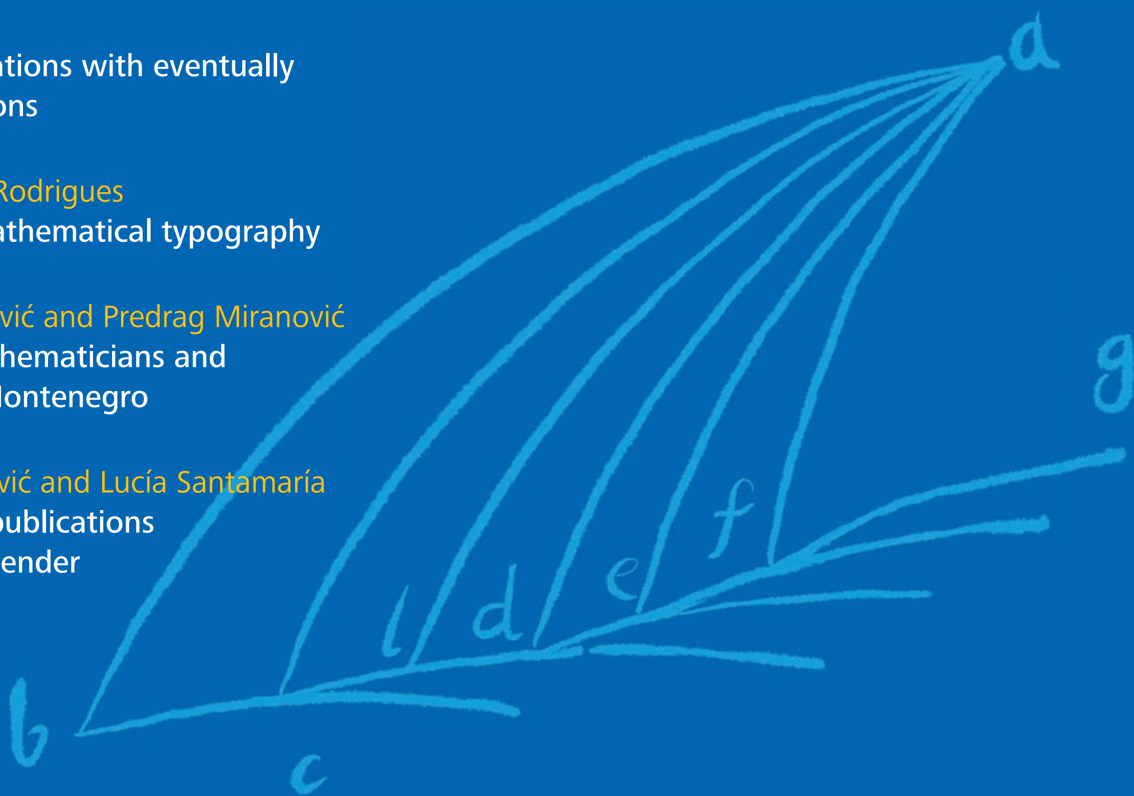
European Mathematical Society Magazine

ISSN 2747-7894 / Issue 123 / March 2022

EMS Magazine



EUROPEAN MATHEMATICAL SOCIETY

European Mathematical Society Magazine

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 art is a reconstruction of a sketch by the Portuguese mathematician Pedro Nunes (1502–1578) in the 1573 edition of his book "De Arte Atque Ratione Navigandi" ("The Art and Science of Navigation"). It shows the approximation of a loxodromic curve (or rhumb line) by arcs of geodesics (bc, cl, etc.). We recall that a rhumb line is a line on the globe defined by keeping to a constant bearing θ on the compass. The angle θ determines a geodesic from the given initial point b. Moving along this geodesic for a length of 1 degree leads to a point c. From c, θ determines the next geodesic arc, and so on, to obtain the sequence of arcs $bcldefg$ which approximates the rhumb line. This is essentially the spherical analogue of Euler's method, two centuries before Euler would use it to compute approximate solutions for differential equations.

António B. Araújo

A message from the president

Dear EMS members,

The editorial page in this issue is different from usual – because so is the year 2022! Firstly, we are once again attempting to celebrate the 30th anniversary of the EMS, and we truly hope that this can take place on site in Edinburgh on March 31. All are invited to participate: keep an eye on euromathsoc.org for details. Secondly, the EMS council is planned for June 25–26 in Bled, Slovenia. At that meeting, we will elect the next president, one vice president, the treasurer, and potentially a member-at-large. We will also discuss EMS plans for two new initiatives: the creation of an EMS youth academy, and the creation of topical activity groups. After that, the ICM in St. Petersburg will be the next major event. Please find a statement by the EMS executive committee below.

A further initiative on the part of the EMS is the creation of a new dynamic website concerning predatory journals; for more on this, see the announcement below.

With the pandemic still continuing and another political crisis keeping diplomats busy all over the world, I wish you all a happy, healthy, and peaceful 2022.

Volker Mehrmann, President of the EMS

A message from the executive committee

Azat Miftakhov is a young, promising Russian mathematician, who has already authored several serious mathematical publications.[1, 2, 3] The "Miftakhov case" has been featured prominently in the mathematical and general academic community. The facts and allegations about the case can be found in several openly accessible links,[4, 5] and concern about this case has led to the publication of statements by many mathematical societies, including members of the European Mathematical Society.[6, 7, 8, 9, 10] We wish to share the widespread unease our community feels about this affair. In this year where the IMU has chosen "Mathematics Unites" as its theme for the International Day of Mathematics, we invite the Russian authorities to contribute to uniting mathematicians in St. Petersburg for the ICM 2022.

EMS executive committee

A message regarding predatory publishing practices

An expansion of so-called predatory practices in dissemination of results is a major problem faced by today's research community, in particular mathematicians. It is also a major concern for the European Mathematical Society. There are journals and proceedings that neglect scientific quality checks via peer review in the sole interest of cashing in on publication fees from the authors and their home institutions. The success of predatory publishing is strongly tied to existing incentives, notably publication lists being measured by crude or sophisticated but ill-suited bibliometric data. Numbers being an essential part of any kind of mathematics, mathematicians are also well equipped to avoid numbers when not necessary. We believe that the mathematics community has enough tradition and means to protect itself from such threats, but best editorial and publication practices and the methods adopted by predatory journals that deviate from them must be recalled frequently. Once every author is aware that publishing in predatory journals can only harm her/his reputation and count negatively in any kind of evaluation, the problem should vanish.

This is why the EMS, upon the initiative of its *ethics committee*[11] and *publication and electronic dissemination committee*[12] has decided to launch an awareness campaign about predatory practices in the broad domain of mathematics. It is important for all researchers, both young and experienced, to know what is going on; they must be capable of recognising predatory practices when they encounter them as authors or editors.

To initiate this campaign, the EMS will soon create a dynamic webpage containing hints on how to avoid being the victim of predatory journals. The EMS will ask member societies to link this page to their websites, and is looking forward to involving community members in discussion on the issue.

Thierry Bouche, Stefan Jackowski, Betül Tanbay

The indices on this page correspond to hyperlinks which can be found in the online version of the EMS Magazine.

Evolution equations with eventually positive solutions

Jochen Glück

We discuss linear autonomous evolution equations on function spaces which have the property that a positive initial value leads to a solution which initially changes sign, but then becomes – and stays – positive again for sufficiently large times. This eventual positivity phenomenon has recently been discovered for various classes of differential equations, but so far a general theory to explain this type of behaviour exists only under additional spectral assumptions.

1 Evolution equations and positivity

To set the stage, we start with a reminder about linear evolution equations whose solutions are positive whenever the initial value is.

Linear ODEs and positivity

For a matrix $A \in \mathbb{R}^{d\times d}$, the linear and autonomous initial value problem

$$\begin{cases} \dot u(t) = Au(t) & \text{for } t \in [0,\infty), \\ u(0) = u_0, \end{cases}$$

where $u_0 \in \mathbb{R}^d$, is well-known to be solved by the function

$$u\colon [0,\infty) \ni t \mapsto e^{tA}u_0 \in \mathbb{R}^d.$$

We say that the matrix family $(e^{tA})_{t\in[0,\infty)}$ is *positive* if $e^{tA}u_0 \geq 0$ for all $t \in [0,\infty)$ whenever $u_0 \geq 0$; equivalently, $e^{tA} \geq 0$ for all $t \in [0,\infty)$. Here, we use the notation ≥ 0 for a vector or a matrix to say that all its entries are ≥ 0.

Remark 1. There is some terminological inconsistency in the literature with respect to this notion: in matrix analysis and in some parts of PDE theory, it is common to use the word *non-negativity*; we use the notion *positivity* instead, which is more common in functional analysis.

To get an intuition for this concept, it is useful to recall that positivity of the matrix exponential function is easy to characterise in terms of A.

Theorem 2. *For $A \in \mathbb{R}^{d\times d}$, the family $(e^{tA})_{t\in[0,\infty)}$ is positive if and only if every off-diagonal entry of A is ≥ 0.*

Proof. "$\Rightarrow$" For indices $j \neq k$, one has

$$A_{jk} = \lim_{t\downarrow 0}\Big\langle e_j, \frac{e^{tA}-\mathrm{id}}{t}e_k\Big\rangle = \lim_{t\downarrow 0}\frac{1}{t}\langle e_j, e^{tA}e_k\rangle \geq 0,$$

where $e_j, e_k \in \mathbb{R}^d$ are the canonical unit vectors and $\langle\,\cdot\,,\cdot\,\rangle$ denotes the standard inner product on $\mathbb{R}^d$.

"$\Leftarrow$" By assumption, one has, for a sufficiently large number $c \geq 0$, the inequality $A + c\,\mathrm{id} \geq 0$, and hence

$$e^{tA} = e^{-tc}e^{t(A+c\,\mathrm{id})} \geq 0$$

for all $t \in [0,\infty)$, where the inequality at the end follows from the series expansion of the matrix exponential function. ∎

A typical situation where positivity of matrix exponential functions occurs is the study of Markov processes on finite state spaces.

Example 3. Assume that all off-diagonal entries of $A \in \mathbb{R}^{d\times d}$ are ≥ 0 and that all rows of A sum up to 0. Then $(e^{tA})_{t\in[0,\infty)}$ is positive, and the vector $\mathbb{1} \in \mathbb{R}^d$ whose entries are all equal to 1 satisfies $A\,\mathbb{1} = 0$ and thus $e^{tA}\,\mathbb{1} = \mathbb{1}$ for all $t \in [0,\infty)$. This shows that each of the matrices e^{tA}, $t \geq 0$, is row stochastic, so $(e^{tA})_{t\in[0,\infty)}$ describes a continuous-time Markov process on the finite state space $\{1,\dots,d\}$.

Infinite-dimensional equations

In infinite dimensions, we are still interested in initial value problems of the form

$$\begin{cases} \dot u(t) = Au(t) & \text{for } t \in [0,\infty), \\ u(0) = u_0, \end{cases}$$

but this time, u_0 is an element of a Banach space E, and $A\colon E \supseteq \operatorname{dom}(A) \to E$ is a linear operator which is defined on a vector subspace $\operatorname{dom}(A)$ of E. The initial value problem is well-posed if and only if A is a *generator* of a *C_0-semigroup* $(e^{tA})_{t\in[0,\infty)}$. Such a C_0-semigroup is a family of bounded linear operators on E which is a suitable infinite-dimensional substitute of the matrix exponential function and has similar properties, but it is not given by an exponential series in general. The solution u to the initial value problem is then given, again, by the formula $u(t) = e^{tA}u_0$ for $t \in [0,\infty)$. The generator and the C_0-semigroup determine each other uniquely, and the relation between semigroup and generator can in general be expressed by the formula

$$\operatorname{dom}(A) = \Big\{v \in E : \lim_{t\downarrow 0} \frac{1}{t}(e^{tA} - \mathrm{id})v \text{ exists in } E\Big\},$$
$$Av = \lim_{t\downarrow 0} \frac{1}{t}(e^{tA} - \mathrm{id})v.$$

The following notion will be used several times later on. For a linear operator $A\colon E \supseteq \operatorname{dom}(A) \to X$ on a Banach space X, the quantity

$$s(A) \coloneqq \sup\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\} \in [-\infty, \infty],$$

where $\sigma(A)$ denotes the spectrum of A, is called the *spectral bound of A*. If A generates a C_0-semigroup, then $s(A) < \infty$ (see [20, Theorem II.1.10 (ii)]). More information about C_0-semigroup theory can be found, for instance, in the monographs [20, 31].

Let us briefly illustrate the concept of a C_0-semigroup by two very classical examples.

Examples 4. (a) Let $p \in (1,\infty)$ and let the operator A be the Laplace operator on the space $L^p(\mathbb{R}^n)$, i.e.

$$\operatorname{dom}(A) = W^{2,p}(\mathbb{R}^n),$$
$$Av = \Delta v \coloneqq \sum_{j=1}^{n} \partial_j^2 v \quad \text{for } v \in \operatorname{dom}(A).$$

Then A generates a C_0-semigroup $(e^{tA})_{t\in[0,\infty)}$ on $L^p(\mathbb{R}^n)$ that is given by the formula

$$(e^{tA}u_0)(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} \exp\Big(-\frac{\|x-y\|_2^2}{4t}\Big) u_0(y)\, \mathrm{d}y$$

for $u_0 \in L^p(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$. The semigroup is called the *heat semigroup* since it describes the solutions to the heat equation

$$\dot u(t) = \Delta u(t).$$

Similar observations can be made on the space $L^1(\mathbb{R}^n)$, but the domain of the Laplace operator cannot be chosen to be a Sobolev space in that case, due to the lack of elliptic regularity.

(b) Let $p \in [1,\infty)$ and let the operator A be the negative first derivative on $L^p(0,\infty)$, given by

$$\operatorname{dom}(A) = \{v \in W^{1,p}(0,\infty) : v(0) = 0\}, \quad Av = -v'.$$

Then A generates the so-called *right shift semigroup* $(e^{tA})_{t\in[0,\infty)}$ on $L^p(0,\infty)$ given by

$$(e^{tA}u_0)(x) = \begin{cases} u_0(x-t) & \text{if } t \le x, \\ 0 & \text{if } t > x, \end{cases}$$

for $u_0 \in L^p(0,\infty)$. The mapping $u\colon [0,\infty) \ni t \mapsto e^{tA}u_0 \in L^p(0,\infty)$ is a so-called *mild solution* to the transport equation

$$\begin{cases} \dot u(t,x) = -\partial_x u(t,x) & \text{for } t, x > 0, \\ u(0,x) = u_0(x) & \text{for } x > 0, \\ u(t,0) = 0 & \text{for } t > 0; \end{cases}$$

see [20, Definition II.6.3] for the definition of mild solutions. This example is an easy illustration of the general principle that boundary conditions of a PDE are encoded in the domain of the corresponding operator A.

Positive C_0-semigroups

In order to discuss *positive* C_0-semigroups, one needs an order structure on the underlying Banach space E. This can be for instance a partial order induced by a general closed convex cone, or more specifically the order structure of a Banach lattice. To facilitate the exposition here, we will restrict our attention to the illustrative case of function spaces, most importantly to L^p-spaces (over σ-finite measure spaces).

For a function $f \in L^p$, we write $f \ge 0$ to indicate that $f(\omega) \ge 0$ for almost all ω. In accordance with the terminology used above, we call a function f *positive* if it satisfies $f \ge 0$. A C_0-semigroup $(e^{tA})_{t\in[0,\infty)}$ on L^p is called *positive* if $e^{tA}u_0 \ge 0$ for all $t \in [0,\infty)$ whenever $u_0 \ge 0$. Equivalently, each of the operators e^{tA} is positive – which we denote by $e^{tA} \ge 0$ – in the sense that it maps positive functions to positive functions.

We have already encountered two examples of positive C_0-semigroups: as is easy to see, both semigroups in Examples 4 are positive.

2 Positivity for large times

Let us now proceed to a more surprising situation, where positive initial values lead to solutions which might change sign at first, but again become – and stay – positive for sufficiently large times. In this section, we illustrate by means of two easy examples that this kind of behaviour can indeed occur; a more systematic account is presented in the subsequent section.

A matrix example

Let us start with a simple three-dimensional example.

Example 5. Consider the orthonormal basis $\mathcal{B}$ of $\mathbb{R}^3$ consisting of the three vectors

$$v_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}, \quad v_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

Let $A \in \mathbb{R}^{3\times 3}$ be such that its representation matrix with respect to the basis $\mathcal{B}$ is given by

$$R := \begin{bmatrix}0 & 0 & 0\\ 0 & -1 & -1\\ 0 & 1 & -1\end{bmatrix},$$

i.e., we let $A = VRV^{-1}$, where $V \in \mathbb{R}^{3\times 3}$ consists of the columns v_1, v_2, v_3. A direct computation shows that A has some strictly negative off-diagonal entries, so $(e^{tA})_{t\in[0,\infty)}$ is not positive according to Theorem 2. On the other hand, A has the eigenvalue 0 (with eigenvector v_1) as well as the further eigenvalues $-1 \pm i$, so e^{tA} converges to the matrix $v_1 \cdot v_1^{\mathsf{T}}$, whose entries are all equal to $1/3$, as $t \to \infty$; this shows that e^{tA} is a positive matrix for all sufficiently large times t.

A fourth order PDE

Let us now discuss an infinite-dimensional example where eventual positivity occurs.

Example 6. Let us consider the biharmonic heat equation with periodic boundary conditions on $L^2(0,1)$. It is given by

$$\dot{u}(t) = Au(t) \quad \text{for } t \in [0,\infty),$$

where $A\colon L^2(0,1) \supseteq \operatorname{dom}(A) \to L^2(0,1)$ has domain

$$\operatorname{dom}(A) = \{v \in H^4(0,1) : v^{(k)}(0) = v^{(k)}(1) \text{ for } k = 0,1,2,3\}$$

and is given by $Av = -v^{(4)}$ for each $v \in \operatorname{dom}(A)$. The C_0-semigroup $(e^{tA})_{t\in[0,\infty)}$ is not positive; this can for instance be seen by associating a sesquilinear form to $-A$ and using the so-called *Beurling–Deny criterion* [30, Corollary 2.18].

However, we can prove positivity for large times. To this end, note that the operator A is self-adjoint, and its spectrum consists of isolated eigenvalues only since $\operatorname{dom}(A)$ embeds compactly into $L^2(0,1)$. The largest eigenvalue of A is 0, and the constant function $\mathbb{1}$ spans the corresponding eigenspace. Hence we conclude, for instance from the spectral theorem for self-adjoint operators with compact resolvent, that

$$e^{tA}u_0 \to \langle u_0, \mathbb{1}\rangle\, \mathbb{1} := \int_0^1 u_0(x)\,\mathrm{d}x \cdot \mathbb{1} \quad \text{in } L^2(0,1)$$

for each $u_0 \in L^2(0,1)$ as $t \to \infty$. Since A is self-adjoint, the operators e^{tA} have the property that for $t > 0$ they map $L^2(0,1)$ into $\operatorname{dom}(A)$ and thus into $L^\infty(0,1)$. Moreover, they are even continuous from $L^2(0,1)$ to $L^\infty(0,1)$ (this follows for instance from the closed graph theorem), so for $u_0 \in L^2(0,1)$, we even have

$$e^{tA}u_0 = e^{1\cdot A}e^{(t-1)A}u_0 \to \langle u_0, \mathbb{1}\rangle\, e^{1\cdot A}\,\mathbb{1} = \langle u_0, \mathbb{1}\rangle\,\mathbb{1}$$

as $t \to \infty$, where the convergence takes place with respect to the norm in $L^\infty(0,1)$. This implies that if $u_0 \geq 0$, then $e^{tA}u_0 \geq 0$ for all sufficiently large times t.

3 A systematic theory

After the previous ad hoc examples, we now present a few excerpts of a more systematic account of eventual positivity.

Eventually positive matrix semigroups

Example 5 already gives quite a straightforward idea of how to obtain a sufficient condition for a matrix exponential function to be eventually positive: if a matrix $A \in \mathbb{R}^{d\times d}$ has a simple real eigenvalue that dominates the real parts of all other eigenvalues and if the corresponding eigenvectors of A and the transposed matrix A^{T} have strictly positive entries only, then we expect e^{tA} to be positive – and in fact to even have strictly positive entries only – for all sufficiently large t. A bit more surprising is the Perron–Frobenius-like fact that the converse implication also holds. This was proved by Noutsos and Tsatsomeros in [29, Theorem 3.3], who thus obtained the following theorem (in a slightly different form; see [17, Theorem 6.1] for the following version of the result).

Theorem 7. *For a matrix $A \in \mathbb{R}^{d\times d}$, the following assertions are equivalent.*

- (i) *There exists a time $t_0 \geq 0$ such that all entries of e^{tA} are strictly positive for all $t > t_0$.*
- (ii) *The spectral bound $s(A)$ is a geometrically simple eigenvalue of A and strictly larger than the real part of every other eigenvalue of A. Moreover, both A and A^{T} have a strictly positive eigenvector for $s(A)$, respectively.*

Here, a *strictly positive* vector means a vector whose entries are all strictly positive.

Individual vs. uniform behaviour

In infinite dimensions, there is a subtlety that we have not properly discussed yet. Let $(e^{tA})_{t\in[0,\infty)}$ be a C_0-semigroup on a function space E. If, for every $0 \leq u_0 \in E$, there exists a time $t_0 \geq 0$ such that $e^{tA}u_0 \geq 0$ for all $t \geq t_0$, it is natural to call the semigroup *individually eventually positive* since t_0 might depend on u_0. If in addition t_0 can be chosen to be independent of u_0, then we call the semigroup *uniformly eventually positive*.

In finite dimensions, the two concepts can be easily seen to coincide (just apply the semigroup to all canonical unit vectors), but in infinite dimensions, there exist semigroups which are individually but not uniformly eventually positive [17, Examples 5.7 and 5.8].

Conditions for eventual positivity in infinite dimensions

The arguments given in Example 6 show individual eventual positivity of the semigroup, and the same argument can easily be generalised to a more abstract setting. There is one important issue to note, though: if the leading eigenfunction is not bounded away from 0, but might be equal to 0 on the boundary of the underlying domain (as in the case of Dirichlet boundary conditions), then it no longer suffices for the argument that $e^{1\cdot A}L^2$ be contained in L^∞; instead, one needs the condition that every vector in $e^{1\cdot A}L^2$ is dominated by a multiple of the leading eigenfunction. This property is closely related to Sobolev embedding theorems, and can be used to give a characterisation of a certain *strong* version of individual eventual positivity that is reminiscent of Theorem 7.

On the other hand, giving conditions for uniform rather than individual eventual positivity is more subtle. It requires a domination condition not only on the vectors in the image of $e^{1\cdot A}L^2$, but also on the image of the dual operator. If the semigroup is self-adjoint, though, this dual condition becomes redundant and one ends up with the following sufficient condition for uniform eventual positivity.

Theorem 8. *Let (Ω, μ) be a σ-finite measure space, let $(e^{tA})_{t\in[0,\infty)}$ be a self-adjoint C_0-semigroup on $L^2 \coloneqq L^2(\Omega,\mu)$ which leaves the set of real-valued functions invariant, and let $u \in L^2$ be a function which is strictly positive almost everywhere. Assume that the following assumptions hold.*

(1) *The spectral bound $s(A)$ is a simple eigenvalue of A, and the corresponding eigenspace contains a function v satisfying $v \geq cu$ for a number $c > 0$.*
(2) *There exists a time $t_1 \geq 0$ such that the modulus of every vector in $e^{t_1 A}L^2$ is dominated by a multiple of u.*

Then $(e^{tA})_{t\in[0,\infty)}$ is uniformly eventually positive.

The really interesting part in the conclusion of the theorem is the word *uniformly*, and this is more involved than the argument presented in Example 6. Two different proofs of the theorem are known: the first one is based on an eigenvalue estimate and the theory of Hilbert–Schmidt operators [24, Theorem 10.2.1] (the assumptions in the reference are slightly different, but the same argument works under the assumptions presented above); the second one employs a duality argument and can thus be generalised to non-self-adjoint semigroups on more general spaces [14, Theorem 3.3 and Corollary 3.5]. This reference also shows that the theorem can be adjusted to even yield a characterisation of a stronger type of eventual positivity.

Theorem 8 implies the non-trivial observation that the semigroup in Example 6 is even uniformly eventually positive.

Spectral properties

Positive semigroups are known to have surprising structural properties, in particular with regard to their spectrum. For some of these properties, it can be shown that they are shared by eventually positive semigroups, though some of the proofs are different from the classical proofs for the positive case. Here are two examples.

- If the spectrum of the generator A of an individually eventually positive semigroup is non-empty, then it follows that the spectral bound $s(A)$ is itself a spectral value [17, Theorem 7.6].
- For uniformly eventually positive semigroups on L^p-spaces, the spectral bound $s(A)$ coincides with the so-called *growth bound* of the semigroup (see e.g. [20, Definition I.5.6] for a definition); this was recently shown by Vogt [35, Theorem 2]. The same can be shown, even for individually eventually positive semigroups, on spaces of continuous functions [6, Theorem 4].

More results on the spectrum of eventually positive C_0-semigroups can be found in [5].

4 More examples

The biharmonic heat equation

Example 6 can be adjusted in the following way: we replace the unit interval with a ball B in $\mathbb{R}^d$, the fourth derivative with the square Δ^2 of the Laplace operator, and the periodic boundary conditions with so-called *clamped plate* boundary conditions, which require both the function and its normal derivative to vanish at the boundary. On $L^2(B)$, this yields the operator A given by

$$\operatorname{dom}(A) = H^4(B) \cap H_0^2(B),$$
$$Av = -\Delta^2 v,$$

where $H^4(B)$ and $H_0^2(B)$ denote Sobolev spaces. The operator A is self-adjoint and has negative spectral bound. It thus generates a C_0-semigroup $(e^{tA})_{t\in[0,\infty)}$ which describes the solutions to the so-called *bi-harmonic heat equation*

$$\dot{u}(t) = Au(t) \quad \text{for } t \in [0,\infty).$$

We have the following result.

Theorem 9. *The bi-harmonic heat semigroup $(e^{tA})_{t\in[0,\infty)}$ on $L^2(B)$ is uniformly eventually positive.*

Rough outline of the proof. Since B is a ball, the inverse operator $(-A)^{-1}$ – or rather its integral kernel, the so-called *Green function* of A – can be computed explicitly, and this was in fact done by Boggio over a hundred years ago [9] (see also [27, Section 2]). The explicit formula shows that $(-A)^{-1}$ maps positive functions to positive functions, and even strengthens their positivity in an appropriate sense. Hence, by a Krein–Rutman type result, we obtain that the leading eigenfunction of A is strictly positive inside B. Given the specific boundary conditions, it is not too surprising that we also get that assumptions (1) and (2) of Theorem 8 are satisfied if we choose $u = d^2$, where $d\colon B \to [0,\infty)$ describes the distance of each point in B to the boundary ∂B. Hence, Theorem 8 gives the desired eventual positivity. ∎

For more details, we refer to [16, second subsection of Section 6] and [14, third subsection of Section 4]. A few comments are in order.

Remark 10. (a) The argument sketched above breaks down for general domains in $\mathbb{R}^d$, since the inverse $(-A)^{-1}$ need no longer be positive in this case. This is a very well-studied topic in PDE theory; see for instance the surveys [33] by Sweers and [11] by Dall'Acqua and Sweers for more information.

(b) However, if we replace B with a domain which is sufficiently close to a ball, we still obtain the same result. The main point here is that positivity of $(-A)^{-1}$ or, under slightly larger perturbations, at least positivity of the leading eigenfunction of A remains true on such domains as shown by Grunau and Sweers in [27, Theorem 5.2]. So Theorem 9 holds on this more general class of domains, too.

(c) Theorem 9 remains true on general L^p-spaces rather than on L^2; see for instance [14, Theorem 4.4].

(d) If we replace the clamped plate boundary conditions with so-called *hinged* boundary conditions, which require $u = \Delta u = 0$ on the boundary, the situation becomes much easier because the operator can then be written as minus the square of the Dirichlet Laplace operator. In this case, we have eventual positivity of the semigroup on general domains; on the space of continuous functions, this example is worked out in [16, Theorem 6.1].

Non-local boundary conditions

Let us now go back to the unit interval and consider the Laplace operator, i.e. the second spatial derivative. If we impose local boundary conditions – such as for instance Dirichlet, Neumann or mixed Dirichlet and Neumann boundary condition, the Laplace operator is well-known to generate a positive semigroup (also on general domains in arbitrary dimension); see for instance [30, Corollary 4.3]. However, let us consider an example of non-local boundary conditions instead. More specifically, we consider the operator A on $L^2(0,1)$ given by

$$\operatorname{dom}(A) = \{v \in H^2(0,1) : v'(0) = -v'(1) = v(0) + v(1)\},$$
$$Av = v''.$$

This is a self-adjoint operator; the operator, and in particular its relation to the Dirichlet and the Neumann Laplace operator, is discussed in more detail in [2, Section 3]. The spectral bound of A is negative, and the inverse $(-A)^{-1}$ can be computed explicitly [16, proof of Theorem 6.11 (i)]; from this formula and the spectral theory of positive operators, we can conclude that $s(A)$ is a simple eigenvalue and that there is a corresponding eigenfunction which is strictly positive on the closed interval $[0,1]$; see [16, Theorem 6.11] for details. Moreover, we have $e^{1\cdot A}L^2(0,1) \subseteq \operatorname{dom}(A) \subseteq L^\infty(0,1)$, so the assumptions of Theorem 8 are satisfied for $u = \mathbb{1}$, and we obtain the following result.

Theorem 11. *The semigroup $(e^{tA})_{t\in[0,\infty)}$ on $L^2(0,1)$ generated by the Laplace operator with the non-local boundary conditions given above is uniformly eventually positive.*

Compare also [4, Section 4.2] for a related discussion. An example of eventual positivity for different non-local boundary conditions which lead to a non-self-adjoint realisation of the Laplace operator can be found in [14, Theorem 4.3].

Further examples

Today, eventual positivity, and closely related properties as for instance *asymptotic positivity*, are known for various further C_0-semigroups, including the semigroup generated by the Dirichlet-to-Neumann operator on the unit circle for various parameter choices [12] (which was the initial motivation for the development of the general theory), several delay differential equations ([17, Section 6.5], [24, Section 11.6] and [14, Theorem 4.6]), the semigroup generated by a bi-Laplacian with certain Wentzell boundary conditions [19, Section 7], various semigroups on metric graphs ([26, Proposition 3.7], [25, Section 6] and [8, Proposition 5.5]) and semigroups generated by Laplacians coupled by point interactions [28, Proposition 2].

5 Unbounded domains and local properties

The biharmonic heat equations on unbounded domains

A major drawback of Theorem 8 is that it can only be applied if the leading spectral value is even an eigenvalue of the operator A. This makes it impossible to apply the theorem to various differential operators that live on unbounded domains. For instance, consider the biharmonic operator A on $L^2(\mathbb{R}^d)$ given by

$$\operatorname{dom}(A) = H^4(\mathbb{R}^d),$$
$$Av = -\Delta^2 v.$$

The spectrum of A, which is the set $(-\infty, 0]$, does not contain eigenvalues, so Theorem 8 cannot be applied. Still, the semigroup $(e^{tA})_{t\in[0,\infty)}$ exhibits a certain local eventual positivity property: for every compact set $K \subseteq \mathbb{R}^d$ and every initial value $0 \le u_0 \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, the solution $u\colon t \mapsto e^{tA}u_0$ to the biharmonic heat equation $\dot{u}(t) = Au(t)$ becomes eventually positive on K. This was proved, under slightly different assumptions on u_0 in [23, Theorem 1 (i)] and [22, Theorem 1.1 (ii)] by explicit kernel estimates; for more general powers of Δ, a similar result was recently shown in [21, Theorem 1.1]. Under the assumptions described above, the result was proved by Fourier transform methods in [18, Theorem 2.1].

If we replace the whole space $\mathbb{R}^d$ with an infinite cylinder – for instance of the form $\mathbb{R} \times B$, where $B \subseteq \mathbb{R}^{d-1}$ is a ball – and again impose clamped plate boundary conditions, the same local eventual positivity result remains true. The proof is technically more involved, though, and relies on a detailed analysis of the specific partial differential equation under consideration; see [18, Theorem 2.3 and Section 4].

However, despite the successful analysis of the aforementioned concrete differential equations, an abstract and general theory as outlined in Section 3 for operators with leading eigenvalue is not yet in sight for the case without eigenvalues.

Open problem. Develop a theory of locally eventually positive C_0-semigroups $(e^{tA})_{t\in[0,\infty)}$ which is applicable in situations where the generator A does not have a leading eigenvalue.

Eigenvalues revisited

Getting back to operators which do have a leading eigenvalue, results such as Theorem 8 might still not be applicable in some cases due to conditions (1) and (2) which are sometimes particularly subtle at the boundary of Ω (if Ω is, say, a domain in $\mathbb{R}^d$ and A a differential operator). When all functions are restricted to compact subsets of Ω, though, conditions of the type (1) and (2) might still be satisfied.

This motivates the development of a theory of locally eventually positive semigroups for generators that do have a leading eigenvalue with strictly positive eigenfunction. Such a theory was presented by Arora in [3]. An application of the theory to certain fourth order operators with unbounded coefficients on $\mathbb{R}^d$ (which sometimes have eigenvalues due to the growth of the coefficients) was given in [1, Section 3.2].

6 Related topics and results

We close the article by discussing a few related concepts.

Perturbation theory

If A generates a positive C_0-semigroup on a function space E, it is quite easy to see that if B is a positive and bounded linear operator on E and M is a bounded and real-valued multiplication operator on E, then the perturbed semigroup $(e^{t(A+B+M)})_{t\in[0,\infty)}$ is positive, too: if $M = 0$, this follows for instance from the so-called *Dyson–Phillips series representation* of perturbed semigroups [20, Theorem III.1.10], and if M is non-zero, it follows from the previous case by using the formula

$$e^{t(A+B+M)} = e^{-tc}e^{t(A+B+M+c\,\mathrm{id})}$$

for a real number $c \geq 0$ that is sufficiently large to ensure that $M + c\,\mathrm{id}$ is positive.

For eventual positivity, though, the situation is much more subtle. Under quite general conditions, one can show that eventual positivity of a semigroup cannot be preserved by all positive perturbations of the generator. This was proved in [15, Theorem 2.3]; related results in finite dimensions had earlier been obtained in [32, Theorem 3.5 and Proposition 3.6]. On the other hand, sufficiently small positive perturbations can be shown not to destroy eventual positivity under appropriate assumptions [15, Section 4].

Maximum and anti-maximum principles

One abstract way to formulate that a linear operator $A\colon E \supseteq \mathrm{dom}(A) \to E$ on a function space E satisfies a *maximum principle* is to require that $(-A)^{-1}$ be a positive operator, i.e. maps positive functions to positive functions. If 0 is in the spectrum of E, or more generally if the spectral bound of A satisfies $s(A) \geq 0$, it is often more natural to consider the *resolvent* $(\lambda\,\mathrm{id} - A)^{-1}$ for real numbers $\lambda > s(A)$. If the resolvent at one such point λ_0 is positive, then the same is true for all $\lambda \in (s(A), \lambda_0)$, too, and we say that A satisfies a *maximum principle*. More precisely, this is a *uniform* maximum principle, while we say that A satisfies an *individual maximum principle* if, for each $0 \leq f \in E$, there exists an (f-dependent) number $\lambda_0 > s(A)$ such that $(\lambda\,\mathrm{id} - A)^{-1} f \geq 0$ for all $\lambda \in (s(A), \lambda_0)$.

Similarly, it is common to say that A satisfies a *uniform anti-maximum principle* if $s(A)$ is, say, an isolated spectral value and for all λ in a left neighbourhood of $s(A)$ the resolvent $(\lambda\,\mathrm{id} - A)^{-1}$ maps positive functions to negative functions. Likewise, we can define an *individual anti-maximum principle* (and clearly, the same concepts can be defined at isolated spectral values different from $s(A)$, too).

Anti-maximum principles have a considerable history and have, for instance, been studied for various elliptic differential operators; see e.g. [10] for a seminal paper on this topic. For biharmonic and polyharmonic operators the validity of (anti-)maximum principles is closely related to the boundary conditions and the geometry of the underlying domain, as explained in Remark 10.

The argument sketched after Theorem 9 can be generalised (and partially reversed) to obtain a correspondence between the following three types of properties:

(a) eventual positivity of the semigroup $(e^{tA})_{t\in[0,\infty)}$,
(b) spectral properties of A and positivity of the leading eigenfunction,
(c) an individual (anti-)maximum principle for A.

This correspondence was discussed in [16, Sections 3–5], where the terminology *eventual positivity and negativity of the resolvent* was used to describe maximum and anti-maximum principles. Indeed, equivalence between the three properties (a)–(c) is true under a number of technical restrictions which have been analysed in more detail in [13].

Uniform (anti-)maximum principles are more difficult to analyse than their individual counterparts – a phenomenon that occurs, as pointed out above, for semigroups, too, but becomes even more pronounced when studying (anti-)maximum principles. An abstract operator theoretic approach to uniform anti-maximum principles was first presented by Takáč in [34, Section 5], and recent progress on the topic was made in [7]. As a sample result, let us discuss the following special case of [7, Corollary 5.4] for self-adjoint operators on L^2.

Theorem 12. *Let* (Ω, μ) *be a σ-finite measure space, and let* $A\colon L^2 \supseteq \mathrm{dom}(A) \to L^2$ *be a real and self-adjoint operator on* $L^2 \coloneqq L^2(\Omega,\mu)$. *Let* $u \in L^2$ *be a function which is* > 0 *almost everywhere, and assume that there exists an integer* $m \geq 0$ *such that every vector in* $\mathrm{dom}(A^m)$ *is dominated in modulus by a multiple of* u. *Assume moreover that* $\lambda_0 \in \mathbb{R}$ *is an isolated spectral value of* A *and a simple eigenvalue whose eigenspace contains a function* v *that satisfies* $v \geq cu$ *for a number* $c > 0$.

If $\mu_1 > \lambda_0$ *is in the resolvent set of* A *and* $(\mu_1 \,\mathrm{id} - A)^{-1} \geq 0$, *then the following assertions are equivalent.*

(i) *One has* $(\mu \,\mathrm{id} - A)^{-1} \leq 0$ *for all* μ *in a left neighbourhood of* λ_0.

(ii) *There exists a real number* $d > 0$ *such that*

$$(\mu_1 \,\mathrm{id} - A)^{-1} f \leq d \,\langle f, u\rangle u \quad \text{for all } 0 \leq f \in L^2.$$

The assumption that A be a *real operator* means that the domain $\mathrm{dom}(A)$ is spanned by real-valued functions and that A maps real-valued functions to real-valued functions. Assertion (i) of the theorem is a uniform anti-maximum principle, while assertion (ii) can be considered as an upper kernel estimate for the resolvent (in other words: as an upper Green function estimate) of A. Simple consequences of this theorem are the classical results that the Dirichlet Laplace operator on an interval does not satisfy a uniform anti-maximum principle, while the Neumann Laplace operator on an interval does (see [7, Proposition 6.1 (a) and (b)] for a few more details). More involved examples where the theorem (or more general versions thereof) can be applied are discussed in [7, Section 6].

References

[1] D. Addona, F. Gregorio, A. Rhandi and C. Tacelli, Bi-Kolmogorov type operators and weighted Rellich's inequalities. arXiv: 2104.03811v1 (2021)

[2] K. Akhlil, Locality and domination of semigroups. *Results Math.* **73**, Paper No. 59 (2018)

[3] S. Arora, Locally eventually positive operator semigroups. *J. Oper. Theory*, to appear

[4] S. Arora, R. Chill and J.-D. Djida, Domination of semigroups generated by regular forms. arXiv:2111.15489v1 (2021)

[5] S. Arora and J. Glück, Spectrum and convergence of eventually positive operator semigroups. *Semigroup Forum* **103**, 791–811 (2021)

[6] S. Arora and J. Glück, Stability of (eventually) positive semigroups on spaces of continuous functions. arXiv:2110.04581v1 (2021)

[7] S. Arora and J. Glück, An operator theoretic approach to uniform (anti-)maximum principles. *J. Differential Equations* **310**, 164–197 (2022)

[8] S. Becker, F. Gregorio and D. Mugnolo, Schrödinger and polyharmonic operators on infinite graphs: parabolic well-posedness and p-independence of spectra. *J. Math. Anal. Appl.* **495**, Paper No. 124748 (2021)

[9] T. Boggio, Sulle funzioni di Green d'ordine m. *Rend. Circ. Mat. Palermo* **20**, 97–135 (1905)

[10] P. Clément and L. A. Peletier, An anti-maximum principle for second-order elliptic operators. *J. Differential Equations* **34**, 218–229 (1979)

[11] A. Dall'Acqua and G. Sweers, On domains for which the clamped plate system is positivity preserving. In *Partial differential equations and inverse problems*, Contemp. Math. 362, Amer. Math. Soc., Providence, RI, 133–144 (2004)

[12] D. Daners, Non-positivity of the semigroup generated by the Dirichlet-to-Neumann operator. *Positivity* **18**, 235–256 (2014)

[13] D. Daners and J. Glück, The role of domination and smoothing conditions in the theory of eventually positive semigroups. *Bull. Aust. Math. Soc.* **96**, 286–298 (2017)

[14] D. Daners and J. Glück, A criterion for the uniform eventual positivity of operator semigroups. *Integral Equations Operator Theory* **90**, Paper No. 46 (2018)

[15] D. Daners and J. Glück, Towards a perturbation theory for eventually positive semigroups. *J. Operator Theory* **79**, 345–372 (2018)

[16] D. Daners, J. Glück and J. B. Kennedy, Eventually and asymptotically positive semigroups on Banach lattices. *J. Differential Equations* **261**, 2607–2649 (2016)

[17] D. Daners, J. Glück and J. B. Kennedy, Eventually positive semigroups of linear operators. *J. Math. Anal. Appl.* **433**, 1561–1593 (2016)

[18] D. Daners, J. Glück and J. Mui, Local uniform convergence and eventual positivity of solutions to biharmonic heat equations. arXiv: 2111.02753v1 (2021)

[19] R. Denk, M. Kunze and D. Ploß, The bi-Laplacian with Wentzell boundary conditions on Lipschitz domains. *Integral Equations Operator Theory* **93**, Paper No. 13 (2021)

[20] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*. Grad. Texts in Math. 194, Springer, New York (2000)

[21] L. C. F. Ferreira and V. A. Ferreira, Jr., On the eventual local positivity for polyharmonic heat equations. *Proc. Amer. Math. Soc.* **147**, 4329–4341 (2019)

[22] A. Ferrero, F. Gazzola and H.-C. Grunau, Decay and eventual local positivity for biharmonic parabolic equations.

[23] F. Gazzola and H.-C. Grunau, Eventual local positivity for a biharmonic heat equation in $\mathbb{R}^n$. *Discrete Contin. Dyn. Syst. Ser. S* **1**, 83–87 (2008)

[24] J. Glück, Invariant sets and long time behaviour of operator semigroups. PhD thesis, Universität Ulm (2016)

[25] F. Gregorio and D. Mugnolo, Bi-Laplacians on graphs and networks. *J. Evol. Equ.* **20**, 191–232 (2020)

[26] F. Gregorio and D. Mugnolo, Higher-order operators on networks: hyperbolic and parabolic theory. *Integral Equations Operator Theory* **92**, Paper No. 50 (2020)

[27] H.-C. Grunau and G. Sweers, The maximum principle and positive principal eigenfunctions for polyharmonic equations. In *Reaction diffusion systems (Trieste, 1995)*, Lecture Notes in Pure and Appl. Math. 194, Dekker, New York, 163–182 (1998)

[28] A. Hussein and D. Mugnolo, Laplacians with point interactions – expected and unexpected spectral properties. In *Semigroups of operators – theory and applications*, Springer Proc. Math. Stat. 325, Springer, Cham, 47–67 (2020)

[29] D. Noutsos and M. J. Tsatsomeros, Reachability and holdability of nonnegative states. *SIAM J. Matrix Anal. Appl.* **30**, 700–712 (2008)

[30] E. M. Ouhabaz, *Analysis of heat equations on domains*. London Math. Soc. Monogr. Ser. 31, Princeton University Press, Princeton, NJ (2005)

[31] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*. Appl. Math. Sci. 44, Springer, New York (1983)

[32] F. Shakeri and R. Alizadeh, Nonnegative and eventually positive matrices. *Linear Algebra Appl.* **519**, 19–26 (2017)

[33] G. Sweers, When is the first eigenfunction for the clamped plate equation of fixed sign? In *Proceedings of the USA-Chile Workshop on Nonlinear Analysis (Viña del Mar-Valparaiso, 2000)*, Electron. J. Differ. Equ. Conf. 6, Southwest Texas State Univ., San Marcos, TX, 285–296 (2001)

[34] P. Takáč, An abstract form of maximum and anti-maximum principles of Hopf's type. *J. Math. Anal. Appl.* **201**, 339–364 (1996)

[35] H. Vogt, Stability of uniformly eventually positive C_0-semigroups on L_p-spaces. arXiv:2110.02310v2 (2021)

Jochen Glück is a postdoc at Universität Passau, Germany. His research focusses on order structures in functional analysis, operator semigroups and their applications to differential equations. In 2021 he was awarded the Jaroslav and Barbara Zemánek Prize for achievements in functional analysis, with special emphasis on operator theory.

jochen.glueck@uni-passau.de

T. Tao and the Syracuse conjecture

English translation of the paper "T. Tao et la conjecture de Syracuse" published in *La Gazette des Mathématiciens*, Number 168, April 2021

Jean-Paul Allouche

We will first recall the Syracuse conjecture (also known as the "$3n+1$" problem) and give a very quick overview of the known results on the subject. Then we will attempt to give some of the ideas behind a remarkable recent result of T. Tao on this conjecture.

1 Introduction

The Syracuse conjecture, also called Collatz conjecture, Kakutani conjecture or $3x+1$ problem (there is even a paper by B. Thwaites entitled *My conjecture*), is one of those extraordinarily attractive mathematical questions whose simplicity of statement is matched by their difficulty of proof, to the extent that many (most) of these questions are still open. One can think for example of the Goldbach conjecture or of the Fermat(–Wiles) theorem. A common characteristic of these very difficult conjectures is that their simple statements attract many amateurs, who are certainly well-intentioned, but who are sometimes difficult to convince that their approach is as naive as it is false. One can, however, hardly blame them, since even "professional mathematicians" are regularly seduced by these conjectures, and then realise that their own attempts towards a proof are unsuccessful. They probably do not know that P. Erdős once said to J. C. Lagarias, referring to this conjecture: "Hopeless. Absolutely hopeless", which is … not very encouraging.

Let us recall the statement of the Syracuse–Collatz–Kakutani–$(3x+1)$–Thwaites problem.

Conjecture. *Let f be the function defined on the positive integers by*

$$f(n) = \begin{cases} \dfrac{3n+1}{2} & \text{if } n \text{ is odd,} \\ \dfrac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

Then all the orbits of f are ultimately equal to $(1, 2, 1, 2, 1, 2, \dots)$.

In other words, defining f^k as the k-th iterate of f, the orbit of every integer n under f, i.e., the sequence $(n, f(n), f^2(n), \dots)$, contains the number 1, from which it alternatively takes the values 1 and 2.

Remark. (i) This conjecture is clearly equivalent to the following one: *Let $v_2(n)$ be the 2-adic valuation of the integer n (in other words, the largest integer a such that 2^a divides n) and g the function defined on the odd integers by* $\mathrm{Syr}(n) = (3n+1)/2^{v_2(3n+1)}$. *Then, for any odd integer n, there exists an integer ℓ such that* $\mathrm{Syr}^\ell(n) = 1$.

(ii) The history of this conjecture and practically all the results before the one by T. Tao which is the subject of this paper can be found in the book by J. C. Lagarias [9].

We propose here to first recall the results that have been proved so far, and then to try to summarise Tao's result, which is stated as follows in [12].

Theorem (Tao). *Almost all orbits of f contain an almost bounded element.*

2 First steps

The simplicity of the statement of this conjecture is likely to impel us to "play" with it and to do experiments. If we compute the iterates of f on sufficiently small integers, we soon see that the orbits reach 1, and are therefore ultimately equal to $(1, 2, 1, 2, 1, 2, \dots)$. We also see that the pre-period (i.e., the part before the periodic part) of the orbit of 27 is (curiously?) much longer than that of the smaller integers.

Another very general observation is that applying f to "half of all integers", namely the even ones, results in a value smaller than the starting one. Furthermore, if we consider integers of the form $(4m+1)$, successive applications of f give $4m+1 \to 6m+2 \to 3m+1$. Since $3m+1 < 4m+1$ (at least for $m > 0$), we thus obtain that "a quarter of the integers" leads, after just two iterations of f, to a number smaller than the starting one. So in fact at least "$3/4$ of all integers" belong to the set $S := \{n > 0 \mid \exists k, f^k(n) < n\}$. More precisely, the natural density of a set A of integers is by definition the limit, if it exists, of $\#\{n \in A \mid n \le x\}/x$ (and we define the upper and lower density by replacing the limit by the upper and lower limit). Thus what we just saw is equivalent to the statement that the lower density of the set S is larger than or equal to $3/4$. If we now proceed by considering the integers of the form $8m + j$ with $j \in [0, 7]$, then the integers in the residual classes modulo 2^a for $a = 4, 5, \ldots$, we successively obtain lower bounds for the lower density of S which are closer and closer to 1. The author of this paper carried out these experiments in the second half of the 1970s with one of the first programmable pocket calculators (a TI 58): after keeping the machine running for forty-eight hours or more, we obtained values so close to 1 that it was tempting to try to prove this result, in the hope that it would be simpler than the initial conjecture.

It is now time to *do some mathematics* – simple for the moment – by stating the following result.

Theorem. *The lower density of S is equal to* 1*, and the same holds for the density of S.*

The proof is based upon the study of residual classes modulo 2^a. The above sketch for integers modulo 2 or 4 is easily generalised by induction on a.

Proposition. (i) *Let* $\alpha(a, j) \coloneqq \#\{v \in [0, a-1] \mid f^v(j) \text{ odd}\}$. *Then*

$$f^a(2^a n + j) = 3^{\alpha(a,j)} n + f^a(j).$$

(ii) *For all i in* $[0, a]$*, we have*

$$\#\{j \in [0, 2^a - 1] \mid \alpha(a, j) = i\} = \binom{a}{i}.$$

We see from this proposition that the study of $f(n)/n$ for $n \le x$ will involve truncated sums of binomial coefficients. In order to estimate these sums, we use the following lemma.

Lemma. *For all* $\varepsilon \in (0, 1)$*, there exists an* $\eta \in (0, 1)$ *such that*

$$\frac{1}{2^a} \sum_{|i - a/2| > \varepsilon a} \binom{a}{i} \le \eta^a$$

for all sufficiently large values of a.

A proof of this lemma suggested by G. Tenenbaum uses the relation

$$\sum_{i=0}^{m} \binom{a}{i} = (a - m)\binom{a}{m} \int_1^2 t^m (2 - t)^{a-m-1}\, dt \quad \text{for } m \le a - 1$$

which can be proved by finite induction on m and which is, by the way, one of the exercises in the beautiful book by L. Comtet (see [4, Exercise 12, p. 91]). For more details, one may refer to a paper by the author published in the *Séminaire de Théorie des Nombres de Bordeaux* [1]. The result on density 1 was also proved independently by C. J. Everett in 1977 [5], H. Möller in 1977 [10] and E. Heppner in 1978 [6], and also by R. Terras in his papers of 1976 and 1979 [13, 14], either for the initial problem or for the generalisation due to H. Hasse, see the remark at the end of this section. These different papers written independently at about the same time suggest that the result on density 1 is not very difficult. It relies on a non-trivial limit, which is still a reasonably easy exercise. They also reveal the absence of electronic tools at that time (even if *Zentralblatt* existed in paper form as well as *MathSciNet* which was called *Mathematical Reviews*). I remember, however, that M. Mendès France told me about Heppner's and/or Möller's paper afterwards and put an "anonymous" letter with easily recognisable handwriting in my locker: *Forget about this problem. A friend who wishes you well* – which confirms Erdős' opinion as reported by Lagarias.

Why doesn't this density result provide a proof of the conjecture? It is the remaining term in the limit for the density computation that spoils the party – the result we actually obtain can be stated as follows:

$$\#\{n \le x \mid \exists k, f^k(n) < n\} = x + O(x^{1-\delta})$$

for some δ in $(0, 1)$, and in fact even if we had $O(1)$, we would still only obtain a weak form of the conjecture, namely that any orbit will eventually "loop" (i.e., is ultimately periodic[1]), but there could be more than one loop. One can even refine this by making the above k depend on n (logarithmically), but this still remains far from the actual conjecture even in its weak form.

However, what the author had overlooked was that the result he had obtained in [1] (actually a bit more precise than that stated above) could yield something more, namely that the set $S_c \coloneqq \{n \mid \exists k, f^k(n) < n^c\}$ is of density 1 for $c > 0.8691$. It was I. Korec who mentioned in 1994 in [7] that this was pointed out by the referee of his paper. Korec improves the 0.8691 value to $\log 3/\log 4$, which is about 0.7925 (see the review MR1290275 by Lagarias on *MathSciNet*).

Remark. A more general formulation of the conjecture, due to H. Hasse, is to replace "multiply by 3 and add 1, then divide by 2, or divide by 2, depending on the value modulo 2 of the starting integer", by "multiply by m and add a suitable residue in a complete system of fixed residues, then divide by d, or divide by d, depending on whether the starting integer is not or is divisible by d". The conjecture is then that there exists a function Φ such that if $m < \Phi(d)$, then all orbits are ultimately periodic and there is a finite number of possible periods, and if $m > \Phi(d)$, then there exists at least one non-ultimately periodic orbit. Möller [11] proposes the function $\Phi(d) = d^{d/(d-1)}$. (Note that $m = 3$ is "just" under the threshold $d^{d/(d-1)}$ for $d = 2$.) Some of the authors quoted above also give density results for this generalisation.

[1] Let us recall that a sequence $(u_n)_{n \ge 0}$ is said to be ultimately periodic if it is periodic for large enough indices, i.e., if there exist $n_0 \ge 0$ and $T \ge 1$ such that, for all $n \ge n_0$ and for all $k \ge 0$, we have $u_{n+kT} = u_n$.

3 Going a bit further

We now give some brief indications on other results (the book by Lagarias mentioned above is very complete and includes in particular an impressive annotated bibliography). We will not discuss the numerical results which produce huge integers N such that the conjecture is true for all integers $n \leq N$, nor to those which show that the number of elements of any possible period apart from $(1, 2)$ for the orbits of the function f is necessarily fantastically large.

An interesting theoretical question is: can we say anything about the integers n for which there exists k such that $f^k(n) = 1$? In other words, can we ask "how many" pre-images of 1 exist under the iterates of f? Because our lack of understanding of this function f, we cannot even say that this set has density 1. The best result obtained thus far is a lower bound of the type

$$\#\{n \leq x \mid \exists k, f^k(n) = 1\} > x^d$$

for sufficiently large x, with one of the most recent values for d being $d = 0.84$, see [8].

4 The result obtained by T. Tao

4.1 Introductory remarks

As is the case for the classical conjectures recalled at the beginning of this overview, whenever we fail to obtain a desired result, we always try to obtain at least a weaker form. A typical example is Goldbach's conjecture, for which J. R. Chen proved a weaker form in [3]: *Any sufficiently large even number is the sum of a prime number and a number having at most two prime factors.* Or think of the Fermat–Wiles theorem, for which an attempt was made to restrict to the "first case" (if $x^p + y^p = z^p$, then $xyz \equiv 0 \bmod p$). Some results of this kind have been described above, in particular the one which states that the density of the set $\{n \leq x \mid \exists k, f^k(n) < n^c\}$ is equal to 1 for $c > 0.7925$.

This last result can be expressed as follows: *Almost all integers have in their orbit by f an element smaller than the* (0.7925)*-th power of the considered integer.* This is of course a (convenient) abuse of language since density is not a probability on the integers. Using this terminology, one can ask which "best function" B could be obtained to replace n^c in the set $\{n \mid \exists k, f^k(n) \leq n^c\}$, while keeping the natural density of this set equal to 1. In other words, B should be such that almost all integers n have an element $\leq B(n)$ in their orbits by f. A *caveat* is necessary: as pointed out for example in Tao's paper, one cannot "improve" $n \to n^c$ by "iterating". Indeed, even if it is true that, for almost all integers n, there exists an element $n' = f^k(n)$ with $n' \leq n^c$, one cannot apply the result of "almost all" to n' to obtain an element $n'' = f^\ell(n')$ such that $n'' \leq (n')^c$, and thus $f^{k+\ell}(n) = f^\ell(n') = n'' \leq (n')^c \leq (n^c)^c = n^{c^2}$, because n' could very well belong to the set of exceptions of the "almost all" and have no associated n''. Let us also note that we do not know how to obtain $B(n) = 1$, since we do not know (end of the previous section) that the Syracuse conjecture is true for almost all integers. Tao's "tour de force" is to prove, *up to replacing the natural density by the logarithmic density* (see below), that one can take for B *any function tending to infinity, as slowly as one wants* for an infinitely large argument. For example, Tao writes, perhaps as a wink to estimates "à la Erdős", $B(n) = \log\log\log\log n$. He summarises this in a figurative way, stating that we can take an "almost bounded" B.

Definition. A set of integers $A \subset \mathbb{N}$ is said to have a logarithmic density equal to δ if the limit, when x tends to infinity, of

$$\frac{1}{\log x} \sum_{n \leq x,\, n \in A} \frac{1}{n}$$

exists and equals δ.

Remark. If the natural density of a set of integers exists, its logarithmic density also exists and is equal to the natural density. The converse is not true.

4.2 T. Tao's theorem

As we have seen above, Tao stipulates in his theorem a striking, even mediatic (in the non-pejorative sense of the term …) statement: *Almost all orbits of f contain an almost bounded element.* This means that, for any function B which tends to infinity, the *logarithmic* density of the set $\{n \mid \exists k, f^k(n) < B(n)\}$ is equal to 1. We will try to describe (as Tao himself does in the first pages of his paper) **in a heuristic way**, yet avoiding technical details (the paper has 49 pages), the steps of the proof and the small improvements that it suggests.

(1) Studying the function f is classically equivalent to studying the function that Tao calls Syr. Let $\nu_2(n)$ be the 2-adic valuation of the integer n, that is $\nu_2(n) = a$ if 2^a divides n and 2^{a+1} does not divide n. We define the function Syr on odd integers by

$$\text{if } n \text{ is odd, then } \mathrm{Syr}(n) = \frac{3n+1}{2^{\nu_2(3n+1)}} = f^{\nu_2(3n+1)}(n).$$

Of course, the Syracuse conjecture is that, for any odd integer n, there exists an integer k such that $\mathrm{Syr}^k(n) = 1$. And Tao's original statement is equivalent to: *Let B be a function defined on the odd integers that tends to infinity at infinity. Then, for almost all integers N, there exists an integer k such that* $\mathrm{Syr}^k(N) < B(N)$ (here "almost all integers" means that the set of odd integers for which the property is true, is of logarithmic density $1/2$ in the set of all integers).

(2) How do we compute the iterates of Syr for an odd integer N? Note that $a_1(N) \coloneqq v_2(3N+1)$, $a_2 \coloneqq v_2(3\,\mathrm{Syr}(N)+1)$, $a_3 \coloneqq v_2(3\,\mathrm{Syr}^2(N)+1)$, etc. Clearly,

$$\begin{aligned}\mathrm{Syr}(N) &= (3N+1)2^{-v_2(3N+1)} \\ &= 3.2^{-a_1(N)}N + 2^{-a_1(N)}, \\ \mathrm{Syr}^2(N) &= (3\,\mathrm{Syr}(N)+1)2^{-v_2(3\,\mathrm{Syr}(N)+1)} \\ &= 3^2.2^{-a_1(N)-a_2(N)} + 3.2^{-a_1(N)-a_2(N)} + 2^{-a_2(N)}, \\ &\text{etc.}\end{aligned}$$

and therefore,

$$\mathrm{Syr}^n(N) = 3^n 2^{-a_1(N)-a_2(N)-\cdots-a_n(N)}N + F_n(a_1(N), a_2(N), \dots, a_n(N)),$$

where

$$\begin{aligned}&F_n(a_1(N), a_2(N), \dots, a_n(N)) \\ &\quad \coloneqq 3^{n-1}2^{-a_1(N)-a_2(N)-\cdots-a_n(N)} \\ &\qquad + 3^{n-2}2^{-a_2(N)-a_3(N)-\cdots-a_n(N)} \\ &\qquad + \cdots + 3^1 2^{-a_{n-1}(N)-a_n(N)} + 2^{-a_n(N)}.\end{aligned}$$

This formula can be compared with the one seen above for f: $f^a(2^a n + j) = 3^{\alpha(a,j)}n + f^a(j)$, which essentially says that we can estimate the values of successive images of an integer in a class modulo 2^a from the images of a representative of this class, until a number of iterations equal to a (this number is thus of the order of the logarithm of the considered integer if we have chosen the representative in $[0, 2^a)$).

(3) Take as above $a_j(N) = v_2(3\,\mathrm{Syr}^{j-1}(N)+1)$. Then, heuristically, *for a "typical" odd integer N large enough, and n much smaller than* $\log N$*, the vector* $(a_1(N), a_2(N), \dots, a_n(N))$ *behaves like a geometric random vector of size n and parameter* $1/2$*, i.e., an n-tuple of independent random variables, all geometric with parameter* $1/2$. More precisely, the "behaves like" has to be taken in the sense of a small distance between random variables, where the distance between two discrete random variables X and Y taking their values in the same discrete space R is the total variation

$$\sum_{r\in R} |\mathbb{P}(X=r) - \mathbb{P}(Y=r)|.$$

A proposition proved in Tao's paper states that the heuristic property in italics at the beginning of step (3) is justified if N is uniformly distributed modulo 2^m for m slightly larger than $2n$. This gives a good control of $\mathrm{Syr}^n(N)$ for almost all N and for n of the order of $\gamma \log N$ with a small constant γ. Since one heuristically has an estimate like $\mathrm{Syr}^n(N) \approx (3/4)^n N$, in fact $\mathrm{Syr}^n(N) = e^{O(\sqrt{n})}(3/4)^n N$ (by the central limit theorem or by the Chernoff bound), one can in this way already obtain again Korec's result recalled above: the density of the set $S_c \coloneqq \{n \mid \exists k, f^k(n) < n^c\}$ is 1 for $c > 0.8691$. As Tao points out, a result of this kind is somewhat analogous to "almost sure local wellposedness results" for evolutionary partial differential equations, in which one has good short-time control for almost all initial conditions. Additionally, the theorem we want to prove is similar to an "almost sure almost global wellposedness" result. Now how to get from "local" to "global"?

(4) The last and most difficult step is to answer the above question by introducing a function that further "accelerates" the maps f and Syr seen above. This "first passage" function Pass is defined as follows: for $x \geq 1$ and any odd integer N, we write

$$T_x(N) \coloneqq \inf\{n \in \mathbb{N} \mid \mathrm{Syr}^n(N) \leq x\}$$

with the usual convention that $T_x(N) \coloneqq +\infty$ if $\mathrm{Syr}^n(N) > x$ for all n. The first passage function is then defined by

$$\mathrm{Pass}_x(N) \coloneqq \mathrm{Syr}^{T_x(N)}(N).$$

Tao is then inspired by a work of J. Bourgain [2] who goes from a local almost everywhere to a global almost everywhere, thanks to the construction of an invariant probability measure. Alas! This is impossible here, but the author gets around this issue by introducing a *family* of probability measures which are approximately transported one to the other by iterating Syr a variable number of times. This is what will permit the use of an iterative argument, which was not feasible "directly" as we pointed out at the beginning of Section 4.1 with n^c and $(n^c)^c$.

We will not go any further in this attempt to demystify Tao's beautiful proof, whose high technicality, but above all its inventiveness, have been barely touched. To try to summarise it – too schematically – let us start by describing a temptation shared both by the professional mathematician who discovers the Syracuse conjecture and by the amateur: basically, iterating the application f from the beginning seems to consist *roughly* of replacing n every second time (when n is odd) by approximately $3/2 \cdot n$, and of replacing n every other time (when n is even) by $1/2 \cdot n$; in other words, applying f^2 amounts to multiplying n approximately by $3/4$. For example (with a "reasonably chosen" initial integer),

$$17 \xrightarrow{f} 26 \xrightarrow{f} 13 \xrightarrow{f} 20 \xrightarrow{f} 10 \xrightarrow{f} \cdots,$$

that is to say

$$17 \xrightarrow{f^2} 13 \xrightarrow{f^2} 10 \xrightarrow{f^2} \cdots.$$

Thus the orbit of a typical integer seems to be obtainable approximately by a sequence of multiplications by $3/4$. It is this temptation, which obviously does not constitute a proof, that Tao, at the cost of unprecedented effort and technicality for such an apparently innocent statement, has transformed into a proof for almost all integers. There should not be any misunderstanding about the purpose of this remark: to go from "we multiply roughly by $3/4$" to Tao's proof and its half a hundred pages is at least as difficult as transforming a frog or a toad into a charming princess or prince.

5 So, what now?

Now what can we expect for this conjecture? Tao indicates that, by further refining his approach, it should be possible to replace "almost all in logarithmic density" with "almost all in natural density". But he leaves little hope that the function tending to infinity as slowly as one likes in his statement can be replaced by a constant. In other words, even the statement "the orbit of almost any integer is ultimately periodic" is still totally out of reach.

Acknowledgements. The author thanks Sophie Grivaux and Damien Gayet for convincing him to present T. Tao's paper and also for their enthusiasm. He also thanks the two referees for their valuable comments which helped to improve the first version of this text.

The *EMS Magazine* thanks *La Gazette des Mathématiciens* for authorisation to republish this text, which is an English translation of the paper entitled "T. Tao et la conjecture de Syracuse" published in *La Gazette des Mathématiciens*, Number 168, April 2021. The author would like to thank Miriam Gellrich Pedra and Jean-Bernard Bru for their translation of the original paper.

References

[1] J.-P. Allouche, Sur la conjecture de "Syracuse–Kakutani–Collatz". In *Sém. Th. Nombres, Bordeaux*, CNRS, Talence, Exp. No. 9, 15 (1979)

[2] J. Bourgain, Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys.* **166**, 1–26 (1994)

[3] J. R. Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica* **16**, 157–176 (1973)

[4] L. Comtet, *Analyse combinatoire. Tome I*. Collection SUP: "Le Mathématicien", Presses Universitaires de France, Paris (1970)

[5] C. J. Everett, Iteration of the number-theoretic function $f(2n) = n$, $f(2n+1) = 3n+2$. *Adv. Math.* **25**, 42–45 (1977)

[6] E. Heppner, Eine Bemerkung zum Hasse-Syracuse-Algorithmus. *Arch. Math. (Basel)* **31**, 317–320 (1978/79)

[7] I. Korec, A density estimate for the $3x+1$ problem. *Math. Slovaca* **44**, 85–89 (1994)

[8] I. Krasikov and J. C. Lagarias, Bounds for the $3x+1$ problem using difference inequalities. *Acta Arith.* **109**, 237–258 (2003)

[9] J. C. Lagarias, *The ultimate challenge: the* $3x+1$ *problem*. Amer. Math. Soc., Providence, RI (2010)

[10] H. Möller, F-Normalreihen. *J. Reine Angew. Math.* **289**, 135–143 (1977)

[11] H. Möller, Über Hasses Verallgemeinerung des Syracuse-Algorithmus (Kakutanis Problem). *Acta Arith.* **34**, 219–226 (1977/78)

[12] T. Tao, Almost all orbits of the Collatz map attain almost bounded values. arXiv:1909.03562 (2020)

[13] R. Terras, A stopping time problem on the positive integers. *Acta Arith.* **30**, 241–252 (1976)

[14] R. Terras, On the existence of a density. *Acta Arith.* **35**, 101–102 (1979)

Jean-Paul Allouche is *Directeur de recherche* emeritus at CNRS. He is working at IMJ-PRG, Sorbonne, Paris on subjects relating number theory and theoretical computer science, including the so-called *automatic sequences.*

jean-paul.allouche@imj-prg.fr

Portuguese mathematical typography: A brief overview from 1496 until 1987

José Francisco Rodrigues

The publication in Leiria, in 1496, of the *Almanach perpetuum*, with the astronomical tables of the Sephardi scholar Abraham Zacut (1452–1515), initiated the use of scientific typography in Portugal. The *Almanach*'s tables were calculated in Salamanca in the Alphonsine tradition with reference to the period 1473–1476 [4]. It was translated and edited by the Portuguese José Vizinho and printed by the Jewish typographer Abraham d'Ortas. It appeared just a few years after a first edition of the *Torah* of 1487, in a Hebrew printing press in Faro in the south of Portugal, and the *Tratado de Confissom* of 1489, which is the first Christian text in Portuguese and was printed in Chaves, in the north of the country [3]. It should be noted that the Gutenberg Bible, the first book ever printed in Europe, dates from 1455, and the first printing of Euclid's *Elements* in 1482 was done by the printer Erhard Ratdolt in Venice, in a Latin edition containing the first geometric diagrams ever printed.

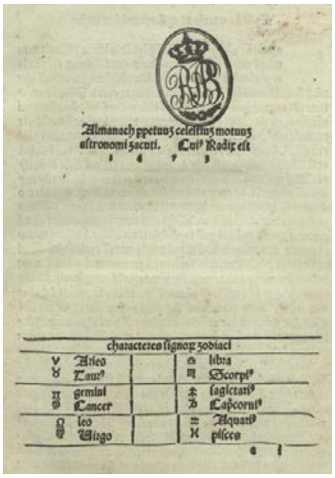
The *Almanach perpetuum* was the main incentive for the recent book exhibition *Tipografia Matemática Portuguesa: 1496–1987*, whose first edition [13] took place from July to October 2021, in the six-centuries-old *Moinho de Papel* in Leiria, a city in the centre of Portugal with an ancient 12th century castle, residence of kings and setting of several *cortes* (medieval parliaments). The city gave its name to a famous pine forest (*Pinhal de Leiria*), which supplied wood for the ships used in Portuguese navigation in the 15th and 16th centuries. The exhibition was an initiative of the city of Leiria in partnership with the *Centro Internacional de Matemática* and the Polytechnic Institute of Leiria. The *Moinho de Papel* is a historical building on the river Lis, and housed the first paper mill established in Portugal in 1411; this may well be relevant to the fact that Leiria was also one of the first Portuguese cities to have a printing press, and indeed the city where the first scientific book – which was also instrumental for navigation in the age of discoveries – was printed in Portugal 526 years ago.

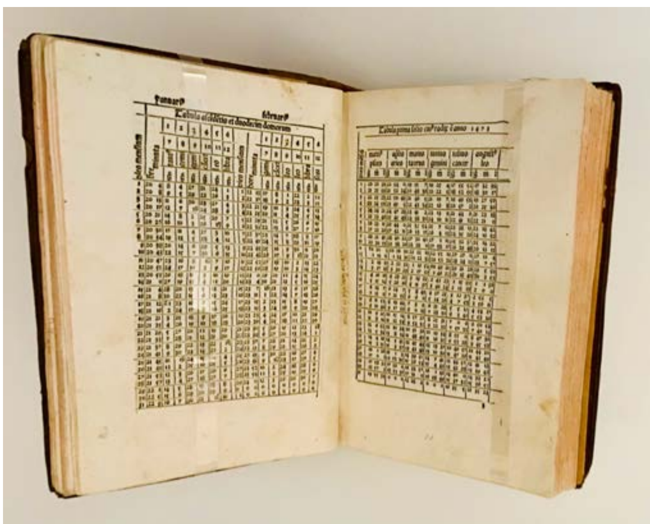
The second date of the title of this unprecedented exhibition corresponds to the publication, coincidentally 500 years after the first book printed in Portugal, of the first volume of *Portugaliae Mathematica* to be electronically composed in $\TeX$. With 32 significant original works in Portuguese mathematical typography, many of them rare books containing unknown or still barely known yet highly interesting pages, the exhibition presents nine sections that traverse the whole history of the mathematical sciences in Portugal, from the age of ocean navigation and the first European globalisation through military engineering, essential in the wars of Restoration (1640–1668) after the end of the Iberian Union, to the successive reforms of Colleges, Military Academies and Universities (1772 and 1911) and finally the scientific research of the 20th century.

Mathematical tables for navigation – arithmetical rules for overseas trade

The *Almanach perpetuum* is a landmark of the beginning of the culture of mathematical sciences in Portugal through the influence and use of the art and knowledge of navigation, namely in the first ocean voyages of Vasco da Gama to India and Pedro Álvares Cabral to Brazil [1]. It had several editions in the 16th century and it was used in the preparation of the *Reportórios dos Tempos*, a set of popular time calendars and almanacs also used in astrology; a certain Gaspar Nicolas participated in the elaboration of the one from the year 1518.

In 1519 in Lisbon, this Portuguese mathematician published his *Tratado da pratica Darismetyca*, a book of a technical and utilitarian nature about the rules of arithmetic, also "for overseas trade"; it went through ten new editions over the course of the next two centuries. In his preface, Nicolas acknowledges practical arithmetic as "very necessary in these Portuguese kingdoms and territories for the sake of the flourishing of the merchant trades of India and Persia and Arabia and Ethiopia and other major parts that have come to us". Between 1472 and 1519, about forty books on arithmetic were published in Europe, including Luca Pacioli's *Summa*, printed in Venice in 1494, which Gaspar Nicolas read and used in his own treatise.





Mathematics of Navigation, Sky Surveying and Cartography

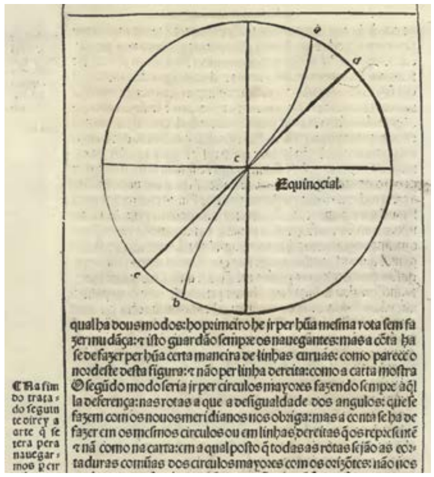
Printing also played an important role in the development of the mathematical theory of navigation through the pioneering works of Pedro Nunes (1502–1578) [2]. It started with the edition in Portuguese of his *Tratado da Sphera*, printed in Lisbon in 1537 by Germão Galharde, the same typographer who typeset the *Tratado da pratica Darismetyca*. In two original chapters of that book, Nunes presented the first mathematical conceptualisation of the rhumb line, later called the loxodrome, as a new spherical spiral distinct from the great circles and illustrated by a polar projection with azimuth ±45° and ±67° 30′.

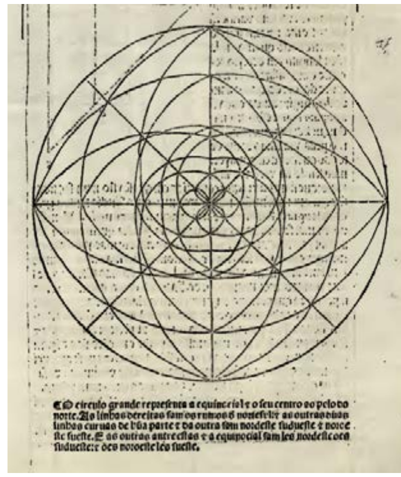
Subsequently this Portuguese mathematician suggested the rectification of the rhumb line for the maritime chart and developed original methods, in particular for the approximation of the rhumb line, which laid the foundations for the elaboration of nautical tables and the cartographic projection made in 1569 by Mercator in his famous *mappa mundi* [6]. The broken line invented by Pedro Nunes to approximate the loxodrome was composed of small sections of great circles. It was first published in Latin in his *Opera*, in 1566 in Basel; it is now called the *noniodrome* and was chosen for the logo of the Leiria exhibition. It is the natural idea of approximating the loxodrome that corresponds to Euler's modern method for integrating differential equations, and it was used numerically by Edward Wright in the secants' method to construct nautical tables in 1596 [10]. The re-edition in Coimbra of his *Opera*, in 1573, of Nunes' fundamental work *De arte atque ratione navigandi* is a reference of Portuguese typography in the age when the scientific revolution started in the sixteenth century.

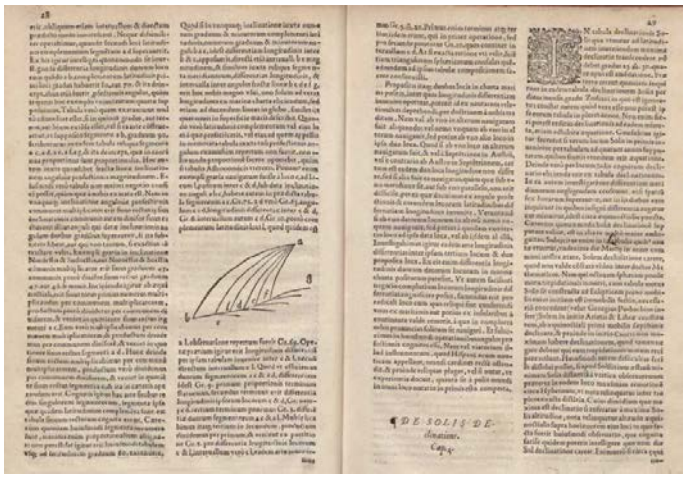
Practical and military use of mathematics before and after the Restoration

Printing was instrumental also in the teaching and practical use of mathematics for military architecture, navigation and artillery in Europe, and in particular in Portugal with the *Methodo Lusitanico* (1680) and *O Engenheiro Portuguez* (1728), both printed in Lisbon, after the Restoration of Portuguese independence in 1640 which ended the Iberian Union. Luís Serrão Pimentel (1613–1679), the posthumous author of the *Methodo*, was the Royal Cosmographer and Engineer who initiated the teaching and practices of military architecture in Portugal, in particular, using logarithms to solve trigonometry problems.

Mathematics book printing continued in Portugal for the navy and the army schools throughout the 18th and 19th centuries among others, with the Portuguese translations of the important book by Lagrange, *La Théorie des fonctions analytiques* being printed in Lisbon in 1798 just one year after its publication in Paris, and the two volumes of the *Traité élémentaire du calcul différentiel et du calcul intégral* by Lacroix at the *Impressão Régia* (Royal Printing







House) in 1812 and 1814 respectively in Rio de Janeiro, which was the capital of the Portuguese kingdom from 1808 until 1821. Other original texts, such as the interesting *Carta Físico-Mathematica sobre a theórica da pólvora em geral e o comprimento das peças em particular*, written in 1769 by José Anastácio da Cunha and printed posthumously in Porto in 1838, show the important role played by the military in the history of mathematics in Portugal.

Teaching at the *Colégio dos Nobres*, *Universidade* and *Colégio de São Lucas*

Following the reforms of Marquis of Pombal, head of the government during the period of the Portuguese Enlightenment, the printing press would serve teaching with the publication of higher mathematics textbooks. These were initially translations of foreign authors, such as Euclid's *Elementos*, printed in 1768, for the *Colégio dos Nobres* in Lisbon, several of Bézout's textbooks for the first Faculty of Mathematics of the University of Coimbra, reformed in 1772, and in the following century, the *Curso Completo de Mathematicas Puras* (1838 and 1839) by Francoeur, which were printed at the University Press in Coimbra.

An exceptional book is the original and remarkable *Princípios Mathematicos*, the printing of which was completed in Lisbon only in 1790, by José Anastácio da Cunha (1744–1787), the military mathematician and "lente penitenciado", a professor at the University of Coimbra from 1773 to 1778 who was then imprisoned by the Inquisition from 1978 to 1781. The *Princípios*, which was translated into French (Bordeaux, 1811) consisted of 21 chapters with 18 prints, and was intended to be the basis of all of mathematics, including geometry and arithmetic, algebra and series, differential and integral calculus. It also contains new and original theoretical contributions, such as the first rigorous definition of the convergence of a series, and the new concept of infinity and infinitesimal as variable quantities [5], as well as a new theory of the exponential and logarithmic functions as convergent power


THEORICA
DAS FUNÇÕES ANALYTICAS,
QUE CONTEM
OS PRINCIPIOS DO CALCULO DIFFERENCIAL
LIVRES DE TODA A CONSIDERAÇAÕ DE QUANTIDADES INFINITAMENTE PEQUENAS OU DE DESVANECENTES, DE LIMITES OU DE FLUXÕES, E REDUZIDOS Á ANALYSE ALGEBRICA DAS QUANTIDADES FINITAS.
POR M. LA GRANGE
E DE ORDEM DE
SUA ALTEZA REAL
O PRINCIPE
NOSSO SENHOR
TRADUZIDA DO FRANCEZ
POR
MANOEL JACINTO NOGUEIRA DA GAMA
Cavalleiro Profeſſo na Ordem de S. Bento de Aviz, Bacharel Formado em as Faculdades de Mathematica, e Philoſophia pela Univerſidade de Coimbra, Capitaõ Tenente da Armada Real, e Profeſſor de Mathematica na Academia Real da Marinha.
PRIMEIRA PARTE.
LISBOA,
Na Offic. de JOAÕ PROCOPIO CORREA DA SILVA,
Impreſſor da Santa Igreja Patriarcal.
ANNO M. DCC. XCVIII.
Impreſſo por ordem de Sua Mageſtade.


TRATADO ELEMENTAR
DE
CALCULO DIFFERENCIAL,
E
CALCULO INTEGRAL
POR
Mr. LACROIX;
POR ORDEM
DE
SUA ALTEZA REAL,
TRADUZIDO EM PORTUGUEZ PARA USO DOS ALUMNOS DA REAL ACADEMIA MILITAR DESTA CORTE,
POR
FRANCISCO CORDEIRO DA SILVA TORRES,
Sargento Mór do Real Corpo de Engenheiros, e Lente da meſma Academia.
PARTE 1.ª
CALCULO DIFFERENCIAL.
RIO DE JANEIRO.
NA IMPRESSÃO REGIA.
1812.
Por Ordem de S. A. R.


PRINCIPIOS MATHEMATICOS
PARA INSTRUCÇAÕ
DOS ALUMNOS DO COLLEGIO
DE
SAÕ LUCAS,
DA REAL CASA PIA DO CASTELLO
DE
SAÕ JORGE:
OFFERECIDOS
AO SERENISSIMO SENHOR
D. JOAÕ,
PRINCIPE DO BRAZIL:
COMPOSTOS PELO DOUTOR
JOSE' ANASTACIO DA CUNHA,
DE ORDEM DO DESEMBARGADOR DO PAÇO
DIOGO IGNACIO DE PINA MANIQUE,
Intendente Geral da Policia da Corte, e Reino, &c, &c, &c.
LISBOA
NA OFFIC. DE ANTONIO RODRIGUES GALHARDO,
Impreſſor do Eminentiſſimo Senhor Cardeal Patriarca,
ANNO M. DCC. XC.
Com licença da Real Meza da Commiſſaõ Geral ſobre o Exame, e Cenſura dos Livros.


PRINCIPIOS
MATHEMATICOS.
LIVRO VIIII.
Definiçaõ I.
SERIE convergente chamam os Mathematicos àquella, cujos termos ſaõ ſemelhantemente determinados, cada hum pelo numero dos termos precedente, de ſorte que ſempre a ſerie ſe poſſa continuar, e finalmente venha a ſer indifferente o continua-la ou naõ, por ſe poder deſprezar ſem erro notavel a ſomma de quantos termos ſe quizeſſe ajuntar aos já eſcritos ou indicados: e eſtes ultimos indicam-ſe eſcrevendo &c. depois dos primeiros dois, ou tres, ou quantos ſe quizer: he porém neceſſario que os termos eſcritos moſtrem como ſe poderia continuar a ſerie, ou que iſſo ſe ſaiba por outra via.
Propoſiçaõ I.
SErie de proporcionaes continuos he convergente, ſe o primeiro he maior que o ſegundo. Sejam A e B duas quaeſquer grandezas homogeneas, mas $A>B$; e $A+B+B\frac{B}{A}+B\frac{B}{A}\times\frac{B}{A}+B\frac{B}{A}\times\frac{B}{A}\times\frac{B}{A}+$ &c. a ſerie. Eſcreva-ſe c em lugar de $\frac{B}{A}$, a ſerie ſerá $A\,(1+c+cc+ccc+cccc+$ &c.$)$. Seja O huma grandeza $<A$, e tal que ſe poſſa deſprezar ſem erro notavel. Na ſerie 1 +


PRINCIPIOS
MATHEMATICOS.
LIVRO XV.
Definiçoens.
I. SE huma expreſſaõ admittir mais de hum valor, quando outra expreſſaõ admitte hum ſó, chamarſe-ha ella conſtante, e aquella, variavel.
II. A variavel que podér ſempre admittir valor maior que qualquer grandeza que ſe proponha chamarſe-ha infinita; e a variavel que podér ſempre admittir valor menor que qualquer grandeza que ſe proponha, chamarſe-ha infiniteſſima.

194 PRINCIPIOS MATHEMATICOS
III. $\Gamma(x, z)$ denota huma funcçaõ de x e z; por exemplo x^3z^5; $x^4-a^2xz+bz^3$ ſaõ funcçoens de x e z.
Definiçoens.
IIII. Eſcolhida qualquer grandeza, homogenea a huma raiz x, para ſe chamar fluxaõ deſſa raiz, e denotada aſſim dx; chamarſe ha fluxaõ de Γx, e ſe denotará aſſim, $d\Gamma x$, a grandeza que faria $\frac{d\Gamma x}{dx}$ conſtante, e $\frac{\Gamma(x+dx)-\Gamma x}{dx}-\frac{d\Gamma x}{dx}$ infiniteſſimo ou cifra, ſe dx foſſe infiniteſſimo, e conſtante tudo o que naõ depende de dx.
V. Toda a grandeza ſe chama fluente da ſua fluxaõ, e ſe denota eſcrevendo eſte final $\int$ immediatamente antes da fluxaõ.
VI. $d(dx)$ chama-ſe ſegunda fluxaõ de x; $d(d(dx))$, terceira fluxaõ de x, e aſſim por diante.

series and the modern notion of differential, more than three decades before Cauchy, making J. A. da Cunha one of the eminent predecessors of the reform of infinitesimal calculus carried out in the 19th century [14, 15].

Memoirs on the usefulness and on the foundations of mathematical sciences

The Lisbon Academy of Sciences, created in late 1779 and endowed with a printing press, began publishing its *Memórias* in 1797 with an article on applied mathematics, worthy of its motto "Nisi utile est quod facimus stulta est gloria" (If what we do is not useful, glory is in vain). Following the period of *Regeneração* in the 1850s which attempted to modernise the country and develop it economically, it continued with a new series that included some contributions to mathematics [11]. The new series included two memoirs by Daniel da Silva (1814–1878), in particular the remarkable *On the rotation of forces around the points of application* (1850, in Portuguese), which anticipated by a good quarter of a century a famous memoir by G. Darboux on the foundation of Statics.

The Academy also created the first Portuguese scientific journal, the *Jornal de Sciencias Mathematicas Physicas e Naturaes* (1866), which included a few mathematics articles, including an original article in Portuguese by Daniel da Silva on the amortisation of pensions in the main Portuguese survivors' funds (1867), and another one devoted to a comparative study of the population in Portugal (1869).

Periodical publications with articles in the mathematical sciences

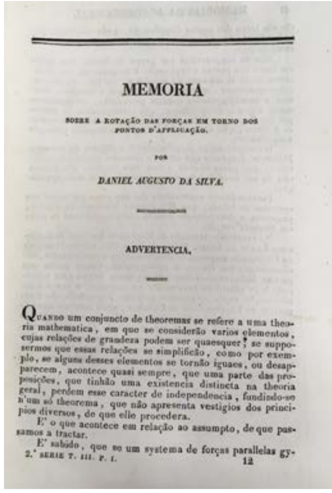
In 1853, the Coimbra University Press started publishing the scientific and literary periodical *O Instituto*, which included few mathematics articles mainly for didactic purposes. Although it was not


MEMORIAS
DA
ACADEMIA REAL
DAS SCIENCIAS DE LISBOA
DESDE O SEU ESTABELECIMENTO
EM 1780 ATÉ 1788.

SOLUÇÃO GERAL
DO
PROBLEMA DE KEPLER
Sobre a Medição das Pipas, e Toneis,
POR JOSÉ MONTEIRO DA ROCHA.

(1.) AS vasilhas, que ordinariamente servem para guardar, e transportar toda a especie de licores, são construidas de maneira, que podem sensivelmente tomar-se por solidos de revolução, compostos de dous troncos iguaes, e semelhantes, os quaes

Tom. I. A se


MEMORIA

SOBRE A ROTAÇÃO DAS FORÇAS EM TORNO DOS PONTOS D'APPLICAÇÃO.

POR

DANIEL AUGUSTO DA SILVA.

ADVERTENCIA.

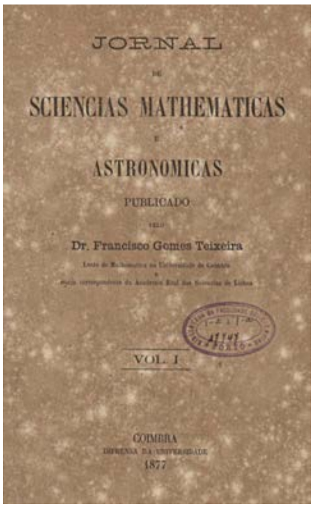
QUANDO um conjuncto de theoremas se refere a uma theoria mathematica, em que se considerão varios elementos, cujas relações de grandeza podem ser quaesquer; se suppozermos que essas relações se simplificão, como por exemplo, se alguns desses elementos se tornão iguaes, ou desapparecem, acontece quasi sempre, que uma parte das proposições, que tinhão uma existencia distincta na theoria geral, perdem esse caracter de independencia, fundindo-se n'um só theorema, que não apresenta vestigios dos principios diversos, de que elle procedera.

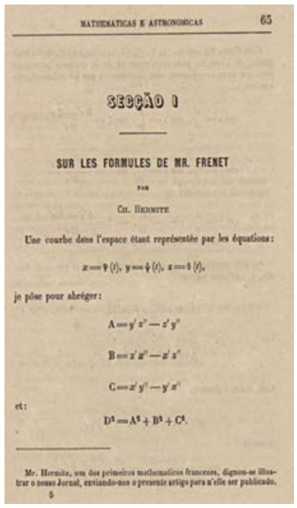
E' o que acontece em relação ao assumpto, de que passamos a tractar.

E' sabido, que se um systema de forças parallelas gy-

2.ª SERIE T. III. P. I. 12


JORNAL
DE
SCIENCIAS MATHEMATICAS
PHYSICAS E NATURAES
PUBLICADO SOB OS AUSPICIOS
DA
ACADEMIA REAL DAS SCIENCIAS DE LISBOA
TOMO I
NOVEMBRO DE 1866–DEZEMBRO DE 1867
A.3800.
LISBOA
TYPOGRAPHIA DA ACADEMIA
1868


JORNAL
DE
SCIENCIAS MATHEMATICAS
E
ASTRONOMICAS
PUBLICADO
PELO
Dr. Francisco Gomes Teixeira
VOL. I
COIMBRA
IMPRENSA DA UNIVERSIDADE
1877


MATHEMATICAS E ASTRONOMICAS 65

SECÇÃO I

SUR LES FORMULES DE MR. FRENET

PAR

CH. HERMITE

Une courbe dans l'espace étant représentée par les équations:

$x=\varphi(t),\ y=\psi(t),\ z=\theta(t),$

je pose pour abréger:

$A=y'z''-z'y''$

$B=z'x''-x'z''$

$C=x'y''-y'x''$

et:

$D^2=A^2+B^2+C^2.$

Mr. Hermite, um dos primeiros mathematicos francezes, dignou-se illustrar o nosso Jornal, enviando-nos o presente artigo para n'elle ser publicado.

5


UNIVERSIDADE DE LISBOA

REVISTA
DA
FACULDADE DE CIÊNCIAS

2.ª SÉRIE
A—CIÊNCIAS MATEMÁTICAS
VOL. IV

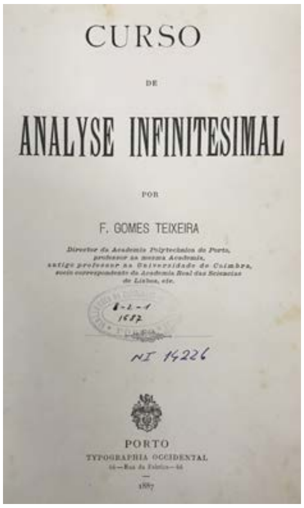
1954-1955
BIBLIOTECA DA FACULDADE DE CIÊNCIAS
Rua da Escola Politécnica
LISBOA

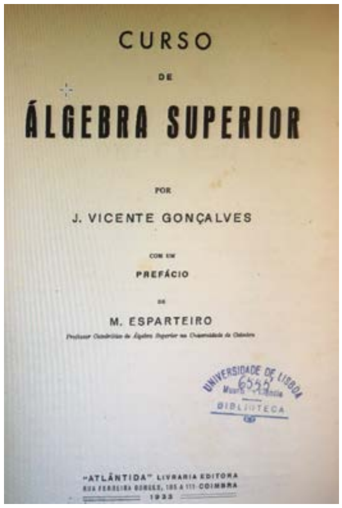
a mathematical journal and did not contain any relevant mathematics research contributions, it did publish the interesting short but deep text *Ensaio sobre os Princípios da Mechanica*, written in the 1780s by J. A. da Cunha, in which he proposed a clear separation between the Physical and the Rational Mechanics, based on the axiomatic method.

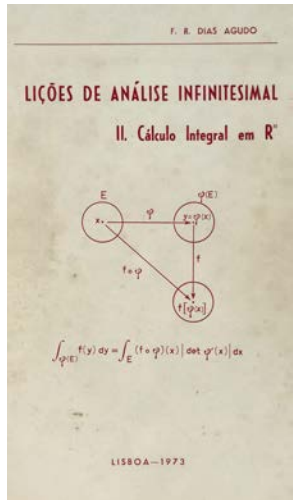
The first Portuguese periodical exclusively dedicated to mathematics, the *Jornal de Sciencias Mathematicas e Astronomicas* by Francisco Gomes Teixeira (1851–1933), started publication in 1877 in Coimbra; it was printed at the University Press but it was independent of the University. It published fifteen volumes, and in addition to internationally promoting the mathematical activity of Portuguese mathematicians, it included international contributions, in particular by the French mathematician C. Hermite. Later its publication was transferred to Porto, and in 1905 it was integrated into the *Annaes Scientificos da Academia Polytechnica do Porto.* Only more than thirty years later did Portugal get a second mathematical journal, the *Portugaliae Mathematica*. However, following the Faculties of Science of the Universities of Porto, of Coimbra and of Lisbon, each of which created their own scientific journal containing some mathematical papers (in 1927, 1931 and 1937 respectively), in 1950 the *Revista da Faculdade de Ciências de Lisboa* started their 2nd series – A, exclusively dedicated to mathematics [8]. It was directed by J. V. Gonçalves until 1966.

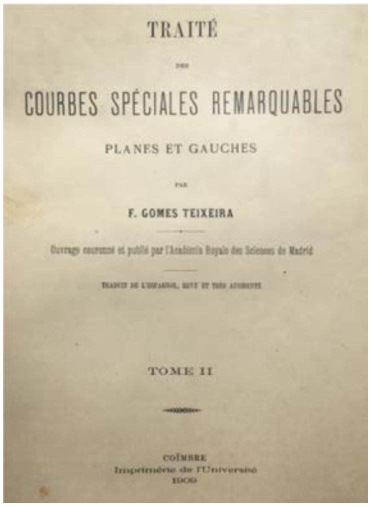
Higher mathematics textbooks in the transition from the 19th to the 20th century

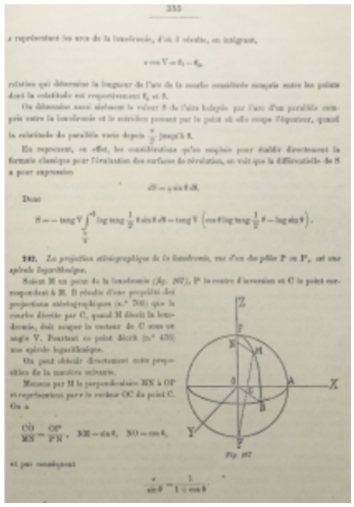
The publication of original university textbooks by Portuguese authors would only continue a century later with the *Curso de Analyse Infinitesimal* by Francisco Gomes Teixeira, which was first published in 1887 in Porto. Written for the students of the *Academia Polytechnica do Porto*, it was later re-edited and expanded to become the reference Portuguese treatise of Mathematical Analysis at the beginning of the 20th century. It was later republished as Volumes III











(1906) and VI (1912) of Teixeira's *Obras*, published by the Coimbra University Press.

That university practice of writing textbooks had resumed only in the next century, examples being the classical *Curso de Álgebra Superior* by J. Vicente Gonçalves, whose first edition was printed in Coimbra, in 1933, and the modern *Lições de Análise Infinitesimal* by F. R. Dias Agudo, printed by the *Tipografia Matemática* in Lisbon in 1973 among several other textbooks.

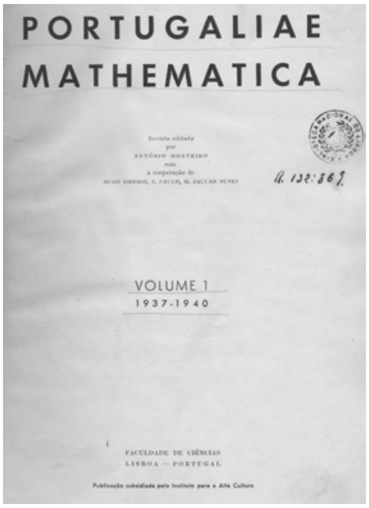
Historical surveys, expository and popularisation books

Francisco Gomes Teixeira was the most active and prolific Iberian mathematician of the 19th century. He corresponded with numerous European mathematicians of his time, and was the first rector of the University of Porto, from 1911 to 1918. He was also the author of the remarkable and unsurpassed *Traite des Courbes Spéciales Remarquables Planes et Gauches*, which went through several editions and occupies three of the seven volumes of his *Obras* with a total of more than 1300 pages.

During the short decade 1937–1946, mathematical activity flourished in Portugal. For example, the remarkable little book *Conceitos Fundamentais da Matemática*, by Bento de Jesus Caraça (1901–1948), a Mathematics Professor at the Economics Institute in Lisbon who was expelled from the University in 1946 for political reasons, had a first edition in 1941. This was the first Portuguese book aimed at the popularisation of mathematics written in view of the "cultura integral do indivíduo" with a historical and materialistic philosophy.

Journals of the Portuguese Mathematics Society

The foundation of the scientific journal *Portugaliae Mathematica* in 1937 in Lisbon, by António Aniceto Monteiro (1902–1980), who had returned the year before from Paris where he had completed



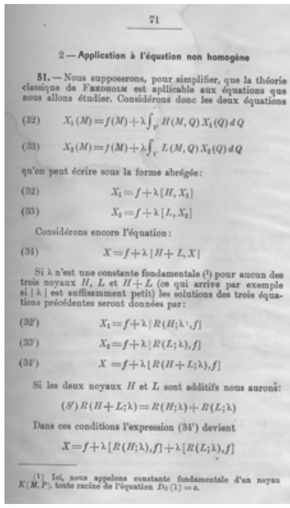
PORTUGALIAE
MATHEMATICA

Revista editada por ANTÓNIO MONTEIRO com a cooperação de HUGO RIBEIRO, J. PAULO, M. ZALUAR NUNES

VOLUME 1
1937-1940

FACULDADE DE CIÊNCIAS
LISBOA — PORTUGAL

Publicação subsidiada pelo Instituto para a Alta Cultura



71

2 — Application à l'équation non homogène

51. — Nous supposerons, pour simplifier, que la théorie classique de Fredholm est applicable aux équations que nous allons étudier. Considérons donc les deux équations

Considérons encore l'équation:

Si les deux noyaux H et L sont additifs nous aurons:

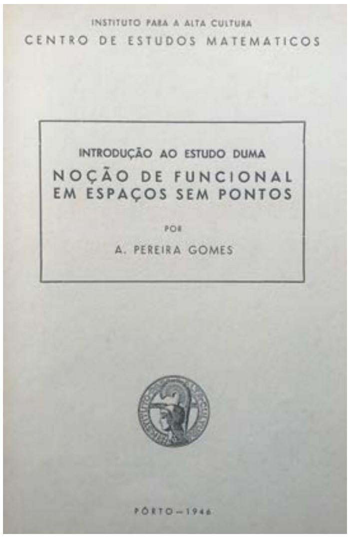
Dans ces conditions l'expression (34') devient

GAZETA
DE
MATEMÁTICA

PUBLICADA POR

A. MONTEIRO, B. CARAÇA, H. RIBEIRO, J. PAULO, M. ZALUAR

GAZETA
DE
MATEMÁTICA



INSTITUTO PARA A ALTA CULTURA
CENTRO DE ESTUDOS MATEMATICOS

INTRODUÇÃO AO ESTUDO DUMA
NOÇÃO DE FUNCIONAL
EM ESPAÇOS SEM PONTOS

POR
A. PEREIRA GOMES

PÔRTO — 1946

PORTUGALIAE
MATHEMATICA

VOLUME 9
1 9 5 0

Publicação subsidiada por
Publication subventionnée par
Publication sponsored by

JUNTA DE INVESTIGAÇÃO MATEMÁTICA • SOCIEDADE PORTUGUESA DE MATEMÁTICA

Edição de
«GAZETA DE MATEMÁTICA, LDA.»

PORTUGALIAE MATHEMATICA
Rua Serpa Pinto, 17, 4.º-E.
LISBOA (PORTUGAL)

HERMANN & C.ie, Éditeurs
6, Rue de la Sorbonne
PARIS (5me)

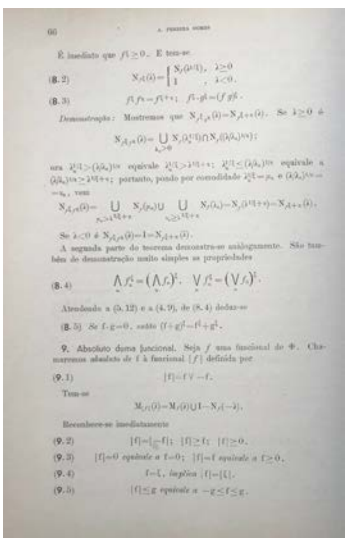
PORTUGALIAE
MATHEMATICA

ISSN 0032-5155

VOLUME 44
1987

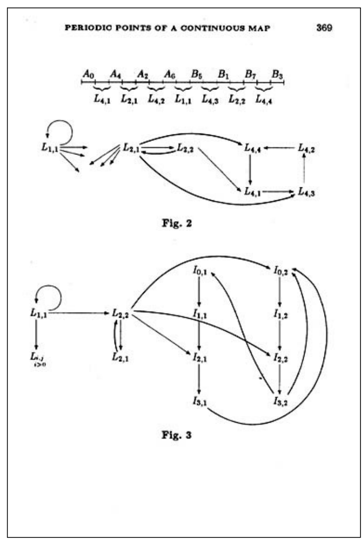
Edição da
SOCIEDADE PORTUGUESA DE MATEMÁTICA

PORTUGALIAE MATHEMATICA
Av. da República, 37 - 4.º
1000 LISBOA — PORTUGAL



66 A. PEREIRA GOMES

108 J. SEBASTIÃO E SILVA



PERIODIC POINTS OF A CONTINUOUS MAP 369

Fig. 2

Fig. 3

his doctoral degree under Maurice Fréchet, started a modernist movement also in Science with a certain "mathematical effervescence" that lasted for a short decade in Portugal [9]. It was followed by the *Gazeta de Matemática*, in 1940, which was later printed by the *Tipografia Matemática*, and by the creation of the first research mathematical centre in the country, the *Centro de Estudos Matemáticos de Lisboa*, the same year. That unique typography was established in 1945 and had a remarkable, although limited, activity in the Portuguese mathematical press for more than three decades until 1977.

Finally, we mention some examples of research publications, such as the two doctoral theses printed at the *Tipografia Matemática*: the *Publicação #18* of the *Centro de Estudos Matemáticos do Porto*, by A. Pereira Gomes (1919–2006), which was the first modern PhD thesis written in a Portuguese university, and *As Funções Analíticas e a Análise Funcional*, by José Sebastião e Silva (1914–1972), both published in *Portugaliae Mathematica*, respectively in 1946 and in 1950. In his work, J. Sebastião e Silva began work on some deep contributions to Functional Analysis, which would later lead him to introduce in 1955 an important class of locally convex spaces as inductive limits of an increasing sequence of normed spaces with compact inclusions [12], later called the Silva LN*-spaces.

After five centuries of existence, the mathematical typography using movable type, with its specific and distinctive aspects such as tables, figures, diagrams and mathematical formulae, gave way to electronic publishing driven by the TeX program created by Donald Knuth in 1978 [7]. This electronic composition system was adopted by the American Mathematical Society five years later, and was used in 1987 by the *Sociedade Portuguesa de Matemática* (SPM) for the publication of the fiftieth anniversary volume 44 of *Portugaliae Mathematica*, which was chosen for the last book of the Leiria exhibition in order to complete this five-century-long retrospective of Portuguese mathematical typography, from 1496 until 1987.

References

[1] L. Albuquerque, *Introdução*, to the *fac-simile* edition of the *Almanach Perpetuum*. Imprensa Nacional-Casa da Moeda, Lisboa (1986)

[2] B. Almeida and H. Leitão, *Pedro Nunes (1502–1578) mathematics, cosmography and nautical science in the 16th century*, pedronunes.fc.ul.pt/index.html (2009)

[3] A. Alselmo, *Origens da imprensa em Portugal*. Imprensa Nacional-Casa da Moeda, Lisboa (1991)

[4] J. Chabàs and B. R. Goldstein, *Astronomy in the Iberian Peninsula: Abraham Zacut and the transition from manuscript to print*. Transactions of the American Philosophical Society 90.2, American Philosophical Society, Philadelphia, PA (2000)

[5] A. J. Franco de Oliveira, Anastácio da Cunha and the concept of convergent series. *Arch. Hist. Exact Sci.* **39**, 1–12 (1988)

[6] J. A. Gaspar and H. Leitão, How Mercator did it in 1569: from tables of rhumbs to a cartographic projection. *Eur. Math. Soc. Newsl.* **99**, 44–49 (2016)

[7] D. E. Knuth, Mathematical typography. *Bull. Amer. Math. Soc. (N.S.)* **1**, 337–372 (1979)

[8] J. F. Rodrigues, Portuguese mathematical journals: some aspects of (almost) periodical research publications. In *The Practice of Mathematics in Portugal*, edited by L. Saraiva and H. Leitão, Acta Univeritatis Conimbrigensis, Coimbra, 601–627 (2004)

[9] J. F. Rodrigues, António Monteiro and his modernist essay. *CIM Bulletin* **42**, 3–10 (2020)

[10] J. F. Rodrigues, The Nunes' spiral and the prehistory of the infinitesimal calculus (in Portuguese). In *33rd Portuguese National Seminar on the History of Mathematics (October 2020)*, Boletim SPM (in press)

[11] L. M. R. Saraiva, Mathematics in the *Memoirs* of the Lisbon Academy of Sciences in the 19th century. *Historia Math.* **35**, 302–326 (2008)

[12] J. Sebastião e Silva, Su certe classi di spazi localmente convessi importanti per le applicazioni. *Rend. Mat. e Appl. (5)* **14**, 388–410 (1955)

[13] *Tipografia Matemática Portuguesa: 1496–1987* – Catalogue (Portuguese), ipleiria.pt/sdoc/tipografia-matematica-portuguesa-1496-1987/

[14] A. P. Youschkevitch, J. A. da Cunha et les fondements de l'analyse infinitésimale. *Rev. Histoire Sci. Appl.* **26**, 3–22 (1973)

[15] A. P. Youschkevitch, C. F. Gauss et J. A. da Cunha, *Rev. Histoire Sci. Appl.* **31**, 327–332 (1978)

José Francisco Rodrigues is professor at the Mathematics Department/ Faculty of Sciences of the Universidade de Lisboa, where he graduated in 1978 and obtained his PhD in 1982. He had studied also at the Université de Paris VI and spent a short post-doc visit at the Universität Bonn. As vice-president and delegate of the Portuguese Mathematical Society, he participated at the 1990 founding meeting of the EMS in Madralin. He was director of the Portuguese Centro Internacional de Matemática, and he is a member of the Academia das Ciências de Lisboa. His field of research is in nonlinear analysis and PDEs, mainly free boundary problems. He has also interests in the relations of mathematics with other subjects, including history, music, architecture and culture.

jfrodrigues@ciencias.ulisboa.pt

Remembering Andreas Griewank

Tsou Sheung Tsun
with help from Juan Carlos De los Reyes, Michel Théra and Andrea Walther

The Committee for Developing Countries was very sad to learn of the untimely and sudden death of long-term member Andreas Griewank. A lot of regrets were poured through the CDC network on learning the news.

- *Andreas has been an active member of this committee for many years, his enthusiasm was contagious and his activities widely spread in the world. He will be missed by many.*
- *I have such good memories of Andreas, his energy and enthusiasm, and his hospitality in Berlin for CDC.*
- *Many people among the actual and past CDC members know him well and testified to his worldwide actions, his enthusiasm, the nice person he was. He will be missed by the mathematical community.*
- *I feel very sad to learn that Andreas is no more. He was a very friendly and active member of our committee for many years. I liked his sense of humour. He was very efficient when he organised a meeting of CDC in Berlin in 2014. I miss him.*

And even some future plans:

- *… we were even talking about visiting the Ecuadorian university some time ago.*
- *…*

So we connected Andreas with good humour, activities all over the world, enthusiasm, and we miss him a lot!

Let us first give voice to the long-term CDC member who was closest to Andreas in research: Michel Théra, professor emeritus at the University of Limoges, and Adjunct Professor at the Federation University Australia.

Photograph of Adreas Griewank taken in Ecuador shortly before his sudden death

Andreas obtained his PhD "Analysis and Modification of Newton's Method at Singularities" from the Australian University, Canberra under the direction of R. P. Brent and M. R. Osborne in 1980. Then he was a postdoc for one year at the University of Cambridge, became an assistant professor at Southern Methodist University, Dallas until 1987. After this, he moved to Argonne National Laboratory as a Senior Scientist, and then joined the Technische Universität Dresden as a Professor before moving to Humboldt University of Berlin. When he retired in 2015, he became Dean of the School of Mathematical Sciences and Information Technology at Universidad de Investigación de Tecnologia Experimental Yachay in Quito.

During his academic career, Andreas' research focused on algorithmic differentiation and iterative methods for nonlinear optimisation, for which in 2017 he was made a Fellow of the Society for Industrial and Applied Mathematics.

For non-specialists, let us briefly describe Andreas' main research interest. Many derivative arrays such as gradients, Jacobians, or Hessians are essential tools for computational purposes in various areas. This is the case for instance in numerical optimisation when using a gradient descent method, or Newton's method. When computing derivatives by hand or using computer algebra systems, one may observe that the formula for the derivative grows very rapidly in a combinatorial way. Andreas was one of the leaders of what is called automatic differentiation (AD), which is a technique that allows the computation of derivatives of any order for a function specified by a computer program. By avoiding the manipulation of complex function formulas with exponential growth,

Andreas chairing the round table at the ICIAM at Berlin in 2007

AD is very useful in various areas of applied mathematics, including numerical optimisation and also computational fluid dynamics, atmospheric sciences, and engineering design optimisation or more generally when working on computational simulation models.

Andreas spent a sabbatical year as researcher at INRIA Sophia Antipolis, Antibes, France in 1988–1989.

Among other honours, he won the Max-Planck Research Prize for International Collaboration in 2001.

In 2006, as a member of the jury of the habilitation thesis of Darinka Dentcheva, Andreas welcomed me very warmly to his home in Berlin. I often recall this stay with pleasure.

For those of us whose mathematical interests are far removed from optimisation, this gives us a more rounded picture of our friend and colleague. Thank you, Michel T.!

Prof. Andrea Walther from the Department of Mathematics, Humboldt-Universität zu Berlin, kindly sent us the official obituary, of which the following excerpts give some more details and further insight into Andreas Griewank's mathematics and achievements.

Andreas was born in Kassel on 26 January 1950. He passed away suddenly and unexpectedly on 16 September 2021, at the age of 71.

In addition to his contributions to AD, Andreas continuously made important contributions to the design and analysis of nonlinear optimisation algorithms. Of the many accomplishments with regard to very different aspects of nonlinear optimisation, just a few are mentioned here. First, the idea of partial separability was developed by Andreas jointly with Prof. Philippe Toint. This structural property is ubiquitous in optimisation problems, and can be exploited to greatly improve the efficiency of algorithms. A second example is his contribution to the convergence theory of Newton and quasi-Newton methods in different settings, including the infinite dimensional setting and the degenerate setting, in which the Hessian is singular the optimum. These topics were the subject of a series of papers that are still frequently cited, and are the basis of ongoing research. Third, the "Griewank function" serves as an academic test function in global optimisation. This function continues to be used widely in the global optimisation community, and is the subject of renewed interest because of the need for non-convex optimisation to minimise objectives in data analysis applications, including deep learning. Andreas' scientific work was always marked by an abundance of ideas and infectious enthusiasm.

Andreas made also significant service contributions to several different communities. He organised numerous conferences and workshops worldwide. Often, the aim of these meetings was also to connect academic researchers with practitioners in the areas of AD and nonlinear optimisation.

Promoting young researchers was always close to Andreas' heart. He supervised 23 doctoral students and numerous masters students from all over the world. He had a special interest in promoting and supporting mathematical education in developing countries.

At some point, we found out that Andreas was very good at organising round tables, and he was thus roped in to organise a couple of them during the EMS annual meetings in Amsterdam and in Krakow, and also during our own CDC meetings. We found that his secret was meticulous preparation beforehand, and he also got his then graduate student Levis Eneya (from Malawi) to help him with much of the correspondence, chasing-up of invitees and so on. He must have been pleased to see Levis elected president of SAMSA in 2012.

For those of us who attended the CDC annual meeting in 2014, organised and invited by Andreas, it was a most pleasant time and everything was well organised. We had a round table discussion in the headquarters of the IMU, and we were invited by him and his wife to a nice visit to his family home (with excellent food of course).

EMS-CDC meeting at ICTP at Trieste

In meetings, Andreas was quite the "trouble-maker", since he questioned many proposals and quite often disagreed with others. But he always had a good reason for doing so, and he was of course a great respecter of democracy, for after a heated discussion, he usually ended up by saying that "if you all agree, of course I'll go along". He said also: "So I'll step down from my nitpicking Oberlehrer soapbox," in one email after some exchanges of opinion. Another exchange of emails was for the written report of a (not EMS) conference; in response to "Just a private word to say how I liked your report (candid and all!)", he replied "I have a hard time orienting myself in a world of superficial politeness. I hope you got the full report …."

One heroic thing he did for CDC was to collect and distribute free mathematical books. It was at the EMS conference in Krakow 2012, and as usual, there were booksellers exhibiting their books, and very often their representatives are more than willing to give away the unsold books after the conference. So Andreas collected 100 books and took them back to Berlin, which he was able to do because he had come in a car with some others. With the help of his secretary Jutta Kerger, these books were catalogued, and he used his ingenuity (writing to embassies, etc.) to distribute them to various developing countries, sometimes by taking them in his luggage and delivering them by hand. It was really a success, so much so that he did it again after another big international conference (on applied mathematics).

Of course, his activities were not confined to those of CDC. For a number of years, he was the German representative with CIMPA (Centre International de Mathématiques Pures et Appliqueés). After he went to Ecuador, he organised a CIMPA School with Marc Lassonde in 2017 at Yachay Tech. Didier Aussel, also a CDC member, was the CIMPA representative responsible for the school. His comment was "we had a good time together during this school. Andreas was so enthusiastic!!" At some point, Andreas was also involved in the work of AIMS, particularly with their new centre in Senegal. And being German, he was also active in the work of DAAD. He was one of the few colleagues who visited Cuba several times, and even heard *viva voce* some of Castro's famous marathon speeches!

Let us end with a warm and moving account from his Ecuadorian colleague Juan Carlos De los Reyes, Professor and Director, Centro de Modelización Matemática (MODEMAT), Escuela Politécnica Nacional de Ecuador.

Andreas landed in Ecuador in 2015 to take part in a new university oriented mainly towards scientific research: Yachay Tech. Apart from his academic interests, Andreas was motivated to support a socially oriented left-wing government, very close to his personal convictions. In fact, from the beginning, Andreas dedicated a significant part of his salary to supporting low-income students through scholarships.

EMS-CDC meeting at Limoges (Andreas is second left in back row)

Andreas' personality was felt from day one. He was fully involved in the university development as Dean of the Faculty of Mathematical and Computational Sciences. During this period, the Mathematics curriculum, with a strong computational component, was established, as well as the Computer Science curriculum, oriented towards modern topics in machine learning and artificial intelligence. Andreas also made a significant effort to establish doctoral programs in the institution, a project which unfortunately he did not succeed in realising. Giving direction to a brand-new university is not easy, and even less so in a country with a lot of political turbulence. Andreas faced opposition from various sectors and had to fight quite a few battles for what he believed to be the best for the future of Yachay Tech.

His period in Ecuador undoubtedly left a strong trace on several of his students, who turned to mathematics with the passion that only Andreas knew how to sow. His time in South America was also productive in terms of research, leaving as a legacy his rigorous work on optimality conditions and automatic differentiation for optimisation problems with piecewise smooth functions.

What a moving portrait of the Andreas we knew and now miss!

Tsou Sheung Tsun (B.Sc. Hong Kong, MA Oxon, D. es Sc. Geneva) is a retired faculty member of the Mathematical Institute of the University of Oxford. She was Chair of the Committee for Developing Countries of the EMS, and President of CIMPA. She was a member of the Strategy Group for Developing Countries of the IMU, and of the Foreign Affairs Committee of the London Mathematical Society. Her research interests are in mathematical physics, in particular gauge theories including the standard model of particle physics. Her most recent work is on CP violation in particle physics, especially from a mathematical point of view. She is the editor, together with Jean-Pierre Françoise and Greg Naber, of a 5-volume Encyclopaedia of Mathematical Physics.

tsou@maths.ox.ac.uk

Report: EWM panel discussion on gender balance in mathematics at the European Congress of Mathematics

Eugénie Hunsicker

A European Women in Mathematics-sponsored online panel on Gender Balance in Mathematics took place on 22 June, 2021 as part of the ECM. As a long-time activist and advocate for diversity in mathematics, I was delighted to have been asked to chair the panel. The panelists were also a group of committed diversity advocates in a range of leadership roles within our community:

- Dr. Shabnam Beheshti (Director of Education, Department of Mathematics, Queen Mary University of London and researcher in mathematical relativity),
- Prof. Klavdija Kutnar (8ECM Organizing Committee Deputy Chair, Professor of Mathematics and Rector of University of Primorska, and researcher in algebraic graph theory),
- Prof. Volker Mehrmann (president of EMS, and researcher in numerical analysis),
- Prof. Jill Pipher (past-president of AMS, and researcher in harmonic analysis and PDE),
- Prof. Carola-Bibiane Schönlieb (former EWM convenor, and researcher in applied PDE, inverse problems and mathematical imaging).

The topics we discussed in the panel covered the span of issues of policy, practice and culture that research suggests are relevant for driving the gender-biased outcomes we observe. These include topics such as recruitment, promotion and the impact of family commitments, but also structural barriers in the shape of mathematical careers and cultural biases in the ways that success is judged within mathematics – topics that we need to discuss in our community if we are to make further progress. I hope the ideas presented in the panel encourage further discussion.

Chair: *Tell us about the current status of gender equity in mathematics in your country. Where has there been progress over the past 10 years? Where are the biggest challenges remaining?*

Shabnam: One of the most significant improvements I have seen in the UK is the recognition that decisions about policies and practices have measurable consequences in terms of diversity, and that relevant data should be analysed and benchmarked by discipline to understand this. There needs to be a broader awareness of the value of these analyses in getting people on board at the policy level.

However, although the number of women in senior levels is increasing in the UK, it isn't happening rapidly enough. One problem is that underrepresented groups are under tremendous pressure to be visible and mentor others, which can be in direct tension with the promotion and progression of the mentors because of the way we place value for promotion on different criteria and activities. Another factor is that people on teaching-based versus research-based contracts get siloed away from each other, which is unhelpful.

Klavdija: I am from Slovenia, where there is gender balance in terms of students studying mathematics at bachelors and masters level. The situation begins to skew towards men at PhD level, and is even worse for women employed in mathematics departments. At my university, one quarter of department members are female; the proportion is even lower at other universities in Slovenia. There are only four female full professors in the whole country; this is progress – three of these full professors were promoted in the past 10 years! Also in the mathematics, physics and astronomy society in Slovenia, we have established an active committee for women. The biggest challenge is that we have quite a few women with excellent research records who have the title of Assistant Professor but only the salary of a teaching associate.

Volker: In Germany, political policies have been set so that the gender balance in mathematics should in principle increase drastically and quickly, through rules about gender bias and inviting people to interview. However, the number of women professors has not improved very much. This is mainly because the number of female applicants is still very low – likely due to the terrible timeline in hiring. Professors are usually hired when they are between 35 and 45 years old. Before that, you must keep moving between temporary positions, which is particularly difficult for women with families. Another challenge is that if you follow the rules on committee representation strictly, the few women professors we have are overburdened with committee work. So, although the national policies are good, given the hiring situation at universities, I am not too optimistic that gender equality will improve quickly.

In the EMS, the executive committee is now 40 % female, and together with the Women in Mathematics Committee, we have come up with an explicit plan to substantially increase female representation in editorial boards over the next few years. The organising committees for this congress have also done quite a good job with invited speakers, continuing the trend towards improvement of recent years. However, we are still in a bad situation with respect to the prize winners, where equity has not improved. This is something where I don't have a clue what to do. We can certainly make sure that more women sit in the organising and prize committees, but that would be another burden on them, so there is a catch.

JILL: I want to start by saying this is a really timely moment for discussion of these issues. There is already a fair amount of research that shows that the effects of the pandemic are especially hard on women in science. In fact, a recent article in Nature showed that submissions by male authors to the arXiv increased during the pandemic[1], unlike submissions by women.

To talk about the situation in the US, I am going to back up a bit. Nationally, in 1977, 13 % of mathematics PhDs were awarded to women. By 1993, this had climbed to 28 %, so real progress was made over this time. Since then, it increased to 33 % in the early 2000s, but now the numbers are again at 29 % – in the past 30 years there has been really no progress. Why is that? We know that work/life issues have a greater effect on women. The career timeline in the US, starting from the PhD, is that you are expected in about 7 years to reach a tenured or permanent position, and these are the same years during which women are creating families as well as careers. Even the adoption of gender-neutral family leave practices in academia has been shown in some disciplines to privilege men[2]. As Dean Alison Davis-Blake in Michigan, puts it, "Giving birth is not a gender-neutral event." Finally, a 2020 National Academy of Sciences report[3], had a really sobering conclusion: bias, discrimination, and harassment are major drivers of underrepresentation. I think it is really difficult for well-intentioned, caring people in our profession to face such a conclusion, but face it we must, or I feel that we'll make no further progress.

CAROLA: Since I moved to the UK about 10 years ago, many actions in favour of women mathematicians have emerged in the UK, such as the national Athena Swan Charter. Many student societies have also popped up. Moreover, hiring started to change so that search committees actively encourage female candidates they want to apply. So, the situation has changed in terms of actions and awareness, but there is certainly not equity yet. There is a long way to go, particularly the higher you go up in the profession, and this starts already at PhD level. One of the biggest challenges is that family and housework are not considered equal endeavours between two people in a couple.

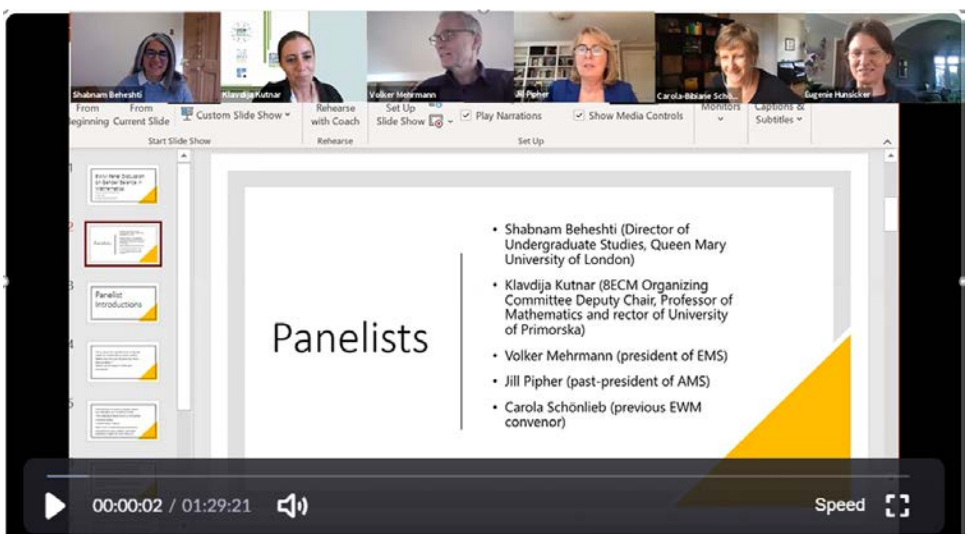
On the European level, I am the former Convenor of the EWM, and we have AWM in the US. The fact that these associations still exist and are super active is a sign that there is still a lot of work to do and that young female mathematicians are looking for networks to exchange experiences. When Elena Resmerita and I were active as Deputy Convenor and Convenor, we also tried to collect gender ratios in mathematics by countries, which was a sobering exercise[4].

[1] nature.com/articles/d41586-020-01294-9

[2] iza.org/publications/dp/9904

[3] nap.edu/read/24994/chapter/1

[4] For statistics from 2005, see womenandmath.wordpress.com/past-activities/statistics-on-women-in-mathematics/



From left to right, panellists Shabnam Beheshti, Klavdija Kutnar, Volker Mehrmann, Jill Pipher, Carola-Bibiane Schönleib, and panel chair, Eugénie Hunsicker.

Chair: *Interventions to improve gender equity can take place at a variety of levels: at the individual department or institution, in national policy, or in mathematics culture. What sorts of interventions have been attempted in your country, and what evidence is there for their success?*

Shabnam: At the individual level, it is the right time for a conversation on intersectionality between gender issues and issues around decolonisation of the curriculum. For example, even if the development of quantum mechanics was influenced directly by the works of Bohr and Schrödinger, there are many people on the planet now who are working in the field who don't "look like them". Our students need to see dotted lines of work going into the future and see themselves as participating in a diverse, living discipline.

At the department level, what we have done well, and which has become easier now with online teaching, is to allow schedules to accommodate personal lives. Also, be aware of administrative allocations being made with a gender bias, such as putting women primarily in teaching-related roles, which are then not valued as highly in promotion applications. This is a place where just looking at data helps.

At a university or professional society level, ring-fence funding to help researchers keep their research going after a particular gap, not just childcare; this may also help address two-body problems. Learned societies have already helped to tackle this. For example, I was a recipient of the LMS Grace Chisholm Young fellowship, which supports mathematicians who have had to take a break from formal employment for any of a variety of reasons. As a recipient of this Fellowship, I was able to establish myself as a mathematician and find a new position after relocating to the UK for family reasons. Finally, professional and learned societies also have a role to play in advocacy and data gathering for policy-making at national level.

Klavdija: Individual departments are critical. I reached my position of Rector because the mathematics department at my university supported me, first to the level of Dean, then to the level of Rector. I would like to see other universities in Slovenia follow suit. I am now involved in making national policies, for example concerning national scientific awards. We don't have enough women candidates to select a balanced set of awardees between women and men as we are required by law to do. Also, we are currently preparing a new act for research and innovation, involving several committees at national level, which again must be gender-balanced. The policies are in place – departments now need to recognise female researchers and suggest them for these awards and positions.

Volker: At the institutional level, we have to give stronger support to the lower levels in the leaky pipeline. For example, we are organising days for girls at our university, to get more girls interested in studying mathematics and science. Then we need departments and universities to support women at the PhD level so they have a stronger track record when they apply for positions, for instance with extra grants so that they can go on research trips and so on, I think that has a positive effect. We have tried this in Berlin.

In terms of national policy, it has had a positive effect that about 9 % of the budget of each university depends on meeting gender balance quotas, which is an incentive to hire women. The countereffect is that universities that focus primarily on the humanities have greater success in reaching this goal than technical universities that focus primarily on science and engineering subjects.

Jill: In terms of mathematics culture, I want to focus a little on the stereotypical assumptions about mathematics careers and mathematicians. There are a lot of studies that show that the perception that women are not as gifted or talented at maths, and that this is what is required to succeed, is very damaging. We also know from many studies that role models matter. This causes its own problems, however, because the same senior women are called on over and over to serve as role models, which hurts their careers.

Like Klavdija, I have been strongly supported by my department in obtaining leadership roles in my department and university. The experience of being supported is phenomenally important.

Concerning strategies at a policy level, I think that we need leaders who acknowledge the importance of addressing the problem of gender underrepresentation and resources that are allocated to address equity issues, such as for recruitment, teaching, training, creating transparency, collecting data, and so on. Actions like recognising implicit bias and reading and writing letters of recommendation with an eye for gendered language all require effort and attention and training. Mathematicians, like all academics, need reminding that this is necessary and important and they need support for doing it.

Carola: As both Klavdija and Volker mentioned, you can't give prizes to women if they aren't nominated. It is crucial to work on increasing the number of nominations of women mathematicians for the prizes of national, European, and international societies, which need to be active in encouraging this. The LMS is very aware of this and has really improved things. On the European level, both the EMS Women in Mathematics Committee and the EWM are working hard to encourage nominations of strong women mathematicians. The AWM also has a prize committee that makes similar efforts, as many societies should.

Shabnam: I would like to raise here the topic of mentorship. People tell me, "You need a mentor to do X," and I say, "I don't need a mentor. What I need is an advocate who is in the room when people are making decisions and who will stand up for the quality of my work." Mentorship has its time and place, but I wonder if we have to move from local practice to policy when we are talking about people who are in permanent positions but don't progress.

Volker: I fully agree about the importance of advocates. Being in committees all the time, I see that people advocate for their own students and field, but not so much for young women in other fields. If you are lucky, you get enough nominations of women for prizes, and even better, a few EMS prizes to women. We need to appeal to the community to think beyond subdisciplines, and to work on looking for the best women to whom to award prizes, independently of field, rather than focusing on the fifteenth person in line in your own subarea.

Jill: I couldn't agree more that there are policy issues and that mentorship will not solve everything. The AMS prize oversight committee was founded particularly to accomplish this – to look at our processes and develop ways to arrive at a more diverse set of prize winners. There are undoubtedly things that occur in all processes that we don't see if we don't look at the data that come out of them. We need to take a look at the data on outcomes for prizes, hiring, promotion and so on and think about whom we are leaving out.

Chair: There are reports coming out now that focus on mid-career as a place where there is an issue[5]. For instance, in the UK, the pipeline in mathematics is roughly 23 % women from PhD up to professor, then it plummets to 13 % among professors. There is a tendency to focus policy and support on people early in the pipeline, to try to get it right for the future, while forgetting those further down the pipeline. We need to think about how to help people who are struggling now wherever they are, rather than always trying to start over to get it right with the next generation.

Chair: *The "Matthew Effect"[6], whereby funding tends to accrue to individuals who have had early and continuous funding success, has been seen to disadvantage women especially. How can we start to tackle this problem in Europe? Do we need to change funding or peer review culture, or both?*

Shabnam: There is certainly the idea that "success begets success". I would propose that funding agencies allow more people to keep their research ticking over to keep the mathematical community alive, instead of devoting large amounts of funding to only a few individuals, who are then supposed to become leaders in their field. I am very curious to see what uptake you would get in terms of gender balance if instead of having half-a-million-pound grants, they were divided by four and funded more people.

Klavdija: I am very familiar with the concrete implications of the Matthew Effect. It took me quite a while to be successful in obtaining funding for a project at the Slovenian Research agency. I was finally awarded a grant three years ago, the first maths project led by a female.

For Europe, I think the funding situation will change because, as you probably know, every institution that wants to apply for EU grants needs to prepare an action plan to improve gender balance and equal opportunities. Universities are therefore forced to consider this problem at the institutional level in order to apply for EU grants.

Volker: The European Research Council (ERC) is a wonderful example of the Matthew Effect. The money is going to the strongest individuals and only to those. I think it is also very common in many countries. This is not primarily about gender. We have a saying in Germany, "The devil always shits on the biggest pile". This is the statement that describes the situation very well. I see a colleague at the panel from the ERC online, and maybe she can say something.

Maria Gonzalez (ERC): I am a woman mathematician, but I left mathematics research and for the past seven years have been a scientific officer for the ERC working on panel coordination. Based on my observations, I believe that culture is the main thing you have to change. There are certain beliefs about what pure or applied mathematicians do or what statisticians do, and sometimes the mathematics community forgets that you are a single community that should be working together rather than fighting with each other.

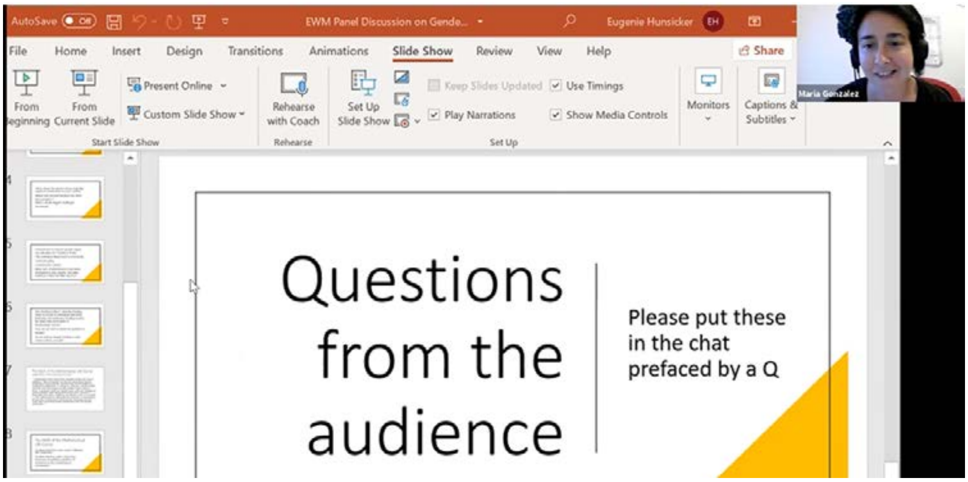
Some of the problems you mention are built into the culture, such as desired trajectories: if someone is a speaker here or there, or publishes here or there, it gives them a certain status. In the past 2–3 years, we have had more women on panels, and I have heard more discussion about different ways to evaluate applications and candidates. I think this is a good sign. I believe in working from the inside.

At the ERC, it is the panel [of mathematicians] that decides on the funding – if there is such a thing as the Matthew Effect, it is coming from something in the culture of the community. But I also think the ERC and the community are jointly responsible for spotting these problems and then working to improve them. Maybe I am overly optimistic, but I think if we start with you, the community of mathematicians, and we forget about fighting among our subcommunities and forget about our biases and accept this as a common responsibility, then things may change.

Chair: We hear the same comments in the UK from the EPSRC Mathematical Sciences panel: members of the mathematical community are the ones making the decisions. But I think there is also confusion in the community about to what extent and how we can change the way we do reviewing and run panels. We have

[5] chronicle.com/article/the-associate-professor-trap; advance-he.ac.uk/knowledge-hub/mid-career-academic-women-strategies-choices-and-motivation-final-report-activity

[6] dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/studien/studie_gender_effects.pdf; pnas.org/content/115/19/4887



Maria Gonzalez, from the European Research Council.

always done it in a certain way, which we assume is okay because we have always done it that way. If we do it in a new way, is that okay, or will we go against some agency policy? How members of the mathematics community can shape the way that the funding policies are enacted is an important question.

Jill: In the US, the majority of funding for mathematics comes from the NSF, and within the NSF, the division of mathematical sciences competes for its funding with other sciences. So the pressure to fund large and new initiatives and award big money to specific places is tremendous. However, the Division of Mathematical Sciences (DMS) devotes a significant part of its budget to funding 6 or 7 mathematics institutes in the country, and through these institutes, it is possible to disperse smaller awards to many, many mathematicians, bringing them to programs, conferences and workshops. I think the institute program within the DMS has been a remarkably successful one, especially since the number of institutes has expanded from three in 1999 to over twice that now.

In terms of peer review, I think that even in situations where it seems difficult or unlikely to succeed, we should really consider double-blind review, not just for review of grants, but also for articles. The argument against this is that people put everything on the arXiv and therefore reviewers know who the author is within a few pages. But I would say that even so, if your first impression is the mathematics, not the person, not where they are from, this creates a different place to start. Even if you know after five pages who the author is, still you started in a different place.

Carola: The Matthew Effect is to some degree related to the way that your track record will influence your future opportunities, which exists in all professions and is not a totally bad thing. But it does cause problems for people whose career path is not standard and, for example, start their careers later than usual. I think we need to change reviewing procedures and how we are educating reviewers to take account of nonstandard careers.

Chair: *In "Women Becoming Mathematicians", Margaret Murray*[7] *talks about "The Myth of the Mathematical Life Course", which is characterised by early recognition of talent and an uninterrupted path to professional success. To what extent has your career followed this trajectory? To what extent is such a trajectory necessary to achieve a position of eminence in the mathematical community?*

Shabnam: My career has not followed this trajectory in any way, shape or form. But here's the important thing. Even though it hasn't, on paper I can make it look like it has, which has been important for successfully navigating through each next stage in my career. I've come to the realisation, from the comfort of a permanent position, that the further my trajectory moved from the "standard one", the harder I had to fight to retain my identify as a mathematician. It wasn't that I had difficulty *being* a mathematician, but rather I had difficulty staking my own claim on what a mathematician looked like.

Carola: Early on, when I started a PhD, things went really, really badly. I got very discouraged and almost quit. I would have done so, if our head of department hadn't said, this isn't about you or your skills, it is just the wrong situation for you. Leave, go somewhere else. I lost quite a bit of time. But from the time I started my PhD again in the right community, I have been really lucky, and things have gone very well.

Volker: Let me be a little provocative. I haven't followed this trajectory, but I am a man and an applied mathematician. I am not a gold medallist of the mathematical Olympiad, and I haven't been

[7] ams.org/notices/200107/rev-green.pdf

in the career hot-spot driving seat, as many of the colleagues we see [at the Congress] are. So maybe we should get rid of the Olympiads and prizes altogether in order to avoid having this hotspot-straight-on-career idea still drive what we do in mathematics.

Chair: I am fully in favour of being provocative! This is a question about the culture – how do we change the cultural expectations? What you implied is that it may not be necessary to follow this trajectory to become a mathematician, but most of those who have become prize winners of eminence have done so. This suggests that when we select people to nominate for prizes, the expectation is that nominees will have had careers like this, which is a very gendered trajectory. So maybe we also need to change what we are looking for in prize winners.

Jill: The notion that mathematicians are somehow born to do this or must be on a track on an early age keeps a lot of people out of the profession. I can point to a lot of well-known and successful people who have not been on this trajectory. I was just at a reception for the Joan Birman Fellowship for mid-career women that she endowed at the AMS. She got her BSc in 1948, and her PhD in 1968: a very different trajectory, followed by a fantastic mathematician. We really do have to do something about this cultural construct that there is a single path to success in mathematics.

—

Eugénie Hunsicker is a Reader in Mathematics working at Loughborough University in the UK. After studying mathematics at Haverford College, she received her PhD in Mathematics from the University of Chicago, in the US. She has recently finished six years as the Chair of the London Mathematical Society's Women and Diversity in Mathematics Committee, and she is current Chair of the Athena Forum, an organisation of STEMM professional societies in the UK collaborating on equity, diversity and inclusion (EDI) in UK academia, hosted by the Royal Society. She has also recently been appointed as Honorary Officer for Equality, Diversity and Inclusion for the Royal Statistical Society, and has advised a wide range of professional and other organisations on EDI matters.

e.hunsicker@lboro.ac.uk

Mathematics publications and authors' gender: Learning from the Gender Gap in Science project

Helena Mihaljević and Lucía Santamaría

The number of women pursuing higher education in STEM areas (science, technology, engineering and mathematics) is steadily growing, yet their presence progressively decreases when it comes to high-level academic positions. The COVID-19 crisis has aggravated preexisting gender inequalities so that, according to the most recent Global Gender Gap Report of the World Economic Forum, "another generation of women will have to wait for gender parity" [9]. The pandemic has hit female academics in STEM particularly hard along multiple dimensions, such as productivity, boundary setting and control, as well as the ability to engage in collaborations and network building [6]. In order to fully understand the gender gap in academia and its development, for instance to assess and counteract the effects of crises such as pandemics, fine-grained data are needed, but these are unfortunately not generally collected and analyzed in a consistent manner.

1 The Gender Gap in Science project

The three-year project "A Global Approach to the Gender Gap in Mathematical, Computing, and Natural Sciences: How to Measure It, How to Reduce It?"[1] was funded by the International Science Council (ISC) in 2017–2020. It brought together eleven scientific organizations, led by the International Mathematical Union (IMU) and the International Union of Pure and Applied Chemistry (IUPAC). The goal of this interdisciplinary, cross-national project was to study the situation of women in mathematics, computing and natural sciences. The work was articulated around three central themes aiming to identify common and discipline-specific issues that might require interventions. Two of the tasks were the launch of a Global Survey of Scientists and the creation of a Database of Good Practices. Moreover, a data-driven study was conducted in order to *examine the situation of academic authors and their publication practices in different fields across world countries and regions with respect to the scientists' gender.*

The interest in bibliometric analyses is rooted in the importance of published papers for academic careers. Scientific publications are not only the major outlet for scholarly communication, they are regarded as a proxy for a researcher's scientific credibility and play a key role in achieving and maintaining a successful career in academia. Decisions on tenure and other academic promotions are mostly based on evaluations of the candidate's research portfolio that pay special attention to publications such as journal articles, in addition to grants, conference presentations, and how visible or well-recognized a scholar is. Thus, *the understanding of publication practices, obtained through measurable data on research output, is of great interest* to academic institutions, science policymakers, and researchers alike.

Multiple studies based on bibliometric data have focused on gender. The literature even comprises discipline-specific findings pertaining to mathematics and physics, albeit in small numbers. Much of the existing scientometric research builds on cross-discipline corpora such as Scopus and accordingly does not highlight individual fields. Research with a topical focus on a particular discipline or subfield is typically limited to a selection of journals or conferences or a narrow time period. For the work executed within the Gender Gap project, we decided to analyze the most comprehensive data sources in terms of content and temporal coverage. Those collections happen to be managed by community organizations and curated by experts: data for mathematics came from zbMATH; for theoretical physics, we used arXiv preprints enriched with Crossref; for astronomy and astrophysics, we resorted to ADS.

For the field of mathematics, we analyzed zbMATH's full collection of publications by authors with a main research focus in mathematics from 1970 until July 2019. This *data set comprises more than 3 million documents corresponding to more than 5.2 million authorships* (pairs of author and document), yielding an average of 1.7 authors per article. We inferred the gender of these authorships from the authors' names via various statistical name-gender databases and services. The resulting gender breakdown was approximately 70 % men, 10 % women, and 20 % undetermined. Omitting the latter, women accounted for about 12 % of the male plus female authorships. These[2] in turn belong to ca.

[1] https://gender-gap-in-science.org

[2] Not all authorships can be assigned to a unique author, in particular if the author's name is frequent.

65,000 and ca. 260,000 distinct authors labeled as women and men, respectively, which yields around 21 % distinct women among all recorded authors in zbMATH in the past 50 years, growing from less than 10 % in the 1970s to over 27 % today. Currently, the rate of *new authors being added to the database is more than 14,000 per year, which means that ca. 4,000 women enter the field of mathematics annually.*

One key aspect of publishing relates to journals and their perceived quality. In mathematics, research is predominantly driven by scholarly journals, and they are a crucial vehicle for the forging of academic careers. Publishing in highly renowned venues is a powerful determinant of tenure and an important predictor of professional success. Therefore, any bibliometric study on publication practices has to take into account their impact in the making of academic careers. Here we present some mathematics-related findings from the work done within the Gender Gap in Science project. Further results plus additional context information, e.g. on the data processing algorithms that were employed, can be found in the final project report [7].

Representation of women in renowned mathematical journals

Previous research [5] showed that authorships by women in mathematics are vastly underrepresented in journals with a high reputation in terms of two common ranking methods, the manually compiled Australian ERA indicator and the Thomson Reuters journal impact factor (JIF). In this project, we intended to offer the scientific community the opportunity of examining gender distributions in journals of particular relevance to them or their subfield. We made this possible via a dedicated web interface[3] that allows readers to filter specific publication venues of their interest.

In addition, we have specifically analyzed *selected journals of particular renown in concrete subfields as well as several journals published by mathematical societies*. Figure 1 illustrates the percentages of authorships from women in said selected journals, which are predominantly constrained below 20 %. Only half of the society journals show a rising tendency over the past decades. No evident increase in women representation can be seen in the *Bulletin de la Société Mathématique de France* or the *Journal of the European Mathematical Society*, with both stagnating below 10 %. Even lower, at 5 % and with no sign of improvement, is the presence of women in the *Journal of the American Mathematical Society*. When it comes to discipline-specific topics, as is the case of the three journals on the bottom right that focus on applied mathematics, we observe a rising development and at least 10 % of women's presence. Apart from the *Journal of Differential Geometry*, all specialized journals reveal a moderate positive trend. The prestigious journals in pure mathematics *Inventiones Mathematicae* and *Annals of Mathematics* stand out with percentages of women authorships predominantly in the single-digit range. For more details, see [4].

Several factors can be hypothesized to be potential causes for the measured underrepresentation, but these cannot be confirmed by the bibliographic data alone. As a complementary data source, we have leveraged the 2018 Global Survey of Mathematical, Natural, and Computing Scientists that was conducted within the Gender Gap Project. The survey reached almost 10,000 mathematicians, physicists, and astronomers who were questioned about their submission practices to top-ranked journals in their disciplines. More precisely, the following was asked: "During the last five years, how many articles have you submitted to journals that are top-ranked in your field?" Respondents were expected to provide a number between 0 and 30; larger values were clustered together. According to the responses obtained, women and men self-report to have submitted similar numbers of articles in the past 5 years, with no statistically significant differences in subgroup analyses broken down by disciplines or world regions.

In order to determine the most important predictors for the number of paper submissions, we trained a statistical model that took into account not only the gender but also other factors like discipline, country, access to childcare, as well as various aspects of the academic career and the professional activity, such as the number of grant applications and supervised graduate students. We found out that gender plays a minor role in the model; far more important are aspects that signal career advancement such as having a broad network and strong research activity, which are indirectly linked to gender to some extent.

In conclusion, the self-report of perceived submission practices does not support the hypothesis that women are underrepresented in prestigious journals because they submit fewer manuscripts than men. Considering the importance of publications in renowned journals on the one hand and the conflicting bibliometric analysis on the other, this prompts the question of the role of the peer review process. We note that the refereeing system in mathematics lacks homogeneity and relies substantially on the authors' credit and the level of trust between editors and reviewer(s). In this regard, we stress that *there are hardly any systematic studies on peer review in mathematics* [1], a need that very much ought to be addressed.

By publishing (analyses of) acceptance rates broken down by gender and other sociodemographic or career-related aspects, publishers of high-impact journals could make an important contribution to the evaluation of the fairness of the publication process. Self-assessed data by one of the major publishers in physics suggests at least a gender- and workplace-based bias in physics [3]. Unfortunately, this type of study is far from being a common practice among scientific publishers. A noteworthy exception is *The British Journal for the Philosophy of Science*, which provides acceptance rates broken down by author gender and affiliation

[3] https://gender-publication-gap.f4.htw-berlin.de/

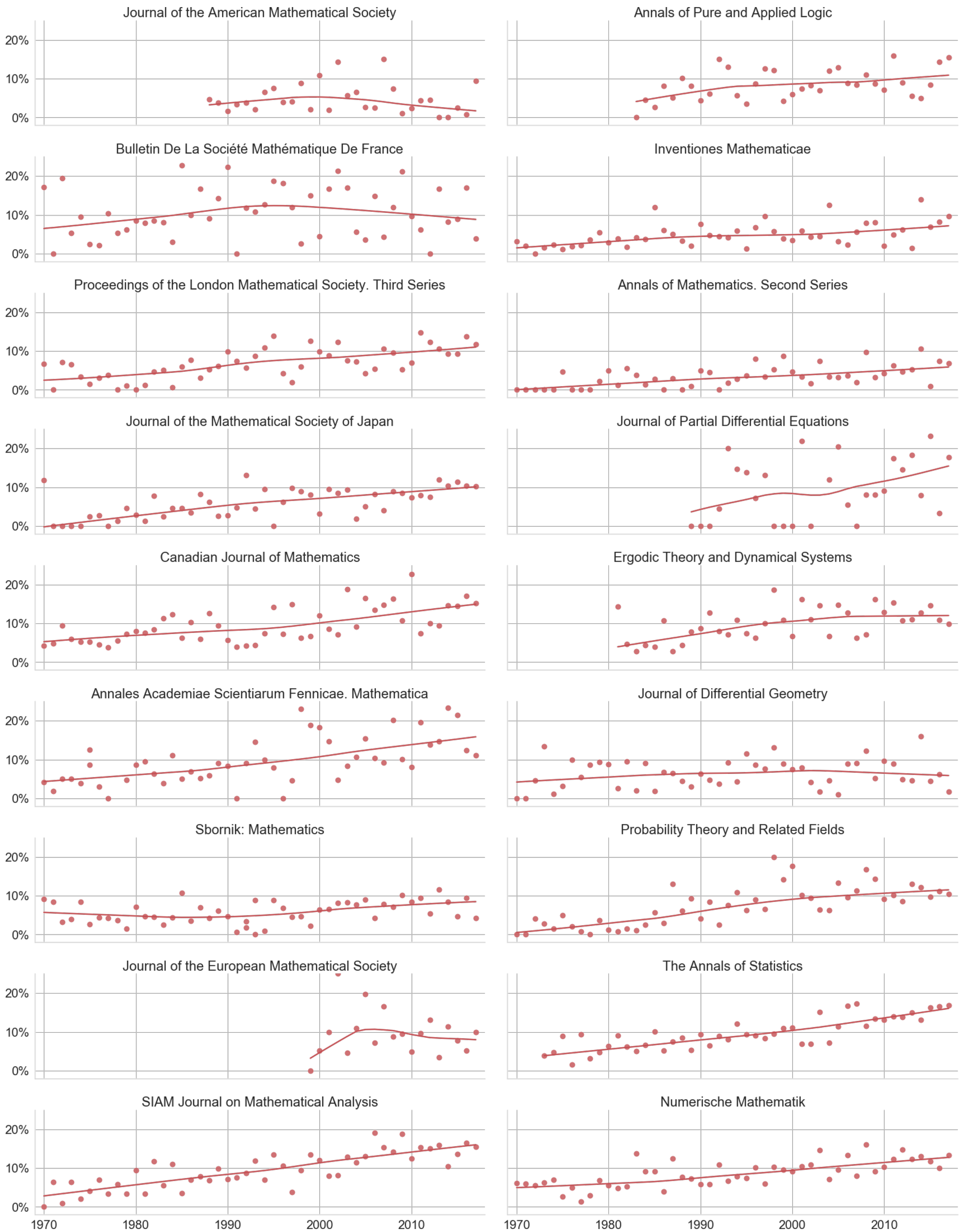


Figure 1. Percentage of authorships from women in renowned mathematics journals per year between 1970 and 2017 [4].

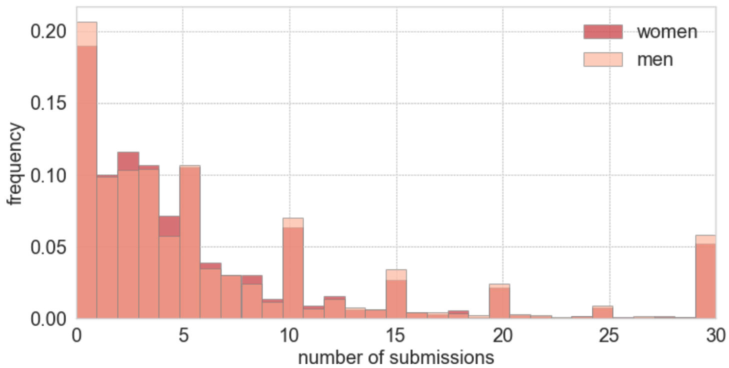


Figure 2. Histogram of the number of publications submitted to top-ranked journals in the last 5 years as self-reported in the Global Survey. Dark and light bars encode answers from women and men, respectively [5].

country. Interestingly, papers by women as sole authors have higher acceptance rates [8].

In relation to the impact of COVID-19 in publication practices, the journal Isis from the History of Science Society has evaluated its submission and acceptance rates, noting an "alarming drop in manuscript submissions by female scholars" in the first half of 2020 [2]. A literature search for similar evaluations of mathematical publishers remains unsuccessful. Along the same lines, the search for studies on the effects of the pandemic for female mathematicians only leads to results based on preprint repositories such as the arXiv, indicating the absence of data from publishers' self-assessments.

Learning and perspectives

There are various aspects to consider when speaking of a gender gap. In the ISC Gender Gap project, we have provided insights on the gap defined by the proportional presence of women as authors of core mathematics publications, we have investigated whether there is a gender gap in the dropout rates that affect the span of mathematicians' academic careers, and we have focused on the gender gap in high-impact mathematical journals.

Consistent with the global trend in higher education, we observe *increasing proportions of women entering the field of mathematics with each passing year*. The understanding of the extent to which those newcomers will progressively attain senior academic positions is crucial to address the "leaky pipeline" phenomenon. Thanks to our cohort analysis based on zbMATH publication data, we are able to provide insights on this issue. We show that dropout rates of mathematicians after their postdoctoral stage, which used to be higher for women than for men, are progressively converging. These data certainly offer optimistic prospects regarding the eventual closure of this particular aspect of the gender gap.

On the other hand, our analysis of women's presence in renowned journals is a good measure of the gender gap in relation to the achievement of a prestigious academic career. In this regard, a non-negligible number of *the prominent mathematical journals under consideration show a meager representation of female authors*. All other factors being equal, the expectation is that the proportion of women among all authors should roughly resemble the percentage of established female mathematicians in the profession, a number that has been steadily growing and that is estimated to be around 25 %. Remarkably, several of the analyzed journals publish very few articles authored by women and have exhibited no signs of turnaround over the last couple of decades. An explanation for this fact might lie in the characteristics of the peer review process in mathematics, which favors close interactions and trust relationships between editors and reviewers and opens the door to conscious and unconscious biases. Regarding subfields, applied areas display a better situation for women than pure ones, which in itself introduces a number of questions regarding the intrinsic differences among subfields of mathematics.

The above remarks provide a compelling starting point for future research. Is the increasing number of young female mathematicians enough to stop the pipeline from leaking? Which factor in the retention of women in academia is played by the professional atmosphere in pure versus applied mathematics? What is the importance of informal academic networks in helping a mathematician's career to thrive? Is the lack of double-blindness in peer review hindering women and other underrepresented groups in mathematics? *It would be excellent to discuss our data-backed findings with experts from the respective subfields in the mathematical community*, with the goal of formulating plausible hypotheses that could explain the observations found by our work in the Gender Gap project.

References

[1] C. Geist, B. Löwe and B. V. Kerkhove, Peer review and knowledge by testimony in mathematics. In *PhiMSAMP. Philosophy of Mathematics: Sociological Aspects and Mathematical Practice. Including Selected Papers of the 3rd PhiMSAMP Conference "Is mathematics special?" (Vienna, 2008)*, edited by B. Löwe et al., Texts in Philosophy 11, College Publications, London, 155–178 (2010)

[2] History of Science Society, Updated report on Isis submissions and gender during the pandemic, hssonline.org/report-on-isis-submissions-and-gender-during-the-pandemic/

[3] IOP Publishing, Diversity and inclusion in peer review at IOP Publishing, IOP Publishing, ioppublishing.org/wp-content/uploads/2018/09/J-VAR-BK-0818-PRW-report-final2.pdf (2018)

[4] H. Mihaljević and L. Santamaría, Authorship in top-ranked mathematical and physical journals: Role of gender on self-perceptions and bibliographic evidence, *Quant. Sci. Stud.* **1**, 1468–1492 (2020)

[5] H. Mihaljević-Brandt, L. Santamaría and M. Tullney, The effect of gender in the publication patterns in mathematics. *PLOS ONE* **11**, e0165367 (2016)

[6] National Academies of Sciences, Engineering and Medicine, *The impact of COVID-19 on the careers of women in academic sciences, engineering, and medicine*. The National Academies Press, Washington, DC (2021)

[7] M.-F. Roy et al., *A global approach to the gender gap in mathematical, computing, and natural sciences: how to measure it, how to reduce it?* DOI 10.5281/zenodo.3882609 (2020)

[8] The British Journal for the Philosophy of Science, BJPS Stats 2020. BSPS, thebsps.org/auxhyp/bjpsstats2020/ (2021)

[9] World Economic Forum, *Global gender gap report 2021*. World Economic Forum, www3.weforum.org/docs/WEF_GGGR_2021.pdf (2021)

Helena Mihaljević is professor of data science at the Hochschule für Technik und Wirtschaft (HTW) Berlin, in association with the Einstein Center Digital Future in Berlin. The focus of her teaching and research activities is on statistical data analysis, applied machine learning, and information retrieval. Previously, she worked as a senior data scientist at The Unbelievable Machine Company in Berlin, as editor and deputy head at zbMATH, and as a research associate at Christian-Albrechts-Universität zu Kiel. She studied mathematics at the Georg-August-Universität Göttingen and received her PhD in mathematics from the University of Liverpool in 2009 on topological dynamics of entire transcendental functions.

helena.mihaljevic@htw-berlin.de

Lucía Santamaría is an applied scientist and instructor at Amazon's Machine Learning University, where she develops and teaches a broad range of ML courses, from mathematical foundations to deep learning methods. One of her research interests is the study of the gender gap in STEM, following her prior work in named entity recognition and author identification at zbMATH. Lucía holds a PhD in physics from the University of Potsdam for her research on the modelling of black hole binaries as sources of gravitational-wave searches with LIGO, conducted at the Max Planck Institute for Gravitational Physics.

lucia.santamaria@ymail.com

Society of Mathematicians and Physicists of Montenegro

Milojica Jaćimović and Predrag Miranović

ДРУШТВО МАТЕМАТИЧАРА И ФИЗИЧАРА ЦРНЕ ГОРЕ

The Montenegrin Society of Mathematicians and Physicists is an association that brings together primary and secondary school teachers, as well as university teachers of mathematics and physics, and other individuals engaged with these sciences or dealing with pedagogical-methodological issues related to mathematics and physics in Montenegro. The society unites about 550 members.

Activities on the renewal of various scientific and professional institutions and organisations in Yugoslavia, as well as the formation of the new ones, began right after the end of World War II. Yugoslav professional associations, as well as the Association of Mathematical and Physical Societies of Federal People's Republic of Yugoslavia, were organised as unions of the corresponding associations of the federal states.

In Montenegro, the smallest Yugoslav republic, there were at that time no institutions of higher education. The establishment of the Higher Pedagogical School in Cetinje, the historical capital of Montenegro, in October 1947, marked the beginning of organised higher education, and can also be considered as the beginning of scientific research in the fields of mathematics and physics in Montenegro. Primary school teachers were trained in that institution, as well as teachers of mathematics and physics at higher levels. The first teachers in the mathematics/physics study programme were mathematician *Lazar Karadžić* and physicist *Petar Jovanović*. Starting in 1963, the teacher-training studies were moved to the Pedagogical Academy in Nikšić).

Following the creation of the Higher Pedagogical School, the Natural Society of Montenegro was formed, and it included a section of mathematicians and physicists. Among others, professors Lazar Karadžić and Petar Jovanović joined the Executive Board of the Society. The section had the task of improving education in these disciplines, contributing to the development of teachers' professional skills, providing help to the educational authorities in conducting educational reforms, initiating and encouraging scientific work in these areas, and popularising the sciences among gifted students.

In November 1949, the First Congress of Mathematicians and Physicists of Yugoslavia was held in Bled (Slovenia). A seven-member delegation from Montenegro participated in the congress. During the congress, the decision was made to establish the Association of Societies of Mathematicians and Physicists (and Astronomers) of Yugoslavia (only after 1980 were the words "and Astronomers" deleted from the name of the Society). Both the plenum and the Executive Board of the alliance were elected. Mathematician *Dušan Gvozdenović* and physicist *Petar Jovanović* from Montenegro became members of the plenum. On that occasion, the Yugoslav Association stated that the Association of Mathematicians and Physicists of Montenegro should be established without delay. The Montenegrin Society of Mathematicians and Physicists was finally created in 1959, with goals similar to those of the Mathematical-Physical Section of the Natural Society.

Professors of the Higher Pedagogical School in Cetinje, *Dušan Gvozdenović* and *Momčilo Kosmajac* were elected president and secretary of the association. The first competition of high school students was organised in 1967; however, it did not include all schools in Montenegro. After the first competition, several more competitions in mathematics and physics (and later in programming) were organised at the level of primary and secondary schools. Since there was no scientific or professional journal in the field of mathematics and physics in Montenegro at that time, both teachers and students from Montenegro contributed to professional journals published outside Montenegro in the Serbo-Croatian language.

The University of Montenegro, the first institution of higher learning in Montenegro, was established in 1974. The Institute of Mathematics and Physics was established in 1978, and gathered under its roof all the study programmes of mathematics and physics taught at the University. The first generation of students enrolled in the study programmes of mathematics and physics in 1980. The first director of the institute was a professor at the University of Montenegro: the mathematician *Predrag Obradović*, who was also the first doctor of mathematical sciences in Montenegro (having received his PhD from the University of Zagreb in 1974). Later, with the opening of the Biology Department, the Faculty of Natural Sciences and Mathematics was founded. We should also mention the creation of the study programme for computer sciences, initially as a part of the study programme in mathematics, but subsequently as an independent subject in itself.

The founding of the institute, the enrolment of the first generation of students of mathematics and physics, and the hiring of new faculty galvanised the work in the fields of mathematics and physics in Montenegro. The institute became the centre of all important activities within the mathematical and physical community in Montenegro. Cooperation with state educational institutions gained momentum; seminars for teachers of mathematics and physics in primary and secondary schools were held regularly; cooperation was established with faculties of mathematics and physics in other university centres; young associates received professional training (in particular from Moscow State University within a broader framework of cooperation with this university), student competitions and scientific meetings were organised. The Association of Mathematicians and Physicists of Yugoslavia entrusted the Society of Mathematicians and Physicists of Montenegro with the organisation of the VII Congress of Mathematicians and Physicists of Yugoslavia, which was held in 1980 in the coastal town of Bečići (Montenegro). The society organised the congress together with the Montenegrin Institute of Mathematics and Physics. Around 1,500 mathematicians and physicists took part in the various sections of the congress. The most populous section was "teaching mathematics". Very lively discussions took place, addressing a range of teaching issues at all levels of education. The congress showed that the number of disciplines studied by mathematicians in Montenegro, and in Yugoslavia as a whole, expanded significantly at that time, and that mathematicians in Yugoslavia, as well as Montenegrin mathematicians, were dealing with modern mathematical disciplines and problems.

In 1993, the first scientific mathematical journal in Montenegro, *Mathematica Montisnigri*, was founded by the Faculty of Natural Sciences and Mathematics of the University of Montenegro and the Association of Mathematicians and Physicists of Montenegro. Professor of the University of Montenegro *Žarko Pavićević* was appointed editor-in-chief of the journal.

With the dissolution of Yugoslavia, the Federation of the Yugoslav Societies of Mathematicians and Physicists ceased to exist. At that time, in 1994, the Societies of Serbia and Montenegro founded the new Federation of Mathematical Societies of Yugoslavia and Federation of Yugoslav Physical Societies. The federation entrusted the Association of Mathematicians and Physicists of Montenegro with the organisation of the IX Congress of Mathematicians of Yugoslavia. The congress was held in Petrovac, Montenegro in 1995, with the participation of about 500 mathematicians.

In 2006, after the referendum, Montenegro became an independent state. The Federation of Yugoslav Societies no longer existed, and the Montenegrin Society of Mathematicians and Physicists was admitted to the European Mathematical Society in 2007, and afterwards to the International Mathematical Union. The first congress of mathematicians and physicists of Montenegro was held in 2010 in Petrovac, organised by the Montenegrin Society of Mathematicians and Physicists and the Faculty of Natural Sciences and Mathematics of the University of Montenegro. About 300 teachers of mathematics from the primary and secondary schools as well as from the university units attended the congress. It included the following sections: teaching mathematics, physics and informatics in Montenegro, mathematics, physics, computer science.

Since its foundation, the Montenegrin Society of Mathematicians and Physicists has been working to create better conditions for the development of mathematics and physics in Montenegro. Members of the society participate in European and world congresses and international scientific conferences, and also organise scientific conferences in Montenegro, while students from Montenegro are given support to participate in international mathematical competitions. The society contributes to all activities of educational and scientific institutions related to mathematics and physics, supports modernisation of teaching/curricula and engages in popularisation of these sciences in Montenegro.

Milojica Jaćimović, mathematician, is a retired professor at the University of Montenegro where he taught different courses. At the Institute of Mathematics and Physics (now the Faculty of Natural Sciences and Mathematics), he created courses in the field of optimisation. He is a member of the Montenegrin Academy of Sciences and Arts (MASA). His research interests lie in the field of optimisation methods, methods of solving variational and quasivariational inequalities, regularised method of solving ill-posed optimisation problems, and in modeling of population dynamics. He was editor-in-chief of the journal Proceedings of the Section of Natural Sciences of MASA.

milojica@jacimovic.me

Predrag Miranović, physicist, president of the Society of Mathematicians and Physicists of Montenegro, is a professor and dean of the Faculty of Mathematics and Natural Sciences of the University of Montenegro. He was rector of the University of Montenegro (2006–2012), and president of the Union of Societies of Physicists of Yugoslavia (2004–2006). He is member of the Montenegrin Academy of Sciences and Arts and Secretary of its Section of Natural Sciences. He teaches quantum mechanics, theoretical mechanics and theory of phase transitions. He research concerns the transport and thermodynamic properties of superconductors.

pedjam@ucg.ac.me

Transition Year in the Irish school system: Challenges and opportunities for teaching and learning mathematics

Ronan Flatley

This article gives an overview of the Irish secondary school Transition Year programme and in particular how the teaching and learning of mathematics is situated within it.

1 Introduction

Second-level education in Ireland caters for students between 12 and 18 years of age. It consists of two cycles, the Junior Cycle and Senior Cycle. The 3-year Junior Cycle prepares students for state examinations in the core subjects of Mathematics, Irish and English, as well as in students' individually chosen subjects. Following the Junior Cycle, students have the option of enrolling on a 1-year Transition Year (TY) programme or progressing directly to the Senior Cycle. The 2-year programme of the Senior Cycle culminates in the terminal state examination, the Leaving Certificate, a qualification which heavily influences the destination of students for further education and training.

In this article, we give an overview of the TY programme and consider the potential for introducing some interesting mathematics within it. TY is an optional extra year of study that is taken up by the majority of students before embarking on two years of study for the Leaving Certificate[1]. The official mission of TY is [6, p. 3]:

> To promote the personal, social, educational and vocational development of pupils and to prepare them for their role as autonomous, participative, and responsible members of society.

2 The Transition Year curriculum

It is government policy that each school design its own TY programme to meet the needs of its students. However, the Department of Education (the government ministry responsible for education) has produced guidelines as to what such a programme should entail. Four areas of learning are suggested [8]:

1. *Core*: This includes Mathematics along with English, Irish, Information Technology and Physical Education
2. *Subject sampling*: These may be taster modules in subjects to help students decide on subject choices for Leaving Certificate. Students may also be exposed to subjects that they have not studied before e.g. Environmental Studies, Japanese, Arabic, Drama, Photography, Global Development
3. *TY-specific*: Study and activities which allow students to pursue their own interests and to contribute to their local communities e.g. social entrepreneurship, charitable activities
4. *Calendar*: Work experience, foreign exchange visits, sporting activities, visiting speakers, drama/musical production

Over the years the programme has been very well received amongst students, parents, guardians and teachers alike and is considered unique in that there is no similar programme offered in education systems of other countries [2]. In its 2020 report on education in Ireland, the OECD commended the programme as it "allows students to sample different subjects and undertake work experience and other projects, and helps guide them in choosing their upper secondary education subjects and future career path" [7].

For students, TY may be viewed as an academically low-risk interlude between state examinations because there is no state-level assessment of student performance. Many schools do issue individual assessment certificates that rate student attainment across the programme but these marks generally do not count towards future progression.

3 Scope of TY mathematics

There is no prescribed syllabus for mathematics in TY and so, theoretically at least, teachers have great flexibility regarding the mathematics they present to their students. Government does not give specific recommendations as to what should be taught

[1] According to [2], about 55 % of students took TY in 2010/11. Although there is consensus that there is an upward trend in participation rates, no data is currently available to confirm this.

but a 2011 circular called for schools "to provide innovative learning opportunities and increased mathematics teaching hours to the extent feasible" in Transition Year [3]. Up to 2012, there had been very little information available on the content of TY mathematics programmes being delivered in schools, how they were delivered and the time spent on delivery. Then, as part of PISA 2012, a national survey of TY mathematics teachers was conducted.

3.1 PISA 2012 National Survey of TY Mathematics Teachers

Mathematics teachers were invited to complete questionnaires which focussed on concerns about the quality of TY mathematics offerings. In total, 1321 questionnaires were completed, an 80 % response rate, and the findings were reported in [5]. In the survey, teachers were asked to rate their level of agreement or disagreement with statements about the purpose of TY mathematics. It was found that only 31 % of teachers strongly agreed that one purpose of TY mathematics is to increase students' confidence in their problem-solving ability, and only 34 % strongly agreed that one purpose was to encourage greater interest in mathematics. Almost 40 % disagreed or strongly disagreed that one purpose of TY mathematics is to familiarise students with the history of mathematics. These results were certainly out of kilter with government guidelines, as set out in [4], which state that these particular aspects of mathematics should be emphasised during TY. More importantly perhaps, they indicate a strong lack of agreement amongst teachers as to the purpose of teaching mathematics during TY.

3.2 Teaching resources and challenges

The national survey found that there was a great deal of variation in terms of resources that TY maths teachers used. Some teachers used textbooks from the Junior Cycle to consolidate previously covered content while others mined the Internet for engaging and challenging material. There was also evidence of teaching the Leaving Certificate syllabus, something the government guidance expressly discourages.

In the period from 2012 to 2018, we have had the opportunity of meeting with many TY teachers during outreach activities in secondary schools across Ireland. Teachers regularly commented on their struggle to engage students with mathematics, especially during TY. They reported that one view commonly held by students was that maths would not be relevant to their lives once their secondary schooling came to an end. Teachers agreed on the need for a book or module which would focus on the importance of the subject in the everyday world, introduce genuinely interesting ideas and also include some history of mathematics. The challenge of stoking the interest of teenagers who were sceptical about the global importance of maths (as well as a concern for students with genuine mathematical curiosity), encouraged us to write a book of projects to engage the interest of TY students[2]. Exercises would need to be recreational as well as challenging in order to enthuse the target audience. In the next section, we present one of the topics presented in the book.

4 Divisibility of integers – an interesting topic for TY

At the teacher training college where we work, it sometimes happens that a student teacher will express their delight to learn for the first time, during an elementary number theory course, that divisibility of a number by 3 can be determined by checking if the sum of its digits is divisible by 3. They may wonder why they had not learned this rule earlier in their careers or why their school knowledge was restricted to rules for 2, 5 and 10. In any case, the topic of divisibility appeared to be a good one for TY students and student teachers alike.

Like the rule for 3, there exists a similar rule for the divisibility of a number by 9 and it forms the basis of a wonderful trick called *Mindreader*, which appeared in [1] and is used in a project on number tricks in our book.

> *Mindreader*
>
> The "mindreader" asks a volunteer to
> 1. pick any five-digit number where not all digits are the same and keep this number secret;
> 2. jumble up the digits to get a new number;
> 3. subtract the smaller of these numbers from the larger;
> 4. hide one of the digits, H, in the answer, making sure $H \neq 0$;
> 5. call out the other digits in the answer.
>
> The "mindreader" pronounces the hidden digit to be H.

In order to know what the hidden digit H is, we need to understand some basic ideas.

Definition 1. The *digital root* of a number is the sum of all its digits.

Example 1. The digital root of 3046178 is 29 because $3+0+4+6+1+7+8=29$.

[2] *MIghTY Maths* was published in October 2021 by Curriculum Development Unit, Mary Immaculate College, Limerick. Each topic can be investigated either in teacher-led lessons or in a guided project by the individual student or groups of students. The book is available for purchase at www.curriculumdevelopmentunit.com.

There follow two mathematical facts.

Lemma 1 (Divisibility rule for 9). *Take any number x. Let d be the digital root of x. Then x is a multiple of 9 $\Leftrightarrow$ d is a multiple of 9.*

Lemma 2. *For any number x, with digital root d, $x-d$ is a multiple of 9.*

Example 2. The digital root of 31572869 is 41. We see that $31572869 - 41 = 31572828$ has digital root $36 = 4 \times 9$.

Armed with Lemmata 1 and 2, we can now look at understanding the details of how the trick *Mindreader* works. We follow the given steps for the volunteer:

1. Pick any five-digit number where not all digits are the same. Let x have digits $ABCDE$, so

$$\begin{aligned} x &= 10000A + 1000B + 100C + 10D + E \\ &= (A+B+C+D+E) + 9999A + 999B + 99C + 9D \\ &= d + 9k, \quad \text{where } d \text{ is the digital root of } x \\ &\qquad\qquad \text{and } k = 1111A + 111B + 11C + D. \end{aligned}$$

   Indeed, Lemma 2 tells us that any number equals its digital root plus some multiple of 9.
2. Jumble up the digits to get a new number. Call the new number y. Let y have digits $CEDAB$, for example, so

$$\begin{aligned} y &= 10000C + 1000E + 100D + 10A + B \\ &= (C+E+D+A+B) + 9999C + 999E + 99D + 9A \\ &= (A+B+C+D+E) + 9\ell \\ &= d + 9\ell, \quad \text{where } d \text{ is the digital root of } y \text{ (and } x) \\ &\qquad\qquad \text{and } \ell = 1111C + 111E + 11D + A. \end{aligned}$$

3. Subtract the smaller of these numbers from the larger. Let us suppose that $x > y$, so

$$x - y = (d + 9k) - (d + 9\ell) = 9(k - \ell).$$

   N.B. The number $x - y$ is divisible by 9. Therefore, by Lemma 1, its digital root is a multiple of 9.
4. Hide one of the digits in your answer, H, making sure the hidden digit is not a zero.
5. Call out the remaining digits of the number $x - y$.
   At this point, the "mindreader" has all but one of the digits that make up the digital root. All they need to do is add them and see what digit is required to sum to the next greatest multiple of 9. For example, if they are told the digits are 0, 1, 3 and 8. These sum to 12. The next highest multiple of 9 is 18 so the hidden digit H is $18 - 12 = 6$. Or, if they are told the digits are 7, 5, 6 and 9. These sum to 27, already a multiple of 9. Since it was specified that zero should not be hidden, they conclude that the hidden digit H is 9.

Example 3. Here is a worked example. The volunteer

1. picks the number $x = 47523$;
2. jumbles the digits of x to get the number $y = 37254$;
3. computes $x - y = 47523 - 37254 = 10269$;
4. hides the digit 2;
5. calls out the digits 1, 0, 6 and 9.

Now the "mindreader" knows that $x - y$, being a multiple of 9, has digital root d a multiple of 9. They know that $d - H = 1 + 0 + 6 + 9 = 16 \Rightarrow H = d - 16$. Since the next multiple of 9 greater than 16 is 18, they conclude that $H = 18 - 16 = 2$. The "mindreader" pronounces that the hidden digit is 2.

Remark 1. The *Mindreader* trick shown here uses 5-digit numbers. Is there anything special about 5 or would the trick work for any number of digits?

Finally, we give Table 1 as a sample detailed lesson plan for teaching the content of *Mindreader*.

5 Conclusion

The Transition Year programme in Irish secondary schools is a highly regarded programme. Maintaining the position of mathematics, a core subject in the programme, requires a great effort on the part of schools because there is no curriculum and guidelines on teaching content are rather vague. Individual teachers and schools go to great lengths to provide useful and interesting material, yet gaps clearly remain. We give an example of divisibility of integers as a topic that may engage TY students' interest and we encourage the use of stimulating and recreational mathematical exercises and projects across the programme.

References

[1] C. Budd and C. Sangwin, *Mathematics galore!* Oxford University Press (2001)

[2] A. Clerkin, Personal development in secondary education: the Irish Transition Year. *Education Policy Analysis Archives* **20.38** (2012)

[3] Department of Education, Circular 0058/0011, education.ie/ (2011)

[4] Department of Education, Transition Year programmes – guidelines for schools, education.ie/en/Schools-Colleges/Information/Curriculum-and-Syllabus/Transition-Year-/ty_transition_year_school_guidelines.pdf (1993)

[5] G. Moran et al., Mathematics in Transition Year: insights of teachers from PISA 2012, erc.ie/documents/p12ty_maths_report.pdf (2013)

[6] National Council for Curriculum and Assessment, Transition Year programmes – guidelines for schools, ncca.ie/en/resources/ty_transition_year_school_guidelines/ (2021)

[7] Organisation for Economic Co-operation and Development, Education in Ireland: an OECD assessment of the senior cycle review, oecd.org/ireland/education-in-ireland-636bc6c1-en.htm (2020)

[8] Professional Development Service for Teachers, TY Curriculum, pdst.ie/TY/curriculum (2021)

Ronan Flatley is Lecturer in Mathematics at Mary Immaculate College, Limerick, Ireland. He currently serves on the Committee of the Irish Mathematical Society. He is the author of MIghTY Maths, a book of maths projects for Transition Year students.

ronan.flatley@mic.ul.ie

Table 1. Sample lesson plan for *Mindreader*

Name:	Anna Lisus	Date: 08.09.2021	Lesson Theme:	Divisibility by 9; *Mindreader* trick
Group:	TY2	Time: 12:25 – 13:05		
Lesson Aim:	To learn how to perform the *Mindreader* trick and to understand the mathematics underlying it.			

Phases	Content focus	Social format	Medium	Goals
Presenting *Mindreader* trick (7′)	Two volunteers V1 and V2 at the board. Teacher facing the class with their back to board. (1.) V1 writes a 5-digit x on board (2.) V2 scrambles x to give y (3.) V1 computes $\lvert x-y \rvert$ (4.) Class checks arithmetic (5.) V2 circles a **nonzero** digit of answer (6.) V1 calls out to Teacher the other digits (7.) Teacher proclaims circled digit.	Teacher & two volunteers	Board	Demonstrate *Mindreader*
Background theory (5′)	What divisibility rules do students know (certainly 2, 5, 10)? Divisibility rule for 9: for arbitrary $x \in \mathbb{Z}$, $9 \mid x \Leftrightarrow 9 \mid d$, where d is the **digital root** of x. Example: $9 \mid 4023$. Non-example: $9 \nmid 73486$.	Teacher presents	Board	Teach divisibility rule for 9
Exercises I (3′)	Decide if 9 divides each of 621, 1284, 9187263.	Student solo	Pencil & paper	Apply divisibility rule for 9
Details/context (6′)	Fact: For any $x \in \mathbb{Z}$, $9 \mid (x-d)$. Example: Put $x = 925$. Then $x - d = 925 - 16 = 909 = 101 \times 9$ (or digital root of 909 is $18 = 2 \times 9$).	Teacher presents	Board	Teach the fact
Exercises II (5′)	Show that $9 \mid (x-d)$ for $x \in \{342, 53124, 6173289\}$.	Student pairs	Talk, pen & paper	Apply the fact
Exercises III (12′)	Students perform *Mindreader* trick in pairs. One student chooses a 5-digit number to start, etc.; the other works out the hidden digit.	Student pairs	Talk, pen & paper	Students perform *Mindreader*
Homework:	(Medium) Find a proof of the divisibility rule for 9. (Hard) Prove that $9 \mid (x-d)$ for any $x \in \mathbb{Z}$.	(Easy) Show that $9 \mid (x-d)$ for $x \in \{456, 6312, 8083236\}$. Invent your own trick which uses divisibility.		

ERME column

regularly presented by Jason Cooper and Frode Rønning

In this issue, with a contribution by
Anna Baccaglini-Frank, Carla Finesilver and Michal Tabach

In this contribution we introduce three classical theoretical stances within the field of mathematics education regarding representations. Our aim is to highlight what we consider to be an interesting shift in how representations are conceived and studied in the field of mathematics education, and how this could impact both the practice of teaching and learning mathematics, and on further theorizing mathematical representation. We also indicate potential directions in which to develop ways to talk about newer forms of dynamic interactive representation.

Representations of mathematical concepts constitute an "integral part of the doing of mathematics" [16] and, therefore, they are also an integral part of teaching and learning mathematics. Indeed, the theme of representation has for some time been a crucial topic in research in mathematics education – for instance, in PME groups (Psychology in Mathematics Education), and in special issues of the prestigious journals Educational Studies in Mathematics and ZDM Mathematics Education. The authors of this paper are currently co-leaders of the Thematic Working Group "Representations in Mathematics Teaching and Learning" of the 12th Congress of the European Society for Research in Mathematics Education (CERME12), and have been involved in the discussions of this working group ever since it was founded at CERME10 in 2017 [17]. The working group has continued its discussions over the years (e.g., [2]) focusing on many pedagogical and theoretical aspects of mathematical representations. Some recurring themes in the discussions have been around the effective uses of different types of representation, imagery and visualization in mathematical problem solving, and how teachers can help learners to make connections between different representations of the same mathematical object.

Another recurring theme in many of these discussions is the advocacy of working with different forms of representation, and the valuing of non-standard forms. Group discussions have pointed to pressures that exist across many educational contexts for teachers to privilege particular standardized forms of representation over alternatives, in order to push students to acquire as swiftly as possible selected so-called "efficient" ways to produce answers [7]. Since these pressures may prematurely curtail students' creativity and intuitive approaches when engaging in problem solving [4], discussions in our working group have focused on how to support teachers' use of more diverse representational forms and formats that enable wider inclusivity, providing all learners with opportunities to engage more meaningfully with mathematical activity and knowledge.

In support of this position we believe it is pertinent that creative mathematical thinking needs incubation time, and that it is very frequently supported by non-standard representations, developed as personal cognitive tools to implement or demonstrate particular objects or reasoning processes. For example, Maryam Mirzakhani, winner of the 2014 Fields Medal, was well known to "doodle" as a central part of her mathematical research process, repeatedly drawing and re-drawing figures (for example, those reproduced in Figure 1) on large sheets of paper spread out on the floor.

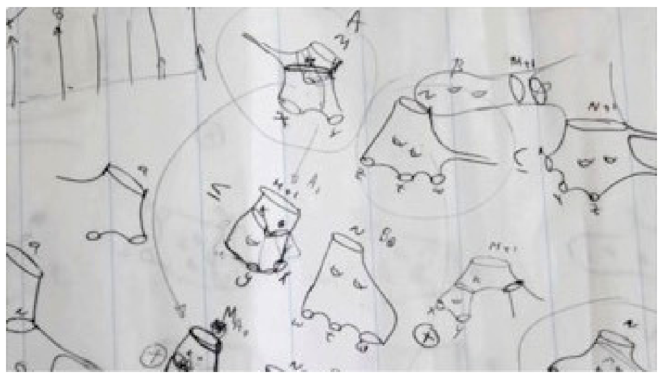


Figure 1. A drawing by 2014 Fields medal, Maryam Mirzakhani. Video still from "Maryam Mirzakhani". © 2014 International Mathematical Union, via *Quanta* magazine.

A second example is shown in the drawing from a brief comic book by Saharon Shelah, designed for a presentation of a recent result; the sketch in Figure 2 illustrates the notion of isomorphism.

As mentioned above, another recurring theme from our CERME group concerns the theoretical foundations and languages through

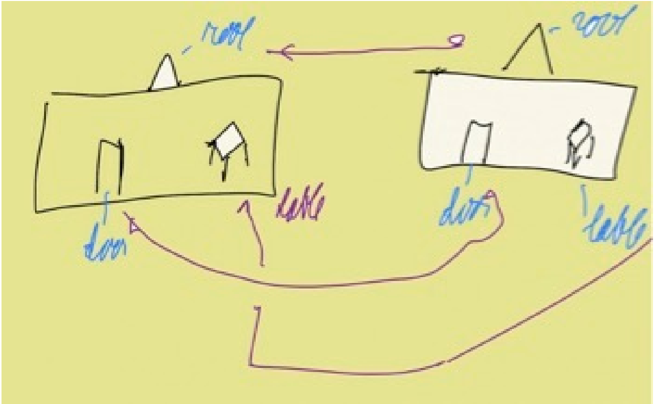


Figure 2. Drawing which illustrates the notion of isomoprhism from a divulgative comic book by Saharon Shelah. By courtesy of the author. © Sharon Shelah 2021.

which representations are thought about, viewed, designed and discussed. We focus on this theme in the rest of this contribution, to explain the significance for mathematics education of having different ways of "talking about" representations. This is an ongoing process: we are still developing appropriate concepts and vocabulary for researching certain kinds of representations, for example those that have a dynamic and interactive nature, or that are multimodal/multimedia, or co-created through collaborative activity. Such representations have become more frequently seen with the advent of digital technology in educational contexts, and because of educators' increasing attention to fostering meaningful mathematical experiences in a variety of physical and digital contexts. More specifically, in this contribution we introduce three classical theoretical stances within the field of mathematics education. These are used to highlight the shift in how representations are conceived and studied in mathematics education, and its impacts on further developing both pedagogy and theory in this field.

First, we need to make explicit the context – both past and present – in which we are writing this contribution. All three authors were educated, and currently live, in cultures that assume (either explicitly or implicitly) that mathematical objects have a Platonic nature. By this we mean that they are commonly taken as existing in some not directly accessible reality from which "shadows" are cast; such shadows are the imperfect forms with which we can access the "real" perfect objects behind them, in order to talk and think about them. Indeed, the verb "to represent" comes from the Latin word "repraesentare", formed from the prefix "re-" expressing intensive force, or reiteration, and the verb "praesentare" that means "to present". So *to represent* entails the idea of something being "out there", and that this something may be realized *again* through one or more representations that manifest some aspects of it. Coherently with this metaphor, we learn to talk and think about mathematical objects by interacting with their representations. However, frequently, as expert mathematicians it happens that we become so comfortable with particular representations that we forget that they are not actually *the* object they stand for (see, for example, Figure 3). This can cause significant difficulties in the teaching and learning of mathematics.

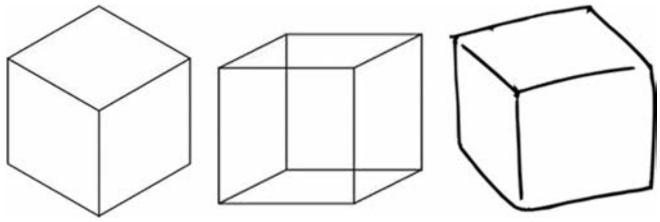


Figure 3. Three configurations which students might be expected to recognize as representations of the 3-dimensional mathematical object "cube", but which are not the actual mathematical object, only 2-dimensional representations of certain aspects of "cube" (each of which in fact contradicts other defining aspects of cubes).

The Platonic philosophical stance outlined above is at the basis of much of the research on mathematical representations, which intersects with mathematics education from various fields of research. Various theories on learning mathematics hold it to be key to appropriately use mathematical representations and master their mutual relationships by *understanding* the mathematical objects they represent, by somehow tapping on their "true" meanings. However, the relationships between mathematical objects, their meanings and their representations are conceptualized and operationalized differently by different researchers and theoretical frameworks. In the following paragraphs we briefly highlight: (1) three key positions from the diversity of theoretical frameworks that have been conceived and used to study mathematical representations, and (2) a shift in how representations are conceived and studied in the field of mathematics education, moving towards the importance of how we talk about (representations of) mathematical objects. We then discuss how such a shift could have an impact on both the practice of teaching and learning mathematics, and on theorizing representations. The three key positions selected are those of Goldin, Duval and Sfard.

Position of Goldin

One influential and essentially pragmatic view of mathematical representations in educational contexts is Goldin's, which firstly distinguishes external from internal representations. The former are often visible or tangible productions such as graphs, arrangements of concrete objects or manipulatives, words, formulas, etc. (although could also include, e.g., communications in speech or gesture) that encode, stand for, or embody mathematical ideas or relationships [12], and aim to communicate them to others or to one's future self. Collected external representations of these types and many more form much of the data used in empirical research by members of our group and others. We cannot (yet!) observe anyone's internal mathematical representations directly, but we may make inferences about learners' internal representations on the basis of their interaction with, production of, or discourse regarding external representations, and to some extent, descriptions, for example, of their mental imagery while problem-solving. These forms might include the mental manipulation of systems of *verbal/syntactic*, *imagistic*, and *formal notational* configurations (or other less frequently discussed forms, such as auditory and/or kinaesthetic rhythmic patterns), which while invisible to the observer, may be inferred [13]. Further, Goldin's view is that any mathematical representation cannot be understood in isolation, but only as part of an interconnected structure of meanings, ideas, systems and practices, which refer to each other in multiple and complex ways.

The relationships between internal and external representations must clearly be bidirectional (i.e. one can recreate and manipulate previously seen imagery in the mind's eye, or recreate and develop one's mental imagery on paper or computer screen, for example); this interaction between internal and external representation is fundamental to effective teaching and learning [11, 13]. Teaching mathematics is thought to happen most effectively "when we understand the effects on students' learning of external representations and structured mathematical activities" [13, p. 19] – yet to do this, it is vital to discuss students' internal representations and how these are connected to one another. The conclusion is that the fundamental goals of mathematics education must include the development of coherent internal systems of mathematical representation that interact effectively with established external systems.

One further point that we would highlight from Goldin's work over the years on internal-external representational relationships is its relation to the pedagogic perspectives of behaviourism and constructivism, which are often presented as diametric opposites or, at least, in conflict. Behaviourist principles exclude any inference about the internal. Resulting pedagogies focus on instructional programmes for shaping learners' behaviour through conditioning [9] – essentially, their acquiring, reproducing and carrying out of procedures with certain external representational forms according to a prescribed set of rules, with clearly measurable results. Constructivist principles, in contrast, strongly emphasize the internal – in particular the radical constructivist movement, according to which any individual only has access to their own perceived experiences, not to any definitive "real world" [10]. Resulting pedagogies focus on learners' discoveries and conceptualizations, often through solo or group problem-solving activity. Research in mathematics education which draws on Goldin's view, then, by centring the interactions between a variety of internal and external representations, has potential to include insights and elements of both perspectives. In terms of pedagogy, this would mean emphasizing "skills and correct answers as well as complex problem solving and mathematical discovery, *without seeing these as contradictory*" [13, p. 8].

Position of Duval

Duval's position stems from the assumption that mathematics is epistemologically different from any other discipline because, as discussed above, mathematical objects are not directly accessible: they can be accessed only indirectly through their representations. Unlike in the case of a person, where any representation of aspects of her could be directly compared to her actual physical form, in the case of mathematical objects no juxtaposition between a representation and the object itself is possible [5]. Therefore, it is extremely difficult to distinguish representations of objects from the objects they represent, but also to (learn to) recognize that multiple different representations may refer to the same mathematical object. This is especially true in cases in which the representations make use of very different *units of meaning*: for example, "$f(x) = 4x^2 - 1$", in which the units of meaning are determined by the algebraic inscriptions on the left and right of the "$=$" sign, and a graph such as the one shown in Figure 4, in which the units of meaning are graphical elements such as the points of the parabola, its intersections with the x-axis, and so on.

To overcome this situation, and thus gain new knowledge about the mathematical objects referred to and solve problems, it is necessary to (learn to) transform one representation into another.

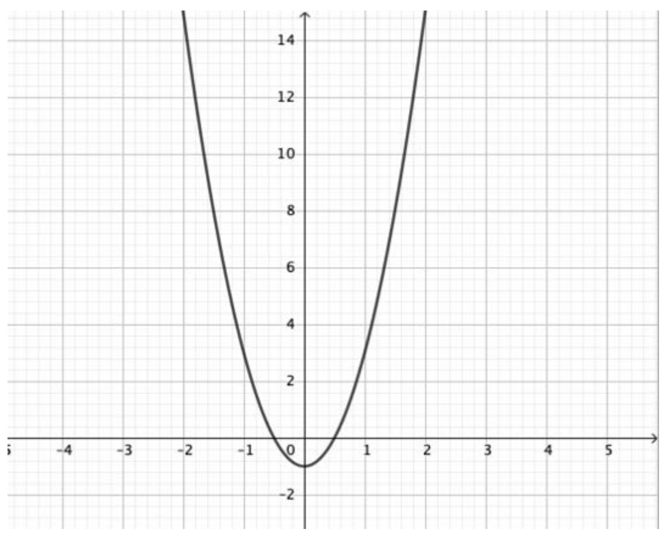


Figure 4. A graphical representation of the function $f(x) = 4x^2 - 1$.

Duval introduces the notion of *register of semiotic representation* [5] to discuss and analyse this transforming activity that lies at the heart of doing mathematics. In order to be a register, a semiotic system (a system of signs) needs to allow the production of representations that provide (indirect) access to mathematical objects, "explore all that is possible" with such signs, and "open a field of specific operations that allow transforming the produced representations into new representations" [5, p. 68].

For example, a register of semiotic representation that Duval has discussed extensively in his work, is the *register of figures*, used heavily in geometry, and developed in order to produce representations that allow us to gain insight and reason about geometrical objects. Just like for any other register, using the register of figures is based on specific cognitive operations, specifically: recognizing at a glance the shapes in the figure, recognizing the figure as being similar to the shapes of real objects, realizing that there are several ways to interpret the shapes or the *figural units*. A key property of figural units is their dimension. Duval argues that "seeing" geometrically means operating dimensional deconstruction of the shapes, and being able to shift quickly from units of one dimension to those of another, to recognize the relationships between the various figural units. So, within the register of figures, one representation can be transformed into another through dimensional deconstruction and reorganization of the figural units.

More generally, a register of semiotic representation has its specific *units of meaning* (which in the case of the register of figures would be figural units) and a representation can be transformed into another in the same register through processes of *treatment* (e.g., dimensional deconstruction in the register of figures). However, according to Duval, the only way to distinguish representations of an object from the object itself is to use at least two registers and to be able to *convert* from one to the other [5, 6]. In the case of geometry, but also the other subfields in mathematics, another fundamental register is that of natural language. Algebra and Analysis make use of the register of symbolic expressions and the graphical register of the Cartesian plane (e.g., Figure 4).

Formally, a semiotic representation is denoted by the couple: (register used, merged meaning units) [6, p. 724]. Understanding in mathematics, according to Duval, means being able to coordinate registers, in his words: "understanding mathematical concepts presupposes awareness of the cognitive one-to-one mapping operation between relevant meaningful units of two registers at least" [6, p. 726]. In the example of the function in Figure 4, treatments in the algebraic register could be rewriting the algebraic expression as $(2x-1)(2x+1)$ to highlight the function's "zeros" (obtained solving $(2x-1)(2x+1)=0$). Conversion into the graphical Cartesian coordinates register that corresponds to such algebraic treatments could correspond to a dimensional deconstruction of the graph (treatment), to visualize the two intersections with the x-axis: $(-\frac{1}{2},0)$, $(\frac{1}{2},0)$.

Position of Sfard – a shift in perspectives

Taking a Vygotskian socio-constructivist perspective, and following Wittgenstein, Sfard [19, 20] sees mathematical objects as no longer residing in some hyper-reality, but in discourse itself, being part of an autopoietic system, a system that defines its own objects. Hence, their meaning stems from the ways in which *realizations* of a mathematical object are used discursively; an implication is that the term 'representation' is inappropriate, as she rejects the Platonic view of mathematical objects existing "out there" and being re-presented in discourse. Rather, under her Commognitive Framework mathematical objects "come to life" as part of a discourse of human communities.

Another difference between realizations and representations can be found in how 'realization' acquires a psychological component: the same physical (graphical, tangible or gestural) production may be a realization of a mathematical object for one person and not for someone else, depending on the phase each person is at in their discursive construction of the mathematical object in question. For example, for an expert $f(x)=3+2x$ may be a realization of a real valued function (of which another realization might be its graph on the Cartesian plane), but for a learner who is not yet familiar with the discourse about either real functions or complex numbers, it is just a strange equality that mixes letters and numbers. Sfard [19] used the term *realization* as follows: "Realization of the signifier S is a perceptually accessible thing S' so that every endorsed narrative about S can be translated according to well defined rules into an endorsed narrative about S'" [19, p. 154]. Sfard sees the relations signifier-signified (between S and S') as symmetrical. So, for example, while the graph shown in Figure 4 could be a signifier of the symbolic expression $y=(2x+1)(2x-1)$, one could also talk about that graph as signifying the symbolic expression. A realization of a signifier can be accomplished through several discourses. For example, the graph shown in Figure 4 and the symbolic expression $y=(2x+1)(2x-1)$ are two realizations of a quadratic function, the first being a signifier in a visual-graphical discourse while the second is a signifier in a symbolic discourse. Generally, for an expert, a quadratic function as a signifier could be realized (or signified) via a table of numbers, symbolic expressions, graphical drawing and more. The way we talk about tables of values, graphs or algebraic expressions is different, as each of them belongs to a different discourse. A learner needs to be able to participate in these different discourses, but also to "same" them into a unified discourse about quadratic functions. The richness of realizations for a signifier can be captured by a realization tree (Figure 5).

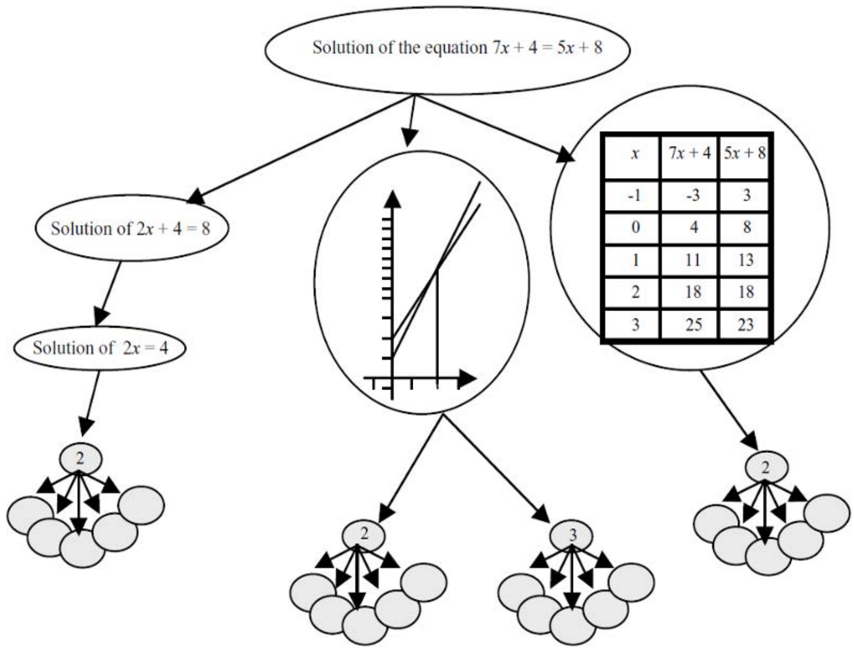


Figure 5. A realization tree for the signified solution of the equation $7x + 4 = 5x + 8$ with three signifiers – symbolic, graphic and numeric. Taken from [19, p. 165]. © 2008 Cambridge University Press. Reproduced with permission of Cambridge University Press through PLSclear.

Discussion: The shift in perspectives and some implications

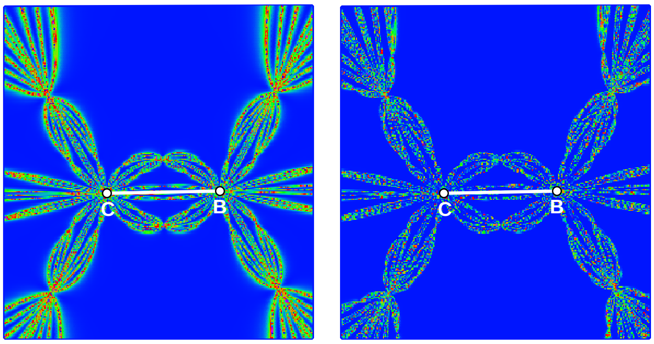
In all three perspectives, representations (or realizations) of mathematical objects are essential in mathematical thinking, teaching and learning, to the extent that no mathematical understanding (or discourse) is possible without them! In Goldin's perspective there is a key dialectic between internal and external representations: teaching mathematics most effectively happens when we understand students' learning of external representations and structured mathematical activities and effectively make use of such an understanding to influence their internal representations. For Duval a fundamental and necessary process in mathematical learning and understanding is that of conversion from one register of semiotic representation to another. Moreover, Duval's theory explicitly stands on the assumption that mathematical objects are not directly accessible, which suggests their existence in some inaccessible-to-us reality. As discussed earlier, this is a typical philosophical stance that is arguably present in other theoretical perspectives, including that of Goldin. In Sfard's approach, however, an important shift seems to occur: mathematical objects no longer exist anywhere other than in discourse itself. Therefore, to "know" a mathematical object means to be able to talk about it through narratives accepted within a community of mathematicians, and through discursive practices, we learn to recognize and express realizations of such an object.

Therefore, in Sfard's theory, a very important process consists in coming to see two "things" that we previously saw as different as the same, that is, as realizations of the same discursive object. A way into understanding students' mathematical learning, in this perspective, is through their discourse, and by the identification of patterns in what is said and done. This perspective opens new avenues of research, providing analytical tools for observing teaching and learning practices not only in contexts in which canonical representations are "presented" to the students, but also in settings in which students are invited to "invent" their own [3], or make sense of feedback stemming from interactions with a range of physical or digital artefacts. In line with this thinking, as educators we need to stay open to multiple creative realizations, and not lock the curriculum to a narrow selection of standardized representations, while disallowing or "hiding" others. This is particularly important when considering the diversity of the learner population, who to different extents may need to access different kinds of realizations in "non-standard" ways. As examples, think of the ways in which a blind student might realize function through non-visual sensory forms, or how mathematical discourse would be different under the grammars of signed compared to spoken language.

Moreover, we see some similarities between Sfard's shift away from a Platonic conception of mathematical objects and their representations, and the position advanced by Schoenfeld [18] and Li [14], who in their Teaching for Robust Understanding Framework take an Aristotelic stance, arguing in favour of "empirical" mathematics. That is, mathematics can and should be seen as a set of products created through experience (as opposed to pre-existing in an inaccessible realm). This perspective allows for what they (and we) see as a necessary focus on students' experience, in which pedagogy is not conceived of as "what should the teacher do" so much as "what mathematical experiences should students have in order for them to develop into powerful thinkers?" [14, p. 8]. For mathematical experiences to accomplish this, Li and Schoenfeld argue that they need to provide not only opportunities for making sense of the mathematics at stake, but also – and perhaps more importantly because education has focused less on this – they need to involve *sense-making* processes [15], highlighting "the importance for students to experience mathematics through creating, designing, developing, and connecting mathematical ideas" [14, p. 6]. Many educational experiences of this sort involve the use of physical or (more recently) digital artefacts that provide interactive and/or dynamic representations (which may be or become realizations of mathematical objects for the students).

As an example, consider the following task, explored by Sinclair [21]: take the three vertices of a triangle ABC and reflect them each across the opposite side of the triangle to obtain a new "reflex" triangle DEF; then iterate the process applying the reflection to triangle DEF, and so on. This problem can be approached in many different ways, involving different representations. We argue that working in a dynamic geometry environment (like Geometer's Sketchpad, Cabri Géomètre, GeoGebra, Desmos, etc.) can offer many students access into mathematical reasoning through sense-making processes. In this problem, for example, Sinclair explains how she used The Geometer's Sketchpad to explore a typical conjecture, that is, that the reflections eventually converge to an equilateral triangle. The software allowed her to iteratively reflect an arbitrary triangle and compute its measurements, effortlessly. Dragging vertex A led her to quickly realize that the conjecture was false: DEF can degenerate into a straight line. Moreover, she noticed that "DEF seemed to change in a very chaotic way" as she dragged A. Eventually, choosing a measure for how close to being equilateral each triangle was (in this case perimeter squared over area, which has a minimum for equilateral triangles), and creating overall pictures of the changing measurement like those in Figure 6, she found confirmation of the chaotic behaviour. The splitting of the "branches" confirms that small changes in the position of point A can give rise to radically different reflex triangles. But the symmetries of the branches also show regularity in the chaos, leading to new conjectures to be proved.

These sorts of experiences, that rely heavily on the interactive representations produced by software, are valuable for learners across the spectrum of mathematical capacities and needs. We note that in particular they may offer the possibility to "open doors" into participation in mathematical discourse: indeed, these tools offer students concrete-enough "things" to interact with and make sense of, allowing them to meaningfully start participating to mathematical discourse, without the need of formal language from the start. Recent studies suggest that through these means students who otherwise would remain excluded from mathematical discourse, actually find insightful ways to start participating to it (e.g., [1, 8]).



(a) (b)

Figure 6. Maps after two (a) and four (b) iterations of the "reflex triangle problem". Points A in the plane are coloured according to how close the reflex triangle is to equilateral.

This takes us back to the pressing need to conceive theoretical languages that allow consistent "talking about" dynamism in representations [1, 2]. Imagine, for example, a theoretical language through which we could differentiate between representing/realizing a mathematical phenomenon through dragging a finger over a touch screen, versus representing/realizing the same mathematical phenomenon with ones' whole body – or recalling those embodied experiences in one's mind when later encountering that mathematical idea in a different form.

We are not arguing that the theoretical lens of Commognition is *the* solution to this open problem; indeed, much research is still needed, and some is being carried out as we write. For example, a special issue of the journal Digital Experiences in Mathematics Education (in preparation) has been devoted to research "supporting transitions within, across and beyond digital experiences for the teaching and learning of mathematics", in which a variety of theoretical approaches are used to describe and study the three types of transition (within, across and beyond digital experiences). However, Sfard's perspective seems to embody an important shift that leaves behind the contradictory binary of the inaccessible-to-us world of perfect mathematical objects and the "real world" with its messy experiences in which we learn to recognize and produce realizations. Instead, it puts discourse, i.e. what is said and done by the community of all those who do mathematics, right at the forefront.

Acknowledgements. We wish to graciously thank Al Cuoco for the feedback he has generously offered us through numerous discussions on drafts of this paper.

References

[1] A. Baccaglini-Frank, To tell a story, you need a protagonist: how dynamic interactive mediators can fulfill this role and foster explorative participation to mathematical discourse. *Educ. Stud. Math.* **106**, 291–312 (2021)

[2] A. Baccaglini-Frank, C. Finesilver, S. Okumus and M. Tabach, Introduction to the work of TWG24: representations in mathematics teaching and learning. In *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education (Utrecht, 2019)*, Utrecht, Netherlands, hal-02394695 (2019)

[3] A. A. Cuoco and F. R. Curcio, *The roles of representations in school mathematics*. NCTM, Reston, VA (2001)

[4] E. Deliyianni, A. Monoyiou, I. Elia, C. Georgiou and E. Zannettou, Pupils' visual representations in standard and problematic problem solving in mathematics: their role in the breach of the didactical contract. *European Early Childhood Education Research Journal* **17**, 95–110 (2009)

[5] R. Duval, *Understanding the mathematical way of thinking – the registers of semiotic representations*. Springer International Publishing AG (2017)

[6] R. Duval, Registers of semiotic representation. Mathematical representations. In *Encyclopedia of Mathematics Education*, 2nd ed., edited by S. Lerman, Springer, 724–727 (2020)

[7] C. Finesilver, Low-attaining students' representational strategies: tasks, time, efficiency, and economy. *Oxford Review of Education* **43**, 482–501 (2017)

[8] C. Finesilver, L. Healy and A. Bauer, Supporting diverse approaches to meaningful mathematics: from obstacles to opportunities. In *Enabling Mathematics Learning of Struggling Students: International Perspectives*, edited by Y. P. Xin, R. Tzur and H. Thouless, Springer (in press)

[9] F. Furinghetti, J. M. Matos and M. Menghini, From mathematics and education, to mathematics education. In *Third International Handbook of Mathematics Education*, edited by M. A. (Ken) Clements, A. J. Bishop, C. Keitel, J. Kilpatrick and F. K. S. Leung, Springer, 273–302 (2013)

[10] E. von Glasersfeld, An exposition of constructivism: why some like it radical. In *Constructivist Views on the Teaching and Learning of Mathematics*, edited by R. B. Davis, C. A. Maher and N. Noddings, National Council of Teachers of Mathematics, Reston, VA, 19–29 (1990)

[11] G. A. Goldin, Systems of representations and the development of mathematical concepts. In *The Roles of Representations in School Mathematics*, edited by A. A. Cuoco and F. R. Curcio, NCTM, Reston, VA, 1–23 (2001)

[12] G. A. Goldin, Mathematical representations. In *Encyclopedia of Mathematics Education*, 2nd ed., edited by S. Lerman, Springer, 566–572 (2020)

[13] G. Goldin and N. Shteingold, Systems of representations and the development of mathematical concepts. In *The Roles of Representation in School Mathematics*, edited by A. A. Cuoco and F. R. Curcio, NCTM, 1–23 (2001)

[14] Y. Li and A. H. Schoenfeld, Problematizing teaching and learning mathematics as "given" in STEM education. *Int. J. Stem. Educ.* **6**, 1–13 (2019)

[15] W. McCallum, Sense-making and making sense, blogs.ams.org/matheducation/2018/12/05/sense-making-and-making-sense/ (2018)

[16] N. C. Presmeg, Preface, In *Representations and Mathematics Visualization*, edited by F. Hitt, North American Chapter of the International Group for the Psychology of Mathematics Education, ix–xvi (2002)

[17] E. Robotti, A. Baccaglini-Frank, G. Sensevy and T. Fritzlar, TWG24: representations in mathematics teaching and learning. Introduction to the papers of TWG24. In *Proceedings of CERME 10 (Dublin, 2017)*, hal-01950557 (2017)

[18] A. H. Schoenfeld, Classroom observations in theory and practice. *ZDM Mathematics Education* **45**, 607–621 (2013)

[19] A. Sfard, *Thinking as communicating: human development, the growth of discourses, and mathematizing*. Cambridge University Press (2008)

[20] A. Sfard, Commognition. In *Encyclopedia of Mathematics Education*, edited by S. Lerman, Springer, Cham (2018)

[21] N. Sinclair, Computer-based technologies and plausible reasoning. In *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*, edited by M. Carlson and C. Rasmussen, Mathematical Association of America, 233–244 (2008)

Anna Baccaglini-Frank is an associate professor at the Department of Mathematics of the University of Pisa, Italy. Her main research interests are in technology-enhanced learning at all grade levels. Among other activities, she serves on the editorial board of Educational Studies in Mathematics, is an associate editor of Digital Experiences in Mathematics Education, and leads a the Thematic Working Group 24: "Representations in Mathematics Teaching and Learning" at the congress of the European Society for Research in Mathematics Education (ERME).

anna.baccaglinifrank@unipi.it

Carla Finesilver is a senior lecturer at King's College London. Her research interests are in mathematics education and inclusive education, in particular improving accessibility for diverse learners. She is a co-editor of the special issue of Digital Experiences in Mathematics "Supporting transitions within, across and beyond digital experiences for the teaching and learning of mathematics", currently under development, and co-leads the CERME Thematic Working Group 24.

carla.finesilver@kcl.ac.uk

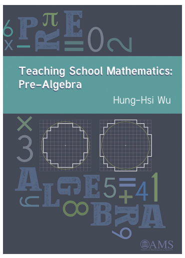
Michal Tabach is professor at the School of Education, Tel Aviv University, and researcher in mathematics education. Her main research interests are integrating technology to enhance teaching and learning of mathematics at all levels, and knowledge shifts in the mathematics classroom. She has been Secretary of PME, the international group of Psychology of Mathematics Education. She is currently an associate editor of Educational Studies in Mathematics. She co-leads the CERME Thematic Working Group 24.

tabachm@tauex.tau.ac.il

Book reviews

Teaching School Mathematics: Pre-Algebra by Hung-Hsi Wu

Reviewed by António de Bivar Weinholtz



This is the second book of a series of six covering the entire K-12 curriculum, as an instrument for the mathematical education of school teachers. It follows a first volume entitled "Understanding Numbers in Elementary School Mathematics", which covers a substantial part of the mathematical curriculum of grades K-6 (namely, numbers and operations), but also some topics more likely to be part of the curriculum of grades 7 and 8, the reason being that clearly the mathematical knowledge of elementary school teachers should go beyond what they actually have to teach.

I strongly recommend readers of this review to start by reading the review [1] of that first volume.[1] There, can one find the reasons why I deem this set of books a milestone in the struggle for a sound mathematical education of schoolchildren. I shall not repeat here all the historical and scientific arguments that support this claim, but regarding the second volume, I have to state that once again, although it is written for school teachers as a tool for their mathematical education (both during pre-service years and for their professional development while in activity), and also as a resource for textbook authors, the set of its potential readers should not be restricted to those for which it was primarily intended. Indeed, exactly as I wrote regarding the first volume, I claim that its readers may and should include anyone with the basic ability to appreciate the beauty of the use of human reasoning in our quest to understand the world around us and the capacity and the will to make the necessary effort required here, as it is for any enterprise that is really worthwhile.

This volume and the one the author explicitly refers to as its companion, namely the third volume in the series "Teaching School Mathematics: Algebra", both address the mathematics generally taught in grades 6–8 (as explained in the third volume), but also including topics that can be part of the curriculum of grades 5 and 9 (as explained in this volume). The first three chapters deal with numbers and operations, starting with fractions and including rational (relative) numbers, and also some other specific topics such as finite probability, the Euclidean algorithm, and some of its applications. The final two chapters give an introduction to geometry and geometric measurement.

As this volume together with its companion are meant to be self-contained as far as the mathematics middle school teachers have to teach is concerned, the chapters dealing with numbers and operations needed to have some overlap with the corresponding chapters in the first volume of this series, with the sole exception of the section on finite probability. What is given here, however, is a new presentation of the different subjects, although along the same lines. Therefore, much that was written in the review of the first volume still applies to the corresponding chapters in this second volume. However, we can find some new ideas in the details, and the topics are presented in a somewhat more synthetic form. It is a pleasure to revisit these fundamental ideas that can be considered, in a certain sense, the main core of school mathematics, and even someone who has thoroughly read the first volume can benefit from this new synthesis. Of the new details, let me refer for instance to the motivation of the concept of the product of fractions. Not only is the priority in this volume no longer given, as it was in the first volume, to the formula for the area of a rectangle, but rather to the concept of "a fraction of a fraction", but also, in this volume, when presenting the latter concept as the basic way to introduce the definition of multiplication (which was only presented as a second possible equivalent definition in the first volume), the author shows that the resulting product is the only possible one if one wants to have an operation that is associative and at the same time extends directly to fractions the meaning of "multiplying by a whole number". It is also understood, of course, that the definition of division by a whole number is exactly the same as the one already adopted for whole numbers, once we

[1] In the printed version of the review [1], there is an unfortunate misspelling of the name of the author of the reviewed book, both in the title and in the text; the correct spelling of the name of the author is, of course, Hung-Hsi Wu.

have the concept of multiplication of fractions by such numbers (defined, as expected, as an iterated addition). In this way, a new independent argument is added for the way one should extend multiplication to fractions, leading to the simple rule that, sadly, is so often presented as something students just have to learn by rote. One can, of course, devise similar alternative ways of motivating this definition; for instance, one can make use of the requirement of commutativity to obtain the rule to multiply a whole number by a fraction, having first figured out what it should naturally mean to multiply any fraction by a whole number, and consequently what it should mean to divide any fraction by a whole number. Then, once one notices that this forces us to define multiplication of a whole number by an fraction exactly as division of the whole number by the denominator of the fraction followed by multiplication of the result by the numerator, one can naturally extend this rule, as a definition, from whole numbers to fractions in general. While one can argue whether this slightly less synthetic (and in a certain sense less "logically compelling") argument, or some similar one, would be more or less appealing to students of a specific grade, what is essential is the requirement that, in general, operations with fractions should be understood and presented as natural extensions of the corresponding operations with whole numbers, and that a clear distinction should be made between "definitions" and "rules that can be justified using definitions". With this set of books at hand, there is no excuse for school teachers, textbook authors and government officials to persist in the unfortunate practice of trying to serve to students this fundamental part of school mathematics in a way that is in fact unlearnable …

Another example of a new formulation is a somewhat more explicit wording of the "Fundamental Assumption of School Mathematics" regarding irrational numbers.

The two final chapters, on geometry and geometrical measurement, are completely new with respect to the first volume. This part of the book will thus give rise to a somewhat more extensive commentary here. Before engaging with the subject itself, the author explains the main goals to be attained, and describes the problems one faces when trying to attain these goals, including an analysis of the situation of school geometry in the last few decades. The core of the geometrical topics to be treated in middle school is defined as a working knowledge of similar triangles, as a tool to set up the intuitive foundation for a more precise discussion of the concepts of congruence and similarity in high school geometry; the role of similar triangles in the study of linear equations of two variables is emphasized, and we are reminded of the continuing crisis in the teaching of school geometry over the last four decades or more, and of the essential discontinuity between the middle school and high school curricula all along that extensive period. While this historical analysis is mainly valid for the USA, it can certainly apply, at least partially, to many other countries around the world. The tension between two extremes, i.e., on one side a purely axiomatic approach and on the other side an exclusively "intuitive" consideration of geometric entities with scarcely any concern for even minimal organization, strongly relying on "manipulatives" and computer software, has done extensive damage to the teaching of geometry in schools, and the author points out the urgent need to get past this unfortunate situation. With these considerations in mind, he sets up a progression starting with some advice on the use of free hand drawing as a way to gain some initial necessary geometric intuition, followed by a rich set of geometric constructions using plastic triangles, ruler and compass; finally, he addresses the problem of reaching the main results on congruence and similarity in a way that introduces students to sound deductions in geometry without the undue burden of a strict axiomatic construction. The chosen main ingredients for such a path are the basic isometries in a plane, namely, translations along a vector, reflections across a line and rotations around a point, complemented by dilations, for the definition of similarity; a carefully weighed equilibrium between mathematical precision and geometric intuition in order to attain the prescribed goals is aimed at and successfully reached. The author advocates the use of transparencies to supply students with the needed geometrical intuition on the assumptions that have to be made at this stage regarding the basic isometries, and on the use of these assumptions for proper justification of the basic geometric theorems to be taught. The manner in which this can be successfully accomplished is described and illustrated in complete detail.

Unlike the case of the topics on numbers and operations, in the case of geometry there may be a wider choice of different ways to attain the same goals. Other paths respecting the same general basic principles have been proposed, even in recent curricular reforms. Just to give a hint of some of the other ways of dealing with the difficult equilibrium between mathematical precision and intuition, let me point out a few items that can be used as a basis for an alternative path. One could use a property that, in a particular case, justifies the construction used in this book to draw an angle with amplitude equal to a given angle, on a given half-line (i.e., having that half-line as one of its sides) as a definition of "equality of amplitude for angles". Admitting the effectivity and consistency of a slightly generalized form of this construction is, in a certain sense, equivalent to the usual axiom that in some axiomatic constructions of elementary (Euclidean or even absolute) geometry replaces the classical SAS criterion of congruence of triangles. This could be a basis for a justification of the other congruence criteria for triangles. Parallelism can be dealt with using, as a criterion for two lines in a plane to be parallel, the equality of the corresponding angles determined in the pair of lines by an intersecting third line. This same property can also be used to justify the construction of a line parallel to a given line and passing through a given point, like the one illustrated in this volume. These two stated criteria, admitted without proof (part of what is formulated in the second one is in fact equivalent, in a certain sense, to the parallel axiom), can be made "intuitive" by the fact that they provide practical ways of obtaining the "transport" of an angle and the construction of a parallel to a given line through a given point respectively, in each case using only a finite number of points and the transport of distance between two points. As for the concepts of congruence and similarity, one could directly use the manner in which the author defines the scale drawing of a figure as a definition of similarity, and define congruence to be the case where the scaling factor is equal to 1. The basic isometries in the plane, and vectors, can be introduced in a second stage, and their properties similarly studied with a good equilibrium between precision and intuition.

We are perhaps in a situation where we still lack the practical experience that allows us to decide if among the various different proposals for "paths of presentation" of school geometry, we should clearly prefer one or another, considering only those that satisfy the basic principles stated in this book. We can try to examine what has been done in the more or less ancient past in schools around the world, and what the evolution has been over recent decades and in some cases even over the last few years; there is still perhaps some room for a confrontation of different hypotheses. I am nevertheless convinced that the general principles stated in this book, and the diagnosis of the disastrous evolution of school geometry in many countries in the last half-century should serve as a compelling guide in this matter. At any rate, this set of volumes is pioneering in that it gives a full presentation of a path that can be followed from middle school to high school, in a manner that respects all of these sound principles for a renewed teaching of school geometry. It is to be hoped that this or perhaps other equivalent ways of meeting these requirements will have the opportunity to prove their effectiveness in schools, thus reversing the sad diagnosis of the present situation made in this volume.

As expected, the final chapter on geometrical measurement is also an excellent example of balance between mathematical precision and controlled intuition, leading to the usual formulas for length, area and volume of basic geometric entities, with a unified vision of geometrical measure.

As in the first volume of this series, the author provides the reader with numerous illuminating activities for every topic, as well as an excellent choice of a wide range of exercises.

Hung-Hsi Wu, *Teaching School Mathematics: Pre-Algebra.*
American Mathematical Society, 2016, 383 pages, Hardback ISBN 978-1-4704-2720-7, eBook ISBN 978-1-4704-3009-2.

References

[1] A. de Bivar Weinholtz, Book review: Understanding numbers in elementary school mathematics. *Eur. Math. Soc. Mag.* **122**, 66–67 (2021)

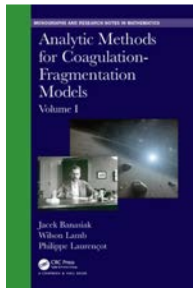
António de Bivar Weinholtz is a retired associate professor of mathematics of the University of Lisbon Faculty of Science, where he taught from 1975 to 2009. He was a member of the scientific coordination committee of the new curricula of mathematics for all the Portuguese pre-university grades (2012–2014).

abivar@sapo.pt

DOI 10.4171/MAG-76

Analytic Methods for Coagulation-Fragmentation Models, Volumes I+II
by Jacek Banasiak, Wilson Lamb and Philippe Laurençot

Reviewed by Barbara Niethammer



Coagulation and fragmentation are of fundamental importance in a vast variety of applications such as aerosol physics, polymerization, blood agglomeration or grouping of certain species, to name just a few. The first approach to describe such processes goes back to Marian von Smoluchowski, who in 1916 developed the first deterministic model for coagulation in a colloidal gold solution. The main assumption in his approach was that only two clusters combine at a time and that the rate at which this happens depends only on the product of the number-density of the respective cluster sizes, a hypothesis similar to the molecular chaos assumption in the Boltzmann equation. This leads to an infinite system of coupled differential equations for the density of cluster sizes and involves a so-called rate kernel which depends on the microscopic details of the specific coagulation process. Later the model was extended in various ways; in particular, different forms of fragmentation were added, allowing a cluster to fragment into two or more smaller ones. While the original model is discrete, allowing only for integer cluster sizes, the continuous version of the model is also of interest, in particular if one is interested in the behaviour of the system over large times.

It should be emphasized that coagulation-fragmentation equations are not only relevant from the point of view of applications; they are also fascinating due to the rich structure that solutions can display. Indeed, depending on the rate kernels, coagulation-fragmentation equations feature the possibility of loss of mass which can happen in finite or infinite time. In pure coagulation, for example, if the kernel grows faster than linearly, infinitely large particles are created in finite time, a phenomenon known as gelation. This can be directly linked to gelation in polymers where polymer chains can become so long such that suddenly in time the viscosity increases significantly and the polymer displays completely different properties. Similarly, in pure fragmentation mass can be lost due to the formation of zero sized particles, also called dust, a phase transition called scattering. In some combined coagulation-fragmentation equations, the loss of mass might also occur in infinite time, in the sense that the solution converges to an equilibrium that has lower mass than the solution for any finite time. This feature, typically termed nucleation, is prominent for example in the case of the well-known Becker–Döring system, that can describe the formation of liquid particles in a supersaturated gas.

Even though Smoluchowski's theory was developed more than 100 years ago, a thorough mathematical theory of coagulation-fragmentation equations going beyond explicitly solvable models started only in the 1980s. The first and until recently only book on the mathematical analysis of deterministic models appeared in 1994, written by Pavel Dubovski. However, in particular since the turn of the century the analysis of these equations has experienced another boost and a corresponding progress in new methods.

Thus the monograph under review is a very welcome, useful and state-of-the-art addition to the literature, written by three experts who have made fundamental contributions to the analysis of coagulation-fragmentation equations. The book takes up new analytical developments and provides an exhaustive treatment in particular of the two main methods that have been further developed in the last twenty years, namely semi-group theory and weak compactness methods. The work consists of two volumes, and starts after a short introduction with a description of coagulation and fragmentation processes in important applications, including numerous references, different modeling approaches and some classical results in particular for so-called solvable models. It also addresses many additional aspects, such as gelation and scattering, or the approach to universal self-similar long-time behaviour. Chapter 3 introduces notation and conventions that are used throughout the book, including in particular relevant function spaces, such as weighted L^1-spaces, as well as linear operators.

The remainder of Volume I is entirely devoted to the semi-group approach for linear fragmentation models. Chapter 4 starts with motivating why semi-groups are useful in this context, and gives an introduction to possible difficulties, such as loss of mass due to shattering or the possible non-uniqueness of solutions. The remaining subsections of Chapter 4 provide a comprehensive summary of semi-group theory. It includes the relevant definitions, the major theorems and further aspects such as inhomogeneous problems, semi-linear equations and perturbation methods. This informative section is not only the basis for Chapter 5 but also of interest in itself, and can be recommended to any reader who wants a brief introduction to semi-group theory.

In Chapter 5 the results from semi-group theory are applied to fragmentation equations. There are two main parts: in Chapter 5.1 pure fragmentation is considered and a rather complete theory is presented for the case where the fragmentation coefficients are separable. Chapter 5.2 deals with the case where transport in size space is added to the fragmentation term, an extension that is particularly relevant from the point of view of applications. The semi-group approach leads to a satisfactory theory of the well-posedness of fragmentation equations and their extensions, and gives information on the analyticity and so-called honesty of the corresponding fragmentation semi-group. In addition, topics such as the approximation of solutions by some cut-off are discussed, as well as the long-time behaviour of solutions.

Volume II addresses nonlinear models which in particular include coagulation. In the case where fragmentation dominates coagulation, semi-group theory can be applied to establish existence and uniqueness of solutions. If coagulation is dominant, which includes the case of pure coagulation, weak compactness methods have been particularly successful. In this case, approximate solutions are constructed, typically by a cut-off of the respective kernels. For these approximate solutions, moment and uniform integrability estimates need to be established to ensure weak compactness of the approximating sequences. This flexible method provides existence results, both for solutions to the original equations and for self-similar solutions. Uniqueness and regularity of solutions must be proved separately, however, and typically require further assumptions on the kernels.

The mathematical tools for this approach are summarized in Chapter 7, while Chapter 8 deals with the well-posedness of coagulation-fragmentation equations. Of particular relevance is the case of mass conserving solutions which are obtained if the rate kernels do not grow too quickly for small and large cluster sizes. The results of this chapter are up-to date and contain, for example, well-posedness results for singular coefficients that have been established only rather recently. Chapter 9 discusses the phenomenon of gelation, both instantaneous or for later times, and scattering. In the not explicitly solvable cases, these results are established via integral inequalities.

Chapter 10 is devoted to the important issue of long-time behaviour of solutions. For homogeneous kernels, one expects for pure fragmentation or coagulation respectively that this is universal and solutions converge to a self-similar solution. This issue is quite well understood for the fragmentation equation and for the coagulation equations with solvable kernel, but it is still mostly open for the pure coagulation equation with non-solvable kernels. Self-similar solutions can be constructed using approximation schemes and weak compactness methods, but apart from a few special cases, their dynamic stability has not been established yet, and for particular kernels one may even expect instability. The authors give a complete account of current results on these questions.

Finally, Chapter 11 provides a short introduction into material that cannot be covered in detail in the two volumes, such as the Becker–Döring equations and coagulation-fragmentation equations with diffusion. Numerous references are provided for readers interested in these topics.

To study the book, a basic background in functional analysis is needed, but all further tools that are used are introduced in the two *Mathematical Toolbox* chapters. The theory presented there is rather detailed and complete in the sense that the assumptions on the initial data and the rate coefficients are very general and the proofs are presented in full detail. These two volumes provide an informative, extensive and inspiring introduction to the subject accessible to all researchers from graduate students to experienced scientists.

Jacek Banasiak, Wilson Lamb and Philippe Laurençot, *Analytic Methods for Coagulation-Fragmentation Models, Volumes I+II.* CRC Press, 2019, 676 pages, Hardback ISBN 978-0-367-23544-4 (set).

Barbara Niethammer is a professor at the Institute for Applied Mathematics at the University of Bonn. Her research focuses on the analysis of problems with multiple scales and high-dimensional dynamical systems as well as the study of long-time behaviour in models of mass aggregation.

niethammer@iam.uni-bonn.de

DOI 10.4171/MAG-77

Solved and unsolved problems

Michael Th. Rassias

The present column is devoted to Geometry/Topology.

I Six new problems – solutions solicited

Solutions will appear in a subsequent issue.

252

Prove that the space of unordered couples of distinct points of a circle is the (open) Möbius band. More formally, consider

$$(S^1 \times S^1) \setminus \{(x,x) \mid x \in S^1\}$$

and the equivalence relation on this space $(x,y) \equiv (y,x)$; prove that the quotient topological space is the (open) Möbius band.

Costante Bellettini (Department of Mathematics, University College London, UK)

253

In the Euclidean plane, let γ_1 and γ_2 be two concentric circles of radius respectively r_1 and r_2, with $r_1 < r_2$. Show that the locus γ of points P such that the polar line of P with respect to γ_2 is tangent to γ_1 is a circle of radius r_2^2/r_1.

Acknowledgement. I want to thank the professors who guided me in the first part of my career for giving me the ideas for these problems.

Paola Bonacini (Mathematics and Computer Science Department, University of Catania, Italy)

254

Let $A \subseteq \mathbb{R}^3$ be a connected open subset of Euclidean space, and suppose that the following conditions hold:

(1) Every smooth irrotational vector field on A admits a potential (i.e., it is the gradient of a smooth function).

(2) The closure $\overline{A}$ of A is a smooth compact submanifold of $\mathbb{R}^3$ (of course, with non-empty boundary).

Show that A is simply connected. Does this conclusion hold even if we drop condition (2) on A?

Roberto Frigerio (Dipartimento di Matematica, Università di Pisa, Italy)

255

A *regulus* is a surface in $\mathbb{R}^3$ that is formed as follows: We consider pairwise skew lines $\ell_1, \ell_2, \ell_3 \subset \mathbb{R}^3$ and take the union of all lines that intersect each of ℓ_1, ℓ_2, and ℓ_3. Prove that, for every regulus U, there exists an irreducible polynomial $f \in \mathbb{R}[x,y,z]$ of degree two that vanishes on U.

Adam Sheffer (Department of Mathematics, Baruch College, City University of New York, NY, USA)

256

(Enumerative Geometry). How many lines pass through 4 generic lines in a 3-dimensional complex projective space $\mathbb{CP}^3$?

Mohammad F. Tehrani (Department of Mathematics, University of Iowa, USA)

257

I learned about the following problem from Shmuel Weinberger. It can be viewed as a topological analogue of Arrow's Impossibility Theorem.

(a) A group of n friends have decided to spend their summer cottaging together on an undeveloped island, which happens to be a perfect copy of the closed 2-disk D^2. Their first task is to decide where on this island to build their cabin. Being democratically-minded, the friends decide to vote on the question. Each friend chooses his or her favourite point on D^2. The friends want a function that will take as input their n votes, and give as output a suitable point on D^2 to build. They believe, to be reasonable and fair, their "choice" function should have the following properties:

- (*Continuity*) It should be continuous as a function $(D^2)^n \to D^2$. This means, if one friend changes their vote by a small amount, the output will change only a small amount.
- (*Symmetry*) The n friends should be indistinguishable from each other. If two friends swap votes, the final choice should be unaffected.
- (*Unanimity*) If all n friends chose the same point x, then x should be the final choice.

For which values of n does such a choice function exist?

(b) The friends' second task is to decide where along the shoreline of the island they will build their dock. The shoreline happens to be a perfect copy of the circle S^1. Again, they decide to take the problem to a vote. For which values of n does a continuous, symmetric, and unanimous choice function $(S^1)^n \to S^1$ exist?

These are special cases of the following general problem in topological social choice theory: given a topological space X, for what values of n does X admit a social choice function that is continuous, symmetric, and unanimous? In other words, when is there a function $A\colon X^n \to X$ satisfying

- A is continuous,
- $A(x_1, \ldots, x_n)$ is independent of the ordering of $x_1, \ldots, x_n$, and
- $A(x, x, x, \ldots, x) = x$ for all $x \in X$?

Jenny Wilson (Department of Mathematics,
University of Michigan, USA)

II Open problem

Embeddings of contact domains

by Yakov Eliashberg (Department of Mathematics,
Stanford University, USA)

One of the cornerstones of symplectic topology, *Gromov's non-squeezing theorem*, see [6], asserts that for $n > 1$ the ball of radius $R > 1$ in the standard symplectic space $(\mathbb{R}^{2n}, \omega = \sum_1^n dx_j \wedge dy_j)$ does not admit a symplectic embedding into the domain

$$\{x_1^2 + y_1^2 < 1\} \subset \mathbb{R}^{2n},$$

while there is no volume constraints to do that. Since that time, the theory of symplectic embedding has made a lot of progress (see F. Schlenk's survey [8] for recent results).

The (non-)embedding results in contact geometry, which is an odd-dimensional analogue of symplectic geometry, are rarer; below, we discuss a few open problems.

Recall that a contact structure ξ on an $(2n+1)$-dimensional manifold M is a completely non-integrable hyperplane field. If ξ is defined by a Pfaffian equation $\alpha = 0$ for a differential 1-form α (and such a form can always be found locally, and if ξ is co-orientable even globally) then the complete non-integrability can be expressed by the condition that $\alpha \wedge d\alpha^n$ is a non-vanishing $(2n+1)$-form on M.

In this set of problems, we will restrict our attention to domains in the contact manifold $X \coloneqq \mathbb{R}^{2n} \times S^1$, $S^1 = \mathbb{R}/\mathbb{Z}$, endowed with the contact structure

$$\xi \coloneqq \Big\{dz + \frac{1}{2}\sum_1^n x_j\, dy_j - y_j\, dx_j = 0\Big\}.$$

Given a *bounded* domain $U \subset \mathbb{R}^{2n}$, we set

$$\hat{U} \coloneqq U \times S^1 \subset \mathbb{R}^{2n} \times S^1 = X,$$

and refer to $\hat{U}$ as a *quantized* domain U. We say that a domain $\hat{U}_1$ admits a *contact embedding* into a domain $\hat{U}_2$ if *there is a contact isotopy $f_t\colon \hat{U}_1 \to X$, starting with the inclusion $f_0\colon \hat{U}_1 \hookrightarrow X$ such that $f_1(\hat{U}_1) \subset \hat{U}_2$*. Note that any Hamiltonian isotopy which moves U_1 into U_2 lifts to a contact isotopy moving $\hat{U}_1$ into $\hat{U}_2$. Hence we will refer to the problem of contact embeddings between the domains $\hat{U}_1$ and $\hat{U}_2$ as a quantized version of the corresponding symplectic embedding problem of U_1 to U_2.

Denote by $B^{2n}(R)$ the $2n$-dimensional open ball of radius R and by $P(r_1, \ldots, r_k)$ the polydisk $B^2(r_1) \times \cdots \times B^2(r_n) \subset \mathbb{R}^{2n}$, where $0 < r_1 \le r_2 \le \cdots \le r_n$. It was shown in [3] that *if $\pi r_1^2 < k < \pi r_2^2$ for any integer $k \ge 1$, then $\hat{B}^{2n}(r_2)$ does not admit a contact embedding into $\hat{B}^{2n}(r_1)$*. Another theorem from [3] states that *if $\pi R^2 < 1$, then $\hat{B}^{2n}(R)$ admits a contact embedding into $\hat{B}^{2n}(r)$ for any $r > 0$*. The former result was improved in [2, 5] to show that for any $r_1 < r_2$ with $\pi r_2^2 > 1$ there is no contact embedding of $\hat{B}^{2n}(r_2)$ into $\hat{B}^{2n}(r_1)$. Recall that Gromov's *symplectic width* $\mathrm{Width}_{\mathrm{Gr}}(U)$ of a domain $U \subset \mathbb{R}^{2n}$ can be defined as the supremum of $\pi\rho^2$ such that $B^{2n}(\rho)$ can be symplectically embedded into the domain U. The above results can be slightly generalized to the following statement, see [3].

If, for two domains $U_1, U_2 \subset \mathbb{R}^{2n}$, we have

$$\mathrm{Width}_{\mathrm{Gr}}(U_2) > \mathrm{Width}_{\mathrm{Gr}}(U_1) \quad \textit{and} \quad \mathrm{Width}_{\mathrm{Gr}}(U_2) > 1,$$

then the quantized domain $\hat{U}_2$ does not admit a contact embedding into $\hat{U}_1$.

Very little is known about embeddings of contact domains beyond the above results. Let us formulate a couple of concrete problems concerning quantized versions of some relatively old embedding results in symplectic topology. As we already mentioned above, many new obstructions to symplectic embeddings were found in recent years. It is unknown whether any of them hold in the quantized versions.

*258**

Contact packing problem. *Suppose that $\pi R^2 > \pi r^2 > 1$. Is there a maximal number of quantized balls $\hat{B}^{2n}(r)$ which admit a contact packing into $\hat{B}^{2n}(R)$? And if the answer is "yes", then what is this number?*

Here we say that $\hat{B}^{2n}(R)$ admits a *contact packing* by k quantized balls $\hat{B}^{2n}(r)$ if the disjoint union

$$\underbrace{\hat{B}^{2n}(r) \sqcup \cdots \sqcup \hat{B}^{2n}(r)}_{k} \subset X$$

admits a contact embedding into $\hat{B}^{2n}(R)$. Note that the corresponding *symplectic packing* problem was intensively studied beginning with the seminal paper by Gromov [6], where he proved that the packing of the ball $B^{2n}(R)$ with 2 disjoint balls $B^{2n}(r)$ is possible if and only if $R^2 > 2r^2$. For $n = 2$, the problem was significantly advanced by D. McDuff and L. Polterovich in [7], and then completely solved by P. Biran in [1]. In the contact case, Problem 258* is completely open.

*259**

Rotating quantized polydisks. *For $r < R$, consider the standard contact inclusion $(x, y) \mapsto (-y, x)$. Let $\hat{j} = j \times \mathrm{Id}$, $\hat{\psi} = \psi \times \mathrm{Id}$ be the corresponding contact inclusion $\hat{P}(r, r) \to \hat{P}(R, R)$, and consider the contactomorphism $\psi \times \mathrm{Id} \colon X = \mathbb{R}^{2n} \times S^1 \to \mathbb{R}^{2n} \times S^1 = X$. When are the embeddings $\hat{j}, \hat{\psi} \circ \hat{j} \colon \hat{P}(r, r) \hookrightarrow \hat{P}(R, R)$ contact isotopic?*

Note that a theorem of Floer–Hofer–Wysocki, see [4], states that when $2r^2 > R^2$ the symplectic embeddings $j, \psi \colon P(r, r) \to \mathbb{P}(R, R)$ are not symplectically isotopic.

References

[1] P. Biran, Symplectic packing in dimension 4. *Geom. Funct. Anal.* **7**, 420–437 (1997)

[2] S.-F. Chiu, Nonsqueezing property of contact balls. *Duke Math. J.* **166**, 605–655 (2017)

[3] Y. Eliashberg, S. S. Kim and L. Polterovich, Geometry of contact transformations and domains: orderability versus squeezing. *Geom. Topol.* **10**, 1635–1747 (2006)

[4] A. Floer, H. Hofer and K. Wysocki, Applications of symplectic homology. I. *Math. Z.* **217**, 577–606 (1994)

[5] M. Fraser, Contact non-squeezing at large scale in $\mathbb{R}^{2n} \times S^1$. *Internat. J. Math.* **27**, 1650107 (2016)

[6] M. Gromov, Pseudo holomorphic curves in symplectic manifolds. *Invent. Math.* **82**, 307–347 (1985)

[7] D. McDuff and L. Polterovich, Symplectic packings and algebraic geometry. *Invent. Math.* **115**, 405–434 (1994)

[8] F. Schlenk, Symplectic embedding problems, old and new. *Bull. Amer. Math. Soc. (N.S.)* **55**, 139–182 (2018)

III Solutions

245

We consider a setting where there is a set of m candidates

$$C = \{c_1, \ldots, c_m\}, \quad m \geq 2,$$

and a set of n voters $[n] = \{1, \ldots, n\}$. Each voter ranks all candidates from the most preferred one to the least preferred one; we write $a \succ_i b$ if voter i prefers candidate a to candidate b. A collection of all voters' rankings is called a *preference profile*. We say that a preference profile is *single-peaked* if there is a total order $\lhd$ on the candidates (called the *axis*) such that for each voter i the following holds: if i's most preferred candidate is c and $a \lhd b \lhd c$ or $c \lhd b \lhd a$, then $b \succ_i a$. That is, each ranking has a single "peak", and then "declines" in either direction from that peak.

(i) In general, if we aggregate voters' preferences over candidates, the resulting majority relation may have cycles: e.g., if $a \succ_1 b \succ_1 c$, $b \succ_2 c \succ_2 a$ and $c \succ_3 a \succ_3 b$, then a strict majority (2 out of 3) voters prefer a to b, a strict majority prefer b to c, yet a strict majority prefer c to a. Argue that this cannot happen if the preference profile is single-peaked. That is, prove that if a profile is single-peaked, a strict majority of voters prefer a to b, and a strict majority of voters prefer b to c, then a strict majority of voters prefer a to c.

(ii) Suppose that n is odd and voters' preferences are known to be single-peaked with respect to an axis $\lhd$. Consider the following voting rule: we ask each voter i to report their top candidate $t(i)$, find a median voter i^*, i.e.,

$$|\{i : t(i) \lhd t(i^*)\}| < \frac{n}{2} \quad \text{and} \quad |\{i : t(i^*) \lhd t(i)\}| < \frac{n}{2},$$

and output $t(i^*)$. Argue that under this voting rule no voter can benefit from voting dishonestly, if a voter i reports some candidate $a \neq t(i)$ instead of $t(i)$, this either does not change the outcome or results in an outcome that i likes less than the outcome of the truthful voting.

(iii) We say that a preference profile is *1D-Euclidean* if each candidate c_j and each voter i can be associated with a point in $\mathbb{R}$ so that the preferences are determined by distances, i.e., there is an embedding $x \colon C \cup [n] \to \mathbb{R}$ such that for all $a, b \in C$ and $i \in [n]$, we have $a \succ_i b$ if and only if $|x(i) - x(a)| < |x(i) - x(b)|$. Argue that a 1D-Euclidean profile is necessarily single-peaked. Show that the converse is not true, i.e., there exists a single-peaked profile that is not 1D-Euclidean.

(iv) Let P be a single-peaked profile, and let L be the set of candidates ranked last by at least one voter. Prove that $|L| \leq 2$.

(v) Consider an axis $c_1 \lhd \cdots \lhd c_m$. Prove that there are exactly 2^{m-1} distinct votes that are single-peaked with respect to this axis. Explain how to sample from the uniform distribution over these votes.

These problems are based on references [4] (parts (i) and (ii)), [2] (part (iii)) and [1, 5] (part (v)); part (iv) is folklore. See also the survey [3].

Edith Elkind (University of Oxford, UK)

Solution by the proposer

(i) We can restrict the voters' preferences to the set $\{a, b, c\}$; the reader can check that a restriction of a single-peaked profile to a subset of candidates remains single-peaked. We consider three cases depending on how a, b and c are ordered by the axis $\lhd$.

Case 1: $a \lhd b \lhd c$ or $c \lhd b \lhd a$. Then all voters who prefer a over b have a as their top choice and hence prefer a to c.

Case 2: $b \lhd a \lhd c$ or $c \lhd a \lhd b$. All voters who prefer c over a have c as their top choice and hence prefer a to b; therefore, these voters are in minority.

Case 3: $a \lhd c \lhd b$ or $b \lhd c \lhd a$. This is impossible: all voters who prefer b to c have b as their top choice, so we have a strict majority preferring b over a, a contradiction.

(ii) Suppose the winner under truthful voting is a. Consider a voter i. If $t(i) = a$, then i cannot improve the outcome by lying. So suppose $t(i) \lhd a$ (the case $a \lhd t(i)$ is symmetric). If i reports a or some candidate b with $b \lhd a$, this does not change what the top choice of the median voter is, and hence does not change the outcome. If i reports a candidate c with $t(i) \lhd c$, then the median voter may shift to the right, i.e., further away from i's true top choice; as i's preferences are single-peaked, this does not improve the outcome from her perspective.

(iii) Ordering the candidates by their position, i.e., placing a before b on the axis $\lhd$ if $x(a) < x(b)$ results in an axis witnessing that the input profile is single-peaked. To show that the converse is not true, consider the following four votes:

$$b \succ_1 c \succ_1 a \succ_1 d, \quad c \succ_2 b \succ_2 a \succ_2 d,$$
$$b \succ_3 c \succ_3 d \succ_3 a, \quad c \succ_4 b \succ_4 d \succ_4 a.$$

This profile is single-peaked on $a \lhd b \lhd c \lhd d$. Now, suppose for contradiction that it is 1D-Euclidean, i.e., it admits an embedding x. Consider the positions of the four voters $x(1), x(2), x(3), x(4)$. Assume without loss of generality that $x(b) < x(c)$. Then we have

$$x(1), x(3) < \frac{1}{2}(x(b) + x(c))$$

(as voters 1 and 3 prefer b to c) and

$$x(2), x(4) > \frac{1}{2}(x(b) + x(c))$$

(as voters 2 and 4 prefer b to c). But now consider the point $\frac{1}{2}(x(a) + x(d))$. Voters 1 and 2 have to be on one side of this point and voters 3 and 4 have to be on the other side of this point, because of their preferences over a vs. d. But this is clearly impossible!

(iv) Assume without loss of generality that P is single-peaked with respect to the axis $c_1 \lhd \cdots \lhd c_m$. Clearly, P may contain a vote that ranks c_1 last or a vote that ranks c_m last. But it cannot contain a vote that ranks some c_i with $1 < i < m$ last: if the top candidate in that vote is a c_j with $j < i$, then this voter prefers c_i to c_m, and if the top candidate in that vote is a c_k with $k > i$, then this voter prefers c_i to c_1.

(v) Induction. For $m = 2$, we have $2 = 2^{2-1}$ orders, i.e., $c_1 \succ c_2$ and $c_2 \succ c_1$. Now suppose the claim has been proved for all $m' < m$. A vote that is single-peaked on $c_1 \lhd \cdots \lhd c_m$ may have c_1 in the last position, with candidates in top $m-1$ positions forming a single-peaked vote with respect to $c_2 \lhd \cdots \lhd c_m$ (2^{m-2} options) or it may have c_m in the last position, with candidates in top $m-1$ positions forming a single-peaked vote with respect to $c_1 \lhd \cdots \lhd c_{m-1}$ (2^{m-2} options). For sampling, we can build the vote bottom-up. At the first step, we fill the last position with c_1 or c_m, with probability $\frac{1}{2}$ each. Once k positions have been filled, $1 \leq k \leq m-1$, the not-yet-ranked candidates form a contiguous segment $c_1 \lhd \cdots \lhd c_j$ of the axis. We then fill position $k+1$ from the bottom with c_i or c_j, with equal probability. This sampling is uniform, because the probability of generating a specific ranking is exactly 2^{-m+1}: we make $m-1$ choice, and with probability $\frac{1}{2}$ each choice is consistent with the target ranking.

References

[1] V. Conitzer, Eliciting single-peaked preferences using comparison queries. *J. Artificial Intelligence Res.* **35**, 161–191 (2009)

[2] C. H. Coombs. Psychological scaling without a unit of measurement. *Psychological Review* **57**, 145 (1950)

[3] E. Elkind, M. Lackner and D. Peters. Structured preferences. In *Trends in Computational Social Choice*, edited by U. Endriss, Chapter 10, AI Access, 187–207 (2017)

[4] H. Moulin, *Axioms of Cooperative Decision Making*. Cambridge University Press, Cambridge (1991)

[5] T. Walsh, Generating single peaked votes. arXiv:1503.02766 (2015)

246

Consider a standard prisoners' dilemma game described by the following strategic form, with $\delta > \beta > 0 > \gamma$:

	C	D
C	β, β	γ, δ
D	δ, γ	0, 0

Assume that any given agent either plays C or D and that agents reproduce at a rate determined by their payoff from the strategic form of the game plus a constant f. Suppose that members of an infinite population are assorted into finite groups of size n. Let q denote the proportion of agents playing strategy C ("altruists") in the population as a whole and q_i denote the proportion of agents playing C in group i. We assume that currently $q \in (0,1)$.

The process of assortment is abstract, but we assume that it has finite expectation $E[q_i] = q$ and variance $\mathrm{Var}[q_i] = \sigma^2$. Members within each group are then randomly paired off to play one iteration of the prisoners' dilemma against another member of their group. All agents then return to the overall population.

(a) Find a condition relating q, σ^2, β, γ, δ and n under which the proportion of altruists in the overall population rises after a round of play.

(b) Now interpret this game as one where each player can confer a benefit b upon the other player by individually incurring a cost c, with $b > c > 0$, so that $\beta = b - c$, $\delta = b$ and $\gamma = -c$. Prove that, as long as (i) there is some positive assortment in group formation and (ii) the ratio $\frac{c}{b}$ is low enough, then the proportion of altruists in the overall population will rise after a round of play.

Richard Povey (Hertford College and St Hilda's College, University of Oxford, UK)

Solution by the proposer

(a) We can firstly see that the number of altruists (type C) and non-altruists (type D) in group i after a round of play will be given by

$$n_i' q_i' = \left(\beta\left(\frac{nq_i - 1}{n-1}\right) + \gamma\left(\frac{n(1-q_i)}{n-1}\right) + f\right) n q_i,$$
$$n_i'(1-q_i') = \left(\delta\left(\frac{nq_i}{n-1}\right) + f\right) n(1-q_i).$$

Summing these two equalities yields

$$n_i' = \beta q_i\left(\frac{n(nq_i-1)}{n-1}\right) + \gamma q_i\left(\frac{n^2(1-q_i)}{n-1}\right) + \delta(1-q_i)\left(\frac{n^2 q_i}{n-1}\right) + fn.$$

Given an infinite population and hence an infinite number of groups, the new proportion of altruists in the overall population will be

$$\begin{aligned} q' &= \frac{E[n_i' q_i']}{E[n_i']} \\ &= \left(\beta\left(\frac{nE[q_i^2] - E[q_i]}{n-1}\right) + n\gamma\left(\frac{E[q_i] - E[q_i^2]}{n-1}\right) + fE[q_i]\right) \\ &\quad \cdot \left(\beta q_i\left(\frac{nE[q_i^2] - E[q_i]}{n-1}\right) + n\gamma\left(\frac{E[q_i] - E[q_i^2]}{n-1}\right) + n\delta\left(\frac{E[q_i] - E[q_i^2]}{n-1}\right) + f\right)^{-1}. \end{aligned}$$

Substituting in $E[q_i^2] = \sigma^2 + E[q_i]^2$ and $E[q_i] = q$ gives us

$$\begin{aligned} q' = {} & \left(\beta\left(\frac{n\sigma^2 + nq^2 - q}{n-1}\right) + n\gamma\left(\frac{q(1-q) - \sigma^2}{n-1}\right) + fq\right) \\ & \cdot \left(\beta\left(\frac{n\sigma^2 + nq^2 - q}{n-1}\right) + n\gamma\left(\frac{q(1-q) - \sigma^2}{n-1}\right) + n\delta\left(\frac{q(1-q) - \sigma^2}{n-1}\right) + f\right)^{-1}. \end{aligned}$$

Assuming f is high enough to make the denominator positive, it then follows that

$$\begin{aligned} & q' - q > 0 \\ \iff {} & (1-q)\left(\beta\left(\frac{n\sigma^2 + nq^2 - q}{n-1}\right) + n\gamma\left(\frac{q(1-q) - \sigma^2}{n-1}\right)\right) - nq\delta\left(\frac{q(1-q) - \sigma^2}{n-1}\right) > 0. \end{aligned}$$

After some further rearrangement, we can derive the following:

$$\begin{aligned} & q' - q > 0 \\ \iff {} & \frac{\sigma^2}{q(1-q)} > 1 - \left(\frac{n-1}{n}\right)\left(\frac{\beta}{(1-q)(\beta - \gamma) + q\delta}\right). \quad (1) \end{aligned}$$

Since the right-hand side of (1) must be strictly between 0 and 1, this has the intuitive interpretation that the *inter-group* variance σ^2 must be sufficiently high relative to the *intra-group* variance[1] so that, although altruists do less well relative to non-altruists *within each group*, the concentration of altruists together within *particular* groups is sufficiently strong to confer enough of an evolutionary advantage to offset this and to enable altruists to do better evolutionarily than non-altruists in the overall population.[2]

(b) In the case where $\beta = b - c$, $\delta = b$ and $\gamma = -c$, condition (1) can be rearranged to give

$$\frac{c}{b} < \left(\frac{\sigma^2}{q(1-q)}\right)\left(\frac{n}{n-1}\right) - \frac{1}{n-1}. \quad (2)$$

With random assortment, q_i would be equal to $\frac{X_i}{n}$ where X_i, the number of altruists in group i would have a binomial distribution: $X_i \sim B(n,q)$. Therefore

$$\sigma^2 = \mathrm{Var}\left[\frac{X_i}{n}\right] = \frac{q(1-q)n}{n^2} = \frac{q(1-q)}{n}.$$

With perfect positive correlation between group members, we would get

$$\sigma^2 = \mathrm{Var}\left[\frac{X_i}{n}\right] = \frac{q(1-q)n^2}{n^2} = q(1-q).$$

[1] This is the variance of the Bernoulli variable $B(q)$ which takes a value of 1 of a single individual drawn from the population is an altruist and 0 otherwise, which is $q(1-q)$.

[2] This result was first proved in a general evolutionary context by George R. Price [3, 4].

With some positive assortment generating positive covariance between group members, we may therefore without loss of generality suppose that

$$\sigma^2 = \frac{q(1-q)n + q(1-q)\varepsilon(n^2-n)}{n^2} = \frac{q(1-q)}{n} + \frac{\varepsilon q(1-q)(n-1)}{n},$$

where $\varepsilon \in (0,1)$. Plugging this into (2) and simplifying, we get $\frac{c}{b} < \varepsilon$. So we can see that for any $\varepsilon > 0$ there always exists a value of $\frac{c}{b}$ low enough for altruists to expand as a proportion of the overall population.[3]

References

[1] B. Cooper and C. Wallace, Group selection and the evolution of altruism. *Oxford Economic Papers* **56**, 307–330 (2004)

[2] R. Povey, Punishment and the potency of group selection, *J. Evolutionary Economics* **24** (2014)

[3] G. R. Price, Selection and covariance, *Nature* **227**, 520–521 (1970)

[4] G. R. Price, Extension of covariance selection mathematics, *Ann. Human Genetics* **35**, 485–490 (1972)

[5] E. Sober and D. S. Wilson, *Unto Others*, Harvard University Press, Cambridge, Massachusetts (1999)

247

Consider a village consisting of n farmers who live along a circle of length n. The farmers live at positions $1, 2, \ldots, n$. Each of them is friends with the person to the left and right of them, and each friendship has capacity m where m is a non-negative integer. At the end of the year, each farmer does either well (her wealth is $+1$ dollars) or not well (her wealth is -1 dollars) with equal probability. Farmers' wealth realizations are independent of each other. Hence, for a large circle the share of farmers in each state is on average 1.

The farmers share risk by transferring money to their direct neighbors. The goal of risk-sharing is to create as many farmers with OK wealth (0 dollars) as possible. Transfers have to be in integer dollars and cannot exceed the capacity of each link (which is m).

A few examples with a village of size $n = 4$ serve to illustrate risk-sharing.

- Consider the case where farmers 1 to 4 have wealth
$$(+1, -1, +1, -1).$$
In that case, we can share risk completely with farmer 1 sending a dollar to agent 2 and farmer 3 sending a dollar to farmer 4. This works for any $m \geq 1$.
- Consider the case where farmers 1 to 4 have wealth
$$(+1, +1, -1, -1).$$
In that case, we can share risk completely with farmer 1 sending a dollar to farmer 2, farmer 2 sending two dollars to farmer 3 and farmer 3 sending one dollar to farmer 4. In this case, we need $m \geq 2$. If $m = 1$, we can only share risk among half the people in the village.

Show that for any wealth realization an optimal risk-sharing arrangement can be found as the solution to a maximum flow problem.

Tanya Rosenblat (School of Information and Department of Economics, University of Michigan, USA)

Solution by the proposer

We augment the village graph by adding two auxiliary nodes. The *source node s* is connected to all the farmers with positive wealth $(+1)$ and each of these links has capacity 1. The *sink node t* is connected to all the farmers with negative wealth (-1) and each of these links also capacity 1. We now look for the *maximum flow* from s to t: this is equal to the number of luck/unlucky farmer pairs who can be matched under the best risk-sharing arrangement.

248

This exercise is a continuation of Problem 247 where we studied risk-sharing among farmers who live on a circle village and are friends with their direct neighbors to the left and right with friendships of a certain capacity. Assume that for any realization of wealth levels the best possible risk-sharing arrangement is implemented and denote the expected share of unmatched farmers with $U(n,m)$. Show that $U(n,m) \to \frac{1}{2m+1}$ as $n \to \infty$.

Tanya Rosenblat (School of Information and Department of Economics, University of Michigan, USA)

Solution by the proposer

The solution proceeds in two stages. In Problem 247, we already established that the problem can be understood as a maximum flow problem. We first formulate a particular algorithm that implements this flow. We then use this algorithm to express $U(n,m)$ in closed form.

Risk-sharing as a Maximum Flow Problem. We next describe a matching algorithm which is to run for m rounds. The claim is that this algorithm implements the maximum flow in the augmented graph for any m. For the purpose of this algorithm, we

[3] The literature on this model has further established the results that *multiple* periods of isolation in finite groups acts to *amplify* inter-group variance, so that even with random assortment into groups altruism can evolve [1]. It has also been found that use of punishment strategies in dynamic interactions can act to weaken this group selection mechanism [2]. For an accessible book-length treatment of the topic of group selection in the biological and social sciences, see [5].

call a black agent an agent with positive shock and a white agent an agent with a negative shock.

Step I: Index all agents from 1 to n clockwise (1 is the neighbor of n on a circle). Set the counter i to 1.

Step II: If agents i and $i+1$ are of different colors, label them "matched" and move the counter clockwise to $i+2$. If agent $i+1$ has the same color as agent i, then declare agent i "unmatched" and move the counter clockwise to agent $i+1$. Repeat this step until the counter has reached the first agent again.

Step III: Define a new circle by ordering all the unmatched agents on a circle without disturbing the order of the agents. Essentially, this implies that all the matched agents are simply removed from the circle and any gaps are plugged by connected the closest unmatched agents with each other. Repeat steps I and II for this new circle. Repeat this algorithm m times.

Lemma 1. *The above matching algorithm implements the maximum flow.*

Proof. The Ford–Fulkerson algorithm computes the maximum flow by looking for open paths which can carry positive flow and then constructing a graph with augmented capacities in which the next path is found, etc. Once no more open path exists the max flow has been implemented. The above algorithm implements Ford–Fulkerson using a particular order of selecting open paths. Therefore, it implements the max flow.

Closed form solutions. We next prove the following lemma.

Lemma 2. *Assume we have a circle of size n where the probability that an agent has a neighbor of the same color is α. Then the share of unmatched agents after one round of the above algorithm converges to $\frac{\alpha}{2-\alpha}$ as $n \to \infty$.*

Proof. In each instance of step I of the algorithm, it produces an unmatched agent with probability α and a pair of matched agents with probability $1-\alpha$. The sum of unmatched and matched agents has to be n. Therefore the share of unmatched agents converges to

$$\frac{\alpha}{\alpha+2(1-\alpha)} = \frac{\alpha}{2-\alpha}.$$

The final step in the proof of the result is to derive the probability α_m that an agent is followed by a same-color agent in round m. We know that in round 1 shocks are i.i.d.; therefore $\alpha_1 = \frac{1}{2}$. We start by proving a recursive formula for calculating α_m.

Lemma 3. *If the sequential probability is α_m in round m, then the sequential probability in round $m+1$ satisfies*

$$\alpha_{m+1} = \frac{2-\alpha_m}{3-2\alpha_m}. \tag{1}$$

Proof. Consider an agent i on whom the counter rested at some point in the algorithm and who stays unmatched in the current round. This must be because he has a neighbor $i+1$ of the same color (without loss of generality assume both are black). With probability α_m agent $i+2$ is also black and therefore agent $i+1$ will survive into round $m+1$ as well and be of the same color as agent i (black). With probability $1-\alpha_m$ agent $i+2$ is white. In this case agents $i+1$ and $i+2$ can be matched. Matching can continue for the subsequent pairs of agents $(i+3, i+4)$, $(i+5, i+6)$, etc.; it will only stop if for any of these pairs agents have the same color. This will happen with probability α_m. To figure out it if this process will stop at a "white pair" or a "black pair" (let's call it the "blocking pair") it is crucial to know whether agent $i+2$, $i+4$, $i+6$, etc. (i.e., the agent just prior to the blocking pair) is white or black.

We know that agent $i+2$ is white. What is the probability that agent $i+4$ is the same color (provided $(i+3, i+4)$ is not a blocking pair)? This can only happen if the pair $(i+3, i+4)$ is a black agent followed by a white agent. If it is a white agent followed by a black agent then $i+4$ has a different color from $i+2$. So the probability of a color change is

$$\frac{\alpha_m(1-\alpha_m)}{\alpha_m(1-\alpha_m)+(1-\alpha_m)^2} = \alpha_m.$$

Assume that the probability that the agent prior to the blocking pair is of the same color as $i+2$ is q. With probability α_m the pair $(i+3, i+4)$ is a blocking pair and with probability $1-\alpha_m$ the pair is not blocking. In that case agent $i+4$ has a different color from agent $i+2$ with probability α_m. Because of the recursive nature of the problem, the probability that the agent prior to the blocking pair has the same color as $i+4$ is q. Therefore we know that

$$q = \alpha_m + (1-\alpha_m)[(1-\alpha_m)q + (1-q)\alpha_m].$$

This allows us to calculate q as

$$q = \frac{\alpha_m(2-\alpha_m)}{1-(1-\alpha_m)(1-2\alpha_m)} = \frac{2-\alpha_m}{3-2\alpha_m}.$$

So we know that the agent prior to the blocking pair is white with probability q. The blocking pair is therefore a black blocking pair with probability $q(1-\alpha_m)+(1-q)\alpha_m$. Therefore the total probability that the next unmatched agent after agent i is of the same color (black in this case) is

$$\begin{aligned}\alpha_{m+1} &= \alpha_m + (1-\alpha_m)[q(1-\alpha_m)+(1-q)\alpha_m] \\ &= q = \frac{2-\alpha_m}{3-2\alpha_m}.\end{aligned}$$

We can check that $\alpha_m = 1 - \frac{1}{2m}$ satisfies both the initial condition and the recursive equation 1. This implies that

$$\frac{\alpha_m}{2-\alpha_m} = \frac{2m-1}{2m+1}.$$

Finally, note that the share of unmatched agents (for $n \to \infty$) can be calculated by taking the product of the share of unmatched agents in each round:

$$\lim_{n \to \infty} U(n,m) = \prod_{i=1}^{m} \frac{a_i}{2 - a_i} = \frac{1}{2m+1}.$$

249

In a *combinatorial auction* there are m items for sale to n buyers. Each buyer i has some valuation function $v_i(\cdot)$ which takes as input a set S of items and outputs that bidder's value for that set. These functions will always be monotone ($v_i(S \cup T) \geq v_i(S)$ for all S, T), and satisfy $v_i(\emptyset) = 0$.

Definition 1 (Walrasian equilibrium)**.** A price vector $\vec{p} \in \mathbb{R}^m_{\geq 0}$ and a list $B_1, \ldots, B_n$ of subsets of $[m]$ form a *Walrasian equilbrium* for $v_1, \ldots, v_n$ if the following two properties hold:

- Each $B_i \in \arg\max_S \{v_i(S) - \sum_{j \in S} p_j\}$.
- The sets B_i are disjoint, and $\bigcup_i B_i = [m]$.

Prove that a Walrasian equilibrium exists for $v_1, \ldots, v_n$ if and only if there exists an integral[4] optimum to the following linear program:

$$\begin{aligned} &\text{maximize } \sum_i \sum_S v_i(S) \cdot x_{i,S} \\ &\text{such that, for all } i, \quad \sum_S x_{i,S} = 1, \\ &\qquad \text{for all } j, \quad \sum_{S \ni j} \sum_i x_{i,S} \leq 1, \\ &\qquad \text{for all } i, S, \quad x_{i,S} \geq 0. \end{aligned}$$

Hint. Take the dual, and start from there.

Matt Weinberg (Computer Science, Princeton University, USA)

Solution by the proposer

First, we take the dual of the given LP. We use the dual variable p_j for the constraints involving items, and the dual variable u_i for the constraints involving bidders. Then the dual problem is

$$\begin{aligned} &\text{minimize } \sum_i u_i + \sum_j p_j \\ &\text{such that, for all } i, S, \quad u_i + \sum_{j \in S} p_j \geq v_i(S), \\ &\qquad \text{for all } j, \quad p_j \geq 0. \end{aligned}$$

Walrasian equilibrium implies integral optimum. Now, assume that a Walrasian equilibrium exists, and let it be $p_1, \ldots, p_m$ and $B_1, \ldots, B_n$. Then consider the integral solution to the LP that sets $x_{i,B_i} = 1$ for all i, and all other variables to 0. This solution is clearly feasible for the LP, and has objective value equal to $\sum_i v_i(B_i)$.

Consider also the dual solution $u_i := v_i(B_i) - \sum_{j \in B_i} p_j$, with p_j as given in the Walrasian equilibrium. We claim this is a feasible solution to the dual. To see this, observe first that each $p_j \geq 0$. Also, because $B_i \in \arg\max_S \{v_i(S) - \sum_{j \in S} p_j\}$ by definition of Walrasian equilibrium, we have that

$$u_i := v_i(B_i) - \sum_{j \in B_i} p_j \geq v_i(S) - \sum_{j \in S} p_j \quad \text{for all } S.$$

Therefore, all dual constraints are satisfied, and this is a feasible dual. Moreover, observe that the value of the dual objective is

$$\sum_i u_i + \sum_j p_j = \sum_i \Big(v_i(B_i) - \sum_{j \in B_i} p_j\Big) + \sum_j p_j = \sum_i v_i(B_i).$$

The last equality follows because each item is in exactly one bundle B_i. So we have proved that if $(\vec{p}, \vec{B})$ is a Walrasian equilibrium, then there is an integral feasible point for the LP with objective value $\sum_i v_i(B_i)$, and also a feasible dual solution with value $\sum_i v_i(B_i)$. By LP duality, both feasible solutions are in fact optimal. Therefore, there is an integral optimum for the LP.

Integral optimum implies Walrasian equilibrium. Now, assume that the LP has an integral optimum. Observe that for this integral solution, there must exist disjoint sets $B_1, \ldots, B_n$ such that for each i, $x_{i,B_i} = 1$ and all other variables are 0. The LP value for this solution is $\sum_i v_i(B_i)$. Moreover, observe that if any item $j \notin \cup_i B_i$, we can add j to an arbitrary B_i without decreasing $\sum_i v_i(B_i)$ (because each $v_i(\cdot)$ is monotone). Therefore, if there is an integral optimum to the LP, there exist disjoint $B_1, \ldots, B_n$ such that $\cup_i B_i = [m]$ and $\sum_i v_i(B_i)$ is the optimal solution to the LP.

By Strong LP Duality, there also exists a feasible dual solution $p_1, \ldots, p_m, u_1, \ldots, u_n$ such that $\sum_i u_i + \sum_j p_j = \sum_i v_i(B_i)$. We will claim that $(\vec{p}, \vec{B})$ form a Walrasian equilibrium.

For a proof by contradiction, assume that this is not the case. Then there must be some bidder i such that $B_i \notin \arg\max_S \{v_i(S) - \sum_{j \in S} p_j\}$. In particular, this means that there exists some B_i' such that $v_i(B_i') - \sum_{j \in B_i'} > v_i(B_i) - \sum_{j \in B_i}$. Because u_i is a feasible solution for the dual, we then conclude that

$$u_i \geq v_i(B_i') - \sum_{j \in B_i'} > v_i(B_i) - \sum_{j \in B_i} p_j.$$

We claim that this contradicts the fact that $\sum_i u_i + \sum_j p_j = \sum_i v_i(B_i)$, since

$$\sum_i u_i > \Big(\sum_i v_i(B_i) - \sum_{j \in B_i} p_j\Big) > \sum_i v_i(B_i) - \sum_j p_j.$$

The first inequality holds because $u_i \geq v_i(B_i) - \sum_{j \in B_i} p_j$ for all i, and the inequality is strict for at least one i. Therefore, $(\vec{p}, \vec{B})$ must be a Walrasian equilibrium.

[4] That is, a point such that each $x_{i,S} \in \{0,1\}$.

Wrapup. This concludes the proof. We have shown that there is an integral optimum to the LP if and only if a Walrasian equilibrium exists. The solution to this problem is given by Nisan et al. in [3, Corollary 11.16]; they cite [1, 2].

References

[1] S. Bikhchandani and J. W. Mamer, Competitive equilibrium in an exchange economy with indivisibilities. *J. Econom. Theory* **74**, 385–413 (1997)

[2] S. Bikhchandani and J. M. Ostroy, The package assignment model. *J. Econom. Theory* **107**, 377–406 (2002)

[3] N. Nisan, T. Roughgarden, É. Tardos and V. V. Vazirani, *Algorithmic game theory*. Cambridge University Press (2007)

250

Consider a game played on a network and a finite set of players $\mathcal{N} = \{1, 2, \dots, n\}$.[5] Each node in the network represents a player and edges capture their relationships. We use $\mathbf{G} = (g_{ij})_{1 \le i,j \le n}$ to represent the adjacency matrix of a undirected graph/network, i.e., $g_{ij} = g_{ji} \in \{0, 1\}$. We assume $g_{ii} = 0$. Thus, $\mathbf{G}$ is a zero-diagonal, squared and symmetric matrix. Each player, indexed by i, chooses an action $x_i \in \mathbb{R}$ and obtains the following payoff:

$$\pi_i(x_1, x_2, \dots, x_n) = x_i - \frac{1}{2}x_i^2 + \delta \sum_{j \in \mathcal{N}} g_{ij} x_i x_j.$$

The parameter $\delta > 0$ captures the strength of the direct links between different players. For simplicity, we assume $0 < \delta < \frac{1}{n-1}$.

A Nash equilibrium is a profile $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ such that, for any $i = 1, \dots, n$,

$$\pi_i(x_1^*, \dots, x_n^*) \ge \pi_i(x_1^*, \dots, x_{i-1}^*, x_i, x_{i+1}^*, \dots, x_n^*) \quad \text{for any } x_i \in \mathbb{R}.$$

In other words, at a Nash equilibrium, there is no profitable deviation for any player i choosing x_i^*.

Let $\mathbf{w} = (w_1, w_2, \dots, w_n)'$, $w_i > 0$ for all i (the transpose of a vector $\mathbf{w}$ is denoted by $\mathbf{w}'$), and $\mathbf{I}_n$ the $n \times n$ identity matrix. Define the *weighted* Katz–Bonacich centrality vector as

$$\mathbf{b}(\mathbf{G}, \mathbf{w}) = [\mathbf{I}_n - \delta \mathbf{G}]^{-1} \mathbf{w}.$$

Here $\mathbf{M} := [\mathbf{I} - \delta \mathbf{G}]^{-1}$ denote the inverse Leontief matrix associated with network $\mathbf{G}$, while m_{ij} denote its ij entry, which is equal to the discounted number of walks from i to j with decay factor δ. Let $\mathbf{1}_n = (1, 1, \dots, 1)'$ be a vector of 1s. Then the *unweighted* Katz–Bonacich centrality vector can be defined as

$$\mathbf{b}(\mathbf{G}, \mathbf{1}) = [\mathbf{I} - \delta \mathbf{G}]^{-1} \mathbf{1}_n.$$

(1) Show that this network game has a unique Nash equilibrium $\mathbf{x}^*(\mathbf{G})$. Can you link this equilibrium to the Katz–Bonacich centrality vector defined above?

(2) Let $x^*(\mathbf{G}) = \sum_{i=1}^n x_i^*(\mathbf{G})$ denote the sum of actions (total activity) at the unique Nash equilibrium in part 1. Now suppose that you can remove a single node, say i, from the network. Which node do you want to remove such that the sum of effort at the new Nash equilibrium is reduced the most? (Note that, after the deletion of node i, we remove all the links of node i, and the remaining network, denoted by $\mathbf{G}_{-i}$, can be obtained by deleting the i-th row and i-th column of $\mathbf{G}$.)
Mathematically, you need to solve the *key player problem*[6]

$$\max_{i \in \mathcal{N}} \big(x^*(\mathbf{G}) - x^*(\mathbf{G}_{-i})\big).$$

In other words, you want to find a player who, once removed, leads to the highest reduction in total action in the remaining network.
Hint. You may come up with an index c_i for each i such that the key player is the one with the highest c_i. This c_i should be expressed using the Katz–Bonacich centrality vector defined above.

(3) Now instead of deleting a single node, we can delete any pair of nodes from the network. Can you identify the key pair, that is, the pair of nodes that, once removed, reduces total activity the most?[7]

Yves Zenou (Monash University, Australia) and Junjie Zhou (National University of Singapore)

Solution by the proposer

(1) Suppose that $\mathbf{x}^*(\mathbf{G})$ is a Nash equilibrium. Then we obtain the following optimality equation for player i:

$$x_i^*(\mathbf{G}) = 1 + \delta \sum_{j \in \mathcal{N}} g_{ij} x_j^*(\mathbf{G}).$$

In matrix form,

$$\mathbf{x}^*(\mathbf{G}) = \mathbf{1}_n + \delta \mathbf{G} \mathbf{x}^*(\mathbf{G})$$

or

$$\mathbf{x}^*(\mathbf{G}) = [\mathbf{I}_n - \delta \mathbf{G}]^{-1} \mathbf{1}_n := \mathbf{b}(\mathbf{G}, \mathbf{1}).$$

In other words, the Nash equilibrium effort is exactly equal to the unweighted Katz–Bonacich vector.

Note that, under the assumption that $0 < \delta < 1/(n-1)$, the matrix $[\mathbf{I}_n - \delta \mathbf{G}]$ is invertible, and the inverse matrix has the following infinite sum representation:

$$[\mathbf{I}_n - \delta \mathbf{G}]^{-1} = \mathbf{I}_n + \delta \mathbf{G} + \delta^2 \mathbf{G}^2 + \delta^3 \mathbf{G}^3 + \cdots.$$

For uniqueness, it is obvious.

[5] For an overview of the literature on network games, see [2].

[6] The key player problem has been introduced in [1].

[7] For the analysis of group players, see [1, 3].

(2) Before solving it, we first enrich the baseline model by taking into account heterogeneous individual weights $w_i > 0$,

$$\pi_i(x_1, x_2, \ldots, x_n) = w_i x_i - \frac{1}{2}x_i^2 + \delta \sum_{j \in \mathcal{N}} g_{ij} x_i x_j. \tag{1}$$

The unique equilibrium of this extended model corresponds to the weighted Katz–Bonacich centrality

$$\mathbf{x}^*(\mathbf{G}, \mathbf{w}) = [\mathbf{I}_n - \delta \mathbf{G}]^{-1}\mathbf{w} \coloneqq \mathbf{b}(\mathbf{G}, \mathbf{w}),$$

or, equivalently, for each $i = 1, \ldots, n$:

$$x_i^*(\mathbf{G}, \mathbf{w}) = \sum_{j=1}^{n} m_{ij} w_j = \sum_{j=1}^{n} \sum_{k=0}^{\infty} \delta^k g_{ij}^{[k]} w_j = b_i(\mathbf{G}, \mathbf{w}),$$

where $g_{ij}^{[k]} \geq 0$ gives the number of walks of length $k \geq 1$ from i to j in the network and $b_i(\mathbf{G}, \mathbf{w})$ is the weighted Katz–Bonacich centrality of player i.

The aggregate equilibrium action is then equal to

$$x^*(\mathbf{G}, \mathbf{w}) = \mathbf{1}_n'[\mathbf{I} - \delta \mathbf{G}]^{-1}\mathbf{w} = \mathbf{b}'(\mathbf{G}, \mathbf{1})\mathbf{w}.$$

Intuitively, when w_i increases by 1 unit, $x_j^*(\mathbf{G}, \mathbf{w})$, each player j's equilibrium effort increases by $m_{ji} = m_{ij}$, and the total equilibrium action increases by $b_i(\mathbf{G}, \mathbf{1}) = \sum_j m_{ij} = \sum_j m_{ji}$ (note that $\mathbf{M}$ is a symmetric matrix). Mathematically,

$$\frac{\partial x_j^*(\mathbf{G}, \mathbf{w})}{\partial w_i} = m_{ji} = m_{ij} \quad \text{for all } i, j, \tag{2}$$

$$\frac{\partial x^*(\mathbf{G}, \mathbf{w})}{\partial w_i} = b_i(\mathbf{G}, \mathbf{1}). \tag{3}$$

To solve the key player problem, it suffices to prove that, for $i \in \mathcal{N}$,

$$c_i(\mathbf{G}) \coloneqq \big(x^*(\mathbf{G}) - x^*(\mathbf{G}_{-i})\big) = \frac{[b_i(\mathbf{G}, \mathbf{1})]^2}{m_{ii}}.$$

And the key player is the player i that maximizes $c_i(\mathbf{G})$.

To prove this, we take the following approach. Instead of removing node i (and all its links with others), we reduce the weight of player from $w_i = 1$ to $\hat{w}_i = 1 - \frac{b_i(\mathbf{G}, \mathbf{1})}{m_{ii}}$, while keeping the weights of other players at 1 as in the baseline model, i.e., $w_j = 1$ for all $j \neq i$. (It will be clear why we pick this particular $\hat{w}_i$.)

We claim that, after this reduction in weight, the resulting equilibrium is the same as the one when i is removed from the network.

To see this, we first ask: what is the new equilibrium after this change in w_i? We claim that player i would choose exactly zero action. This is because, by (2), the change in equilibrium action by player i,

$$\Delta x_i^*(\mathbf{G}) \coloneqq x_i^*(\mathbf{G}, w_i = \hat{w}_i) - x_i^*(\mathbf{G}, w_i = 1),$$

is given by

$$\Delta x_i^*(\mathbf{G}) = m_{ii} \times (\hat{w}_i - w_i) = -m_{ii}\frac{b_i(\mathbf{G}, \mathbf{1})}{m_{ii}} = -b_i(\mathbf{G}, \mathbf{1})$$

from the construction of $\hat{w}_i$. Since initially player i chooses $b_i(\mathbf{G}, \mathbf{1})$, we have $x_i^*(\mathbf{G}, w_i = \hat{w}_i) = 0$ and the claim follows. What happens to other nodes? By (2), player j's equilibrium action would change by $m_{ij} \times (\hat{w}_i - w_i)$, and the aggregate equilibrium action, by (3), would change by

$$\begin{aligned} x^*(\mathbf{G}_{-i}) - x^*(\mathbf{G}) &= b_i(\mathbf{G}, \mathbf{1}) \times (\hat{w}_i - w_i) \\ &= -\frac{b_i(\mathbf{G}, \mathbf{1})^2}{m_{ii}} = -c_i(\mathbf{G}). \end{aligned}$$

This completes the proof of our claim. Thus, the key player in a network is the player i who has the highest $c_i(\mathbf{G})$.

(3) For any group $S \subset \mathcal{N}$, we can define the inter-centrality measure

$$d_S(\mathbf{G}) = \mathbf{b}_S'(\mathbf{G}, \mathbf{1})\mathbf{M}_{SS}^{-1}\mathbf{b}_S(\mathbf{G}, \mathbf{1}),$$

where $\mathbf{M}_{SS} = (m_{kl})$, $k, l \in S$, is the submatrix of $\mathbf{M}$, that is, the $|S| \times |S|$ $\mathbf{M}$ matrix of the subnetwork formed by players in S. Similarly, $\mathbf{b}_S(\mathbf{G}, \mathbf{1})$ is a subvector of the unweighted Katz–Bonacich centrality vector $\mathbf{b}(\mathbf{G}, \mathbf{1})$ for indices in the set S. It can be shown that

$$d_S(\mathbf{G}) = x^*(\mathbf{G}) - x^*(\mathbf{G}_{-S}),$$

where $\mathbf{G}_{-S}$ is the network obtained after removing all nodes and their links in S. The proof is similar to the one in part (2), since removing S from the network has the same effect on the equilibrium as changing the weight vector from $\mathbf{w}_S = \mathbf{1}_S$ to $\hat{\mathbf{w}}_S = \mathbf{1}_S - \mathbf{M}_{SS}^{-1}\mathbf{b}_S(\mathbf{G}, \mathbf{1})$, while the weights of the nodes in the complement of S remain fixed at 1 as in the baseline model.

When S is a pair (i, j) with $i \neq j$, we can explicitly express the index as follows:

$$d_{\{i,j\}}(\mathbf{G}) = [b_i(\mathbf{G}, \mathbf{1})\ b_j(\mathbf{G}, \mathbf{1})] \begin{bmatrix} m_{ii} & m_{ij} \\ m_{ji} & m_{jj} \end{bmatrix}^{-1} \begin{bmatrix} b_i(\mathbf{G}, \mathbf{1}) \\ b_j(\mathbf{G}, \mathbf{1}) \end{bmatrix}$$

The *key pair* (i, j) is the pair (i, j) with the largest $d_{\{i,j\}}(\mathbf{G})$.

References

[1] C. Ballester, A. Calvó-Armengol and Y. Zenou, Who's who in networks. Wanted: the key player. *Econometrica* **74**, 1403–1417 (2006)

[2] M. O. Jackson and Y. Zenou, Games on networks. In *Handbook of Game Theory. Vol. 4*, edited by H. P. Young and S. Zamir, Handbooks in Economics, 91–157, Elsevier/North-Holland, Amsterdam (2015)

[3] Y. Sun, W. Zhao and J. Zhou, Structural interventions in networks. arXiv:2101.12420 (2021)

We hope to receive your solutions to the proposed problems and your ideas on the open problems. Send your solutions to Michael Th. Rassias by email to mthrassias@yahoo.com.

We also solicit your new problems with their solutions for the next "Solved and unsolved problems" column, which will be devoted to Differential Equations.

New editors appointed

Jezabel Curbelo is Ramon y Cajal researcher in the Department of Mathematics at Universitat Politècnica de Catalunya, Spain, since July 2020 and member of the Institute of Mathematics (IMTech) of UPC-BarcelonaTech. She obtained her PhD in mathematics (2014) at Universidad Autónoma de Madrid. Her thesis was awarded in the field of nonlinear geophysics with the 2015 Donald L. Turcotte Award (American Geophysical Union). After that, she was a postdoc LabEX LIO at Laboratoire de Géologie de Lyon (CNRS/ENS Lyon/Lyon1) (2014–2016), Juan de la Cierva Formación Postdoc at Universidad Politécnica de Madrid (Apr – Sep 2016), Assistant Professor at Universidad Autónoma de Madrid (Sep 2016 – Jan 2020) and Visiting Assistant Researcher at University of California Los Angeles (Feb – Jun 2020).

She has been awarded with the "2020/2021 L'Oréal-*UNESCO For Women in Science*" award, edition "Spain – National Young Talents Programme 2020" and the "2020 Antonio Valle SEMA Prize for Young Researchers".

Her research focuses mainly on applied mathematics in geophysical fluids and related problems. More specifically in the simulation and modeling of nonlinear processes underlying fluid motion and in the description of transport, mixing and stirring in the ocean and atmosphere from the standpoint of dynamical systems theory. She also investigates fundamental behaviours of fluid processes in planetary mantles, where she analyses convective motions. Her webpage is web.mat.upc.edu/jezabel.curbelo/

Photo by Friederike von Heyden/Bergische Universität Wuppertal

Ralf Krömer studied mathematics and computer science in Saarbrücken (Germany). His PhD thesis concerned the history and philosophy of category theory and was written in a French-German cotutelle between Saarbrücken and Nancy. Postdoc positions included work at Aix-en-Provence, Hannover (Leibniz edition) and Nancy (Poincaré edition). After a time as a teacher of mathematics and music in secondary school in Germany, he was a lecturer in mathematics at Siegen University; since 2013, he is professor for mathematics and its didactics at Wuppertal University.

Ralf Krömer's research is mainly in the history of mathematics from the late 17th to the 20th century. Research interests include, in particular, historical themes around Leibniz, Poincaré, the Bourbaki group, structural mathematics, and category theory, but also philosophical aspects of these topics as well as the use of the history of mathematics in its teaching. He is currently working on a large collective publication on duality in 19th and 20th century mathematical thinking. His webpage is www2.math.uni-wuppertal.de/~kroemer/

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.