's website), and most importantly, some funny quotes! At the end, the book provides a list of symbols, an index and an extensive bibliography.

Let me now give a brief description of the content of each chapter. The first chapter presents (discrete and continuous) random variables and related topics. Chapter 2 introduces basic computational concepts to simulate random variables. The third chapter contains the heart of the theory of SDEs, namely Brownian motion and its main properties. Chapter 4 deals with (mostly Itô) stochastic integrals. These four chapters provide all the tools required to define scalar SDEs, which are duly presented in Chapter 5, along with many examples from the domains of application. One particularly important tool, namely the Itô formula, is described in Chapter 6. The goal of Chapter 7 is to give a brief overview of the Stratonovich form of an SDE. Chapter 8 deals with the simplest and most used numerical scheme for SDEs: the Euler–Maruyama scheme. Chapters 9 and 10 give a sketch of the proofs of weak and strong convergence of the Euler–Maruyama scheme. From this point on, my impression is that the chapters begin touching on more advanced topics. Chapter 11 investigates the mean-square and asymptotic stability of the stochastic θ-method. The aim of Chapter 12 is to use numerical methods in a Monte Carlo setting to compute mean exit times in several applications. Chapter 13 deals with a typical area of application of SDEs: computing various financial quantities. Chapter 14 investigates long-time properties of SDEs and steady states. Chapter 15 presents the multilevel Monte Carlo technique to reduce the computational cost of the classical Monte Carlo method. Chapter 16 introduces SDEs with jumps. Chapter 17 presents a derivation of high-order numerical methods for SDEs (based on stochastic Taylor expansions). Chapters 18 and 19 extend the main concepts seen so far to systems of SDEs. Finally, the last chapter of the book deals with stochastic modeling and simulation of chemical reactions.

The book under review is very well written, accessible, enjoyable to read, not too long, and offers heuristic explanations to key concepts and results on SDEs and their numerical discretisations. It is an ideal book for undergraduate and graduate students in mathematics and statistics, as well as interested students from computer science, engineering, finance, life sciences, or physics, for instance. This book is also recommended for more "senior" scientists who would like to learn the "basics" about numerics for SDEs, for example to give an introductory lecture on the subject.

Desmond J. Higham and Peter E. Kloeden, *An Introduction to the Numerical Simulation of Stochastic Differential Equations.* SIAM, 2012, 293 pages, Hardback ISBN 978-1-6119-7642-7.

David Cohen is professor of mathematics at Chalmers University of Technology. When not teaching or racking his brains investigating numerical methods for stochastic partial differential equations, David enjoys riding his bikes in the forests and roads around Gothenburg and in the Swiss Alps (with a helmet of course!).

david.cohen@chalmers.se

DOI 10.4171/MAG-51

Announcement of the next meeting of the EMS Council Bled, Slovenia, 25th and 26th June 2022

The EMS Council meets every second year. The next meeting will be held in Bled, Slovenia, **25th and 26th June 2022,** at Rikli Balance Hotel (www.sava-hotels-resorts.com/en/sava-hoteli-bled). Registration will be held on 25th June between 12:00 and 13:45. The Council meeting begins on 25th June at 14:00 and ends on 26th June at 12:00.

Delegates

Delegates to the Council shall be elected for a period of four years. A delegate may be re-elected provided that consecutive service in the same capacity does not exceed eight years. Delegates will be elected by the following categories of members.

(a) *Full members*

Full members are national mathematical societies, which elect 1, 2, 3, or 4 delegates according to their membership class. The Council decides the membership class and societies are invited to apply for the class upgrade. However, the *current* membership class of the society determines the number of delegates for the 2022 Council.

Each society is responsible for the election of its delegates. To be eligible to nominate its delegates the society must have paid the corporate membership fee for the year 2021 and/or 2022. It is not compulsory but is highly appreciated that the full member delegates join the EMS as individual members.

The link to the online nomination form for delegates of full members is below. The deadline for nominations for delegates of full members is **4th April 2022.**

(b) *Associate members*

Delegates representing associate members shall be elected by a ballot organized by the Executive Committee from a list of candidates who have been nominated and seconded by associate members, and have agreed to serve. In October 2021, there were three associate members and, according to our statutes, (up to) one delegate may represent these members. Associate members delegates must themselves be members of the EMS and have paid the individual membership fees for the year 2021 and/or 2022. The associate member delegate Susanne Ditlevsen can be re-elected for the second term 2022–2025.

The link to the online nomination form for delegates of associate members (including the ones eligible for re-election) is below. The deadline for nominations for delegates of individual members is **28th February 2022.**

(c) *Institutional members*

Delegates representing institutional members shall be elected by a ballot organized by the Executive Committee from a list of candidates who have been nominated and seconded by institutional members, and have agreed to serve. In October 2021, there were 48 institutional members and, according to our statutes, (up to) four delegates may represent these members. Institutional member delegates must themselves be individual members of the EMS and have paid the individual membership fees for the year 2021 and/or 2022. The delegate whose term includes 2022 is Klavdija Kutnar. The institutional member delegates who can be re-elected for the second term 2022–2025 are David Abrahams, Alex Mielke and Luis Vega.

The link to the online nomination form for delegates of institutional members (including the ones eligible for re-election) is below. The deadline for nominations for delegates of institutional members is **28th February 2022.**

(d) *Individual members*

Delegates representing individual members shall be elected by a ballot organized by the Executive Committee from a list of candidates who have been nominated and seconded, and have agreed to serve. These delegates must themselves be individual members of the European Mathematical Society and have paid the individual membership fees for the year 2021 and/or 2022.

In October 2021 there were 3143 individual members and, according to our statutes, these members may be represented by (up to) 32 delegates.

Here is a list of the current delegates of individual members whose term includes 2022:

- Luis Alseda
- Antonio Campillo
- Carles Casacuberta
- Fernando da Costa
- Heike Faßbender
- Olga Gil-Medrano
- Thierry Horsin
- Kenji Iohara
- Christian Kassel
- Hrvoje Kraljević
- Pierangelo Marcati
- Wacław Marzantowicz
- Vicente Muñoz
- Piotr Oprocha
- Joaquín Pérez
- Carlo Petronio
- Jasna Prezelj
- Armen Sergeev
- Juan Soler

Here is a list of the delegates of individual members who could be re-elected for the second term 2022–2025:

- Jean-Marc Deshouillers
- Alice Fialowski
- Jan Pospíšil
- Primož Potočnik
- Muhammed Uludag

The link to the online nomination form for delegates of individual members (including the ones eligible for re-election) is below. The deadline for nominations for delegates of individual members is **28th February 2022.**

Agenda

The Executive Committee is responsible for preparing the matters to be discussed at Council meetings. Items for the agenda of this meeting of the Council should be sent as soon as possible, and no later than **25th April 2022,** to the EMS Secretary Jiří Rákosník (rakosnik@math.cas.cz).

© Jošt Gantar, www.slovenia.info

Executive Committee

The Council is responsible for electing the President, Vice-Presidents, Secretary, Treasurer and other members of the Executive Committee. The present membership of the Executive Committee, together with their individual terms of office, is as follows:

President:	Volker Mehrmann (2019–2022)
Vice-Presidents:	Betül Tanbay (2019–2022)
	Jorge Buescu (2021–2024)
Secretary:	Jiří Rákosník (2021–2024)
Treasurer:	Mats Gyllenberg (2019–2022)
Members:	Frédéric Hélein (2021–2024)
	Barbara Kaltenbacher (2021–2024)
	Luis Narváez Macarro (2021–2024)
	Beatrice Pelloni (2017–2024)
	Susanna Terracini (2021–2024)

Members of the Executive Committee are elected for a period of four years. The President can only serve one term. Committee members may be re-elected, provided that consecutive service shall not exceed eight years.

The Council may, at its meeting, add to the nominations received and set up a Nominations Committee, disjoint from the Executive Committee, to consider all candidates. After hearing the report by the Chair of the Nominations Committee (if one has been set up), the Council will proceed to the elections to the Executive Committee posts.

All these arrangements are as required in the Statutes and By-laws, which can be found here together with the web page for the Council: http://euromathsoc.org

The online nomination form for **full member delegates:**

https://elomake.helsinki.fi/lomakkeet/114719/lomake.html

The deadline for nominations is **4th April 2022.**

The nomination form for **institutional, associate and individual member delegates:**

https://elomake.helsinki.fi/lomakkeet/114718/lomake.html

The deadline for nominations is **28th February 2022.**

Secretary: Jiří Rákosník (rakosnik@math.cas.cz)
Secretariat: ems-office@helsinki.fi

European Mathematical Society

EMS executive committee

President

Volker Mehrmann (2019–2022)
Technische Universität Berlin, Germany
mehrmann@math.tu-berlin.de

Vice presidents

Betül Tanbay (2019–2022)
Bogazici University, Istanbul, Turkey
tanbay@boun.edu.tr

Jorge Buescu (2021–2024)
University of Lisbon, Portugal
jbuescu@gmail.com

Treasurer

Mats Gyllenberg (2015–2022)
University of Helsinki, Finland
mats.gyllenberg@helsinki.fi

Secretary

Jiří Rákosník (2021–2024)
Czech Academy of Sciences, Praha, Czech Republic
rakosnik@math.cas.cz

Members

Frédéric Hélein (2021–2024)
Université de Paris, France
helein@math.univ-paris-diderot.fr

Barbara Kaltenbacher (2021–2024)
Universität Klagenfurt, Austria
barbara.kaltenbacher@aau.at

Luis Narváez Macarro (2021–2024)
Universidad de Sevilla, Spain
narvaez@us.es

Beatrice Pelloni (2017–2024)
Heriot-Watt University, Edinburgh, UK
b.pelloni@hw.ac.uk

Susanna Terracini (2021–2024)
Universita di Torino, Italy
susanna.terracini@unito.it

EMS publicity officer

Richard H. Elwes
University of Leeds, UK
r.h.elwes@leeds.ac.uk

EMS secretariat

Elvira Hyvönen
Department of Mathematics and Statistics
P.O. Box 68
00014 University of Helsinki, Finland
ems-office@helsinki.fi

Join the EMS

You can join the EMS or renew your membership online at euromathsoc.org/individual-members.

Individual membership benefits

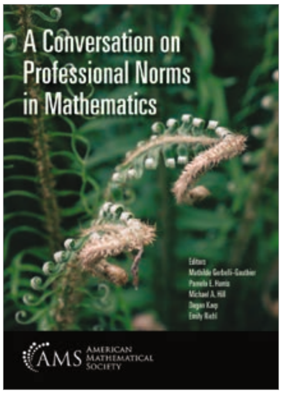
- Printed version of the EMS Magazine, published four times a year for no extra charge
- Free access to the online version of the *Journal of the European Mathematical Society* published by EMS Press
- Reduced registration fees for the European Congresses
- Reduced registration fee for some EMS co-sponsored meetings
- 20 % discount on books published by EMS Press (via orders@ems.press)*
- Discount on subscriptions to journals published by EMS Press (via subscriptions@ems.press)*
- Reciprocity memberships available at the American, Australian, Canadian and Japanese Mathematical Societies

* These discounts extend to members of national societies that are members of the EMS or with whom the EMS has a reciprocity agreement.

Membership options

- 25 € for persons belonging to a corporate EMS member society (full members and associate members)
- 37 € for persons belonging to a society, which has a reciprocity agreement with the EMS (American, Australian, Canadian and Japanese Mathematical societies)
- 50 € for persons not belonging to any EMS corporate member
- A particular reduced fee of 5 € can be applied for by mathematicians who reside in a developing country (the list is specified by the EMS CDC).
- Anyone who is a student at the time of becoming an individual EMS member, whether PhD or in a more junior category, shall enjoy a three-year introductory period with membership fees waived.
- Lifetime membership for the members over 60 years old.
- Option to join the EMS as reviewer of zbMATH Open.

AMERICAN MATHEMATICAL SOCIETY

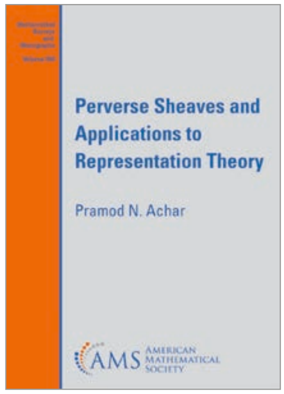


A CONVERSATION ON PROFESSIONAL NORMS IN MATHEMATICS

Edited by Mathilde Gerbelli-Gauthier, Institute for Advanced Study, Princeton et al

The articles in this volume grew out of a 2019 workshop, held at Johns Hopkins University, that was inspired by a belief that when mathematicians take time to reflect on the social forces involved in the production of mathematics, actionable insights result. Topics range from mechanisms that lead to an inclusion-exclusion dichotomy within mathematics to common pitfalls and better alternatives to how mathematicians approach teaching, mentoring and communicating mathematical ideas.

Dec 2021 157pp 9781470467135
Paperback €58.00



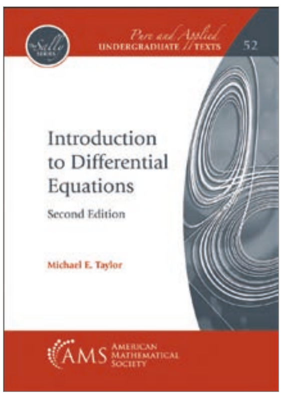
PERVERSE SHEAVES AND APPLICATIONS TO REPRESENTATION THEORY

Pramod N. Achar, Louisiana State University

Since its inception around 1980, the theory of perverse sheaves has been a vital tool of fundamental importance in geometric representation theory. This book gives a comprehensive account of constructible and perverse sheaves on complex algebraic varieties, and show how to put this to work in the context of geometric representation theory.

Mathematical Surveys and Monographs, Vol. 258

Dec 2021 562pp 9781470455972
Paperback €120.00



INTRODUCTION TO DIFFERENTIAL EQUATIONS

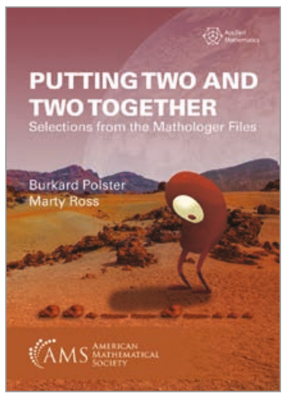
Second Edition

Michael E. Taylor, University of North Carolina

Introduces students to the theory and practice of differential equations, which are fundamental to the mathematical formulation of problems in physics, chemistry, biology, economics, and other sciences. This second edition incorporates much new material, including sections on the Laplace transform and the matrix Laplace transform, a section devoted to Bessel's equation, and sections on applications of variational methods to geodesics and to rigid body motion.

Pure and Applied Undergraduate Texts, Vol. 52

Jan 2022 388pp 9781470467623
Paperback €86.00



PUTTING TWO AND TWO TOGETHER

Selections from the Mathologer Files

Burkard Polster, Monash University & Marty Ross, Monash University

A humorous and quirky collection of unusual, ingenious, and beautiful morsels of mathematics. Authors Burkard Polster (YouTube's *Mathologer*) and Marty Ross delve into mathematical puzzles and phenomena in engaging stories featuring current events, sports, and history, many flavoured with a distinctive bit of Australiana.

Jan 2022 270pp 9781470460112
Paperback €38.00

AMS is distributed by EUROSPAN

Free delivery at **eurospanbookstore.com/ams**

CUSTOMER SERVICES:
Tel: +44 (0)1767 604972
Fax: +44 (0)1767 601640
Email: eurospan@turpin-distribution.com

Prices do not include local taxes.

FURTHER INFORMATION:
Tel: +44 (0)20 7240 0856
Fax: +44 (0)20 7379 0609
Email: info@eurospan.co.uk