Note

The AM-GM Inequality is Equivalent to the Bernoulli Inequality

LECH MALIGRANDA

The AM-GM (= arithmetic-mean/geometric-mean) inequality is sometimes called the *Cauchy inequality* (1821):

$$A_n = \frac{x_1 + \cdots + x_n}{n} \geq (x_1 \cdot \ldots \cdot x_n)^{1/n} = G_n \quad (\text{all}\, x_k > 0). \tag{AG}$$

In the books [1] and [2] there are given 52 and 74 proofs of this inequality, respectively, and mathematicians still find new proofs (!).

Much earlier, Isaac Barrow (1670) and Jacob Bernoulli (1689) proved the inequality, which now has the name *Bernoulli inequality*:

$$x^n \geq 1 + n(x-1) \quad \text{for any}\, x > 0 \text{ and } n \in \mathbb{N}. \tag{B}$$

In the paper [5] we can read about the history, generalizations, and different proofs of the Bernoulli inequality.

In this note I show that these inequalities are in fact equivalent, that is, each follows one from the other. This was observed in [4, Thm 3] (see also [2], p. 213, and [5, Thm 10]), but I present here even simpler arguments.

THEOREM *Inequalities* (AG) *and* (B) *are equivalent.*

PROOF. (B) $\Longrightarrow$ (AG). Since $\frac{A_n}{A_{n-1}} > 0$ it follows from the Bernoulli inequality (B) that

$$\left(\frac{A_n}{A_{n-1}}\right)^n \geq 1 + n\left(\frac{A_n}{A_{n-1}} - 1\right) = \frac{A_{n-1} + nA_n - nA_{n-1}}{A_{n-1}} = \frac{nA_n - (n-1)A_{n-1}}{A_{n-1}} = \frac{x_n}{A_{n-1}}$$

or

$$A_n^n \geq x_n \cdot A_{n-1}^{n-1}. \tag{C}$$

Now, using inequality (C) successively, we obtain

$$\begin{aligned} A_n^n &\geq x_n \cdot A_{n-1}^{n-1} \geq x_n \cdot x_{n-1} \cdot A_{n-2}^{n-2} \\ &\geq \ldots \geq x_n \cdot x_{n-1} \cdot \ldots x_2 \cdot A_1^1 \\ &= x_n \cdot x_{n-1} \cdot \ldots x_2 \cdot x_1 = G_n^n, \end{aligned}$$

and thus $A_n \geq G_n$.

(AG) $\Longrightarrow$ (B). For $n = 1$ we have even equality in (B). If $n \geq 2$ and $0 < x \leq 1 - \frac{1}{n}$, then $x^n > 0 \geq 1 + n(x-1)$, that is, (B) holds. Therefore, we can assume that $n \geq 2$ and $x > 1 - \frac{1}{n}$. Then $1 + n(x-1) > 0$, and now applying assumption (AG) to the n positive numbers

$$1 + n(x-1), \overbrace{1, 1, \ldots, 1}^{(n-1)\text{ times}}$$

we obtain

$$\begin{aligned} x^n &= \left\{\frac{[1 + n(x-1)] + 1 + \ldots + 1}{n}\right\}^n \\ &\geq [1 + n(x-1)] \cdot 1 \cdot \ldots \cdot 1 = 1 + n(x-1), \end{aligned}$$

and the inequality (B) is proved.

From the theorem, it follows that one way to get a simple proof of inequality (AG) is to give a simple proof of inequality (B). Here is one:

$$\begin{aligned} x^n - 1 - n(x-1) &= (x-1)(x^{n-1} + x^{n-2} + \ldots + 1) - n(x-1) \\ &= (x-1)(x^{n-1} + x^{n-2} + \ldots + 1 - n) \geq 0. \end{aligned}$$

The last inequality holds because the expression $x^{n-1} + x^{n-2} + \ldots + 1$ is bigger than n for $x \geq 1$ and smaller than n for $0 < x \leq 1$. Thus (B) is proved, and consequently also (AG). Other proofs of (AG) via inequality (B) were given in [6] and [3].

REFERENCES

[1] P. S. Bullen, D. S. Mitrinović, and P. M. Vasić, *Means and their Inequalities*, Reidel, Dordrecht, 1988.

[2] P. S. Bullen, *Handbook of Means and their Inequalities*, Kluwer, Dordrecht, 2003.

[3] M. D. Hirschhorn, The AM-GM Inequality, *Math. Intelligencer* 29, 7 (2007).

[4] L. Maligranda, Why Hölder's Inequality Should be Called Rogers' Inequality, *Math. Inequal. Appl*. 1, 69–83 (1998).

DOI 10.1007/s00283-011-9266-8

[5] L. Maligranda, Bernoulli Inequality – Over 300 Years of History, in: *Around Bernoullis*, 19th All Polish Conference on History of Mathematics (June 6-10, 2005, Zamość), Lublin Technical University Publisher, Lublin, 2006, 31–62 (in Polish).

[6] D. Rüthing, Proofs of the Arithmetic Mean – Geometric Mean Inequality, *Int. J. Math. Educ. Sci. Technology* 13, 49–54 (1982).

Department of Engineering Sciences and Mathematics
Luleå University of Technology
971 87 Luleå
Sweden
e-mail: lech.maligranda@ltu.se

Note

Remarks on Kreĭn's Inequality

MINGHUA LIN

There are two different notions of angle between vectors in $\mathbb{C}^n$ that are frequently used in the literature. One, Φ_{xy}, is defined by means of the standard formula

$$\cos \Phi_{xy} = \frac{\mathrm{Re}\langle x, y\rangle}{\|x\|\|y\|}, \quad x, y \in \mathbb{C}^n \setminus \{0\};$$

the other one, Ψ_{xy}, is defined by

$$\cos \Psi_{xy} = \frac{|\langle x, y\rangle|}{\|x\|\|y\|}, \quad x, y \in \mathbb{C}^n \setminus \{0\}.$$

Both notions have been found useful. Either one gives a Riemannian distance on the unit ball in $\mathbb{C}^n$ locally. The function Ψ_{xy} is a natural metric on complex projective space (but Φ_{xy} is less natural, for $\Phi_{xy} = 0$ in case $y = ix$). Thus they must obey convenient triangle inequalities.

M. G. Kreĭn [2] in 1969 discovered such an inequality:

Kreĭn's Inequality.

$$\Phi_{xz} \leq \Phi_{xy} + \Phi_{yz} \quad \text{for any } x, y, z \in \mathbb{C}^n \setminus \{0\}. \tag{1}$$

For completeness, I include a known proof [1, p. 56].

PROOF. Without loss of generality, assume that x, y, z are unit vectors. Let

$$\langle x, y\rangle = a_1 + ib_1, \quad \langle y, z\rangle = a_2 + ib_2, \quad \langle x, z\rangle = a_3 + ib_3,$$

where a_j, b_j are real numbers and $|a_j|^2 + |b_j|^2 \leq 1$ for $j = 1, 2, 3$. We have $\cos \Phi_{xy} = a_1$, $\cos \Phi_{yz} = a_2$, and $\cos \Phi_{xz} = a_3$. Since $\cos \alpha$ is a decreasing function of $\alpha \in [0, \pi]$, we need only to prove

$$\cos \Phi_{xz} \geq \cos(\Phi_{xy} + \Phi_{yz}),$$

or equivalently,

$$a_3 \geq a_1 a_2 - \sqrt{1 - a_1^2}\sqrt{1 - a_2^2}.$$

Thus, we need

$$\sqrt{1 - a_1^2}\sqrt{1 - a_2^2} \geq a_1 a_2 - a_3. \tag{2}$$

We are done if the right-hand side of (2) is negative. Otherwise, we need to prove $(1 - a_1^2)(1 - a_2^2) \geq (a_1 a_2 - a_3)^2$ or $1 - a_1^2 - a_2^2 - a_3^2 + 2a_1 a_2 a_3 \geq 0$. Because the matrix $G = \begin{bmatrix} \langle x, x\rangle & \langle x, y\rangle & \langle x, z\rangle \\ \langle y, x\rangle & \langle y, y\rangle & \langle y, z\rangle \\ \langle z, x\rangle & \langle z, y\rangle & \langle z, z\rangle \end{bmatrix}$ is positive semidefinite, its real part is positive semidefinite as well, that is, the matrix $\begin{bmatrix} 1 & a_1 & a_3 \\ a_1 & 1 & a_2 \\ a_3 & a_2 & 1 \end{bmatrix}$ is positive semidefinite. Hence its determinant is nonnegative, and the desired result follows.

It is natural to ask whether the analogous relation holds for Ψ_{xz}, Ψ_{xy}, Ψ_{yz}. The answer is yes.

Note that

$$\Psi_{xy} = \inf_{\alpha, \beta \in \mathbb{C}\setminus\{0\}} \Phi_{\alpha x \beta y} = \inf_{\alpha \in \mathbb{C}\setminus\{0\}} \Phi_{\alpha x y} = \inf_{\beta \in \mathbb{C}\setminus\{0\}} \Phi_{x \beta y}.$$

Using (1) we have, for any $x, y, z \in \mathbb{C}^n \setminus \{0\}$

$$\begin{aligned} \inf_{\alpha, \beta \in \mathbb{C}\setminus\{0\}} \Phi_{\alpha x \beta z} &\leq \inf_{\alpha, \beta \in \mathbb{C}\setminus\{0\}} (\Phi_{\alpha x y} + \Phi_{y \beta z}) \\ &= \inf_{\alpha \in \mathbb{C}\setminus\{0\}} \Phi_{\alpha x y} + \inf_{\beta \in \mathbb{C}\setminus\{0\}} \Phi_{y \beta z}, \end{aligned}$$

so

$$\Psi_{xz} \leq \Psi_{xy} + \Psi_{yz}. \tag{3}$$

The derivation shows that the inequality for Ψ_{xy} is in a sense weaker than that for Φ_{xy}.

The next result is an application of the inequalities (1) and (3). Roughly speaking, it gives the triangle inequalities for the corresponding chordal metrics on the unit ball.

I start with a simple lemma, whose geometric meaning is obvious, and which can be regarded as the real 2-dimensional case of the general result.

LEMMA. Let $\alpha \in [0, \pi]$, β, $\gamma \in [0, \frac{\pi}{2}]$ with $\alpha \leq \beta + \gamma$. Then

$$\sin \alpha \leq \sin \beta + \sin \gamma. \tag{4}$$

PROOF. If $0 \leq \beta + \gamma \leq \frac{\pi}{2}$, obviously,

$$\sin \alpha \leq \sin(\beta + \gamma) = \sin \beta \cos \gamma + \sin \gamma \cos \beta \leq \sin \beta + \sin \gamma.$$

If $\frac{\pi}{2} \leq \beta + \gamma \leq \pi$, then $\beta \geq \frac{\pi}{2} - \gamma$,

$$\begin{aligned} \sin \beta + \sin \gamma &\geq \sin\left(\frac{\pi}{2} - \gamma\right) + \sin \gamma = \cos \gamma + \sin \gamma \\ &= \sqrt{2} \sin\left(\gamma + \frac{\pi}{4}\right) \geq 1 \geq \sin \alpha. \end{aligned}$$

In [3] (see also [4, p. 195]), the following elegant inequality was derived as a tool in proving a trace inequality for unitary matrices:

$$\sqrt{1 - |\langle u, v\rangle|^2} \leq \sqrt{1 - |\langle u, w\rangle|^2} + \sqrt{1 - |\langle v, w\rangle|^2}, \tag{5}$$

where u, v and $w \in \mathbb{C}^n$ are unit vectors. In our present notation, this could be written

DOI 10.1007/s00283-011-9270-z

$$\sin \Psi_{uv} \leq \sin \Psi_{uw} + \sin \Psi_{vw}.$$

We see that (5) follows immediately from (3) and the Lemma. Moreover, we have

PROPOSITION 1. For any unit vectors u, v, and $w \in \mathbb{C}^n$

$$\sqrt{1-(\mathrm{Re}\langle u,v\rangle)^2} \leq \sqrt{1-(\mathrm{Re}\langle u,w\rangle)^2} + \sqrt{1-(\mathrm{Re}\langle v,w\rangle)^2}. \tag{6}$$

PROOF. If $\mathrm{Re}\langle u,w\rangle \leq 0$, we replace u by $-u$; if $\mathrm{Re}\langle v,w\rangle \leq 0$, we replace v by $-v$. That is to say, we may always let Φ_{uw}, $\Phi_{wv} \in [0, \frac{\pi}{2}]$ and $\Phi_{uv} \in [0, \pi]$, thus the condition in the Lemma is satisfied. Therefore $\sin \Phi_{uv} \leq \sin \Phi_{uw} + \sin \Phi_{wv}$, that is, (6) holds.

To end this note, here is an extension of (5) and (6).

PROPOSITION 2. Let $p > 2$, then for any unit vectors u, v and $w \in \mathbb{C}^n$ we have

$$\sqrt[p]{1-|\langle u,v\rangle|^p} \leq \sqrt[p]{1-|\langle u,w\rangle|^p} + \sqrt[p]{1-|\langle v,w\rangle|^p} \tag{7}$$

$$\sqrt[p]{1-|\mathrm{Re}\langle u,v\rangle|^p} \leq \sqrt[p]{1-|\mathrm{Re}\langle u,w\rangle|^p} + \sqrt[p]{1-|\mathrm{Re}\langle v,w\rangle|^p} \tag{8}$$

PROOF. Fix $p > 2$ and set $f(t) = (1-(1-t^2)^{p/2})^{1/p}$ for $t \in [0,1]$. Then

$$\frac{d}{dt} f(t) = (1-(1-t^2)^{p/2})^{1/p-1}(1-t^2)^{p/2-1} t \geq 0.$$

and

$$\frac{d}{dt}\frac{f(t)}{t} = t^{-2}(1-(1-t^2)^{p/2})^{1/p-1}((1-t^2)^{p/2-1}-1) \leq 0.$$

Since $f(t)$ is increasing and $f(t)/t$ is decreasing, if $a, b, c \in [0,1]$ and $0 \leq a \leq b + c \leq 1$, then

$$\begin{aligned} f(a) \leq f(b+c) &= b\frac{f(b+c)}{b+c} + c\frac{f(b+c)}{b+c} \leq b\frac{f(b)}{b} + c\frac{f(c)}{c} \\ &= f(b) + f(c). \end{aligned}$$

Taking $a = \sqrt{1-|\langle u,v\rangle|^2}$, $b = \sqrt{1-|\langle u,w\rangle|^2}$, and $c = \sqrt{1-|\langle v,w\rangle|^2}$, we get

$$\begin{aligned} (1-|\langle u,v\rangle|^p)^{1/p} &= f(a) \leq f(b) + f(c) \\ &= (1-|\langle u,w\rangle|^p)^{1/p} + (1-|\langle v,w\rangle|^p)^{1/p}. \end{aligned}$$

This proves (7). Inequality (8) can be proved similarly.

ACKNOWLEDGMENT

The author would like to thank Gord Sinnamon for helpful discussions.

REFERENCES

[1] K. E. Gustafson, D. K. M. Rao, *Numerical Range*, Springer-Verlag, New York, 1997.

[2] M. Kreĭn, Angular localization of the spectrum of a multiplicative integral in a Hilbert space, *Funct. Anal. Appl*. 3 (1969), 89–90.

[3] B. Wang, F. Zhang, A trace inequality for unitary matrices, *Amer. Math. Monthly* 101 (1994), 453–455.

[4] F. Zhang, *Matrix Theory: Basic Results and Techniques*, Springer-Verlag, New York, 2011.

Department of Combinatorics and Optimization
University of Waterloo
Waterloo, ON N2L 3G1
Canada
e-mail: mlin87@ymail.com

Picturing Gauss

Judith Saunders

In this portrait, etched on an old ten-Mark bill,
you wear a velvet cap and sport white sideburns,
your earnest gaze directed far beyond the world
of getting and spending. Symbols of your legacy
adorn the worthless paper, lending it
antique dignity: a normal curve, a sextant,
the university buildings at Göttingen, where you
calculated the orbit of Ceres, invented the telegraph.

The currency that paid you homage is obsolete.
Your picture no longer makes commercial rounds,
traveling hand to hand or nestling securely
in German wallets. Serenely indifferent
to your displacement, you do not mourn
the Deutsche Mark, that blip on the screen
of numismatic history.

Let each new state
of human affairs, however great or glorious,
be represented on the graph of time as one
small point, insignificant in itself. Now
plot those points, which assume meaning
in the aggregate-emphatic patterns.
Disregarding individual example
(the claims of an Ozymandias) calculate
the distribution of many instances.
Let welter yield to elegant design,
a bell-shaped principle that replicates itself
reliably, outlasts all application, squeezes
sense from recalcitrant data. Stand
by this, the currency of the mind, which remains
negotiable and never ceases to circulate.

Department of English
Marist College
Poughkeepsie, NY 12601
USA
e-mail: judith.saunders@marist.edu

DOI 10.1007/s00283-011-9263-y

Mathematicians in the World of Finance

PETER HAGGSTROM

The Viewpoint *column offers readers of* The Mathematical Intelligencer *the opportunity to write about any issue of interest to the international mathematical community. Disagreement and controversy are welcome. The views and opinions expressed here, however, are exclusively those of the author, and the publisher and editors-in-chief do not endorse or accept responsibility for them.* Viewpoint *should be submitted to one of the editors-in-chief, Chandler Davis and Marjorie Senechal.*

I read with interest the article by Jonathan Korman ["Finance and Mathematics: A Lack of Debate", *Mathematical Intelligencer*, vol. 33 (2011), no. 2, 4–6], which provoked a response from Janusz Konieczny (*ibid*., no. 4). I went back and read Marc Rogalski's article ["Mathematics and Finance: An Ethical Malaise", *Mathematical Intelligencer*, vol. 32 (2010), no. 2, 6–8] and Ivar Ekeland's response to that article (*ibid*., 9–10).

As a nonacademic, mathematically trained person who has worked at the director level in a global asset management business in Australia and also in senior federal government jobs, I offer a different perspective. I understand Korman's visceral reaction to the havoc that mathematically trained people have wreaked in the investment banking world. The bank I worked for was leveraged 55 times at the height of the global financial crisis. This was more than twice the level of borrowing employed by the infamous Long Term Capital Management Fund that exploded in 1998 and vaporized vast amounts of investors' money. Nobel Prize winners Myron Scholes and Robert Merton were involved with that fund, providing early proof that mathematical ability and ability to predict markets consistently are not necessarily related.

What is going on here is basic economics rather than ideology. People go where the money is – it is as simple as that. If the National Security Agency decided to pay cryptographers $1 million per year plus a code-cracking bonus, watch how many people would suddenly become enthusiastic about number theory.

In 2002 I employed a Ph.D. in photonics to do financial work that was the analytical equivalent of street sweeping. He could not get work in his field after the "tech wreck" in 2000, but in asset management he subsequently more than doubled his salary doing quant work for a bank, earning far more than his academic colleagues. Not a matter of ability: these high salaries in finance are a result of a long bull market in equities; but we are now seeing a systemic decline in financial businesses because of the flat equities markets and the unresolved stresses generated by the global financial crisis.

Something interesting happened, though, to perceptions of risk during the last twenty years. Risk management used to be the preserve of very narrowly trained and conservative actuaries who generally had no exposure to stochastic calculus. Then the option pricing formula of outsiders Black and Scholes stripped away the actuarial monopoly on quantifying risk. Now risk quantification could be performed by anyone with an Excel spreadsheet. I can only speak for Australia, but here actuaries' remuneration declined as applied-finance qualified people have competed with them in the financial world. With more people being able to quantify risk, it is no surprise that things generally got riskier. As people see the money to be made, there is a net flow of talent into the area, and there is a general tendency to "push the envelope".

Like Ivar Ekeland, I agree that more incisive mathematical training will enable people to understand risk better. What academics generally miss is how much risk arises from the business models employed. There is a lot of literature about the mathematics of risk, but little about how the business models can render a theoretical analysis largely irrelevant.

Business models are now better understood by regulators throughout the world; it will become harder for businessmen to play the games they used to play with other people's money, because the rules are being changed to make capital much more expensive. I have learned from working in the industry that the people at the top of big financial institutions will not really risk their own capital, but are quite happy to risk that of someone else. The regulator, by making externally-sourced capital more expensive as in the lamented Glass-Steagall Act, can constrain the extent to which Wall Street firms can feed off others. Before the repeal of the Glass-Steagall Act the big investment banks were essentially harmless partnerships simply because they could not use funds of a deposit-taking bank to fund their risk appetites. Small-business

DOI 10.1007/s00283-011-9273-9

people who mortgage their homes to run their businesses appreciate personal risk better than the CEO of an investment bank.

This regulatory trend has already started, and people will behave accordingly. In Australia the market is already providing signals to people that investment banking may not be the growth area it used to be, and young people are now looking at growth areas such as gaming (for the mathematically inclined this provides all sorts of roles). Now is not the moment for an academic mathematician to jump ship into finance.

Although these observations do not exonerate some simply outrageous business behaviour that has occurred at management level in financial institutions worldwide, they may make it explicable.

A final note: At the height of the global financial crisis there were senior people in investment banking "close to the action" who let go their equities in favour of cash to preserve their own capital, yet the businesses they controlled kept pushing equities to their clients. In short, they used their knowledge to minimise risk to their own capital while actively inciting everyone else to take on risk and hence keep the merry-go-round turning. Alas, one does not find that insight in the standard textbooks.

Bondi Beach
Sydney
Australia
e-mail: peter.haggstrom@exemail.com.au

Ivar Ekeland comments:

I agree with Haggstrom. I see limited liability (the fact that you are not subject to the downside risk, only to the upside) as the major problem facing the financial industry, and I am currently working on general models of this situation. Unfortunately, I am not as confident as Haggstrom is that the situation will ever be corrected. Remember the old saying: "Never give a sucker an even break."

Jonathan Korman comments:

Haggstrom says "people go where the money is" as though this were a natural law or a moral imperative. It is neither. Our financial masters may tell us that obedience to the markets is simply human nature, but the issue Rogalski raised is that we have a choice. So let us face up to our responsibility to choose. We who are not committed to the profit motive – in particular, most mathematicians – do not need to play along with the financial industry, let alone glorify it.

The boundless greed of the financial industry has been terribly destructive; Haggstrom promises it will self-regulate, but does not show us any reason to be reassured; Ekeland envisages a policy of regulation that might rein in its reckless behaviour, but does not expect to see this happen. (By the way, let us distinguish models that describe the financial market from models that are part of it. Just as one may study probability without running a casino, one may study risk-taking without working for Wall Street.)

Sketchy Tweets: Ten Minute Conjectures in Graph Theory

ANTHONY BONATO AND RICHARD J. NOWAKOWSKI

Comments by Richard Hamming in his address *You and Your Research* [16] resonated with us. On the one hand, Hamming says,

"*What are the most important problems in your field?*"

which suggests working on hard problems. Yet on the other, he exhorts us to

"*Plant the little acorns from which mighty oak trees grow.*"

Following this advice, we should look over the big questions, then doodle and sketch out some approaches. If you are an expert, then this is easy to do, but most people do not want to wait to become an expert before looking at interesting problems. Graph theory, our area of expertise, has many hard-to-solve questions. Some hark back to the recreational roots of the area yet still keep their mystery. These "acorns" can be planted on the backs of envelopes, on a blackboard, and over a coffee.

Our goal is to collect some of these conjectures—arguably some of the most intriguing—in one place. We present ten conjectures in graph theory, and you can read about each one in at most ten minutes. As we live in the era of Twitter, all the conjectures we state are 140 characters or fewer (so "minute" here has a double meaning). We might even call these *sketchy tweets*, as we present examples for each conjecture that you can doodle on as you read.

Hamming also references *ambiguity*: good researchers can work both on proving and disproving the same statement, so we approach the conjectures with an open mind. He also mentions that a good approach is to reframe the problem, and change the point of view. One example from Vizing's Conjecture (which is discussed as our second-to-last conjecture in the following text), is the three-page paper [2] which, with a new way of thinking, reduced most of the published work of twenty years to a corollary of its main result!

Given the size of modern graph theory, with its many smaller subfields (such as structural graph theory, random graphs, topological graph theory, graph algorithms, spectral graph theory, graph minors, and graph homomorphisms, to name a few), it would be impossible to list all, or even the bulk of the conjectures in the field. We are content instead to focus on a few family jewels, which have an intrinsic beauty and have provided some challenges for graph-theorists for at least two decades. There is something for everyone here, from undergraduate students taking their first course in graph theory, to seasoned researchers in the field. For additional reading on problems and conjectures in graph theory and other fields, see the Open Problem Garden maintained by IRMACS at Simon Fraser University [24].

We consider only finite and undirected graphs, with no multiple edges or loops (unless otherwise stated). We assume the reader has some basic familiarity with graphs and their terminology, including notions such as cycles, paths, complete graphs, complete bipartite graphs, vertex degrees, and connected graphs. We use the notation C_n for the cycle with n vertices, P_n for the path with n vertices, and K_n for the complete graph with n vertices. The complete bipartite graph with m and n vertices of the respective colors is denoted by $K_{m,n}$.

All the background we need can be found in any text in graph theory, such as those of Diestel [9] and West [41], or online (see for example [42]).

The authors gratefully acknowledge support from NSERC, MITACS, and Ryerson.

DOI 10.1007/s00283-012-9275-2

For a graph G, we write $V(G)$ for its *vertex set*, and $E(G)$ for its *edge set*. If two vertices are joined by an edge, then we say they are *adjacent*. The *order* of a graph is the cardinality of its vertex set.

The Conjectures

Some conjectures we present (such as Meyniel's) are less known and deserve more exposure, whereas others (such as Hadwiger's) are better known. We provide no justification for our bias toward one problem over another, so we apologize upfront if your favorite conjecture is missing.

We present each conjecture using minimal technical jargon. To shorten the number of references to partial results, we cite surveys wherever possible. We always cite the original authors of the problem. The conjectures we present have spawned enormous amounts of work on related problems and concepts, which in the present article can only be hinted at. For example, a quick check of Google Scholar or MathSciNet will reveal many thousands of papers related to the topic of one of the ten conjectures we present. Some of these works are discussed in the surveys.

All the conjectures here are considered difficult, having remained unsolved for many years. We do not rank the conjectures in order of difficulty. To show no preference among the problems, we present the conjectures in alphabetical order.

Double the Fun

There is an old puzzle, found in many books that feature "pencil-and-paper" problems, of attempting to trace a diagram without lifting the pen off the paper or retracing any part of the figure. Euler in his famous 1736 solution to the Königsberg bridge problem, essentially found when this can be done. The problem can be restated as covering the diagram (represented by a graph) with a cycle. Such graphs are now called *Eulerian*; the connected Eulerian graphs are those in which every vertex has even degree. (We note that Euler did not prove the characterization of Eulerian graphs; the first proof was published by Hierholzer [18].) Our first conjecture may be thought of as a generalization of this kind of problem to graphs with some vertices of odd degree.

A *bridge* is an edge whose deletion disconnects the graph. A graph with no bridges is *bridgeless*. For example, each edge of a tree is a bridge.

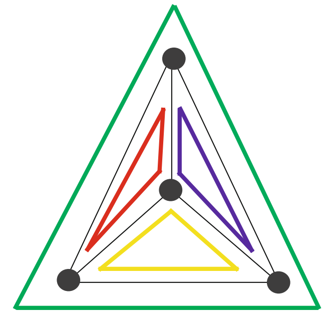


Figure 1. A CDC in K_4, with the cycles in different colors. What would be a CDC of K_5?

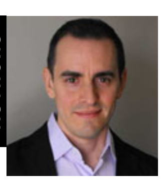
> **Cycle Double Cover Conjecture**: For every bridgeless graph, there is a list of cycles so that every edge appears in exactly two cycles.

A list of cycles as in the conjecture is called a *CDC*. See Figure 1 for an example. Note that the list may have repeated cycles, as is the case with C_n. The conjecture was formulated independently by Szekeres in 1973 [34] and Seymour in 1979 [30]. See the survey [20] and book [45] for additional background and references on the conjecture.

The conjecture has connections to *embeddings of graphs on surfaces*; that is, drawings of graphs on different surfaces so that no two edges cross. The simplest case is the family of *planar graphs*, which have an embedding in the plane (see Hadwiger's Conjecture for more on planar graphs). If each face in the embedding corresponds to a cycle in the graph, then the faces form a CDC as in Figure 1, as is true for all connected, bridgeless, planar graphs. That there is an embedding in some surface where each face corresponds to a cycle is the *Strong Embedding Conjecture*, which is a stronger conjecture than the Cycle Double Cover Conjecture.

In the other direction, much is known about the smallest counterexample, if it exists: every vertex has degree 3, it is not 3-*edge-colorable* (i.e., there is no coloring of the edges with three colors so that no edge is incident with an edge of the same color), it is *cyclically 4-connected* (i.e., every partition of the vertices into two parts with a cycle in each part has at least four edges that join the parts), and the smallest cycle has length at least 10. A connected, bridgeless, cubic graph that is not 3-edge-colorable is called a

AUTHORS



ANTHONY BONATO is Chair of the Department of Mathematics at Ryerson. His research is mostly on graph theory, including its applications to real-world complex networks. He is an Editor-in-Chief of the journal *Internet Mathematics.*

Department of Mathematics
Ryerson University
Toronto, M5B 2K3
Canada
e-mail: abonato@ryerson.ca

RICHARD J. NOWAKOWSKI is a Professor at Dalhousie and is one of the leaders of the Graphs and Games Group there. He studies especially combinatorial game theory—learning, at least, how to lose well—and graph theory.

Department of Mathematics and Statistics
Dalhousie University
Halifax, (NS) B3J 3H5
Canada
e-mail: rjn@mathstat.dal.ca

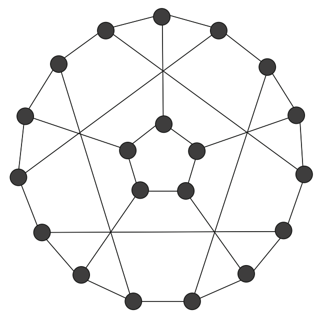


Figure 2. The flower snark J_5.

snark. Another conjecture says that there are not any snarks with the smallest cycle length at least 10. See Figure 2 for an example of a snark.

A stronger conjecture than the Cycle Double Cover Conjecture is the *Small Cycle Double Cover Conjecture*: every bridgeless graph on n vertices has a CDC of size at most $n - 1$.

Party, But Know Your Limits

Frank Ramsey, who died at the early age of 26, wrote a paper [26] in mathematical logic that has gone on to have applications in many fields, including graph theory. To motivate Ramsey numbers, consider a party with six people, some pairs of whom are friends, and some of whom are strangers. It is not hard to show that among six people we can always find three mutual friends or three mutual strangers. Readers should convince themselves that in smaller parties this property is not always satisfied.

Ramsey numbers generalize this setting from three to n mutual friends or strangers. For a positive integer n, define the nth *Ramsey number*, written $R(n)$, to be the minimum integer r such that any coloring of the edges of K_r with red or blue (red joins friends and blue joins strangers) results in a complete subgraph of order n whose edges all have the same color.

It is not immediately clear whether the Ramsey numbers even exist. Calculating them directly is hard; whereas $R(4) = 18$, the value of $R(5)$ is unknown (although it is between 43 and 49; see [25] for a dynamic survey of the known small Ramsey numbers). We must be content with lower and upper bounds. An inductive argument gives $R(n) \leq \binom{2n-2}{n-1}$. In an early application of the probabilistic method, Erdős [11] proved the lower bound

$$(1 + o(1))\frac{1}{e\sqrt{2}} n 2^{n/2} \leq R(n),$$

which has not been substantially improved to this day (Spencer [32] improved the constant $\frac{1}{\sqrt{2}}$ to $\sqrt{2}$). The best known upper bound for $R(n)$, which is far apart from the known lower bounds, was given by Thomason [36]:

$$R(n) \leq n^{-1/2 + c/\sqrt{\log n}} \binom{2n-2}{n-1}.$$

Erdős in 1947 posed the following asymptotic conjecture, and it remains one of the major topics in Ramsey numbers.

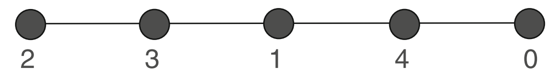


Figure 3. The path P_5 is graceful. Show that all the paths P_n are graceful.

Erdős' Ramsey Number Conjecture: $\lim_{n\to\infty} R(n)^{1/n}$ exists.

Solve this conjecture and you will be awarded \$100. However, this is a ridiculously difficult way to make \$100! From the bounds stated previously, if the limit exists, then it is between $\sqrt{2}$ and 4 (finding the value of the limit is worth \$250). Hence, we can think of the conjecture as a way to understand better which bound for $R(n)$ stated previously is more accurate. For more on Erdős and his questions about Ramsey numbers, see [8].

Saving Grace

We next consider a graph-labelling problem, where vertices or edges are labelled by numbers subject to some constraints. A graph is *graceful* if the vertices can be assigned numbers from among $0, 1, \ldots, m$, where m is the number of edges, so that the differences along the edges are precisely $1, 2, \ldots, m$. Graceful graphs have received ample attention in the literature. See Figure 3 for an example.

Graceful labellings were introduced by Rosa under the name of β-labellings and were renamed "graceful" by Golomb.

The following conjecture is now sometimes called the *Ringel-Kotzig Conjecture* (because, if the conjecture were true, it would imply conjectures of both authors on certain decompositions of complete graphs).

Graceful Tree Conjecture: Every tree is graceful.

More than 200 papers have been written on proving special cases of this conjecture, and a bewildering number of variants on graceful labellings have been proposed and studied. See the dynamic survey of Gallian [13] for further background and references on graceful (and other) labellings. Kotzig labelled the collective work on proving the conjecture a "disease." A few of the classes of trees where we know the conjecture holds are: caterpillars (a caterpillar is a tree whose nonleaf vertices form a path), trees with at most four leaves, trees with diameter at most 5, and trees with at most 35 vertices.

Much of the research on the conjecture tries to settle it in the affirmative. One class of trees where the conjecture remains open are lobsters (those where the removal of the leaves gives a caterpillar).

No Minors Allowed

Coloring has both fascinated and perplexed graph-theorists since the early days of the field. The *chromatic number* of G, written $\chi(G)$, is the minimum number of colors in a vertex labelling such that adjacent vertices receive distinct colors; that is, the minimum k such that G is *k-colorable*. The most famous theorem proved so far in graph theory is the Four-Color theorem [1], which states that every planar graph is 4-colorable. All known proofs of this fact are computer-assisted.

A graph is a *minor* of G if it results by repeatedly performing one of the following operations: i) deleting a vertex, ii) deleting an edge, or iii) contracting an edge (i.e., shrinking an edge to a vertex and preserving adjacencies and nonadjacencies with vertices outside the edge). A beautiful result of Wagner [40] states that a graph is planar if and only if it does not have K_5 or $K_{3,3}$ as a minor. The reader can show that the Petersen graph (see Figure 8) has K_5 as a minor, and hence, is not planar.

Hadwiger's Conjecture, dating back to 1943 [15], relates graph coloring to minors.

Hadwiger's Conjecture: For $m \geq 2$, a graph not having K_m as a minor is $(m-1)$-colorable.

Hadwiger's Conjecture is open for $m \geq 7$. The startling case for small m is $m = 5$, which was shown by Wagner [40] to reduce to the Four-Color theorem. Hence, Hadwiger's Conjecture may be viewed as a broad generalization of that theorem. The case $m = 6$ was settled by Robertson, Seymour, and Thomas [28] by showing that a minimal counterexample to the conjecture is planar after the removal of one vertex (so this also reduces to the Four-Color Theorem).

The cases $m = 2$ and 3 are elementary (for example, a graph without K_2 as a minor has no edges, and a graph bit having K_3 as a minor is a forest). Dirac [10] and Hadwiger [15] proved the case $m = 4$, by showing that graphs not having K_4 as a minor have a vertex of degree at most 2 and, hence, can be 3-colored using a greedy algorithm. Although the case $m = 7$ is open, in 2005 Kawarabayashi and Toft [21] proved that any 7-chromatic graph has K_7 or $K_{4,4}$ as a minor.

X Marks the Spot

As with Hadwiger's Conjecture, our next conjecture also deals with coloring but adds graph products to the mix. All the references in this section can be found in three surveys on the conjecture: [29, 35, 46].

A graph product makes new graphs from old. We consider one of the best known products: the *categorical product* (which is also referred to as the tensor or Kronecker product). For graphs G and H, define $G \times H$ to have vertex set $V(G) \times V(H)$, with (a, b) adjacent to (c, d) if a is joined to c in G, and b is joined to d in H. See Figure 4, which motivates the notation for this product.

Hedetniemi's Conjecture yields a simple formula for the chromatic number of the categorical product; it was posed by him in 1966 [17] while he was a graduate student.

Hedetniemi's Conjecture: For graphs G and H,

$$\chi(G \times H) = \min\{\chi(G), \chi(H)\}.$$

The conjecture was stated independently by Burr, Erdős, and Lovász in 1976. Most experts think the conjecture is

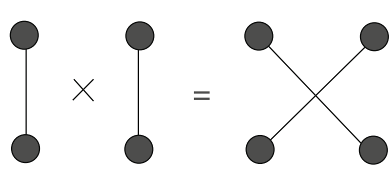


Figure 4. The graph $K_2 \times K_2$.

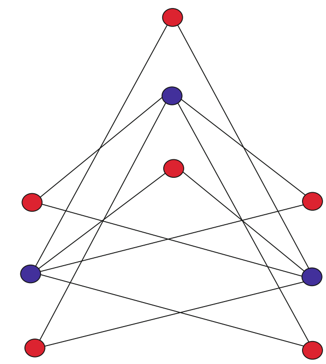


Figure 5. The graph $P_3 \times K_3$ is 2-colorable.

true. For starters, $G \times H$ may be visualized as replacing each vertex v of G by a copy of the vertices of H. Label these vertices as (v, h). Then add the edges (v, h) (w, j) just if v is adjacent to w and h is adjacent to j. See Figure 5. Now take a proper coloring of G. For each vertex v of G, color all vertices (v, h) with the same color as v. Since (v, h) and (v, h') are not adjacent, this is also a proper coloring of $G \times H$. Hence, $\chi(G \times H) \leq \chi(G)$. The same construction, but considering vertices of H, gives $\chi(G \times H) \leq \chi(H)$.

The conjecture has a convenient restatement that is often used. For a positive integer n, let $H(n)$ denote the following statement:

If $\chi(G \times H) = n$, *then either* $\chi(G) = n$ *or* $\chi(H) = n$.

Hedetniemi's Conjecture is equivalent to $H(n)$ being true for all $n \geq 1$, which permits an incremental approach. Indeed, it is not too difficult to show that $H(1)$ and $H(2)$ are true. El-Zahar and Sauer proved in 1985 that $H(3)$ is true, but little is known about $H(n)$ for $n > 3$.

Burr, Erdős, and Lovász in their 1976 paper showed that if G is a graph in which every vertex lies in a complete subgraph of order n, and H is a connected graph with $\chi(G \times H) = n$, then $\min\{\chi(G), \chi(H)\} = n$. (In Figure 5, we have that $n = 2$). This is not too surprising, since the presence of K_n in G is a trivial reason why $\chi(G) \geq n$. Proofs of the conjecture, or the search for a counterexample, therefore must consider graphs that have large chromatic number and small complete subgraphs.

The strangest result arising out of the work on the conjecture has to do with a special case. Define the function

$$g(n) = \min\{\chi(G \times H) : \chi(G) = \chi(H) = n\}.$$

It is known that $g(1) = 1$, $g(2) = 2$, $g(3) = 3$, and $g(4) = 4$. Several authors discovered the striking fact that either g is unbounded or $g(n) \leq 9$ for all n. Since Hedetniemi's Conjecture has received a lot of attention over the past 45 years, if there were a counterexample surely it would have been found by now!

We mention in passing (and without explanation of the jargon!) that Hedetniemi's Conjecture is equivalent to the meet-irreducibility of the complete graphs in the lattice of cores. For this reason, the conjecture is of ample interest not only to experts in graph coloring, but also to those working on graph homomorphisms.

The Long Arm of the Law

Many of us played games such as Cops and Robbers (or other pursuit games) as children, and our next conjecture considers such a game played on graphs. In the graph game of *Cops and Robbers* there are two players, a team of cops and a robber, who move from vertex to vertex along edges in the graph or can pass. The game is played with alternate moves of the players. The cops move first, by choosing some set of vertices for their team to occupy. The robber then chooses a vertex. The cops win if eventually they capture or land on the vertex with the robber; the robber wins if he can indefinitely evade capture. The game has perfect information, in the sense that both players can see and remember each other's moves. Placing a cop on each vertex provides an easy win for the cops. The minimum number of cops needed to win the game is the *cop number* of a graph. The reader may verify that the cop number of the snark J_5 in Figure 2 is 3.

As the cop number of a disconnected graph is the sum of the cop numbers of its components, it is sensible to consider only connected graphs. For functions f and g on positive integers taking positive real values, we write $f = O(g)$ to mean that there is a constant d, such that for large enough n, $f(n) \leq dg(n)$.

Meyniel's Conjecture: If G is a connected graph, then

$$c(G) = O(\sqrt{|V(G)|}.$$

Meyniel's Conjecture states that about a constant multiple of $\sqrt{n}$ many cops are sufficient to capture the robber in a connected graph of order n (and there are examples of graphs needing this many cops). Aigner and Fromme in 1984 proved that the cop number of a planar graph is at most 3.

Meyniel's Conjecture may be one of the lesser known unsolved conjectures in graph theory, but it has received a fair bit of recent attention. For further background on the conjecture, see Chapter 3 of the book [3]. Meyniel's Conjecture was communicated by Frankl [12], who could only prove that

$$c(G) = O\left(n\frac{\log\log n}{\log n}\right).$$

To date the best available general bound is the following, recently discovered by three independent sets of researchers:

$$c(G) = O\left(\frac{n}{2^{(1-o(1))\sqrt{\log_2 n}}}\right).$$

Even to prove that $c(G) = O(n^{1-\varepsilon})$ for some positive ε is open! The conjecture was settled for bipartite graphs with diameter 3, and Andreae proved it is true in graph classes formed by avoiding a fixed graph as a minor (in fact, the cop number is bounded by a constant in such graphs); see Chapter 3 of [3].

House of Cards

The Reconstruction Conjecture has proved to be notoriously difficult and suggests how much more there is to learn about graphs. The *deck* of a graph G is the multiset consisting of all subgraphs of G formed by deleting a vertex. Each such vertex-deleted subgraph is a *card*. See Figure 6 for an example.

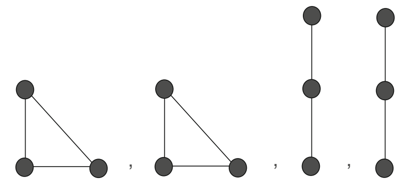


Figure 6. Which graph has this deck?

The conjecture was posed independently by Kelly in 1957 [22] and Ulam in 1960 [38].

Reconstruction Conjecture: If two graphs with at least three vertices have the same deck, then they are isomorphic.

It is easy to see that K_2 and its complement have the same deck: hence the modest requirement on the order of the graph. Given a deck, we immediately know the order of G, and some thought yields the number of edges and the degrees of all the vertices. For references to results on the conjecture, see the survey [4]. Kelly proved that disconnected graphs, trees, and regular graphs are reconstructible from their deck. McKay showed that the conjecture is true for all graphs with at most 11 vertices. The conjecture also holds for outerplanar graphs. Bollobás proved that with probability tending to 1 as n tends to infinity, there exist three cards that determine the graph. Surprisingly, the conjecture remains open for planar graphs.

Go with the Flow

We may view the edges of a graph as a series of pipes transporting some liquid (or electric current, or information) between nodes. Usually edges have a maximum capacity for carrying materials, and what enters a node must equal what must come out. Further, these *flows*, as they are called, usually move in one direction, so some orientation must be assigned to the edges of the network. Flows have deep connections to the Four-Color Theorem, and Tutte's conjecture on flows extends these connections beyond the context of planar graphs.

To be more precise, an *integer flow* on a graph is a pair consisting of an orientation of the graph and an assignment of integer weights to the edges such that for each vertex, the total weight on exiting edges equals the total weight on entering edges. It is a *k-flow* if all weights have absolute value less than k, and it is *nowhere-zero* if weight 0 is never used. See Figure 7. Note that every k-flow is a $(k + 1)$-flow.

Nowhere-zero k-flows were introduced by Tutte [37] as a generalization of face-coloring problems in planar graphs (where we color the faces so that no adjacent faces receive

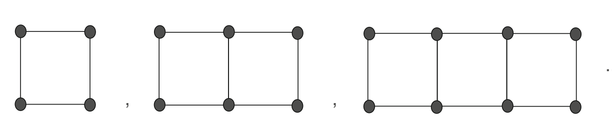


Figure 7. Find nowhere-zero 4-flows for the graphs in this sequence. Do CDCs help?

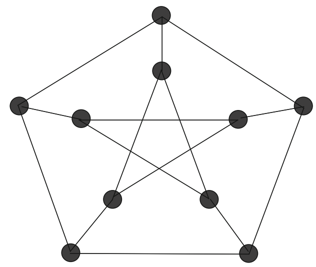


Figure 8. The Petersen graph.

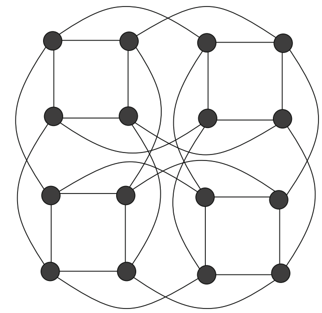


Figure 10. Find a dominating set of size 4 in $C_4 \square C_4$.

the same color). The famous Four-Color Theorem is equivalent to saying that every planar bridgeless graph has a nowhere-zero 4-flow. Unfortunately, this result cannot be extended to arbitrary bridgeless graphs, since the Petersen graph has no nowhere-zero 4-flow. See Figure 8.

Tutte in 1954 therefore considered 5-flows instead and conjectured the following in [37].

Tutte's 5-Flow Conjecture: Every bridgeless graph has a nowhere-zero 5-flow.

The conjecture holds for planar graphs, using the duality between flows and coloring and the Five-Color Theorem (every planar graph is 5-colorable). We may therefore view the 5-flow conjecture as a generalization of the Five-Color Theorem to graphs that are not planar.

Jaeger [19] proved that every bridgeless graph has a nowhere-zero 8-flow. Seymour [31] improved upon this result by showing that bridgeless graphs have nowhere-zero 6-flows. Celmins [6] proved that a smallest counterexample to the conjecture must be a cyclically 5-edge-connected snark with girth at least 7 (see also the Cycle Double Cover Conjecture). The conjecture can be reduced to the 3-regular case, and Steinberg [33] proved the conjecture for graphs that embed on the projective plane.

You Are So Square

We state another conjecture about products, this time related to domination. (No, not the other kind!) For graphs G and H, define the *Cartesian product* of G and H, written $G \square H$, to have vertex set $V(G) \times V(H)$, with (a, b) adjacent to (c, d) if $a = c$ and b is joined to d in H, or if $b = d$ and a is joined to c in G. See Figure 9, which motivates the notation for this product.

In a graph G, a set S of vertices is a *dominating set* if every vertex not in S has a neighbour in S. The *domination number* of G, written $\gamma(G)$, is the minimum size of a dominating set. For example, see Figure 10, where we note that $\gamma(C_4 \square C_4) = \gamma(C_4)\gamma(C_4) = 4$.

The following was proposed by Vizing in 1968 [39].

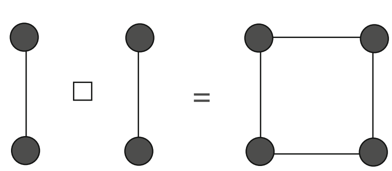


Figure 9. The graph $K_2 \square K_2$.

Vizing's Conjecture: For graphs G and H,

$$\gamma(G \square H) \geq \gamma(G)\gamma(H).$$

This is a conjecture everyone thinks is true. All the references that follow appear in [5].

An important theorem from 1979 was not appreciated (or was completely overlooked) for the next 16 years. During some collaboration of Rall and the second author of this article, they noticed Math Review MR0544028 of a paper by Barcalkin and German [2]. From that brief description, in a couple of days Hartnell and Rall were able to reconstruct the mathematical arguments of the paper (rather than translate the paper, which was in Russian). The results in [2] reduced much of the work from 1968 through 1996 to a corollary of their main theorem!

As with several of the conjectures considered so far, many experts think that if the conjecture were false then a minimal counterexample would have been found. But since a proof has not been found, where would one look for a counterexample? For example, a minimal counterexample to Vizing's conjecture must have domination number larger than 3; adding an edge between two nonadjacent vertices decreases the domination number; and every vertex belongs to a minimum dominating set.

Don't Get Cross

Our final conjecture has its origins in a 1940s labor camp in Budapest. The famous mathematician Turán was imprisoned there, watching trucks move bricks along rails from kilns to storage areas. Every once in a while, two trucks would cross each other's paths and the bricks would come crashing down. No doubt as a kind of liberation from the monotony, Turán began thinking about minimizing the crossings of the trucks, assuming the general situation that there were m kilns and n trucks. For more on the history of the problem see [47].

We may formalize Turán's problem in the following way. The *crossing number* of G, written $\mathrm{cr}(G)$, is the minimum number of pairwise crossings of edges in a drawing of G in the plane. Some readers may recall the *Three Utilities Problem*, which reduces to showing that $K_{3,3}$ has crossing number 1.

Crossing numbers tend to be hard to calculate exactly, because of the exponentially many drawings any given graph may possess. The following conjecture is named after Zarankiewicz who published a flawed proof of it [44], but it is also called *Turán's Brick Factory Conjecture* (see [14] for a survey of the history of the conjecture).

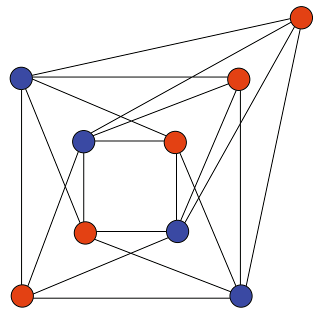


Figure 11. A drawing of $K_{4,5}$ with 8 crossings, which is the minimum number possible.

Zarankiewicz's Conjecture:

$$\mathrm{cr}(K_{m,n}) = \left\lfloor \frac{m}{2} \right\rfloor \left\lfloor \frac{m-1}{2} \right\rfloor \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor.$$

The best exact result on the conjecture was published by Kleitman in 1970 [23] who confirmed it for $n \leq 6$. Kleitman also proved that the smallest counterexample, if it exists, must occur for m and n odd. Woodall [43] computed the crossing numbers $\mathrm{cr}(K_{7,7}) = 81$ and $\mathrm{cr}(K_{7,9}) = 144$. Hence, the smallest unsolved cases are for $K_{7,11}$ and $K_{9,9}$. As $\lfloor\frac{4}{2}\rfloor\lfloor\frac{3}{2}\rfloor\lfloor\frac{5}{2}\rfloor\lfloor\frac{4}{2}\rfloor = 8$, the drawing in Figure 11 achieves the conjectured bound.

A recent result in [7] states that if for a fixed m, the conjecture holds for all values n smaller than some constant depending on m, then the conjecture holds for all n. Hence, for each m there is an algorithm that verifies the conjecture for all n or gives a counterexample.

Epilogue

Now it is your turn to finish this survey and solve (or partially solve) one or more of these conjectures in graph theory. And when you are done with those, we have a few others that might keep you busy, such as Barnette's Conjecture, the Berge-Fulkerson Conjecture, the Erdős-Sós Conjecture, the Middle Levels Conjecture, or Sheehan's Conjecture.

Our version of another Hamming quote [16] provides our parting words:

"Go forth, then, and doodle great work!"

ACKNOWLEDGMENTS

We would like to thank the anonymous referees for several suggestions that improved the quality and readability of the paper.

REFERENCES

[1] K. Appel, W. Haken, J. Koch, Every planar map is four colorable, *Illinois J. Math.* 21 (1977), 439–567.

[2] A. M. Barcalkin, L. F. German, The external stability number of the Cartesian product of graphs, *Bul. Akad. Stiinte RSS Moldoven*, 94 (1979), 5–8.

[3] A. Bonato, R. J. Nowakowski, *The Game of Cops and Robbers on Graphs*, American Mathematical Society, Providence, Rhode Island, 2011.

[4] J. A. Bondy, A Graph Reconstructor's Manual, Surveys in Combinatorics, 1991 (Guildford, 1991), London Mathematical Society Lecture Note Series, 166, Cambridge University Press, Cambridge, 1991, pp. 221–252.

[5] B. Brešar, P. Dorbec, W. Goddard, B. L. Hartnell, M. A. Henning, S. Klavžar, D. F. Rall, Vizing's conjecture: a survey and recent results, *J. Graph Theory* 1(2012), 46–76.

[6] U. Celmins, On cubic graphs that do not have an edge-3-coloring, Ph.D. Thesis, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Canada, 1984.

[7] R. Christian, B. R. Richter, G. Salazar, Zarankiewicz's conjecture is finite for each fixed m, Preprint 2011.

[8] F. R. K. Chung, Open problems of Paul Erdős in graph theory, *J. Graph Theory* 25 (1997), 3–36.

[9] R. Diestel, *Graph Theory*, Springer-Verlag, New York, 2000.

[10] G. A. Dirac, A property of 4-chromatic graphs and some remarks on critical graphs, *J. Lond. Math. Soc.* 27 (1952), 85–92.

[11] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* 53 (1947), 292–294.

[12] P. Frankl, Cops and robbers in graphs with large girth and Cayley graphs, *Discrete Appl. Math.* 17 (1987), 301–305.

[13] J. A. Gallian, A dynamic survey of graph labeling, *Electr. J. Comb.* **DS6** (2002).

[14] R. K. Guy, The Decline and Fall of Zarankiewicz's Theorem, Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968), Academic Press, New York, 1969, pp. 63–69.

[15] H. Hadwiger, Über eine Klassifikation der Streckenkomplexe, *Vierteljschr. Naturforsch. Ges. Zürich* 88 (1943), 133–143.

[16] R. Hamming, *You and Your Research*. Accessed December 14, 2011. http://www.cs.virginia.edu/~robins/YouAndYourResearch.html.

[17] S. Hedetniemi, Homomorphisms of graphs and automata, Technical Report 03105-44-T, 1966.

[18] C. Hierholzer, Ueber die Möglichkeit, einen Linienzug ohne Wiederholung und ohne Unterbrechung zu umfahren, *Mathematische Annalen* 6 (1873), 30–32.

[19] F. Jaeger, Flows and generalized coloring theorems in graphs, *J. Combin. Theory Ser. B* 26 (1979), 205–216.

[20] F. Jaeger, A survey of the cycle double cover conjecture, *Ann. Discrete Math.* 27 (1985), 1–12.

[21] K. Kawarabayashi, B. Toft, Any 7-chromatic graph has K_7 or $K_{4,4}$ as a minor, *Combinatorica* 25 (2005), 327–353.

[22] P. J. Kelly, A congruence theorem for trees, *Pacific J. Math.* 7 (1957), 961–968.

[23] D. J. Kleitman, The crossing number of $K_{5,n}$, *J. Combinatorial Theory* 9 (1970), 315–323.

[24] Open Problem Garden. Accessed August 23, 2011. http://garden.irmacs.sfu.ca/.

[25] S. P. Radziszowski, Small Ramsey numbers, *Electr. J. Comb.*, Dynamic Survey **DS1**.

[26] F. P. Ramsey, On a problem of formal logic, *Proc. Lond. Math. Soc.* 30 (1930), 264–286.

[27] G. Ringel, problem 25, In: *Theory of Graphs and Its Applications, Proc. Symp. Smolenice*, 1963, G/A Czech. Acad. Sci. 162 (1964).

[28] N. Robertson, P. Seymour, R. Thomas, Hadwiger's conjecture for K_6-free graphs, *Combinatorica* 13 (1993), 279–361.

[29] N. Sauer, Hedetniemi's conjecture: a survey, *Discrete Math.* 229 (2001), 261–292.

[30] P. D. Seymour, Sums of circuits, In: *Graph Theory and Related Topics*, J. A. Bondy and U. S. R. Murty (eds.), Academic Press, New York (1979), pp. 341–355.

[31] P. D. Seymour, Nowhere-zero 6-flows, *J. Combin. Theory Ser. B* 30 (1981), 130–135.

[32] J. H. Spencer, Ramsey's theorem—a new lower bound, *J. Combin. Theory Ser. A* 18 (1975), 108–115.

[33] R. Steinberg, Tutte's 5-flow conjecture for the projective plane, *J. Graph Theory* 8 (1984), 277–289.

[34] G. Szekeres, Polyhedral decomposition of cubic graphs, *Bull. Aust. Math. Soc.* 8 (1973), 367–387.

[35] C. Tardif, Hedetniemi's conjecture, 40 years later, *Graph Theory Notes N.Y.*, 54 (2008), 46–57.

[36] A. Thomason, An upper bound for some Ramsey numbers, *J. Graph Theory* 12 (1988), 509–517.

[37] W. T. Tutte, A contribution to the theory of chromatic polynomials, *Canad. J. Math.* 6 (1954), 80–91.

[38] S. M. Ulam, *A Collection of Mathematical Problems*, Wiley, New York, 1960.

[39] V. G. Vizing, Some unsolved problems in graph theory, *Uspehi Mat. Nauk* 23 (1968), 117–134.

[40] K. Wagner, Über eine Eigenschaft der ebenen Komplexe, *Math. Ann.* 114 (1937), 570–590.

[41] D. B. West, *Introduction to Graph Theory*, 2nd edition, Prentice Hall, 2001.

[42] Wikipedia: Glossary of Graph Theory, Accessed December 14, 2011. http://en.wikipedia.org/wiki/Glossary of graph theory.

[43] D. R. Woodall, Cyclic order graphs and Zarankiewicz's crossing number conjecture, *J. Graph Theory* 17 (1993), 657–671.

[44] K. Zarankiewicz, On a problem of P. Turán concerning graphs, *Fund. Math.* 41 (1954), 137–145.

[45] C. Q. Zhang, *Integer Flows and Cycle Covers of Graphs*, Marcel Dekker Inc., New York, 1997.

[46] X. Zhu, A survey on Hedetniemi's conjecture, *Taiwanese J. Math.* 2 (1998), 1–24.

[47] L. Beineke, R. Wilson, The early history of the brick factory problem, *Math. Int.* 32 (2010), no. 2, 41–48.

The proof is in the pudding.

Opening a copy of The Mathematical Intelligencer *you may ask yourself uneasily, "What is this anyway—a mathematical journal, or what?" Or you may ask, "Where am I?" Or even "Who am I?" This sense of disorientation is at its most acute when you open to Colin Adams's column.*

Relax. Breathe regularly. It's mathematical, it's a humor column, and it may even be harmless.

➢ Column editor's address: **Colin Adams,**
Department of Mathematics,
Bronfman Science Center,
Williams College,
Williamstown, MA 01267, USA
e-mail: Colin.C.Adams@williams.edu

The Dog Who Knew Calculus

Colin Adams

Hello, class. I want to welcome you to Math 106 Multivariable Calculus. My name is Elvis. That's right. Just Elvis. No last name. Just call me Dr. Elvis.

Many of you may never have had a dog as a professor before. But I do not want there to be any concern about my qualifications. Just because I am covered in fur and have paws instead of hands and feet does not mean that I am in any way deficient as a professor. I am as qualified as any of the faculty on this campus.

Maybe more qualified. Most of the math faculty here learned calculus the standard way by taking courses. I learned calculus the hard way, in the freezing waters of Lake Michigan. My owner used to stand a short distance from the straight shoreline and throw a ball in the water, aimed at an angle to the shore. My goal, a natural one for a canine, was to minimize the time it took me to reach the ball. Given the speed I could run on the shore and the speed I could swim, the fundamental question was where along the shoreline I should enter the water to minimize the time it took for me to reach the ball. The answer to that question is determined by calculus. As my owner demonstrated in his paper ["Do dogs know calculus?" Timothy Pennings, *The College Mathematic Journal*, vol. 34 (May 2003), no. 3, 178-182], I am quite capable of performing the necessary calculations on the fly, as I am racing toward the ball.

Of course, just being able to solve that one problem does not qualify me to teach calculus. But in fact, I could solve various related-rates problems involving squirrels, I could calculate optimal leaps to minimize the time it takes for a thrown doggie biscuit to reach my mouth, and I could even determine optimal paths for pursuit problems, as for instance might occur when chasing a cat.

So I certainly have the background to teach this course. Given the current budget climate, the university was more than happy to send me to the appropriate "English as a Second Language" classes and then hire me as a calculus instructor. So here I am.

And although it's true I am only six human years old, that translates into 42 dog years. So believe me, I am experienced. In fact, I am already too old to win the doggie equivalent of the Fields medal.

As for my other credentials, my rabies vaccinations are up to date and I have the tag to prove it. So if I ever look rabid in class, it's probably just frustration with your exam results.

DOI 10.1007/s00283-011-9243-2

Now, I hope very much that we can become friends this semester, perhaps best friends. However, I do have a few expectations for your behavior in the classroom.

Number one. Do not scratch my rear end or suggest you would like to, even if I arch my rear end in your direction.

Number two. Do not try to get my attention by saying "Here, boy," or "Who's a good doggie?" I expect to be treated with the same respect you afford the noncanine faculty at this university.

Number three. Do not eat in class. I find it particularly distracting.

Number four. Once in a while, I will lick my nether regions. This is neither intended as a subtle nor not-so-subtle form of sexual harassment. It is just something I need to do. We have had extensive discussions about this within the department and ultimately with the administration, and there is now an understanding that this behavior is an integral part of the cultural milieu within which I grew up and as such, the university community must accept it. So I have special permission from the Dean of the Faculty to do so. That does not, however, mean that you also have such permission.

As many of you already know, TA sessions are already up and running. For those of you who complained that the TA did not show up to the first session last night, I should mention that the TA is a dachshund. Next time, please look under the table at the front of the room.

During the semester, there will be two exams and a final. All exams are closed-book, and will occur down at the river. I recommend bringing a swimsuit.

Okay, now I would like everyone to line up at the front of the class. I need to learn names, so it will be necessary for me to smell each one of you, as you say your name. Then we can get started.

Forcing Nonperiodicity with a Single Tile

JOSHUA E. S. SOCOLAR AND JOAN M. TAYLOR

It is easy to create nonperiodic tesselations of the plane composed of one or a few types of tiles. In most cases, however, the tiles employed can also be used to create simpler, periodic patterns. It is much more difficult to find shapes, or "prototiles," that can fill space *only* by making a nonperiodic structure. We say that such sets are *aperiodic*, or that they "force" nonperiodicity, and there are many open questions about what types of structure can be forced and the prototiles required. In this article we discuss recent progress on the fundamental problem of forcing nonperiodicity using a *single* prototile, jokingly called an *einstein* (a German pun on "one stone"). A new example we found [6] shows one way in which an einstein can work and highlights several issues that arise in posing the problem precisely.

One motivating factor in the search for an einstein comes from condensed matter physics. Local rules for how tiles fit together may represent the energetics of a physical system, which could support self-assembly into an ordered but nonperiodic structure. The discovery of icosahedral and decagonal phases of metallic alloys, in which the atomic structure shares the essential structure of the Penrose tilings, has opened our eyes to the fact that nonperiodic materials can indeed form spontaneously [1, 2]. In materials physics applications, where the tiles may represent clusters of many atoms or larger building blocks, the tiles can have complex shapes or markings that determine how they may be joined. Finding a single shape that can do the job may make the physical realization of such a material easier.

The first example showing that it is possible to force nonperiodicity was Berger's set of 20,426 distinct prototiles [3]. Aperiodic sets with just two prototiles were subsequently discovered, the most famous being the Penrose tiles [4], nicely described by Martin Gardner [5]. Candidates with einstein–like features have been presented before, but there is no precise definition of the einstein problem, and several candidates that could be argued to qualify have not passed the consensus "I know it when I see it" test. There are several issues involved, including the specification of what counts as nonperiodic, what characteristics make for a valid prototile, and what form the local rules must take.

We recently showed that the prototile in Figure 1 is an einstein and determined a number of remarkable properties of the tilings it forces. [6] contains two proofs of the forced nonperiodicity along with derivations of several intriguing properties of the tiling (including a surprising connection to the regular paperfolding sequence [7]). In working out the properties of the forced limit–periodic structure and searching for different ways of encoding the information about how the tiles must fit together, we were led to a series of questions about how the einstein problem should be posed. In the present paper, we discuss the definitions of the terms *local matching rules* and *tile*, and we propose a new definition of *nonperiodic* that emphasizes distinctions that to our knowledge have not been made explicit before. We use our hexagonal prototile throughout to clarify key points, including a 3D version for which the shape alone is sufficient to force nonperiodicity. We also describe some of the more intriguing aspects of the forced structure, some of which require further study. Articulating the criteria satisfied by our hexagonal prototile reveals the sense in which it is the "best" einstein currently known and delineates a precise problem that remains open.

The newly discovered tile, shown together with its mirror image in Figure 1(a), is a regular hexagon with markings that determine how neighboring tiles must be oriented. Adjacent

DOI 10.1007/s00283-011-9255-y

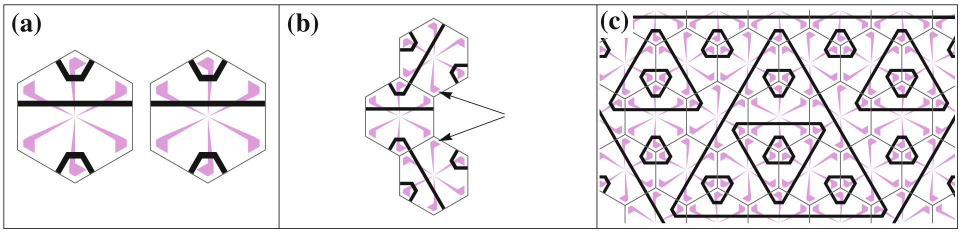


Figure 1. The hexagonal prototile and its mirror image with color matching rules. (a) The two tiles are related by reflection about a vertical line. (b) Adjacent tiles must form continuous black stripes. Flag decorations at opposite ends of a tile edge, such as the indicated flags at opposite ends of the vertical edge, must point in the same direction. (c) A portion of an infinite tiling that respects the matching rules.

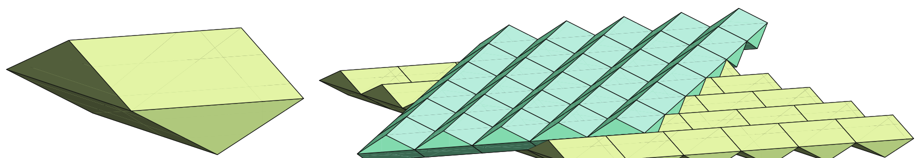


Figure 2. The SCD prototile and the space–filling tiling it forces.

tiles must form continuous black stripes, and flag decorations at opposite ends of each tile edge must point in the same direction. (The arrows in (b) point to the two flags at opposite ends of a vertical tile edge.) Each tile in (c) is a rotation and/or reflection of the single prototile, and the only way to fill space while obeying the rules everywhere is to form a nonperiodic, hierarchical extension of the pattern in (c).

Defining the einstein

Two constructions that could conceivably be counted as einsteins were discovered in 1995. A single prototile that forces a pattern of the Penrose type was presented by Gummelt (with a complementary proof by Steinhardt and Jeong) [8, 9]. But in this case tiles are allowed to overlap and the covering of the space is not uniform. For this reason the prototile is not considered to be an einstein.

The uniformly space-filling, three-dimensional prototile of Figure 2, a rhombic biprism, was exhibited by Schmitt, Conway, and Danzer [10]. To fill space, one is forced to construct 2D periodic layers of tiles sharing triangular faces, with ridges running in the direction of one pair of rhombus edges on top and the other pair below. The layers are then stacked such that each is rotated by an angle ϕ with respect to the one below it, where ϕ is the acute angle of the rhombic base. Any choice of ϕ other than integer multiples of $\pi/3$ or $\pi/4$ produces a tiling that is not periodic, and certain choices permit a tiling in which the number of nearest neighbor environments is finite, so that the prototile can be endowed with bumps and nicks in a way that locks the relative positions of adjacent layers.

Again, however, the universal reaction was "This is not really what we are looking for." The nonperiodicity of the

AUTHORS

JOSHUA E. S. SOCOLAR received his Ph.D. in Physics from the University of Pennsylvania in 1987, with a thesis on quasilattices and quasicrystals. He has been on the faculty at Duke since 1992, and he is affiliated with the Center for Nonlinear and Complex Systems and the Duke Center for Systems Biology. His hobbies include jazz piano, singing, and word games.

Physics Department and Center for
Nonlinear and Complex Systems
Duke University
Durham, NC 27708
USA
e-mail: socolar@phy.duke.edu

JOAN M. TAYLOR took up mathematics in 1991 at age 34 after being inspired by a magazine article on quasicrystals featuring Penrose's rhombus tiling. She began but did not complete a degree, preferring to conduct her own research. Since then she has pursued tiling and related topics in abstract algebra and number theory including original work on constructible polygons. She likes to unwind with knitting and reading.

Post Office Box U91
Burnie, TAS 7320
Australia

tiling does not seem mysterious enough to count; one can immediately grasp the global structure of simple 2D periodic lattices stacked with a twist in the third dimension. We seek structures with long–range correlations that are not immediately evident from the examination of a single tile. Goodman–Strauss emphasizes this point and suggests a classification that distinguishes the SCD structure from the Robinson tiling and Penrose tiling [11]. Goodman–Strauss calls the SCD tile *weakly aperiodic* because it admits a tiling with a cyclic group of symmetries involving finite (and non-zero) translations, in this case the screw operations along the twist axis. We emphasize here an additional feature of the SCD tiling that also weakens the sense in which it may be called nonperiodic: every individual tile in the tiling is a unit cell of a periodic 2D layer. In fact, for the cases with a finite number of nearest neighbor environments, any finite stack of layers is periodic in the two transverse directions, a point that may be important for physical applications. The presence of infinite periodic substructures within the pattern suggests that a formal definition be developed to distinguish different degrees of periodicity or nonperiodicity.

A definition of "nonperiodic"

We offer here a new classification scheme, based on the notion of a "partial translational symmetry," that we believe captures the shared intuitive notion of a nontrivial, nonrepeating pattern. A partial translational symmetry is an operation that maps some subset of a full pattern into itself.

Definition 1 *Let $\mathcal{T}$ be a infinite set of tiles in $\mathcal{R}^N$. A* partial translational symmetry *of $\mathcal{T}$ is an operation of the form $\{\vec{x} \to \mathbf{R} \cdot \vec{x} + \vec{e}\}$ that acts on some infinite subset of $\mathcal{T}$ and leaves it invariant, where $\mathbf{R}$ is a rotation matrix and $\vec{e} \in \mathcal{R}^N$ is a constant, nonzero, displacement vector. The magnitude of $\vec{e}$ is the* spacing *of the partial translational symmetry.*

We say that a tile *participates* in a partial translational symmetry if it is a member of the subset of tiles that is left invariant by the symmetry operation. The tiles participating in a given partial translational symmetry need not form a connected region.

The SCD tiling has many partial translational symmetries. The tiles in any one layer form a subset that is invariant under a 2D lattice of translations (for which $\mathbf{R}$ is the identity). In addition, every tile is an element of a subset that is invariant under a screw operation ($\mathbf{R}$ being a rotation about the stacking axis) that maps one layer into the next.

Another simple example of a system with nontrivial partial translational symmetries will help clarify the sense in which our new tiling is nonperiodic. Figure 3 shows a 1D tiling in which tiles have unit width and the number on a tile indicates its type. The tiling extends infinitely in both directions and the spacing between the nearest tiles of the same type is just twice their numerical value; type x occurs periodically at positions $x(1 + 2n)$ for integer n. (The blank element at the center is not repeated at any finite distance.) Thus, for each value y occurring in the sequence, the subset of tiles with values less than or equal to y is invariant under translation by $2y$. For example, under translation by 8, the set of all tiles with values less than or equal to 4 is invariant, though the set of remaining tiles is not.

Definition 2 *The elements of a set of partial translational symmetries are* independent *if and only if it is not possible to express the displacement $\vec{e}$ for any one of them as an integer linear combination of the others.*

In an ordinary N-dimensional periodic tiling, every tile participates in N independent partial translational symmetries, the displacements $\vec{e}_i$, with $i = 1, \ldots N$, being the basis vectors of the lattice of translations that leave the tiling invariant.

Definition 3 *Let the number of independent partial translational symmetries with spacing less than r that a given tile t participates in be denoted $S_t(r)$. A tiling in $\mathcal{R}^N$ is* nonperiodic *if for any finite (non-zero) r, the fraction of tiles with $S_t(r)$ strictly less than N is finite.*

By this definition, the decagonal quasicrystal structures [12], which are periodic stackings of quasicrystalline layers, are nonperiodic because the only independent partial translational symmetry containing any given tile is the one corresponding to the periodic stacking direction. (There could conceivably be additional screw operations if the tiling has an axis of complete 5-fold rotational symmetry, but they all have the same $\vec{e}$ or integer multiples of it.)

But the SCD tiling is *not* nonperiodic by this definition. As noted previously, every tile is a member of a subset that is invariant under two distinct translation operations in the plane and a subset that is invariant under a screw operation along the stacking direction. To emphasize the high degree of partial translational symmetry in the SCD structure, we might classify it as *heterogeneously periodic*; "heterogeneously" because, unlike familiar periodic structures, the $\vec{e}$'s associated with the symmetries of tiles in different layers are not all the same. Heterogeneously periodic is not synonymous with weakly nonperiodic; some weakly nonperiodic tilings, such as the stacked layers that form decagonal quasicrystals, are nonperiodic by our definition (and by common usage in the physics community).

Many tilings satisfy a more stringent criterion:

Definition 4 *A tiling is* maximally nonperiodic *if and only if it contains no partial translational symmetries.*

The Penrose tilings are maximally nonperiodic, as are many tilings generated by substitution rules or projections from higher-dimensional periodic lattices onto incommensurate subspaces.

...	16	1	2	1	4	1	2	1	8	1	2	1	4	1	2	1		1	2	1	4	1	2	1	8	1	2	1	4	1	2	1	16	...

Figure 3. A limit–periodic pattern with partial translation symmetries.

Following Grünbaum and Shephard, we adopt the following terminology.

DEFINITION 5 *A prototile is (maximally)* aperiodic *if the only space-filling tilings that can be composed from it are (maximally) nonperiodic.*

That is, we call a single prototile "aperiodic" if it can be used to tile an entire space with no overlaps but only in a pattern that is nonperiodic by the above definition.

The question of allowable matching rules

Whether or not an aperiodic prototile exists could hinge on whether one requires that the tile be a simply connected domain, whether tiles are allowed to overlap so that some parts of space are doubly covered, whether the rules must be encoded by tile shape alone as opposed to color-matching rules, and whether mirror-image sets are considered to count as a single prototile or not. A highly restrictive definition of an einstein would demand the following properties:

Rotations only: Reflections of a chiral tile are *not* allowed;
Simply connected tiles: The prototile is a simply connected domain (a topological disk in 2D);
Shape alone: All configurations of tiles that do not contain overlaps are permitted without regard to any colored markings;

The first condition makes what some may view as an arbitrary distinction between rotations and reflections. Nevertheless, we see a meaningful distinction between cases where the tiles could all be manufactured from a single physical mold and cases where a second, mirror-image mold must be built.

The restriction to simply connected prototiles is consistent with the intuitive notion of a tile as a thin, rigid piece of material, as is used in mosaics or floor tiling. On the other hand, there is no obvious reason to insist that a tile cannot be composed of a set of disconnected domains with fixed relative positions [13], and in fact certain types of color matching rules that cannot be enforced by the shape alone using a simply connected 2D prototile can be enforced using a tile consisting of disconnected pieces [14]. Grünbaum and Shephard make a further distinction between tiles with cutpoints (where regions are connected only through a vertex) and tiles with entirely disconnected regions [13]. From a materials physics perspective, tiles may represent complex atomic configurations with low energy, and these may conceivably interpenetrate in ways that could not be represented by simply connected tiles.

The "shape alone" condition requires further comment, as there are several kinds of rules that cannot be encoded in the geometry of the prototile.

Colors required: Instead of bumps and nicks, the rules that force relative orientations of nearby tiles can be encoded as a colored decoration of the prototiles together with rules about how colors must match. Not every rule enforced by color matching can be implemented through shape alone without increasing the number of prototiles. A classic example is Ammann's aperiodic set *A*5 (a square and a 45° rhombus), where rules for how tiles must join at vertices may be implemented either through constraints on colored decorations around the vertex or by introduction of a new tile that must fit at each vertex [13]. Another example is the hexagonal parquet tile of [14], for which the color rules for tile edges (either red or black can match black, but red cannot match red) could be implemented by introducing two new tiles that fit into notched edges.
Non–adjacent, but pairwise: A rule may specify the relative orientations of two tiles separated by some bounded distance but not sharing an edge. In such a case, it is still possible to check whether the tiling satisfies the rules by examining only two tiles at a time, or, as physicists would say, by considering only pairwise interactions between tiles.
Configuration atlases: The set of allowable configurations may be expressed as an atlas of allowed configurations within some ball of finite radius, but not be expressible as a set of pairwise constraints. Examples include the trivial cases presented by Goodman-Strauss in which rectangular tiles are required to form pixellated versions of Robinson square tiles [11] as well as the recent construction of Fletcher in which face-matching rules for an aperiodic set of 21 cubic prototiles are expressed as an atlas of allowable configurations for a single cubic prototile in which the 21 different tile types are encoded as 21 different orientations of a single tile [15].

The einstein

The discovery of the prototile and rules of Figure 1 was initiated by Taylor's observation that a single colored hexagon together with its mirror image could force a structure similar to the one forced by a set of 12 tiles (discovered by Socolar and Goodman-Strauss) that appeared later on Socolar's web page. She had been searching since 1993 for a superposition of matching rules to force nonperiodicity on the hexagon with black stripes, viewing it as an elementary version of the Penrose rhombi, which form a quasiperiodic hierarchy of overlapping, irregular hexagons. Taylor's constructions were based heavily on a complex scheme for generating the tilings through a substitution procedure in which each tile is divided into smaller tiles that respect the same local rules [16]. A note from Taylor requesting feedback led Socolar to refine the set of necessary local rules and construct a simple proof of aperiodicity, which initiated an extended collaboration conducted entirely by email between Tasmania and North Carolina. The conceptual breakthroughs needed to resolve various subtle issues came about through repeated exchange of figures and discussion of details specific to these tilings. There was no clearly generalizable strategy involved, though we hope that our results will lead by example to further discoveries.

The prototile and its mirror image are shown in Figure 1 as regular hexagons decorated with colors that encode rules constraining the relative orientation of nearby tiles. There are two such constraints, or *matching rules*: (**R1**) the black stripes must be continuous across all edges in the tiling; and (**R2**) the

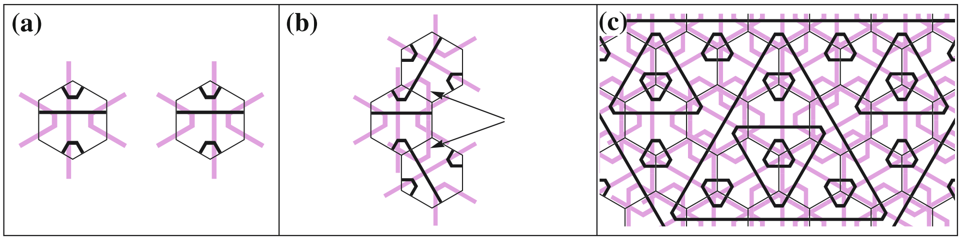


Figure 4. Alternative coloring of the 2D tiles. The arrows indicate stripes on next nearest neighbor tiles that must join to form a continuous line.

flags at the vertices of two tiles separated by a single tile edge must always point in the same direction. The rules are illustrated in Figure 1(b), and a portion of a tiling satisfying the rules is shown in Figure 1(c).[1]

An alternative way to represent the matching rules is to allow decorations that extend beyond the tile edges as shown in Figure 4. **R1** remains the same and **R2** is now enforced by requiring that the purple stripes be continuous. This makes it clear that **R2** and **R1** have identical geometric forms related by a scale factor of $\sqrt{3}$ and a rotation by $\pi/2$. It is the relative positions of the long black and long purple stripe that distinguish the different reflections of the prototile. The purple stripes in Figure 4 form hierarchical triangular structures just like the black stripes, but there are three interpenetrating purple structures. (See Figure 5 that follows.)

The only space-filling tilings allowed by the 2D prototile of Figure 1 are nonperiodic. The proof given in [6] shows that the tiling forms an infinite hierarchy of interpenetrating honeycomb lattices of black rings, and the tiles in lattices with translational symmetry scales larger than r cannot participate in partial translational symmetries with spacing smaller than r. For any r, the density of tiles in larger scale lattices is clearly finite (nonzero), so the tiling is indeed nonperiodic by our definition. Thus, assuming that we allow color matching rules and count mirror images as a single prototile, we have an einstein that requires only pairwise matching rules!

The partial translational symmetries of the black ring structure are immediately clear, but because of the interplay between that and the purple stripes, the partial translational symmetries of the full pattern are more difficult to locate. Figure 5 displays a subset of them. The shaded tiles are a motif that is repeated periodically to form a triangular lattice. For visual clarity, the black stripe decoration is displayed for every tile and the purple stripe decoration of Figure 4 is shown only for a subset of the tiles that form a triangular lattice consisting of one third of all of the tiles. There is no partial translational symmetry in this tiling with a smaller spacing than the ones shown except for the special cases where partial translational symmetries occur along one infinite line in the tiling. The proof involves analysis of the separate symmetries of the black stripe and purple stripe patterns. We omit it here because it is not terribly illuminating and we have not yet solved the general problem of which tiles participate in partial translational symmetries with given spacings.

The matching rules **R1** and **R2** may appear to be unenforceable by shape alone. **R2** necessarily refers to tiles that are not in contact in the tiling and **R1** cannot be implemented using only the shape of a single prototile and its mirror image. Both of these obstacles can be overcome, however, if one relaxes the restriction that the prototile must be a simply connected shape. Figure 6(a) shows how the color–matching rules can be encoded in the shape of a single prototile that consists of several disconnected regions. In the figure, all regions of the same color are considered to comprise a single tile. **R1** is enforced by the small rectangles along the tile edges. **R2** is enforced by the pairs of larger rectangles located radially outward from each vertex. The flag orientations are encoded in the chirality of these pairs. Thus we have an einstein that does not require color matching rules! Figure 6(b) shows a deformation of the disconnected prototile to a prototile with cutpoints; that is, a tile in which all the pieces are connected through vertices and tiles are allowed to overlap at those points. For a beautiful rendering of this construction, see Araki's beetles [18].

Whether you prefer to enforce the matching rules using colors or a disconnected prototile is a matter of taste. Of course you may find both less than fully satisfying, in which case we can offer a third way out—via escape to the third dimension. The tiles of Figure 1 are related by reflection through a line in the 2D plane, but they can also be thought of as related by a rotation in 3D space of 180° about that same line, suggesting that the two mirror-image tiles be thought of as the front and back faces of a single 3D tile. Such a tile is shown in Figure 7.

The colored bars running through the 3D tile are guides to the eye that display the black and purple stripe structure, but they are not required. The continuity of the bars is enforced by the shape of the tile alone. To see how, consider first the flag matching rule **R2**. To enforce this rule, we must have arms extending outward from the basic hexagonal prism to meet with the arms of next-nearest-neighbor hexagons. At each vertex of the hexagonal tiles, three arms must somehow pass through each other. The tile shown in Figure 7 solves this problem by allowing tiles to be staggered at three different heights. The full tiling is divided into three triangular lattices of tiles, each of which contains tiles at one height. The top faces of the tiles in the three different lattices are at heights 0, $h/3$, and $2h/3$, where h is the height (or thickness) of a tile.

[1] We note that this tiling is similar in many respects to a tiling exhibited previously by Penrose [17], but the two are not equivalent [6].

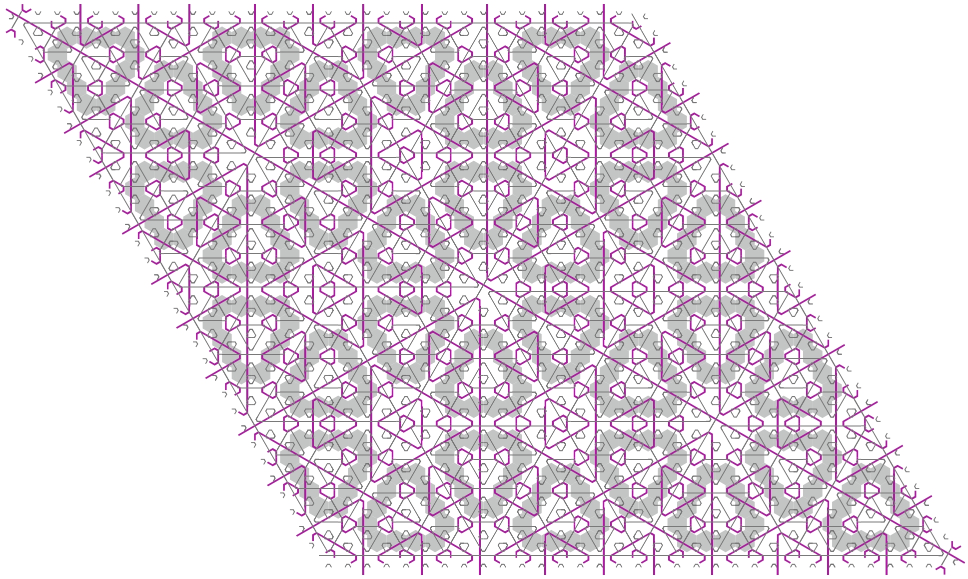


Figure 5. The partial translational symmetry with the smallest spacing. Clusters of 24 shaded tiles (two of each of the twelve tile orientations) are repeated throughout the tiling, forming a triangular lattice. Purple stripes are shown only for a subset of one third of the tiles.

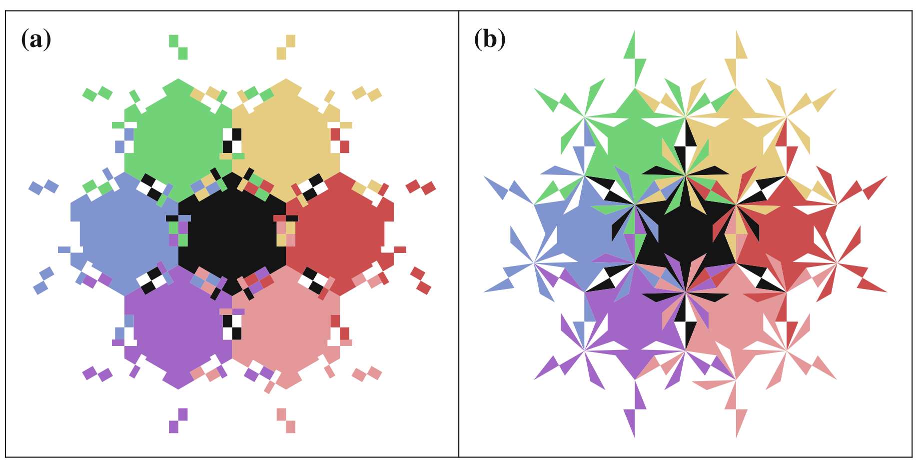


Figure 6. (a) Enforcing by the shape alone with a disconnected 2D tile. All the patches of a single color, taken together, form a single tile. (b) A deformation of the disconnected prototile in (a) to a prototile with cutpoints.

The hexagonal blocks on each arm have thickness $h/3$, allowing the blocks from three crossing arms to make a full column. The six arms on the prototile have outer faces that are tilted from the vertical in a pattern that encodes the chirality of the flags of the 2D tile. Forming one triangular lattice requires that bevels of opposite type be joined, and hence that flags of opposite chirality match in accordance with **R2**.

The small bumps on the tiles and the holes in the arms are arranged such that adjacent tiles can fit together if and only if the black stripes match up properly, as required by **R1**. The three square holes in each arm are positioned so that projections from the faces on neighboring tiles can meet with each other. The holes are all the same; they do not themselves encode the positions of the black stripes. Next, we create two types of plug that can be inserted into a hole. One type consists of two square projections that fill opposite quadrants of the hole; the other type fills the entire hole but only to half its depth. The two types are both invariant under rotation by 180°. Two plugs of the same type can fit together to fill a hole, but plugs of different types cannot. Finally, we place two columns of three plugs each on each of the large vertical faces of the main hexagonal portion of the tile. Each column aligned with a black stripe has plugs of one type, and the other columns have plugs of the other type. (The latter are needed to fill the holes in the arms at those positions.) Three plugs are needed because of the staggered heights of neighboring tiles. If a prototile that is a topological sphere is desired, the plugs can be moved toward the middle of their respective faces so that the left and right side plugs meet and the holes in the arms are converted to U-shaped slots.

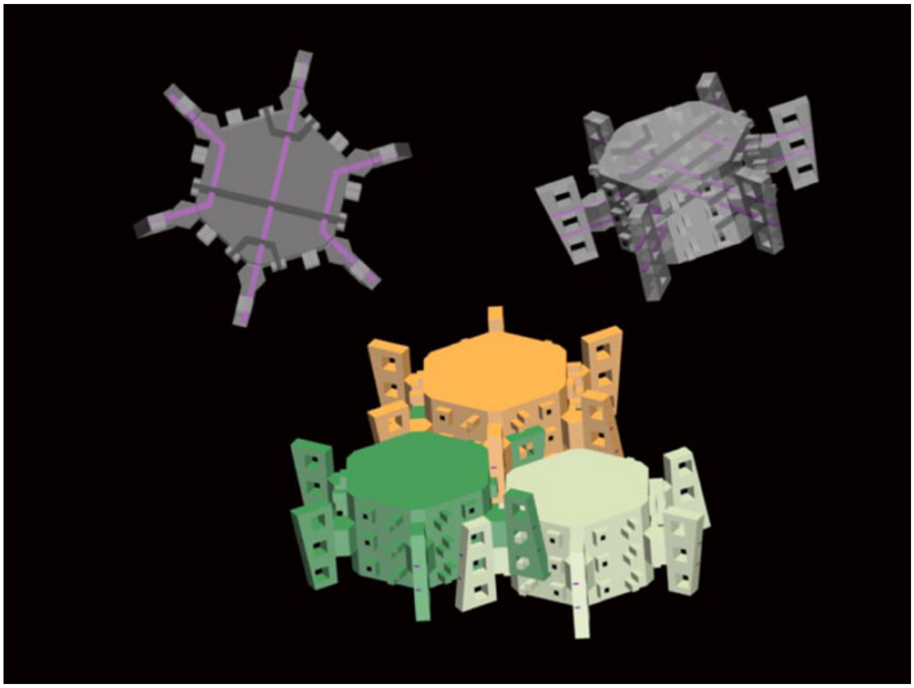


Figure 7. A 3D einstein. Colored bars are included only to clarify the relation of the 3D shape to the matching rules for the 2D tile. A translucent tile is shown in two orientations to emphasize the relation of the colored bars to the shape. Three solid tiles that fit together properly are shown (with the ends of the colored bars visible).

To fill 3D space, the staggered layer can be stacked. Note that in its current form the tiles in a single stacked column do not have to have identical orientations; the 2D tilings constituting successive staggered layers need not coincide, though each must be a version of the nonperiodic tiling. If desired, a bump could be placed on the segments on the top face of the tile directly over the point where purple and black stripes cross, with a matching indentation on the bottom face, so the tiling would be unique (periodic in the stacking direction).

Thus we see that matching rules equivalent to those of the 2D tile can be enforced by the shape of a simply connected *three-dimensional* prototile. The space-filling tiling forced purely by the shape of this tile consists of a corrugated slab isomorphic to the structure forced by the 2D tiles, as shown in Figure 8, which may be stacked periodically to fill the 3D space. This forced structure means that the tile satisfies a rather strict definition of an einstein—the strictest definition currently known to be satisfiable. Though the periodicity in the third dimension makes this a weakly nonperiodic tiling by Goodman-Strauss's definition, it has a very different character from the SCD type of weak nonperiodicity. In particular, the 3D tiling does satisfy our definition of nonperiodic, which indicates that the structure has complex correlations over large scales.

Our 3D prototile, like the SCD prototile, is not isomorphic to its reflection. It is not possible, however, to construct a tiling that contains a mixture of the two enantiomorphs. The shapes of the plugs enforcing the black stripe rule do not allow placing a left–handed tile adjacent to a right–handed one. Thus the prototile is aperiodic even if one does not

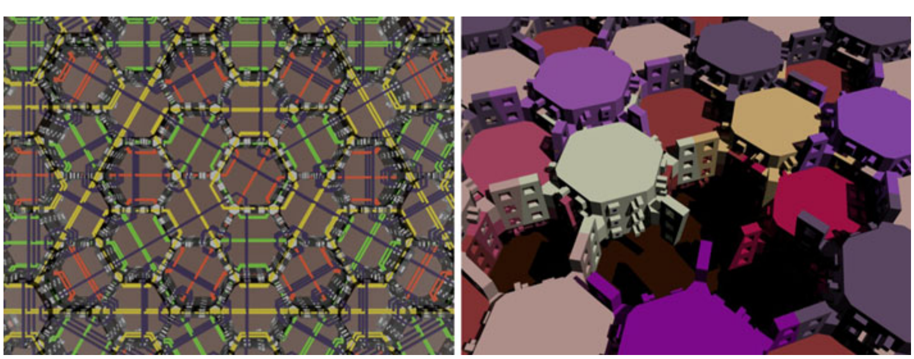


Figure 8. Two renderings of portions of the 3D tiling. For visual clarity, the purple stripes on translucent tiles at different heights are rendered in different colors.

explicitly prohibit reflections. For the SCD rhombic biprism, which is also chiral, reflections allow the construction of a periodic tiling with layers that alternately twist clockwise and counterclockwise. In order to prevent mixing of the two enantiomorphs, one may decorate the SCD tile with chiral plugs, though the prototile then loses the appealing property of convexity.

The einstein pattern

The fundamental structure of the tiling is visually evident in the patterns of black rings and purple rings in Figure 5. The black rings form truncated triangles with side lengths related by powers of two. Any set of triangles of the same size forms a periodic honeycomb pattern equivalent to that of the smallest ones. As expected from the similarity of the black and purple decorations of Figure 4, the purple rings form exactly the same pattern, rotated by $\pi/2$ and scaled up by a factor of $\sqrt{3}$.

Two proofs of aperiodicity are given in [6]. The first begins with an inspection of the possible ways of surrounding a given tile, which quickly reveals that a subset consisting of 3/4 of all the tiles in the plane must be arranged to form the honeycomb of smallest black rings. One then shows that the markings of those tiles induce precisely the same set of rules applied to the remaining tiles, so that 3/4 of those will have to form the truncated vertices of the next largest honeycomb of black triangles. Iterating the reasoning implies that there is no largest honeycomb, so that for any finite r, there will be a finite density of tiles that do not participate in partial translational symmetries with spacings smaller than r.

The second proof makes use of the invariance of the tiling under inflation, a procedure in which tiles are grouped into larger tiles that obey the same matching rules on the larger scale. By identifying seven distinct local environments of each chiral tile type (fourteen environments altogether) and assigning central tiles in them, labels A through G and $\bar{A}$ through $\bar{G}$, it is possible to obtain the tiling from an iterated substitution rule, as shown in Figure 9. The scale factor associated with the substitution rule is 2, which implies that the tiling is limit-periodic (rather than quasiperiodic). [19, 20] A proof that the pattern of tile types can be enforced by a single prototile is given in [6].

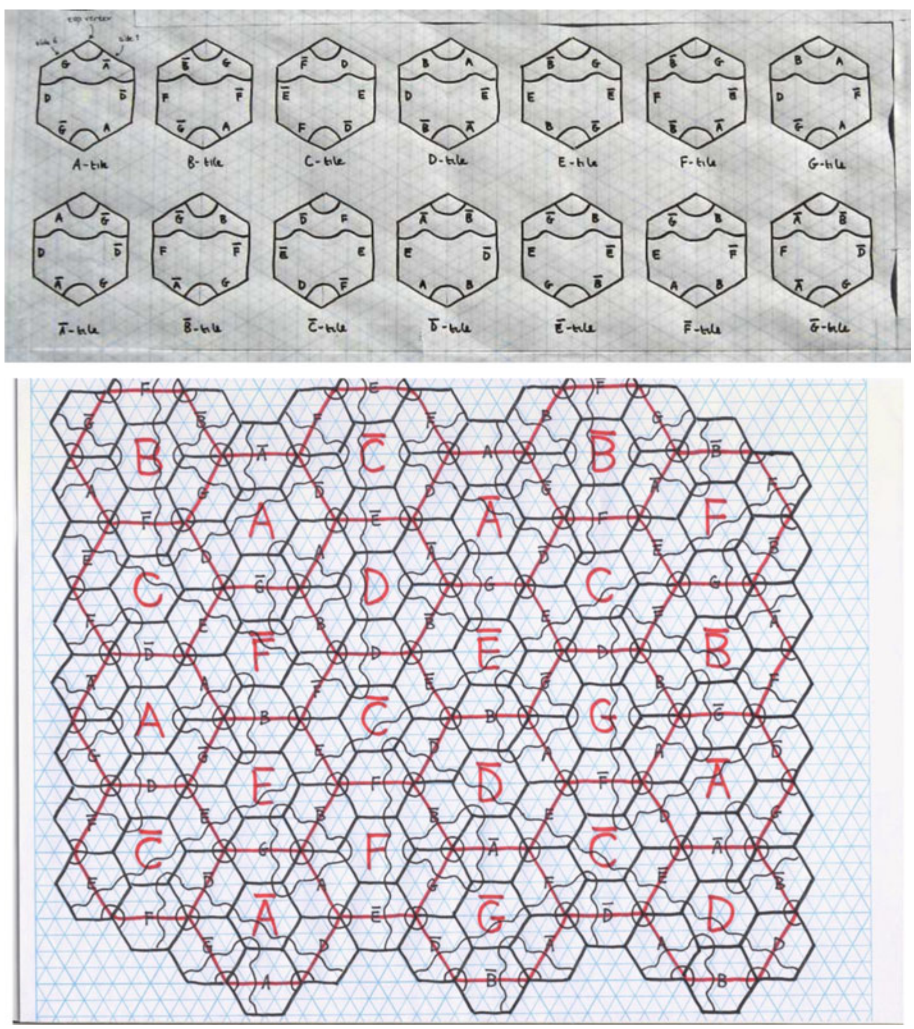


Figure 9. Top: The fourteen-tile set. Bottom: Illustration of the substitution rule. The label of the central black hexagon within each red hexagon is C if the red label is unbarred and $\bar{C}$ if the red label is barred.

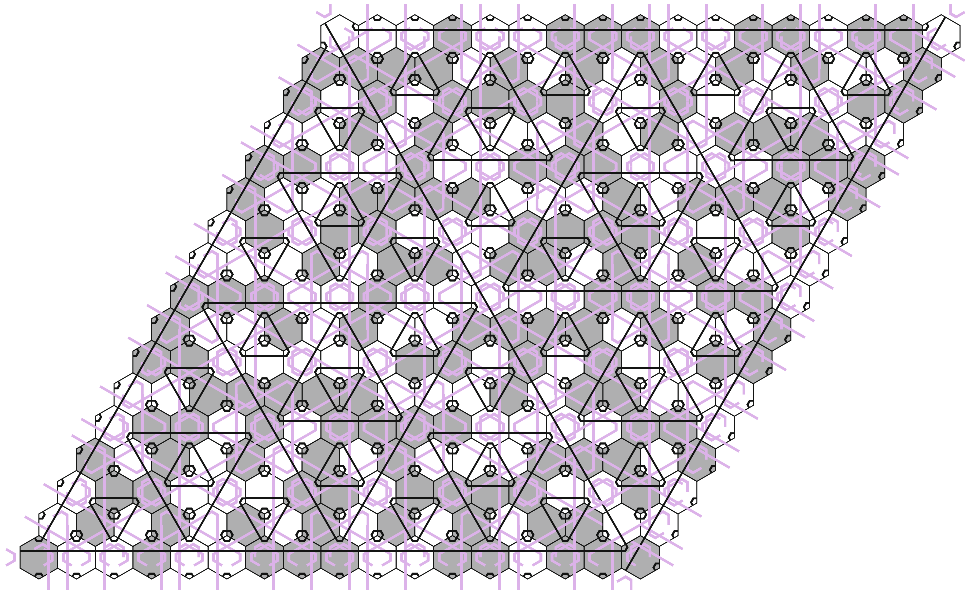


Figure 10. Another coloring of the forced 2D tiling. The purple stripes have been moved outward from the center of each tile to make it easier to see the purple triangles, and the handedness of each tile is encoded in the gray and white tile colors.

A curious feature of the set of forced tilings is that there is a particular arrangement of three tiles around a vertex for which the entire tiling is uniquely determined; that is, there is a local configuration that has a unique extension to the whole plane [6]. This may be surprising, as one might expect the uniqueness of the extension to imply that the tiling must be periodic. Almost every finite patch that appears in a complete tiling appears an infinite number of times and permits an infinite number of distinct extensions to the entire plane. There is, however, one particular tiling (plus its mirror image) that contains a single threefold symmetric vertex that does not appear in any of the other tilings. The situation is analogous to having a decapod defect at the center of a Penrose tiling [5], but the "defect" in the present case does not violate the matching rules in its interior.

A good visualization of the complexity of the tiling is obtained by shading the two mirror images differently, as shown in Figure 10. Figure 11 shows a larger portion of the tiling with one tile type shaded light grey and the mirror image tile shaded dark grey. We have noticed a curious feature of this pattern. There are islands of 13 dark (or light) tiles that are surrounded completely by light (or dark) tiles. We

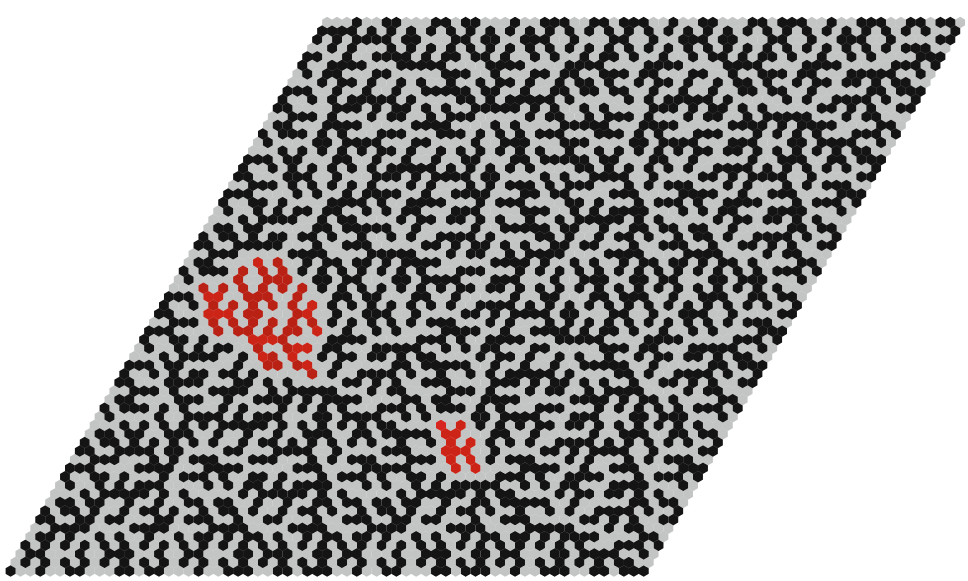


Figure 11. The pattern of mirror image tiles. Red tiles show the islands mentioned in the text.

refer to these as "llamas." (See Figure 11.) Islands of 63 tiles can also be seen in the figure. These are obtained from llamas and a few nearby tiles by application of the substitution rule. We have also found islands of 242 tiles formed by a second iteration of the substitution rule. Those islands each surround one llama, so the total size of the patch is 255 tiles. We have not determined, however, whether islands of arbitrarily large size exist or whether the fraction of tiles that are *not* in an island of some finite size is nonzero.

The parity pattern can be specified completely as a function of tile locations with a closed-form expression [6]. An unexpected feature is the emergence of regular paperfolding sequences (A014577 of the Online Encyclopedia of Integer Sequences). Inspection of the substitution rules for hexagons along certain rays shows that they recapitulate precisely the iterative rule that produces the regular paperfolding sequence. This connection suggests that the full 2D pattern exhibits a rich algebraic structure that holds additional surprises and possibly affords a new window onto the properties of paperfolding and related sequences.

Closing remarks

We have exhibited a tile that lies in a distinct new class—a single tile that forces nonperiodicity in a space-filling tiling—and we have presented a supporting classification scheme that captures certain intuitive distinctions between classes of nonperiodic tilings.

If mirror image tiles are counted as equivalent to the original tile, and if disconnected tiles or tiles with cutpoints are allowed, we have a 2D tile that forces a nonperiodic tiling (in exactly the same sense that the Robinson tilings are nonperiodic). Our 3D construction gives the long-sought simply connected einstein with matching rules enforced by shape alone (and no mirror image tile required). The structure of our 3D aperiodic tile is somewhat complex and does not appear open to simplification, but two elements of the construction suggest new directions in the search for an einstein. First, we use the possibility of rotation in 3D to create a single tile that is equivalent to two different 2D tiles, mirror images in the present case. Second, we use the third dimension to implement rules that require either disconnected tiles or color matching in 2D.

The crucial point, in our view, is that our 3D prototile is the first known to force a nonperiodic structure that cannot be easily anticipated by examination of a single tile. It is still interesting, however, to search for a single, simply connected 2D or 3D prototile that forces *maximal* nonperiodicity by shape alone, or one that does not permit any weakly nonperiodic tilings. As we write, we are aware of several current computer-based searches for an aperiodic topological disk in 2D. The general strategy is to enumerate all possible prototiles consisting of the union of simple triangles (polyiamonds), squares (polyominoes), hexagons (polyhexes), or certain pairs of triangles (polykleins), and, for each prototile, to examine all possible ways of forming small portions of a space-filling tiling. Typically, one quickly finds a portion that can tile the whole plane periodically, or one finds that the prototile does not admit a space-filling tiling at all. In each case, there is a measure of complexity associated with how many prototiles are needed to form the unit cell (the smallest anisohedral number permitted) or how many rings of tiles can be added around a central one before an irresolvable conflict is encountered (the Heesch number). [11, 14, 21, 22] The hope is that these computer searches will turn up a prototile that does not appear to have a finite Heesch number but does appear to have an infinite isohedral number. Such a prototile would then have to be examined analytically to establish that it really does have both properties. To date, the largest anisohedral number discovered is 10, which is achieved for a particular 16-hex [21]. It is not yet clear whether computer search will beat human creativity to finding the elusive unmarked, simply connected, two-dimensional einstein—if such a thing exists at all.

ACKNOWLEDGMENTS

We thank Chaim Goodman–Strauss, Marjorie Senechal, and an anonymous referee for helpful comments on earlier drafts.

REFERENCES

[1] D. P. DiVincenzo and P. J. Steinhardt. *Quasicrystals: the state of the art*, volume 16 of *Direction in condensed matter physics*. World Scientific, second edition, 1999.

[2] C. Janot. *Quasicrystals: a primer*. Oxford University Press, New York, 1997.

[3] R. Berger. The undecidability of the domino problem. *Mem. Amer. Math. Soc.*, 66:1–72, 1966.

[4] R. Penrose. The role of aesthetics in pure and applied mathematical research. *Bull. Inst. Math. Appl.*, 10:266–271, 1974.

[5] M. Gardner. Mathematical games. *Sci. Am.*, 236:110–121, 1977.

[6] J. E. S. Socolar and J. M. Taylor. An aperiodic hexagonal tile. *Journal of Combinatorial Theory: Series A*, 118:2207–2231, 2011.

[7] M. Dekking, M. Mendès France, and A. van der Poorten. Folds. *Mathematical Intelligencer*, 4, no. 4:173–181, 1982.

[8] P. Gummelt. Penrose tilings as coverings of congruent decagons. *Geometriae Dedicata*, 62:1–17, 1996.

[9] P. J. Steinhardt and H. C. Jeong. A simpler approach to Penrose tiling with implications for quasicrystal formation. *Nature*, 382: 431–433, 1996.

[10] L. Danzer. *A family of 3D–spacefillers not permitting any periodic or quasiperiodic tiling*, pages 11–17. In *Aperiodic '94* (G. Chapuis, Ed., World Scientific, Singapore), 1995.

[11] C. Goodman-Strauss. Open questions in tilings. Available at http://comp.uark.edu/~strauss/papers.

[12] A. R. Kortan, R. S. Becker, F. A. Thiel, and H. S. Chen. Real-space atomic structure of a two-dimensional decagonal quasicrystal. *Physical Review Letters*, 64:200–203, 1990.

[13] Branko Grünbaum and G. C. Shephard. *Tilings and patterns*. W. H. Freeman, New York, NY, USA, 1986.

[14] J. E. S. Socolar. Hexagonal parquet tilings: *k*–isohedral monotiles with arbitrarily large *k*. *Mathematical Intelligencer*, 29, no. 2:33–38, 2007. Proper version available at http://arxiv.org/abs/0708.2663.

[15] D. Fletcher. A construction of a nonperiodic tiling with simple atlas matching rules and one prototile. Private communication, December 2009.

[16] D. Frettloh and E. Harriss. Tilings encyclopedia. http://tilings.math.uni-bielefeld.de/.

[17] R. Penrose. *Remarks on tiling: Details of a* $(1+\epsilon+\epsilon^2)$*–aperiodic set*, pages 467–497. In *The mathematics of long-range aperiodic order* (R. V. Moody, Ed., NATO ASI Series C, **489**, Kluwer Academic Publishers, Dordrecht), 1997.

[18] Y. Araki. Mimicry beetles. http://www.gigapan.org/gigapans/49694/, 2010.

[19] F. Gähler and R. Klitzing. *The diffraction pattern of self–similar tilings*, pages 141–174. In *The mathematics of long-range aperiodic order* (R. V. Moody, Ed., NATO ASI Series C, **489**, Kluwer Academic Publishers, Dordrecht), 1997.

[20] C. Godreche. The sphinx: a limit–periodic tiling of the plane. *Journal of Physics A: Mathematical and General*, 22:L1163, 1989.

[21] J. Myers. Polyomino, polyhex and polyiamond tiling. http://www.srcf.ucam.org/jsm28/tiling/, 2009. See table of *n*–hexes with anisohedral number *k*.

[22] C. Mann. Heesch's problem. http://math.uttyler.edu/cmann/math/heesch/heesch.htm.

Viewpoint

Is There Any Conflict Between Evolution and the Second Law of Thermodynamics?

Bob Lloyd

The Viewpoint *column offers readers of* The Mathematical Intelligencer *the opportunity to write about any issue of interest to the international mathematical community. Disagreement and controversy are welcome. The views and opinions expressed here, however, are exclusively those of the author, and the publisher and editors-in-chief do not endorse or accept responsibility for them.* Viewpoint *should be submitted to one of the editors-in-chief, Chandler Davis and Marjorie Senechal.*

In 2010 the Discovery Institute published a volume of essays by Granville Sewell [1], accompanied by a publicity website [2]. This volume forms part of a series whose general thrust is to oppose the idea of Darwinian evolution, and the title of this one emphasizes "Intelligent Design." The volume includes a reprint of an article published in *The Mathematical Intelligencer* (*MI*) more than a decade ago [3], in which, among other matters, it was claimed that the Second Law of Thermodynamics is in opposition to evolution. Very strong criticisms of this paper were made at the time (among others, [4, 5]), but despite these, the work is still cited by the Discovery Institute in their list of publications "supporting the idea of Intelligent Design" [6], and so presumably opposing evolution.

In response to critics of the first *MI* paper, the author provided an expansion of his claim in a later note in this journal [7]. Some of this "*MI* response" has been expanded elsewhere [8], and another version of the argument is available on the Discovery Institute web site [9]. This last version has been expanded to form essay no. 5 in the new volume. Recently, another version was submitted to *Applied Mathematics Letters* (*AML*). The text of this has been made available [10], and a comment about it from the Editor of *AML*, including a link to the text, has appeared in that journal [11]; other comments have appeared on various web sites, which range from the fairly neutral [12] to much more polemical ones, and which can readily be found with a web search. The *AML* version quotes substantially from the original *MI* response, and the analysis in the present article will concentrate on these two versions.

This claim of conflict between the Second Law and Darwinian evolution may seem to some readers to have the endorsement of *The Mathematical Intelligencer*, though it was published with the heading "Opinion." Given the potential importance of such a claim (cf. Zimmerman and Loye [13]), and given that it has been repeated several times throughout the past decade, it is worth taking a look at the basis for it. In the *MI* response [7], a new idea was put forward in support of the claim, and this idea is fundamental to the later versions cited previously. It seems that the author believes that the various contributions to the total entropy of a thermodynamic system can be treated independently of each other, or, as Sewell writes in the *MI* response, "there are many entropies." The abstract of the later *AML* paper makes this particularly clear:

> It is commonly argued that the spectacular increase in order which has occurred on Earth does not violate the second law of thermodynamics because the Earth is an open system, and anything can happen in an open system as long as the entropy increases outside the system compensate the entropy decreases inside the system. However, if we define "X-entropy" to be the entropy associated with any diffusing component X (for example, X might be heat), and, since entropy measures disorder, "X-order" to be the negative of X-entropy, a closer look at the equations for entropy change shows that they not only say that the X-order cannot increase in a closed system, but that they also say that in an open system the X-order cannot increase faster than it is imported through the boundary. Thus the equations for entropy change do not support the illogical "compensation" idea; instead, they illustrate the tautology that "if an increase in order is extremely improbable when a system is closed, it is still extremely improbable when the system is open, unless something is entering which makes it **not** extremely improbable." Thus, unless we are willing to argue that the influx of solar energy into the Earth makes the appearance of spaceships, computers and the Internet not extremely improbable, we have to conclude that the second law has in fact been violated here.

In Sewell's first version, the *MI* response, two particular examples of "X-entropy" were provided, "thermal entropy"

DOI 10.1007/s00283-012-9277-0

and "carbon entropy" (see the following text), and in one of the intermediate versions [8] it is stated:

> Furthermore, there is really nothing special about thermal entropy. We can define another entropy, and another order, in exactly the same way, to measure randomness in the distribution of any other substance that diffuses.

This is the basis of what may appear to be a very powerful argument, which I will address. There is some lack of clarity: Are Sewell's "X-entropies" components of thermodynamic entropy as it is normally understood? Or is his intention more radical, to let the usual entropy be merely one, "thermal entropy," among the various X-entropies, the others being new quantities previously unknown to thermodynamics? I will proceed using the more modest terminology of components. Fortunately, most of my arguments could be applied with the other interpretation.

In both the *MI* response and the *AML* paper, a derivation is provided for a limit on the "flow" of entropy, which is referred to in the previous statements about the equations for entropy change. If Sewell is correct in his partitioning of total entropy into separate and independent X-entropy components, then his derivation applies separately and independently to each of these components. The "illogical compensation idea" to which he is objecting is the argument that was presented, in a rather dismissive fashion, in the first sentence of the abstract above. In passing, we might note that the phrase "anything can happen" is highly tendentious: if one shows that a particular process is not forbidden by the Second Law, that falls far short of showing that it *can* happen. Putting the "compensation" idea into more conventional language, it is frequently argued that although there is a local decrease in entropy associated with the appearance and evolution of life on earth, this is very small in comparison with the very large entropy increase associated with the solar input to earth. This qualitative idea has received quantitative backing from the calculations of Styer [14], and particularly as modified by Bunn [15], which show that the solar contribution is many orders of magnitude larger than any possible decrease associated with evolution.

Sewell comments that this idea is widespread in the literature, but he does not accept it, and in the absence of a refutation of his own argument about "many entropies," he may be justified in this. If he were correct in claiming that these have to be assessed separately, the calculations of Styer and of Bunn would be irrelevant, because they are mixing what he considers to be different quantities, different sorts of entropy. The final sentence of his *AML* abstract, quoted previously, depends on this partitioning into separate and independent X-entropies, as does the closely related claim in the *MI* response that

> the "thermal order" imported from the Sun does not help explain the formation of humans.

If the entropy partitioning he proposes were valid, these claims would need to be taken very seriously, and the Second Law and Darwinian evolution might then be in conflict.

Before considering the partitioning, I want to note the way Sewell treats the terms "entropy," "probability," and "disorder" as if they were freely interconvertible. In the citation mentioned previously, not only is entropy partitioned, but each individual variety of X-entropy has an equivalent negative "X-order" associated with it, "since entropy measures disorder." This is a very loose use of language. As Carnap showed [16], even the connection between probability and entropy is not a simple one. The entropies derived by Boltzmann and by Gibbs, using statistical methods, are generally not equivalent to thermodynamic entropy.

The further connection, between entropy and "disorder," is even more difficult to establish, and assumptions of equivalence between these terms can give rise to substantial confusion [17]. Styer considers that the connection between the two is at best that of a simile [18], and in his analysis of some of the difficulties from a physics perspective he notes three problems. These are that the language of disorder is vague; that there is a substantial emotional content to the word "disorder"; and, most important, that "the analogue between entropy and disorder invites us to think about a single configuration rather than a class of configurations." But entropy is related to disorder only through statistical mechanics and its derivation of thermodynamic laws—a derivation in which it is essential to speak of classes of states, not of individual states only.

From a chemical point of view, the problem is set out succinctly by Levine [19]: "However, order and disorder are subjective concepts, whereas probability is a precise quantitative concept. It is therefore preferable to relate S to probability rather than to disorder." It is apparent from these citations that statements about order and disorder have no necessary connection to thermodynamics; the incorrect assumption of a formal connection may have been the source of the error discussed in the following text.

There must be some suspicion over claims that are presented in the cavalier manner of Sewell's abstract

AUTHOR

BOB LLOYD was Professor of General Chemistry at Trinity College Dublin until his retirement in 2000. His research focused mainly on the chemical and physical information that can be obtained about solids and their surfaces, and molecules adsorbed on these, by techniques involving the energy spectra of ejected electrons, such as angle-resolved photoemission. In his retirement he has been busily reading the classical literature, improving his Biblical and classical Greek, and correcting errors as they come to his attention in several areas, including the environment and classical studies. He is a consumer of live and recorded music, from Hildegard to Hindemith.

School of Chemistry
Trinity College
Dublin 2
Ireland
e-mail: dlloyd@tcd.ie

quoted earlier. Nevertheless, his original statement is in terms of entropy, and the main argument can be examined without assuming that entropy and disorder are equivalent. His argument is that because entropy can be partitioned into separate independent varieties, any restriction that applies to any one of these supposed varieties of entropies, such as the one derived [7, 8, 10], applies equally to all other varieties. For this reason he contends that it is incorrect to "compensate" the entropy change associated with one sort, such as that associated with evolution, by the entropy change associated with the degrading of the solar flux to Earth.

To assess the validity of this partitioning of entropy into separate varieties, let us look at the main example given in the original *MI* response. This concerns the transport of heat and the diffusion of carbon atoms in hot steel. (In the language of materials science, hot steel is a "solid solution," and the carbon atoms in this solution diffuse through the iron in much the same way as solutes in a liquid solution.) A thought experiment is proposed that involves two isolated blocks of steel at high temperatures and a steel rod that can connect them; all three contain the same concentration of diffusing carbon atoms. Discussing the diffusion of the carbon atoms, Sewell writes:

> But it is important to notice that now "entropy" measures the randomness of the distribution of carbon, not heat, so the amount of thermal entropy exported is not relevant to the change in carbon entropy in the solid. For example, if a steel rod of uniform temperature and uniform carbon concentration is placed between two steel blocks of unequal temperatures but carbon concentrations identical to that in the rod, the rod may import "thermal order" (export thermal entropy), but the "carbon order" will be unaffected.

He seems to be asserting that the number densities for the carbon atoms will remain equal— that although connection of the rod to the blocks will generate a temperature gradient along the length, the carbon atom distribution will not change, since "the amount of thermal entropy exported is not relevant to the change in carbon entropy in the solid."

This is a very clear statement that he believes that the different "X-entropies," in the later *AML* formulation, behave independently. This point is central to all of the versions of his argument. It is subject to a number of objections. The particular experiment described for steel at high temperatures would be difficult to carry out in practice, but the principle being claimed, that the "order" of the carbon atoms, the solute in a solid solution, is independent of the temperature gradient, cannot be specific to this particular system. Sewell's argument is general and should apply equally to solutes in liquid solutions. Here also, thermal gradients should have no effect on the concentrations of dissolved species *if* the independent treatment of entropies is valid.

Now the effect of thermal gradients on liquid solutions has been studied in detail, and the Soret or Ludwig-Soret effect, which is found in such experiments, has been known for a century and a half [20, 21]. In these experiments thermal gradients *do* induce concentration gradients of solutes in solvents, and very precise measurements are possible [22]. In a theoretical paper on systems where two solutes are present in a solvent, it has even been shown that it could be possible to generate an almost complete separation of these two solutes, simply by applying a thermal gradient [23]. The authors of this paper make the fundamental point, which is very relevant to Sewell's thought experiment, that "once the temperature difference takes the system away from equilibrium and breaks translational symmetry, there is no reason why the ratio of concentrations should stay constant." In the thought experiment it was assumed that the C atoms in the steel rod are unaffected by the asymmetry of their thermal environment, but this cannot be generally correct. A further example to illustrate this point, which is physically very similar to Sewell's solid state model, is provided by the diffusion of dopant ions in solid ionic conductors. If such solids are placed in a thermal gradient, there is a migration of the dopant ions. This generates a thermoelectric potential difference in the material, which can be measured readily [24]; if Sewell's proposal were correct, there would be no such potential difference. More generally, thermoelectric effects in metals and semiconductors, which depend on the diffusion of holes and/or electrons, should also not exist, but they do, and are technically important.

The error is not merely in the details of the thought experiment. The assumption that the various components of the total entropy can be treated independently has been imposed arbitrarily, without justification. In a block of steel the "thermal entropy" and the "carbon entropy" interact to give one observable quantity, the entropy; similar statements apply to any system. For any particular system in a specified condition it may be possible to define nominally separate *contributions* to the total entropy, but this does not make these contributions *independent* of each other. Such a conceptual separation is sometimes made for convenience, but it is essential to show that it is valid in any particular case. Ignoring this can lead to serious error even in such an apparently simple situation as the melting of a pure solid [25].

Here is a simpler model than the steel-carbon one, which preserves its essential features, and which may help to emphasize the nature of the error. Consider instead a system of two volumes, and a tube that can connect them, all filled with argon to the same pressure at some temperature, so that the argon atom number density is the same in all three, just as the carbon concentrations were in Sewell's thought experiment. With the tube disconnected, the volumes are heated to different temperatures. In the language used by Sewell, if the tube is now connected, a temperature gradient will be created along the tube, and the tube will "export thermal entropy." However, the argon number densities are the same in each volume, so if the entropies were independent as claimed, the "argon entropy" should be unaffected, and argon densities should remain constant along the tube, despite the fact that the higher temperature of one vessel has created a higher pressure at that end. In reality there is only one entropy associated with the system, not two, and when the tube is connected, gas will of course flow as this total entropy of the system increases. The separation of total entropy into "different entropies" has led to an unreasonable conclusion and is invalid.

It is difficult to know where this mistaken idea, that entropy can be separated into independent components, has come from. One possibility is that this comes from assuming a precise equivalence between entropy, to which the formalisms of thermodynamics apply, and disorder, which is too ill-defined for thermodynamics to be applied. There is a curious sentence in the *MI* response: "But just because two things are both improbable does not necessarily mean that the importation of one (say, TV sets) into an open system can explain the appearance there of the other (say, airplanes)." Shortly after this there is a mention of the "order associated with airplanes," so it seems that the message is that these different sorts of objects each have their own specific sort of order. However, no thermodynamic conclusions can be drawn from this. If in fact the sentence is intended to convey a thermodynamic meaning, as the context suggests, this can only be because of a belief that these different sorts of order correlate directly with qualitatively different sorts of entropy. Sewell's idea of "many entropies" may have been generated by assuming this correlation, but there *is* no such correlation. If the proposal of "many entropies" *were* valid, then rather than destroying the case for evolution, it is more likely that this would destroy the case for thermodynamics.

(Sewell does not address the matter of how many X-entropies he would countenance. From the passages quoted previously it would seem that there might be one for 27-inch TVs and one for 42-inch TVs: they are clearly distinguishable. Shall we apply the Second Law separately to each of a myriad of X-entropies? If this were valid, it would render the Law completely useless.)

There is another possible source for the error. Immediately after the description of the steel rod thought experiment in the *MI* response, quoted previously, Sewell writes:

> In the scientific literature, thermal entropy is usually referred to simply as "entropy," but in fact there are many entropies (depending on what we choose to measure: see [2], p. xiii) and many kinds of order: any macroscopic feature or property that is improbable from the microscopic point of view can be considered order.

From this, the citation "[2]" seems to be related to the proposal that there are "many entropies," and no other reference is offered in support of the claim, in this paper or in the *AML* manuscript, so it needs to be followed up. This reference is to Carnap's "Two Essays on Entropy" [16], and given Carnap's strictures in this work, noted previously, on the difficulties of comparing even statistical and thermodynamic entropy, this is a surprising choice. The page citation given (p. xiii) does not actually refer to anything written by Carnap, but to Shimony's critical introduction to Carnap's essays. Early in this introduction Shimony writes:

> The core of Carnap's argument is given in §6. Entropy in thermodynamics is asserted to have the same general character as temperature, pressure, heat, etc., all of which serve "for the quantitative characterization of some objective property of a state of a physical system."

Thus entropy, like temperature, is a function of the state of the system, and there is no mention of constituents of the system. In this framework there is no more reason to discuss "carbon entropy" separately from the total entropy of a system than there is to discuss "carbon temperature" separately from the temperature of the equilibrium system.

The passage cited in the *MI* response comes later in this introduction. The text on the cited page xiii that might be related to "many entropies" is as follows (references omitted, italics original):

> There are many thermodynamic entropies, corresponding to different degrees of experimental discrimination and different choices of parameters. For example, there will be an increase of entropy by mixing samples of O_{16} and O_{18} only if isotopes are experimentally distinguished. Jaynes therefore claims that *"Even at the purely phenomenological level, entropy is an anthropomorphic concept.* For it is a property, not of the physical system, but of the particular experiment you or I choose to perform on it." Carnap's point of view can clearly accommodate a multiplicity of thermodynamic entropies, to each of which the principle of physical magnitudes would be applied when statistical counterparts are sought.

When he cites this passage, Sewell writes of "many entropies," but the significant qualifier "thermodynamic" is omitted. The main point Shimony is making here is that before comparing thermodynamic entropies with statistical entropies, it is essential to define the former precisely, because, even within the thermodynamic context, it is possible to vary the definition of entropy according to the choice of parameters. There can be a multiplicity of thermodynamic entropies simply because many definitions of this concept are possible, and Shimony's point is no more than that. After any particular definition has been chosen, a discussion of entropy can be set up. However there is not the least suggestion here that, within the framework of any particular definition of total entropy, entropy can be partitioned in the way proposed. There is *no* logical connection at all between Shimony's words and the proposal that "carbon entropy" can be separated from "thermal entropy," and no justification for the creation of separate independent "X-entropies" has been offered. Sewell appears to think that Shimony gives support to his erroneous picture, but this can only have come from a basic misunderstanding of the words cited.

The proposal that entropy can be partitioned is relied on in a series of papers, from the original *MI* response to the recent revisiting of this in the *AML* paper, but the proposal has no validity. The illustrative example given in the *MI* response is simply wrong, and these papers rely on erroneous interpretations of basic physics and on a failure of logic, which vitiate the conclusions drawn. In particular, these papers provide no reason whatsoever to suppose that the Second Law makes any statement that denies the possibility of Darwinian evolution, or even makes it improbable. The qualitative point associated with the solar input to Earth, which was dismissed so casually in the abstract of the *AML* paper, and the quantitative formulations of this by Styer and Bunn, stand, and are unchallenged by Sewell's work.

ACKNOWLEDGMENTS

I thank D. A. McDónaill for bringing this problem to my attention and for many useful discussions, J. Corish for pointing out the analogy with diffusion in ionic semiconductors, and M. Lloyd for a careful reading of the final version. The comments from the anonymous referee of an earlier version led to substantial improvements in the presentation of this work, and I am very grateful for these.

REFERENCES

[1] Sewell, Granville, *In the Beginning, and Other Essays on Intelligent Design*, Discovery Institute Press, Seattle, 2010.

[2] http://www.discovery.org/v/1951.

[3] Sewell, Granville, A mathematician's view of evolution, *Math. Intelligencer* 22 (2000), no. 4, 5–7.

[4] Rosenhouse, Jason, How anti-evolutionists abuse mathematics, *Math. Intelligencer* 23 (2001), no. 4, 3–8.

[5] Davis, Tom, The credibility of evolution, *Math. Intelligencer* 23 (2001), no. 3, 4–5.

[6] http://www.discovery.org/a/2640.

[7] Sewell, Granville, Can ANYTHING happen in an open system? *Math. Intelligencer* 23 (2001), no. 4, 8–10.

[8] Sewell, Granville, *The Numerical Solution of Ordinary and Partial Differential Equations*, 2nd Edition, Appendix D, John Wiley & Sons, New York, 2005, also reprinted in Ref. 1.

[9] http://www.discovery.org/a/3122.

[10] http://www.math.utep.edu/faculty/sewell/AML_3497.pdf.

[11] Editorial note, *Appl. Math. Lett.* 24 (2011), 1968.

[12] http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=416514&c=1.

[13] Zimmerman, M., Loye, D., Science and religion: a new alliance to combat the new wave of creationism, *World Futures: J. Gen. Evol.* 67 (2011), no. 1, 1–10.

[14] Styer, Daniel, Entropy and evolution, *Am. J. Phys.* 76 (2008), 1031–1033.

[15] Bunn, Emory, Entropy and the Second Law of Thermodynamics, *Am. J. Phys.* 77 (2009), 922–925.

[16] Carnap, R., *Two Essays on Entropy*, University of California Press, Berkeley, CA, 1977, available at: http://www.ucpress.edu/op.php?isbn=9780520027152.

[17] http://entropysite.oxy.edu/.

[18] Styer, Daniel, Insight into entropy, *Am. J. Phys.* 68 (2000), 1090–1096.

[19] Levine, Ira N., *Physical Chemistry*, 6th Edition, McGraw-Hill, New York, 2008, p. 100.

[20] Ludwig, C., Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen, *Sitzungsberichte Akad. Wiss. Wien Math.-Naturwiss. Kl.* 20 (1856), 539.

[21] Soret, C., Sur l'état d'équilibre que prend au point de vue de sa concentration une dissolution saline primitivement homogène dont deux parties sont portées à des températures différentes, *Ann. Chim*. *(Phys.)* 22 (1881), 293–297.

[22] Zhang, K. J., Briggs, M. E., Gammon, R. W., Sengers J. V., Optical measurement of the Soret Coefficient and the Diffusion Coefficient of Liquid Mixtures, *J. Chem. Phys*. 104 (1996), 6881–6892.

[23] Garriga, A., Kurchan, J., Ritort, F., Strong Soret effect in one dimension, *J. Stat. Phys.* 106 (2002), 109–123. (Also available at: arXiv:cond-mat/0104327 v1 18 Apr 2001).

[24] Corish, J., Jacobs P. W. M., Thermoelectric power of silver chloride, *J. Phys*. *C: Solid State Phys*. 6 (1973), 57–79.

[25] Karasz, F. E., Couchman, P. R., Klempner, D., On the entropy of fusion and its separation into configurational and volume-change contributions, *Macromolecules* 10 (1977), 88–89.

The Mondee Gills Game

SASHA GNEDIN

This column is a place for those bits of contagious mathematics that travel from person to person in the community, because they are so elegant, surprising, or appealing that one has an urge to pass them on.

Contributions are most welcome.

➢ Please send all submissions to the Mathematical Entertainments Editor, **Ravi Vakil,** Stanford University, Department of Mathematics, Bldg. 380, Stanford, CA 94305-2125, USA
e-mail: vakil@math.stanford.edu

To switch or not to switch, that is the question …

A game played in Mondee Gills is not so familiar because of the isolation of the country and peculiarities of the dialect of the Mondees. Many people confuse the game with its rogue variant sometimes played for money with thimbleriggers on the streets of big cities. Mainland historians trace the roots to primitive contests such as coin tossing and matching pennies. However, every Mondee knows that the game is as old as their land, and the land is really old. An ancient tomb unearthed by archaeologists on the western side of Capra Heuvel, the second-highest hill in the country, was found to contain the remains of domesticated cave goats and knucklebones, which supports the hypothesis that the basics of the game go back to farmers of the late Neolithic Revolution.

The rules of the Mondee Gills Game (MGG) were always the same, although the scenery changed with time. Nowadays the Mondee Gills *Monday TV* broadcasts the game as an open-air reality show in which there are three caves, one goat, and two traditional characters called Monte and Connie. The goat is hidden by Monte in one of the caves before the show starts, two other caves are left empty. The second actor, Connie, who is randomly chosen from the audience, is ignorant of what the caves conceal, and will be offered two tries to find the goat. She is asked first to choose one of three caves. Then Monte always proceeds to reveal an empty cave – one of the caves not chosen by Connie – and asks whether she wants to stay with her initial choice or switch the choice to the other remaining unrevealed cave. The goat is won if the final choice is correct.

Most visitors from nearby Laieland who are new to the game, when playing in the role of Connie, are keen on making the first choice at random, but they rarely realize the true meaning of the switch offer when two unrevealed caves remain. Many of them think that once the first choice is equally likely to fall on the cave with the goat, the equality of chances still persists when one empty cave is excluded, so the odds are 1:1 whichever the action. Some of them are reluctant to switch and prefer to stick with the first chosen cave. Others, more experienced in primitive games, flip a fair coin to decide on the second choice. The newbies quickly notice that the coin-flippers win about a half of the goats in the rounds played, whereas those who stay with the initial choice score only one third.

Needless to say, native Connies are superior in the game, managing to win about two thirds of the goats in the rounds played by switching all the time. The advantage of the Mondees, whose strategy also involves choosing the first cave uniformly at random, is due to the local educators, who developed composite teaching programs. Training starts in early childhood with a chocolate goat concealed in one of

DOI 10.1007/s00283-011-9253-0

three bags, but it is problematic at this stage, as the kids cry when they are not given the bag they initially chose or when they are not rewarded. Later on, children practice Mondee Karlo techniques using educational computers and playing the MGG for valuable coins with goat heads and tails, either against cartoon Monte applets or by distributing the roles over two groups of players. Students of secondary schools are dedicated to the history and philosophy of the MGG [17] and have theoretical lessons on the fundamentals of the Doctrine of Chances [2].

The theory of the MGG is a subject by itself, being largely focused on the problems of *utility of switch* and the resolution of the *paradox of odds*. Basic propositions are found in Books of Texts [1, 5, 15, 19]. The reader is referred to the works of Magister [6, 7, 8] for the up-to-date, critical analysis of the major theoretical developments. A rich source for those who have a good command of Mondees language is Quickygnosis [21], the online forum of the Mondee Gills educators. In this essay we shall mention only some theories that will be important for seeing the place of our contribution.The names of the two principal schools of thought can be translated from Mondees as Big-Endians and Little-Endians.

The traditional school of Big-Endians, also known as simplists, originates from prophet Prosto, who taught that paradoxes should be resolved from the convenient end. Having postulated that the first choice gives the goat with probability 1/3, Big-Endians prove that the never-switching strategy wins with this very probability 1/3, and the always-switching strategy wins with probability 2/3 – because the revealed cave is empty, the goat is found either in the cave first chosen or in another one unrevealed. From this they conclude, quite convincingly, that the odds for the second choice cannot be 1:1 with certainty, since always-switching wins with probability strictly higher than 1/2. As a further justification of the advantage of the always-switching strategy, a metaphysical experiment is conducted in which Connie's personality at the moment of the first choice dissociates in two, of which Jekyll-Connie sticks with the first choice, whereas Hyde-Connie switches. A newest branch of the school goes farther to assert that the odds of the second choice do not exist unless it is explicitly specified what kind of random device Monte is using when the first Connie's choice was correct and he is free to decide which of two empty caves he is going to reveal. Despite the internal split, all Big-Endians consider the paradox of odds completely resolved by these arguments, and the utility of switch justified.

Formerly, everyone accepted the simplistic views, but, a few generations in the past, the princess of Mondee Gills had bad luck in playing the game. The legend narrates that she chose cave 1, and when Monte opened cave 3 she switched to cave 2, whereas the golden goat was hidden in cave 1. To console the princess, the duke, a devoted friend of the ruling house, passed a law to increase the number of caves from 3 to 1000. But on the next Monday the princess lost again by switching to cave 2 after all caves numbered from 3 to 1000 had been revealed as empty. Based on conclusions of the investigation commissioned by the duke, the High Court accused Monte of fraud, and Monte, who was in fact the minister of education, was fired. Later studies showed that the minister was right when he murmured that such misfortune is not unlikely to occur to some Connies, given the fact that the MGG is played every Monday, but this theme will lead us away from our main topic... The ruler at that time, father of the princess, overruled the duke's law and issued a decree commanding all to stick with tradition by playing exclusively with 3 caves, and for the host to flip a fair coin for revealing one of two empty caves when Connie's first guess is correct. The decree called into life the Little-Endian school of conditionalistic thinking. The wisdom of the ruler enabled the scholars to define the odds for Connie's second choice with perfect mathematical rigor and eventually to justify the advantage of switching by evaluating the odds in its favor at 2:1 in the particularly important case when cave 1 was chosen by a random pick and cave 3 revealed. More advanced modern methods showed that under this mode of play, the result 2:1 also holds for other selections of two distinct cave numbers.

The schools of Big-Endians and Little-Endians are in a state of constant controversy, each accusing the other of resolving the odds paradox from the wrong end. The Little-Endians claim that their opponents do not provide enough evidence that switching is a better action for Connie *after* she has chosen cave 1 and cave 3 has been revealed. The differences have grown so large as to threaten the country's educational activities, leading the High Court to ban some of the most offensive scholars from Quickygnosis to cool down the mood [22].

One insightful, unorthodox offspring of the Little-Endian paradigm is the Monte Crawl game [18], in which Monte always crawls to reveal the cave with smaller number from the two caves unchosen by Connie, when her initial choice happens to be correct. Under this mode of play, Connie receives a definite signal about the location of the goat when Monte is forced to go to the cave with the larger number, because when he cannot reveal the cave with the smaller number the goat is certainly there. When it comes to Connie's

AUTHOR

SASHA GNEDIN is (in the reverse lexicographic order) a Russian-German-Dutch-British mathematician doing research and teaching in the fields of probability, optimization, games, and combinatorics. Being an adept of Concrete Mathematics he especially likes working on the easy-to-formulate classical problems commonly viewed as "solved". Outside mathematics, he enjoys travel, tennis, and playing blues on guitar.

Department of Mathematics
Utrecht University
Post Office Box 80010
Utrecht, 3508 TA
The Netherlands
e-mail: A.V.Gnedin@uu.nl

second guess, the odds in favor of switching are no longer constant, as with the fair-coin-flipping Monte, rather they stay 1:1 if the cave with smaller number is revealed, but 1:0 if the cave with larger number is revealed. A dual regime is the Monte Haul game, in which Monte is always hauled to the cave with higher number to reveal it when Connie's initial choice happens to be correct, the odds being 1:0 and 1:1, respectively. These instances of Monte's behavior are important for seeing that the odds *can* be 1:1 in some situations, but under no circumstances do the odds strictly favor sticking with the initial choice, provided that the initial guess is correct with probability 1/3 (and assuming that the odds for the second choice are well defined).

One subtle distinction between two major theories, which is rarely stated explicitly, is that the Big-Endianness only admits two of Connie behaviors, which may be called *constant-action* strategies: always switch, or always stick. The Big-Endian solution states the superiority of always-switching over always-sticking by the same initial choice of the cave. In contrast to that, the Little-Endianness admits dependence of the last Connie action on what happened so far, thus allowing for *mixed-action* (or history-dependent) strategies - perhaps sometimes switch and sometimes not; the task of the theory amounts then to showing that always switching is optimal within the larger class of strategies.

In this connection we should warn the reader against possible misinterpretations of the name of our principal character. The name *Connie* should not be associated with "contestant", rather it stems from *Constance*, meaning the actor winning by playing constant-action, always-switching strategies.

A major actual breakthough in the odds paradox is found in the newest writings of Magister [8], where the Little-Endian solution is derived from the Big-Endian with the aid of the principles of symmetry. The present article is devoted to another line of Magister's thought in the direction of viewing MGG as a *game of strategy*, that is, a situation in which two actors interact to achieve their goals within the framework specified by the rules of the game. We credit Professor [5] for the evaluation of Connie's winning probability in the worst case as being 2/3, and for identifying whcih of Monte's strategies is most unfavourable for her, namely, his rolling a symmetric three-sided die to determine the goat's hiding place. This key result and its simple proofs are ranked by Magister [9] as the Holy Grail of the MGG studies, not to say that it sets objective limits to the long-run proportion of goats won from a hostile Monte willing to keep as many goats as possible.

Middle-Endianness: A Simplistic Path

The approach we take here may be called Middle-Endianness, because it is equidistant from the two mainstreams, and it circumvents the probabilistic postulates in the problem of the utility of switching. The new doctrine is an instance of the fundamental *principle of eliminating dominated strategies*; in simplistic terms it can be stated as follows: No matter where Monte hides the goat and how he plays when he can choose between two caves to reveal, each Connie strategy "choose cave Y and stay with it" is outperformed by a strategy "choose cave $Y' \neq Y$, then switch". We believe that this mode of thinking has the potential to provide a convincing explanation of the advantage of the switching action to the people from Laieland, as compared with more sophisticated Big-Endian/Little-Endian and other methods based on computing probabilities. Once Laiemen adopt strategic thinking and realize that there is a two-step action, the comparison of alternatives becomes obvious and, moreover, free of *any* probability considerations.

To conceive the twist in the switch-versus-notswitch dilemma we stay in this introduction with the simplistic scenario, that is, ignore which particular cave is revealed by Monte in the case of an initial correct guess. Let X denote the cave hiding the goat. Consider the following three strategies for Connie:

A: choose cave 1, do not switch,
B: choose cave 1, then switch,
C: choose cave 2, then switch.

Strategy A wins if $X = 1$, whereas strategy B wins if $X \in \{2, 3\}$, so the strategies A and B cannot win simultaneously. The odds are 1:2 against A if the values of X are assumed to be equally likely. More generally, the extended Big-Endian theory assigns arbitrary probabilities to the values of X and leads to the familiar conclusion that B should be preferred to A provided the probability of $X = 1$ is less than the probability of $X \in \{2, 3\}$. Nothing new so far, but now including C into the consideration we observe that strategy C wins for $X \in \{1, 3\}$, so if A wins then C wins too, and there is a situation when C wins while A fails. Thus strategy C is not worse than A, and it is strictly better if cave 3 *sometimes* hides the goat. This provides a general ground to avoid A, and for a similar reason to avoid all other strategies that do not switch in some situation.

Experimentation with strategies A and C exercised simultaneously is possible by complex metaphysics. Before the game starts, Connie's personality splits into Jekyll-Connie to play A and Hyde-Connie to play C, and the dissociated players will not communicate. Monte hides the goat in one of the caves, and then Jekyll-Connie and Hyde-Connie make their first choices according to strategies A and C, respectively. On the next move Monte reveals for Jekyll-Connie an empty cave from caves 2 and 3, and he reveals for Hyde-Connie an empty cave from caves 1 and 3. Based on their initial choices and their private information received from Monte, the dissociated players proceed with a final move according to their strategies. No matter which rule Monte uses to reveal the cave, Hyde-Connie will only lose when the goat is in cave 2, but in this case Jekyll-Connie will lose, too, as she does not switch to cave 2 according to strategy A. Following a suggestion by Magister [9], we may call this way of comparing strategies *coupling* by the initial position (i.e., location of goat). The coupling is analogous to simultaneously playing the same strategy in chess with, say, white pieces against two players on two boards.

We stress that, according to the Big-Endian/Little-Endian theory, discarding A in favor of B must be based on the probabilities of the mutually exclusive events $X = 1$ and $X \in \{2, 3\}$. In the Middle-Endian theory, the advantage of C over

A follows from the fact that the event $X = 1$ is included in the event $X \in \{1, 3\}$, which is much stronger as the win of A implies the win of C.

The dominance extends our horizons in understanding the Mondee Gills Game. For instance, thinking of the superiority of C over A, we do not need to interpret *sometimes* as "having positive probability". Think of a computer program that schedules in a deterministic manner the location of the goat, and suppose that on some Mondays the goat is hidden behind cave 3. Then C will have a strict advantage over A because of these very Mondays. In the following discussion we shall only focus on application of the Middle-Endian paradigm to the antagonistic variant of the game [5, 7].

The Zero-Sum Game

We shall consider the interaction of Monte and Connie as a game of strategy, as opposed to a game against nature. Monte has freedom to choose the cave hiding the goat and to reveal one of the empty caves when he has two options. Connie chooses a cave, and when offered the second choice decides between switch and notswitch. Connie *wants* to win the goat. We shall no longer confine ourselves to the simplistic scenario, rather we allow arbitrary dependence of the actions on the course of the game. A fundamental model of interaction is the actors' pure competition for goats, because this instance reveals what Connie can achieve under the least favorable circumstances. A technical term for this is a zero-sum game, in which at the start of the game the goat belongs to *Monday TV* and Monte *wants* to keep the goat for the program.

Now we need to make the exposition more formal, so as to avoid possible misinterpretations of Middle-Endianness. Probabilities will be introduced into the scene as mixed strategies of two active actors, Monte and Connie. Although this article is mainly self-contained, we expect that the reader is familiar with the basic concepts of strategy, minimax solution, payoff, and common knowledge. For these and the propositions used in the following text, we refer to the online Course on Games [4].

Strategies and the Payoff Matrix

To introduce notation for the possible actions of actors and to formalize the rules, it will be convenient to label the caves 1, 2, 3 in the left-to-right order. The game in extensive form has four moves:

(i) Monte chooses a cave X out of 1, 2, 3 in which to hide the goat. The choice is kept secret.
(ii) Connie picks a cave Y out of 1, 2, 3 and announces her choice. Now both actors know Y, and they label the caves distinct from Y Left and Right in the left-to-right order.
(iii) If $Y = X$, so the choice by Connie fell on the cave with the goat, Monte chooses cave Z to reveal from the Left and Right caves. In the event of the mismatch $Y \neq X$, Monte reveals the remaining cave Z (distinct from X and Y), which is either Left or Right depending on X, Y.
(iv) Connie observes the revealed cave Z and makes a final decision: she can choose between Notswitch or Switch from Y to another unrevealed cave (so distinct from Y and Z). Connie wins if the final choice yields X and loses otherwise.

Monte's action on step (iii), when he has the freedom of choosing to reveal either the Right or Left cave, may depend on Connie's initial choice Y. Connie's final action in (iv) depends on both Y and Z. The rules of the game are part of the common knowledge, which means that everybody knows the rules, and knows that everybody knows, and knows that everybody knows that everybody knows, etc.

To put the game in matrix form, we label the admissible pure strategies of the actors. The pure strategies of Monte are

1L, 1R, 2L, 2R, 3L, 3R.

For instance, according to strategy 2L the goat is hidden in cave $X = 2$, then if the outcome of (ii) is $Y = 2$, Monte will reveal the Left cave (which is cave 1). Otherwise the second move of Monte is forced and the "L" part of the code 2L is irrelevant, in particular if $Y = 1$ he will reveal cave 3 (which happens to be the Right cave), and if $Y = 3$ he will reveal cave 1 (which happens to be the Left cave).

The pure strategies of Connie are

1SS, 1SN, 1NS, 1NN, 2SS, 2SN, 2NS, 2NN, 3SS, 3SN, 3NS, 3NN.

The digit 1, 2, or 3 is a value of Y, whereas SS, SN, NS, NN encode how Connie's second action depends on Y and whether the Left or Right cave is revealed. For instance, 1NS means that cave $Y = 1$ is initially chosen, then Connie plays Notswitch if Monte reveals the Left cave; and she plays Switch if Monte reveals the Right cave.

The game is played as if Monte and Connie have specified their two-step pure strategies before the Mondee Gills show starts. For this purpose they may ask friends for advice or employ random devices such as spinning a roulette wheel or rolling dice. After the choices are made the actors just follow their plans. The choices could be communicated to a referee who announces the then predetermined outcome of the game. For example, if Connie and Monte choose profile (2SN, 1R) the show proceeds as follows:

(i) Monte hides the goat in cave 1.
(ii) Connie picks cave 2, thus the actors label cave 1 as Left and cave 3 as Right.
(iii) Monte observes a mismatch, hence he reveals the remaining cave 3.
(iv) Connie observes revealed cave 3, which is Right, hence she plays Notswitch - meaning that she stays with cave 2 (and loses).

In the zero-sum game the payoff of one actor is the negative of the payoff of the other: Connie wants to win the goat while Monte aims to avoid this. With regard to the satisfaction of actors there are two distinguishable outcomes – so we agree that Connie's payoff is 1 if she wins the goat and 0 otherwise. All possible outcomes of the game are summarized in matrix $\boldsymbol{C}$ showing Connie's payoffs. The matrix has appeared in Professor's Book of Texts [5]:

	1L	1R	2L	2R	3L	3R
1SS	0	0	1	1	1	1
1SN	0	1	0	0	1	1
1NS	1	0	1	1	0	0
1NN	1	1	0	0	0	0
2SS	1	1	0	0	1	1
2SN	0	0	0	1	1	1
2NS	1	1	1	0	0	0
2NN	0	0	1	1	0	0
3SS	1	1	1	1	0	0
3SN	0	0	1	1	0	1
3NS	1	1	0	0	1	0
3NN	0	0	0	0	1	1

The mathematical structure of the game will not be changed by certain transformations of the payoff matrix. For instance, we might replace each 0 by -1. The latter would correspond to the variant of the game in which both Connie and Monte bring to the show one of their own goats, which are then hidden by Monte in the same cave. Both goats go into the possession of Connie, respectively Monte, depending on whether Connie finds the cave with goats or not.

A mixed strategy for Connie is a row vector $\boldsymbol{P}$ of twelve probabilities that are assigned to her pure strategies. Similarly, a mixed strategy for Monte is a row vector $\boldsymbol{Q}$ with six components. When strategy profile $(\boldsymbol{P}, \boldsymbol{Q})$ is played by the actors, the expected payoff for Connie, equal to her winning probability, is computed by matrix multiplication as $\boldsymbol{PCQ}^T$, where ^{T} denotes transposition. This way of computing the winning probability presumes that the actors' choices of pure strategies are independent random variables, which may be simulated by their private randomization devices. Intuitively, the statistical independence is a way to say that cooperation of the actors in zero-sum games is impossible, so that the actors cannot agree on any other kind of joint distribution over the outcomes.

Suppose for a while that Monte is determined to play some given strategy $\boldsymbol{Q}$ and that everybody including Connie knows this. Then her optimal behavior is $\boldsymbol{P}'$, a *strategy of best response* (aka *Bayesian strategy*) against $\boldsymbol{Q}$ to maximize $\boldsymbol{PCQ}^T$:

$$\boldsymbol{P}' = \boldsymbol{P}'(\boldsymbol{Q}), \quad \boldsymbol{P}'\boldsymbol{CQ}^T = \max_{\boldsymbol{P}} \boldsymbol{PCQ}^T.$$

Exchanging the roles of the actors, we suppose that Connie is determined to test a particular $\boldsymbol{P}$ and tells this to everybody, so that Monte can react with a best response $\boldsymbol{Q}'$:

$$\boldsymbol{Q}' = \boldsymbol{Q}'(\boldsymbol{P}), \quad \boldsymbol{PCQ}'^T = \max_{\boldsymbol{Q}} \boldsymbol{PCQ}^T.$$

A strategy profile $(\boldsymbol{P}^*, \boldsymbol{Q}^*)$ is said to be a *game solution* (or *minimax solution*) if the strategies are the best responses to each other,

$$\boldsymbol{P}^*\boldsymbol{CQ}^{*T} = \max_{\boldsymbol{P}} \boldsymbol{PCQ}^{*T} = \min_{\boldsymbol{Q}} \boldsymbol{P}^*\boldsymbol{CQ}^T.$$

The common value is the uniquely determined (so the same for all minimax profiles) *value of the game*, which we denote V. The existence of the solution is the Minimax Theorem (see [2], p. 197 for the history).

Dominance

The search for a solution is facilitated by a simple reduction process based on the Middle-Endian idea of strategic dominance. An actor's strategy A is said to be *weakly dominated* by this actor's strategy B If anything the actor can achieve using strategy A can be achieved at least as well using B, no matter what the opponent does. That is to say, for *every* counterstrategy S of the opponent, the outcome of the game with B played against S is at least as favorable as the outcome of the game with A played against S. If A and B are pure strategies of Connie, the dominance simply means that if row A of the payoff matrix has 1 in some column then row B also has 1 in this very column.

The principle of eliminating dominated strategies takes the form of a theorem, which says that reduction of the game matrix by removal of weakly dominated pure strategies (rows or columns) does not affect the value of the game. It is a good exercise to derive the principle from the Minimax Theorem. This enables us to reduce $\boldsymbol{C}$ by noticing that 1SS dominates 2SN and 2NN:

1SS	0	0	1	1	1	1
2SN	0	0	0	1	1	1
2NN	0	0	1	1	0	0

and that 3SS dominates 2NS:

3SS	1	1	1	1	0	0
2NS	1	1	1	0	0	0

Similarly, all YNS, YNN, and YSN strategies are dominated for $Y = 1, 2, 3$. After row elimination, the original game matrix $\boldsymbol{C}$ is reduced to a smaller matrix:

	1L	1R	2L	2R	3L	3R
1SS	0	0	1	1	1	1
2SS	1	1	0	0	1	1
3SS	1	1	1	1	0	0

Note that the strategies involving Nonswitch action are all gone!

Continuing the reduction process, we observe that columns XR and XL of the reduced matrix are identical for $X = 1, 2, 3$, hence, using dominance, now from the perspective of Monte, the matrix can be further reduced to the square matrix $\boldsymbol{c}$:

	1L	2L	3L
1SS	0	1	1
2SS	1	0	1
3SS	1	1	0

Mismatching the Caves and the Monte Crawl Solution

The matrix c is the structure of payoffs in the Big-Endian theory, concerned with the constant-action strategies. Each actor has only three pure strategies. Monte and Connie simultaneously choose caves; if the choices *mismatch* Connie wins, otherwise there is no payoff.

The game has no solution in pure strategies (saddle point), thus we turn to the actors' mixed strategies p, q, which we write as vectors of size three. One may guess and then check that if Connie plays the mixed strategy with probability vector $p^* = (1/3,\ 1/3,\ 1/3)$, then her probability of winning is 2/3, no matter what Monte does. It is sufficient to check this for three products $p^* c q^T$, where q is one of Monte's pure strategies, (1,0,0), (0,1,0), (0,0,1). Similarly, if Monte plays the mixed strategy $q^* = (1/3,\ 1/3,\ 1/3)$, then Connie's winning probability is 2/3 no matter what she does. Connie can guarantee winning chance 2/3, and Monte can guarantee that the chance is not higher, therefore $V = 2/3$ is the minimax value of the reduced game, that is,

$$\max_p \min_q pcq^T = \min_q \max_p pcq^T = p^* cq^{*T} = 2/3.$$

Instead of guessing the minimax probability vectors p^* and q^* one could use various computational techniques found in Chapter 3 of Course on Games [4]. Here, we mention one insightful transformation that should convince the reader that the solution is correct. Subtracting from c the constant matrix with all entries equal to 1 reduces to the game with diagonal matrix

	1L	2L	3L
1SS	-1	0	0
2SS	0	-1	0
3SS	0	0	-1

This corresponds to the variant of the game in which Connie comes to the show with her goat and risks losing the pet to Monte in the case of match, when her original guess falls on the cave hiding the goat. A similar 2×2 matrix is the familiar game of matching pennies, with Monte winning in the event of match (the Quickygnosis article "Matching Pennies" is a good reference).

Going back to the original matrix C, we conclude that $V = 2/3$ is the value of the game, and that the profile

$$P^* = \left(\frac{1}{3}, 0, 0, 0, \frac{1}{3}, 0, 0, 0, \frac{1}{3}, 0, 0, 0\right),$$

$$Q^*_{1,1,1} = \left(\frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3}, 0\right)$$

is a solution to the game. The subscript of $Q^*_{1,1,1}$ will be explained soon. According to this solution, Monte plays the Monte Crawl strategy: he hides the goat uniformly at random, and he always reveals the Left cave when there is freedom for the second action. Connie selects cave Y uniformly at random and always plays Switch.

A feature of this solution is that the preference of Monte to the Left cave sometimes gives strong confidence to Connie's decision. When Monte reveals the Right cave, he signals that the Left could not be opened, so Connie learns the location of the goat and her Switch action bears no risk.

All Minimax Solutions

The reader has certainly noticed that strategy $Q^*_{1,1,1}$ disagrees with the Little-Endian postulate of Monte's fair-coin-flipping in the event of match, $X = Y$. Monte's behavior *at random* corresponds to the uniform distribution over all possible choices,

$$Q^*_{\frac{1}{2},\frac{1}{2},\frac{1}{2}} = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right).$$

According to $Q^*_{\frac{1}{2},\frac{1}{2},\frac{1}{2}}$, Monte hides the goat uniformly at random, and for the second choice between Left and Right if there is freedom a fair coin is flipped. This strategy is also minimax. Furthermore, instead of crawling to the Left cave Monte could be hauled to the Right cave, when there is a choice option.

What are *all* the minimax strategies of either actor? One way to answer this question is to trace back what was lost in dominated strategy elimination. By column elimination we may delete either of the two pure strategies XL, XR for each $X = 1, 2, 3$. This yields eight minimax solutions $Q^*_{0,0,0}$, $Q^*_{0,0,1}, \ldots, Q^*_{1,1,1}$, where in position $X = 1, 2, 3$ of the index we write 0 if XL is never used, and we write 1 if XR is never used. Each of these solutions is a deterministic behavior – either crawling or hauling – depending on the cave by which the match has occurred; these can be considered as mixed-action strategies of *Monte*, as opposed to his constant-action strategies of always-crawling or always-hauling.

Mixtures of these minimax strategies are again minimax, and each such mixture can be uniquely represented in the form

$$Q^*_{\lambda_1,\lambda_2,\lambda_3} = \left(\frac{\lambda_1}{3}, \frac{1-\lambda_1}{3}, \frac{\lambda_2}{3}, \frac{1-\lambda_2}{3}, \frac{\lambda_3}{3}, \frac{1-\lambda_3}{3}\right)$$

where $0 \le \lambda_X \le 1$. The parameter λ_X has a transparent interpretation: this is the conditional probability that Monte will reveal the Left cave when the goat is hidden behind X and a match $Y = X$ occurs.

The subclass of Monte's strategies with the second action independent of the first given $X = Y$ consists of strategies with equal probabilities $\lambda_1 = \lambda_2 = \lambda_3$. This mode of Monte's behavior appears as *Version Five* in Mondee Book [17]. More general strategies $Q^*_{\lambda_1,\lambda_2,\lambda_3}$, were also considered in the extended Little-Endian framework [21].

We need to check further whether some minimax strategies for Connie were lost in the course of row elimination. This verification is necessary because the deleted dominated strategies YNN, YNS, and YSN are only *weakly* dominated, meaning that in some situations they perform equally well as the strategies Y'SS that dominate them. Examples of games can be given showing that weakly dominated strategies may be minimax (see Course on Games, Section 2.6, Exercise 9).

Recall that a best response is a strategy optimal for an actor knowing which particular strategy the opponent will use. Every minimax strategy P is necessarily a best response to every minimax strategy for Monte, yielding expected payoff equal to the value $PCQ^{*T}_{\lambda_1,\lambda_2,\lambda_3} = 2/3$. Suppose for the time being that minimax strategy $\boldsymbol{P}$ assigns nonzero probability $p > 0$ to the pure strategy 2SN, and let $\boldsymbol{P}'$ be a strategy obtained from $\boldsymbol{P}$ by removing the 2SN-component but adding weight p to the 1SS-component. Recalling the pattern

	1L	1R	2L	2R	3L	3R
1SS	0	0	1	1	1	1
2SN	0	0	0	1	1	1
2NN	0	0	1	1	0	0

we obtain

$$P'CQ^{*T}_{\frac{1}{2},\frac{1}{2},\frac{1}{2}} = PCQ^{*T}_{\frac{1}{2},\frac{1}{2},\frac{1}{2}} + \frac{p}{6} > PCQ^{*T}_{\frac{1}{2},\frac{1}{2},\frac{1}{2}},$$

which means that $\boldsymbol{P}'$ strictly improves $\boldsymbol{P}$ in combat against the minimax strategy $Q^*_{\frac{1}{2},\frac{1}{2},\frac{1}{2}}$. But this is in contradiction with the assumed minimax property of $\boldsymbol{P}$, and thus 2SN cannot have positive probability in $\boldsymbol{P}$. In the same way it is shown that 2NN does not enter $\boldsymbol{P}$, and by symmetry among the caves we conclude that none of the dominated strategies enters $\boldsymbol{P}$. Thus nothing was lost by the row elimination.

A crucial property of $Q^*_{\frac{1}{2},\frac{1}{2},\frac{1}{2}}$ we just used is that this strategy gives nonzero probability to each of the six pure strategies for Monte. A mixed strategy $\boldsymbol{Q}$ may be called *fully supported* if every pure strategy has a positive probability in $\boldsymbol{Q}$. In particular, $Q^*_{\lambda_1,\lambda_2,\lambda_3}$, is fully supported if and only if λ_X is distinct from 0 and 1 for $X = 1, 2, 3$. A best response to a fully supported strategy cannot be weakly dominated, hence minimaxity of *some* fully supported strategy of Monte precludes minimaxity of *every* weakly dominated strategy for Connie.

To compare, let us examine the Monte Crawl strategy $Q^*_{1,1,1}$, which always reveals the Left cave when there is a match. The pure strategy 1NS is a best response to $Q_{1,1,1}$, with the winning chance 2/3, as for any other minimax strategy of Connie. If Connie were ensured that Monte will play $Q^*_{1,1,1}$ then she may, in principle, choose 1NS. However, 1NS versus $Q^*_{1,1,1}$ would be an unstable profile, since Monte will then drop Connie's winning chance by swapping to $Q^*_{\frac{1}{2},\frac{1}{2},\frac{1}{2}}$ (a best response against 1NS).

We summarize our analysis of the zero-sum game in the following theorem:

Theorem *The strategy P^*, which is the uniform mixture of* 1SS, 2SS, 3SS, *is the unique minimax strategy for Connie. Every strategy $Q^*_{\lambda_1,\lambda_2,\lambda_3}$ with $0 \le \lambda_X \le 1$, $(X = 1, 2, 3)$ is a minimax strategy for Monte. The value of the game is $V = 2/3$.*

We see that in the setting of zero-sum games any rational behavior of Monte keeps Connie away from employing strategies with Notswitch action. Professor [5] and Magister [7] already knew that $V = 2/3$ and that throwing symmetric three-sided dice by each of the actors is involved in the solution of the MGG. Our contribution is that the Holy Grail result appears as the first application of the Middle-Endian doctrine of dominance.

The potential of the approach has been further explored by the author [11, 12, 13, 14]. In particular, the dominance implies that it is impossible to attribute probabilities to the variables outside of Connie's control in such a way that *some* strategy utilizing notswitching action will give (strictly) higher winning probability than *every* always-switching strategy.

Afterword

The great mathematical minds of the past were not immune to the fallacies of symmetry. The History of the Doctrine of Chances (see [20], p. 48) tells us of Leibniz, who argued that since both 12=6+6 and 11=5+6 can be achieved in only one way, with two dice it is as feasible to throw a total 12 points as to throw a total 11. The two-child paradox and the two-envelopes paradox (consult Quickygnosis) are also of the kind where symmetry is confusing. Most of these "paradoxes" are resolved by adequately setting up the sample spaces and events.

The mystery of switch seems to go beyond just that, and can be only loosely compared with throwing dice, where the uniform distribution is objectively justified by the physical structure of the artifact. Human behavior is a more complex matter, and the assumption of uniform distribution and randomness at all (in the objective, frequentist sense) needs justification, especially when it comes to interaction. Given the model that the actors only care about whether the goat is found or not, the rationality incorporated in the idea of minimax indeed gives a justification for the uniform randomizations both for hiding the goat and the first guess, but it offers no ground for assuming anything about Monte's behavior when he can choose between two empty caves. Moreover, if for some reason Monte exploits a biased random device to hide the goat, Connie still has an optimal always-switching counter-strategy whose performance does not depend on how Monte reveals one of two empty caves.

The world is wrong about the Mondee Gills Game, some people say. Something certainly goes wrong when Laiemen first do not believe that the odds are unequal, then hasten to explain how stupid one is not to see this. What makes the author sceptical about many such explanations is the amazing story about Paul Erdős [16], where in the chapter "Getting the Goat" we read:

Vázsonyi wrote out a "decision tree," not unlike the table of possible outcomes that vos Savant had written out, but this did not convince him. "It was hopeless," Vázsonyi said. "I told this to Erdős and walked away. An hour later he came back to me really irritated. 'You are not telling me why to switch,' he said. 'What is the matter with you?' I said I was sorry, but that I didn't really know why and that only the decision tree analysis convinced me. He got even more upset." Vázsonyi had seen this reaction before [$\cdots$] but he hardly expected it from the most prolific mathematician of the twentieth century.

This sounds unbelievable, but his *why* and later pages of the memoirs suggest that the major concern of Erdős was the lack of *The Book* proof. Using the dominance we do not need *any* probability assumptions in order to discard notswitching, so we hope with this argument we are some steps closer to what Erdős might have wanted to see.

REFERENCES

[1] Dekking, F. M., Kraaikamp, C., Meester, L. E., and Lopuhaä, H. P., *A modern introduction in probability and statistics: understanding why and how*, Springer, 2005.

[2] Ethier, S. N., *The doctrine of chances: probabilistic aspects of gambling*, Springer, 2010.

[3] Ferguson, T. S., (1989) Who solved the secretary problem? *Statistical Science* **4**, 282–296.

[4] Ferguson T., (2000) *A course on game theory*, http://www.math.ucla.edu.

[5] Häggström, O., *Streifzüge durch die Wahrscheinlichkeitstheorie*, Springer, 2005.

[6] Gill, R., Three door problem...-s, http://arxiv.org/abs/1002.3878, version March 1, 2010.

[7] Gill, R., (2011) The Monty Hall problem is not a probability puzzle (it's a challenge in mathematical modelling), *Statistica Neerlandica* **65**, 58–71.

[8] Gill, R., Monty Hall problem, http://en.citizendium.org/wiki/Monty_Hall_problem, http://statprob.com/encyclopedia/MontyHallProblem.html, http://www.math.leidenuniv.nl/~gill/mhp-statprob.pdf.

[9] Gill, R., The Holy Grail of Monty Hall studies (and other comments) http://www.math.leidenuniv.nl/~gill/.

[10] Gnedin, A., (1994) A solution to the game of googol, *Annals of Probability* **22**, 1588–1595.

[11] Gnedin, A., (2011) The Monty Hall problem in the Game Theory class, `arxiv.org`.

[12] Gnedin, A., (2011) The Doors, http://arxiv.org/abs/1105.5809.

[13] Gnedin, A., (2011) The Mondee Gills Game: strategic dominance for dummies, preprint.

[14] Gnedin, A., (2011) Dominance in the Monty Hall problem, http://arxiv.org/abs/1106.0833.

[15] Grinstead, C. M., and Snell, J. L., *Grinstead and Snell's introduction to probability*. http://www.math.dartmouth.edu/prob/prob/prob.pdf, 2006 (online version of Introduction to Probability, 2nd edition, published by the Amer. Math. Soc.).

[16] Hoffman, P. *The man who loved only numbers*, Fourth Estate, London, 1998.

[17] Rosenhouse, J. *The Monty Hall problem*, Oxford University Press, 2009.

[18] Rosenthal, J. S., (2008) Monty Hall, Monty Fall, Monty Crawl. *Math Horizons*, September issue, 5–7.

[19] Tijms, H., *Understanding probability*, Cambridge University Press, 2007.

[20] Todhunter, I., *A History of the mathematical theory of probability from the time of Pascal to that of Laplace*. Chelsea Publishing, 1865.

[21] http://en.wikipedia.org/wiki/Monty_Hall_problem, http://www.scribd.com/doc/40800646/Recreational-Math.

[22] http://en.wikipedia.org/wiki/Wikipedia:Arbitration/Requests/Case/Monty_Hall_problem.

Math Is My *Femme Fatale*

KATHARINE MEROW

The Viewpoint *column offers readers of* The Mathematical Intelligencer *the opportunity to write about any issue of interest to the international mathematical community. Disagreement and controversy are welcome. The views and opinions expressed here, however, are exclusively those of the author, and the publisher and editors-in-chief do not endorse or accept responsibility for them.* Viewpoint *should be submitted to one of the editors-in-chief, Chandler Davis and Marjorie Senechal.*

It was cold the day I ended it. Cold in a way that had less to do with the bone-chilling damp than the indifference of the Emerald City to my woes. A more sympathetic Seattle would have quelled her midwinter tear fest. She might have mustered a sunbeam or two to boost my dragging spirits. But no. Seattle wept that day. And it wasn't the sort of constant, mild crying a man can learn to ignore. Seattle wept that day with the fitful violence of a jane cast off before she wants to be. And Seattle wasn't even the one I was quitting.

I was quitting mathematics. It was January 4, 2007, and I had dropped out of math graduate school into a scene straight from a Raymond Chandler novel. Periodic glows of rain-streaked lamplight punctuated the gathering darkness as I slogged my way home from the University of Washington campus. The skinny teens and twenty-somethings who, with grimy hair and scary dogs, plead pennies from passersby on the U District's main drag had retreated into recessed doorways, the weather was so vile. Seattleites who did have homes to go to were hurrying toward them, withdrawn into the cavernous hoods preferred in that part of the country to umbrellas.

I credit the dark wetness of that day with landing me in a noirish frame of mind, a mental state where my sorry situation began to make a sort of twisted sense. Maybe I wasn't a private dick, maybe I hadn't killed or blackmailed anyone, but there under the drear Seattle sky I could sympathize with the classic Chandler protagonist. We shared a vulnerability, that hard-boiled detective and I, an unfortunate weakness for a dangerous beauty. Math, it dawned on me the day I left grad school—math was my *femme fatale*.

I can date my fascination with mathematics to October of 2003. I was a sophomore at Swarthmore College. The power of calculus had impressed me in high-school—who would have guessed that a mere teenager could calculate the volume of a solid as curvaceous as a bombshell's hourglass?—but I never raved about a mathematical idea until real analysis, my third semester at Swat. Countability of sets. Cantor's diagonalization argument. Transfinite numbers.

In these I appreciated for the first time the pristine, ethereal beauty mathematicians were always prattling on about. It's easy to overlook, I realized: subdued where a call girl is flashy, retiring where a flirt is forward. I didn't see just beauty in mathematics that Fall, either. Suddenly she was truth, too, and certitude. Every abstract positive I could think of, rolled into one ineffable loveliness. No denying it: I was dizzy with the dame.

I put in papers to become a math major.

Math was not easy, of course. This doll took work, and even work didn't always pay. She was finicky. For hours I'd get nowhere, trying every trick in the book, and then some. Then I'd run a routine that hadn't a prayer in the world...and, BAM, something would click. I couldn't explain how I'd made the breakthrough, couldn't reliably stage another, but the conquest still made me feel like a million bucks.

For an hour or two, anyway. No number of seemingly stroke-of-luck triumphs could dispel my self-doubt, or banish my insecurities. I had trouble sleeping the night before each big date, each test. Wedged into the angle formed by my fourth floor dorm room's knee wall and slanting roofline, I'd toss and turn. I'd wake up in a sweat, feverish, haunted by dreams of mathematics. All the ways I might do her wrong, all the ways she might trick me. For restless hours my head would spin with outlandish, insoluble math problems.

Math had been the death of men, I'd heard. Driven them mad. With a math teacher for a father, I'd grown up on stories of mathematicians, their fateful entanglements with a discipline as perilous as it is beguiling. Cantor got so frustrated when he couldn't manage to put his finger on one particular part of the mathematical edifice that he ended up in a sanatorium. He'd explored the very core of mathematics, penetrating there a paradise where none had gone before, but it wasn't enough. Crazy cat had his heart set on that one

DOI 10.1007/s00283-011-9268-6

unreachable secret. And Gödel, he was so upset when he discovered that math wasn't as perfect as he'd thought that he couldn't cope. Starved himself.

I didn't as a college student have all the brashness of youth, didn't interpret cautionary tales of mathematics-induced insanity as challenges. I didn't feel upon hearing about Gödel and Cantor that I had to prove something, that I had to have a go at the ensnarer, danger be damned. Even so, my dad saw what math was doing to me, and he worried.

He sent me a cartoon at college once: *FoxTrot*. It showed a tow-headed Peter Fox hard at work on a geometry assignment. In the strip's lone frame a triangle hung above the boy's desk, two side-lengths equal and with the included angle, menacingly pointy, seeming to threaten the structural integrity of Peter's skull. "Damocles meets isosceles," the caption read, but my father did not let the cartoon speak for itself. My involvement with mathematics had me on edge, he wrote in the letter that accompanied the clipping, always nervous that, due to the objectivity of the subject, its uncompromising rigor and exactitude, one misstep—one ill-timed brain-freeze—could spell my doom. By registering for a math class each semester, my father implied, I was consigning myself to near-constant anxiety.

I didn't admit even to myself the potential destructiveness of my relationship with mathematics, but I was, in reality, sick about her. I didn't know whether I could measure up or what I would do if I didn't. Sophomore Spring, the same semester I declared my math major, I started throwing up nights. I'd have an episode every few months, sometimes accompanied by an outbreak of hives. I'd wake burping a foul taste into my mouth and hug the cold white of the toilet bowl until morning, squatting on a tile floor in a metal stall under unforgiving fluorescent lights. I'd retch out the previous day's meals, reversed: dinner, an uneasy hour or two of shut-eye; lunch, another fitful rest; then breakfast. By sunrise the intestinal spasms could bring up nothing but a bilious fluid. Bleary-eyed and empty, I'd crave only sleep—the big one or otherwise.

And once I started graduate school, the Fall after I got my BA, the frequency of illness only increased. During the two weeks leading up to my first quarter's finals, I threw up everywhere and often: in the middle of topology class, during the office hours I held as a calculus teaching assistant, several times a night in my apartment. Once, when the irrepressible queasiness swept over me as I trudged home, I left what little food I'd managed to keep down that day splashed on a Seattle sidewalk. The rain would wash it off.

Rain was falling with a fiendish persistence the day I finally ended it, the day I informed the graduate advisor of my decision to take a leave of absence I expected would be permanent. I arrived at my apartment bedraggled and with nothing to cheer me but the coins I had spotted glinting in the gutter and had picked up to fund in some small way my uncertain future. As I struggled in vain to unlock a dead-bolted front door, I cut my hand on the broken innards of my supposedly gale-proof umbrella, rendered razor-sharp by one of the evening's more vicious winds.

Hand bleeding, I navigated the narrow passage between the house and its squat neighbor, sidestepping puddles and stumbling over empties come loose from my landlady's makeshift wall of wine bottles. In the backyard a stand of bamboo loomed above a mildewed kiddie pool. A motion-activated floodlight clicked on as I approached.

I stood moments later in the doorway of a large and largely empty room. A traffic signal of indicators blinked from a battery of consumer electronics sprawled beneath the window, playing eerily over the worn parquet floor. Water pooled at my feet. Blood gummed my fingers, and my pockets jingled with dough that wasn't mine.

After that noirish day, I tried to distance myself from mathematics. I tutored high-schoolers, sure, but the quadratic equations and routine applications of triangle trigonometry I encountered daily in that capacity bore little resemblance to the dynamically beautiful mathematics that had lured me off to grad school. Eventually I moved across the continent to pursue a writing degree, to court a muse a little less exacting. Before the trip back East, I sold many of my math books. I was done with math.

Not really, of course. A *femme fatale* doesn't lose her charm just because you'd like to shake the hold she has on you. I broke with mathematics more than four years ago, but when I meet one of her current lovers—a topology professor, say—I'm jealous of the intimacy he enjoys. I spend public lectures greedily absorbing new insights into the subtleties of mathematics. I file them away for use in future engagements. Logic puzzles and test prep exercises simultaneously attract and repulse me: They remind me of the creative mathematics I came to love and loathe; they disgust me with their rank inferiority. Nothing can compare to math. I need her in my life, and I'll contrive somehow to get her.

Curves like hers, honey, they're to die for.

AUTHOR

KATHARINE MEROW received a master's degree in nonfiction writing from Johns Hopkins University in December, 2011. She is adrift anew and in grave danger, once again, of falling prey to mathematics.

3217 Connecticut Ave. NW
Apartment 45
Washington, DC 20008
USA
e-mail: kmerow@gmail.com

A Letter of Hermann Amandus Schwarz on Isoperimetric Problems

URS STAMMBACH

Years Ago features essays by historians and mathematicians that take us back in time. Whether addressing special topics or general trends, individual mathematicians or "schools" (as in schools of fish), the idea is always the same: to shed new light on the mathematics of the past. Submissions are welcome.

➢ Send submissions to **David E. Rowe,**
Fachbereich 08, Institut für Mathematik,
Johannes Gutenberg University,
D-55099 Mainz, Germany.
e-mail: rowe@mathematik.uni-mainz.de

In 1995 (or thereabouts), Eli Maor, then at Loyola University, discovered a little booklet at a used book fair in Chicago containing a reprint of a paper published by Hermann Amandus Schwarz in 1884 [Sch1]. Interestingly, it also contained an original letter by Schwarz written in the old German script. The letter, dated January 28, 1884, began with the words: *Hochgeehrter Herr Director!* Maor could decipher enough of the text to realize that it dealt with isoperimetric problems, which further awakened his interest. With the help of Reny Montandon and Herbert Hunziker, he contacted Günther Frei, who was able to read and transcribe the old German script. Thanks to his detailed knowledge of the history of mathematics of the 19th century, Frei was also able to provide a number of pertinent remarks. Unfortunately, before completing this task, Günther Frei fell gravely ill. The text of the letter together with his remarks was then given to the present author, who agreed to continue the work.

We begin with some background information about Schwarz and the state of research on isoperimetric problems and minimal surfaces. Schwarz became interested in this area of investigations during his student days in Berlin. We next turn to Schwarz's letter itself, given here in an English translation. To make it easier to understand, we provide a number of supplementary explanations, mostly biographical data about the various people Schwarz mentions. Among them, Edvard Rudolf Neovius merits special attention. In the final section we describe the way this booklet, together with Schwarz's letter, found its way from Berlin to Chicago; this in itself is an interesting and surprising story. In an appendix we present a facsimile of Schwarz's letter (courtesy of Eli Maor) along with a German transcript.

To acknowledge Günther Frei's work and to honor him, it is the wish of all involved to dedicate this paper to him. Reny Montandon, who long followed our work on Schwarz's letter with great interest, unexpectedly died in the spring of 2011; it is sad that he did not live to see the publication of this paper.

Schwarz and Isoperimetric Problems

Hermann Amandus Schwarz (1843–1921) took his doctoral degree with Ernst Kummer in 1864 and his habilitation at the University of Berlin in 1866. The next year, he was appointed associate professor in Halle, and in 1869, at just 26 years of age, he became professor at the *Eidgenössische Polytechnikum* (now called *Eidgenössische Technische Hochschule*, ETH) in Zürich. In 1875, he obtained a professorship in Göttingen, and in 1892 succeeded Karl Weierstrass at the University of Berlin, where he remained until 1917. He died in Berlin in 1921. (See [FS], p. 73.)

Although Kummer was Schwarz's thesis adviser/examiner, his work was more closely related to the analytic interests of Weierstrass, and Schwarz always regarded himself chiefly as Weierstrass's pupil. It was during those years

DOI 10.1007/s00283-011-9267-7

that Weierstrass began his famous program for building a solid foundation for analysis based on rigorous proofs for the many results he felt had been obtained by dubious reasonings. Closely related to this program were various results on isoperimetric problems and minimal surfaces that his former colleague, Jakob Steiner, had proved by appealing to the methods of synthetic geometry. After Steiner died in 1863, Weierstrass took on the additional task of teaching courses on synthetic geometry, upholding the Steinerian tradition in Berlin. It is easy to understand why Weierstrass took a keen interest in providing analytic and fully rigorous proofs of Steiner's results. Nor is it surprising that his pupil Hermann Amandus Schwarz also began to work in this same direction.

Some of Schwarz's later activity at the *Eidgenössische Polytechnikum* in Zürich sheds further light on this unusual situation. When Schwarz came to Zürich in 1869, he succeeded Elwin Bruno Christoffel, thereby assuming the most prestigious professorship in mathematics at the *Polytechnikum*. His principal task was teaching the large course in differential and integral calculus, the introductory mathematics course required for all beginning students. However, each semester he also taught two or even three more specialized courses on, for example, ordinary differential equations or topics of complex function theory. Some were introductory in nature, but others covered rather advanced topics, such as elliptic functions and their applications. Occasionally he also lectured on fields such as number theory, and in the summer term of 1871 he offered a course on *Analytische Geometrie der Raumkurven und der krummen Flächen*. He was apparently quite successful in attracting students to mathematics, despite the fact that most of his auditors were prospective engineers.

Several eminent names in mathematics have been associated with the *Eidgenössische Polytechnikum*, so Schwarz was hardly alone (see [FS], pp. 71–74). Wilhelm Fiedler was responsible for teaching descriptive geometry. This subject played a central role in the education of engineers, so Fiedler's position, like Schwarz's, carried high prestige. Fiedler was a former pupil of August Ferdinand Moebius; compared with his considerably younger colleagues, he was a rather conservative mathematician. Shortly after Schwarz's arrival, Heinrich Weber was appointed to a chair; like Schwarz, he left Zürich in 1875 to take a position in Germany. The Swiss mathematician Ludwig Stickelberger joined the faculty of the *Polytechnikum* in 1873, even before he took his doctorate in Berlin under Weierstrass. After Schwarz left Zürich, Stickelberger became a close collaborator of his successor, Ferdinand Frobenius, another Berlin product.

For Schwarz, there can be little doubt that his single most important colleague was another Swiss, Carl Friedrich Geiser (1843–1934), who was a grandnephew of Jacob Steiner. Geiser was appointed to a professorship at the *Polytechnikum* in 1869. As a student he spent some time in Berlin with his great-uncle before returning to Switzerland. In Bern he studied under Ludwig Schläfli, a Steiner pupil, and in 1866 Geiser became his first doctoral student. Little wonder that Geiser later went on to compile and edit Steiner's unpublished manuscripts and lecture notes. Steiner was also suitably remembered in 1897 at the First International Congress of Mathematicians in Zürich, an event for which Geiser was a major driving force, both as initiator and organizer (see [FS], pp. 11–13 and p. 74).

During the time Schwarz was in Zürich, Geiser regularly taught courses on synthetic geometry, presumably following the spirit of Steiner's courses rather closely. One can easily imagine that Geiser's close relationship to his great-uncle led to some heated mathematical discussions with Schwarz, who was just as devoted to the rigorous analytic approach championed by Weierstrass. These circumstances help shed light on the remark about Geiser in Schwarz's letter, a remark that is very interesting in many respects. There Schwarz refers to Jacob Steiner's Collected Works [St1], which were edited by Weierstrass in 1882, and in particular to a *Mittheilung* on pages 728 and 729 of Volume 2. On pages 727–729 under the heading "Anmerkungen und Zusätze" (Comments and Additions) we there find some comments by Weierstrass on a paper by Steiner, entitled "Aufgaben und Lehrsätze" (Problems and Theorems). In these comments Weierstrass mainly quotes from a related handwritten table – and apparently also some notes – which were given to him by Geiser and which he explicitly attributes to Steiner. With the remark in his letter Schwarz makes clear that the quote (two paragraphs) in the latter part of Weierstrass's comments is not due to Steiner but to Schwarz himself! And he adds that Geiser has taken responsibility for this mistake.

The classical isoperimetric problem,[1] with which Schwarz's letter is concerned, consists in proving that of all

AUTHOR

URS STAMMBACH received his doctoral degree in 1966 at the ETH in Zürich under the supervision of Beno Eckmann. After two years at Cornell University in Ithaca, New York, he acquired a professorship at the ETH in Zürich in 1969, and since 2005 has enjoyed emeritus professor status there. Urs Stammbach is the author and coauthor of numerous research articles in homological algebra and group theory and of several books, among them *A Course in Homological Algebra* (jointly with Peter Hilton) and *Mathematicians and Mathematics in Zurich, at the University and at the ETH* (jointly with Günther Frei). His interests outside of mathematics include history, literature, and music.

Mathematik, LEO D1
ETH-Zentrum
CH 8092 Zürich
Switzerland
e-mail: stammb@math.ethz.ch

[1]For details about the isoperimetric problem and its history, see [B] or [CR], Chapter VII, § 8.

Figure 1. H. A. Schwarz (from Hermann Amandus Schwarz zu seinem 50-jährigen Doktorjubiläum, Springer, 1914) and his interlocutors: Jacob Steiner, Karl Weierstrass, Carl Friedrich Geiser, and Edvard Rudolf Neovius (courtesy of the Hamburg Mathematische Gesellschaft, http://www.math.uni-hamburg.de/home/grothkopf/fotos/math-ges/). H. A. Schwarz was drawn to investigate isoperimetric problems as a student in Berlin. There he learned that the famous synthetic proofs of Steiner were faulty and spoke with Weierstrass about analytic methods for proving them. In Zürich, Schwarz discussed these matters with his colleague, Geiser, a relative of Steiner, as well as with Neovius, a student from Finland who was an expert on minimal surfaces.

curves of fixed length, the circle bounds the largest area in the plane. Many eminent mathematicians tried to solve this problem, but none succeeded completely. In 1841 Jacob Steiner presented his proof that the circle has the required isoperimetric property in two long papers published under the title *Sur le maximum et le minimum des figures dans le plan, sur la sphère et dans l'espace en général* (see [St2]). There Steiner dealt not only with the original isoperimetric problem, but also with several generalizations, including analogous questions for curves on spheres as well as the isoperimetric problem in 3-dimensional space: which surface bounds the largest volume? Soon after their publication, several mathematicians criticized the proofs as incomplete.

Steiner's colleague at the University of Berlin, Peter Lejeune Dirichlet, criticized Steiner for presupposing the existence of a solution for the problems in question. Steiner's proof proceeds by showing that a given curve that is not a circle fails to satisfy the required condition, since it is always possible to construct another curve of the same length that bounds a larger area. Clearly, for such an argument to be rigorously valid, one must somehow ensure that a curve actually exists that maximizes the area it bounds. Many other, perhaps less important, points were raised by Weierstrass, Schwarz, and also by Friedrich Edler and Rudolf Sturm.

During the 1870s, Weierstrass apparently presented a series of lectures that aimed to give rigorous proofs of the isoperimetric property of the circle as well as other results of Steiner. For whatever reasons, though, he never published his proofs, so the problem remained open until 1884 when Schwarz unveiled his paper, *Beweis des Satzes, dass die Kugel kleinere Oberfläche besitzt, als jeder andere Körper gleichen Volumens* (see [Sch1]). This is the paper contained in our booklet. As the title suggests, the larger part of this article is devoted to the isoperimetric problem in dimension three.

In the technical portion of the letter that follows, Schwarz critisizes one of Steiner's proofs. His remarks, however, do not point to an actual mathematical error, such as the one Dirichlet noted many years previously. Schwarz only describes a striking simplification that could be made in one of Steiner's arguments; it is an idea so simple that he was surprised Steiner could have overlooked it.

Schwarz's Letter

The original letter consists of three pages written on a single folded sheet of paper in the old German script (see the facsimile of the letter in the Appendix). It should be noted that Schwarz writes an extremely prolix German; his long and complicated sentences make a literal word-by-word translation practically impossible. To make the English text readable, we have therefore broken up some of the long sentences into two or even more parts.

Goettingen, January 28 1884
(Weender Chaussée 17 A.)[2]

Most honorable director![3]
Herewith I return with my best thanks your script of Steiner's course on maxima and minima, which you have lent me so kindly for such a long time. I take the opportunity, to send you some papers of mine as well as the dissertations of two of my students, the academic work of my former auditor Dr. E. Neovius, who is now Professor at the University in Helsingfors, and two copies of my proof that the sphere has the smallest surface area among all bodies of the same volume, which I have presented to our *Gesellschaft der Wissenschaften*. I have also enclosed two copies of the text by Mr. Edler. Please accept the second copy for the library of your institution. The proof that concerns the sphere is no doubt the first for the theorem in question, that is, the first really rigorous proof, since all the other attempts must be regarded as incorrect, since they are based on unproven facts. I may be allowed to reveal to you that the writer of these lines is the author of the statement in volume 2 of the *Gesammelte Abhandlungen* of Steiner (page 728 at the bottom and 729 at the top), where it is mistakenly attributed to Steiner. Professor Geiser has already admitted that he thereby committed an error.

The truly admirable works of Steiner on maxima and minima contain, unfortunately, in addition to the inaccuracies mentioned by Sturm, quite a number of others, which I have discovered while giving public lectures on Steiner's work, which I have repeatedly done at this university.

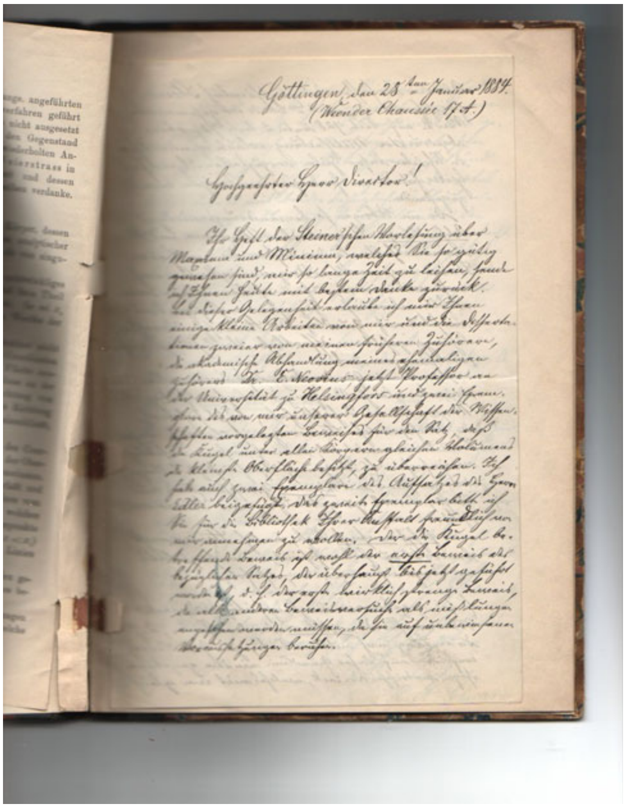
The proof of the theorem that of all triangles the equilateral triangle has the maximal area can be simplified considerably, despite the fact that Steiner has undoubtedly given much attention to his proof. It does not seem acceptable to use two intermediate steps, as Steiner does, when one intermediate step is sufficient.

The very simple thing that Steiner does not seem to have noticed – for otherwise he would have mentioned it – is the following:

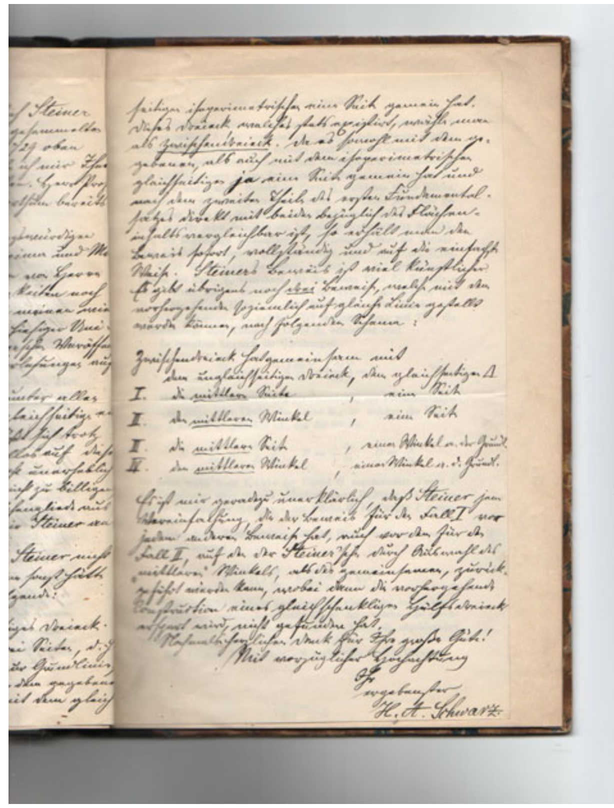
Let a triangle be given which is *not* equilateral. Choose as baseline the intermediate side, that is, the side of intermediate length; on this baseline construct a triangle which is isoperimetric to the given one and which also has a common side with the isoperimetric equilateral triangle. This triangle, which always exists, is chosen as the triangle for the intermediate step in the proof; since it has a common side with the original triangle as well as with the isoperimetric equilateral triangle, the second part of the first fundamental theorem can be applied to directly compare the areas of these triangles. This immediately provides the proof in a complete and simple way. Steiner's proof is much more complicated and it uses additional constructions. By the way, there exist three more proofs that are in the same spirit, following the scheme:

[2]Schwarz gives Weender Chaussée 17 A as his home address. This was the house next to Weender Chaussée No. 17, where Ida Riemann, the only surviving sister of Bernhard Riemann, lived together with Riemann's widow and her daughter, following Riemann's death in 1866.

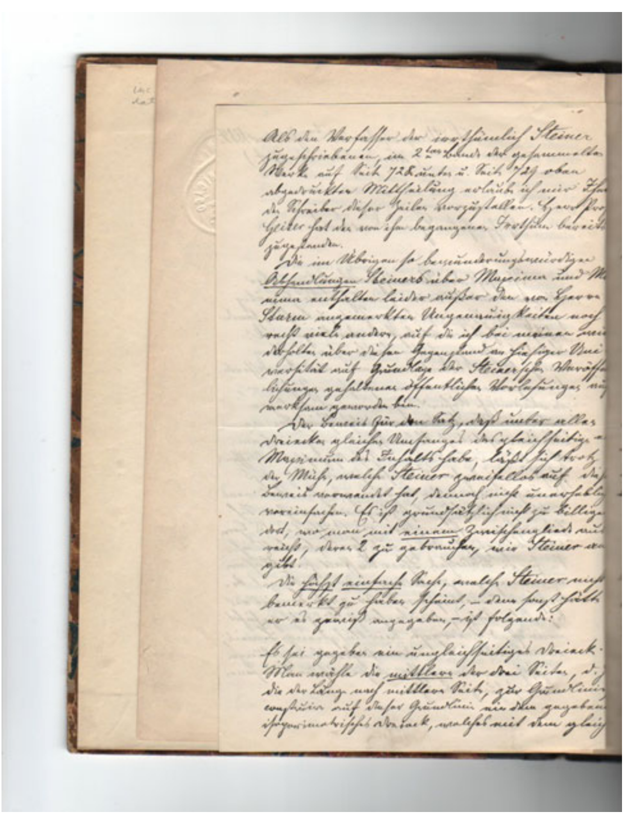
[3]We suspect that the addressee was the director of a *Gymnasium*, but we have not been able to verify this, nor have we been able to identify the person. From the first sentences of the letter, one can assume that the addressee had attended some of the courses on maxima and minima by Steiner.



Göttingen, den 28ten Januar 1884.
(Weender Chaussee 17a)



H. A. Schwarz



Intermediate triangle has in common with

	the nonequilateral triangle,	the equilateral triangle
I.	the intermediate side	one side
II.	the intermediate angle	one side
III.	the intermediate side	one angle at the baseline
IV.	the intermediate angle	one angle at the baseline

I cannot really understand why Steiner has not noticed this simplification which *case I* provides with respect to any other proof, even with respect to *case II*, to which Steiner's proof can be reduced by choosing the intermediate angle as the common one; in this way [in Steiner's proof] the intermediate step of constructing an isosceles triangle can be avoided. Again many thanks for your kindness!

With highest esteem.
Your devoted
H. A. Schwarz

It is interesting to note that in his letter Schwarz does not simply try to promote his own work – although he includes a reprint of his paper [Sch1] – but also calls attention to the work of his students and associates. In particular he mentions the work of two of his doctoral students (whom we were unable to identify) and underscores very prominently the *academic work* of Edvard Rudolf Neovius.

Edvard Rudolf Neovius (see [Elf], [L][4]) lived from 1851 until 1917. His father, Edvard Engelbert Neovius, was a teacher of mathematics at the military school in Hamina (Swedish: Fredrikshamn). Like his older brother Lars, Edvard Rudolf

[4]We are grateful to Professor Olli Lehto for making us aware of these two sources of information, and to Dr. Yvonne Voegeli from the ETH-library for helping us to uncover the traces of Neovius's stay in Zürich.

joined this school and then served in the Russian army, but he was soon dissatisfied with military life and decided to leave service for university studies. Together with his brother Lars he went to Zürich, and during the following three years, from 1871 until 1874, he studied mechanical engineering at the *Eidgenössische Polytechnikum*. After that, he spent a year at the Technical University in Dresden before returning to Zürich for a year between 1875 and 1876 (see [Elf]). Finally, he was awarded a doctoral degree in 1880 at the Kejserliga Alexander Universitetet (Imperial Alexanders-University) in Helsingfors (Helsinki) with his dissertation "Kurvor af tredje och fjerde graden betraktade sø asom alster af tvänne projektiviska involutioner" (Curves of degree three and four considered as a product of two projective involutions). His dissertation is written in Swedish, the official language at the time. From 1883 until 1900 Neovius was Professor for mathematics at the University in Helsingfors.

During his first stay in Zürich, Neovius apparently came into close contact with Schwarz. In a letter from October 2, 1882, to Angelo Genocchi in Turin, Schwarz refers to Neovius as *un de mes anciens élèves de Zurich*. Schwarz and Neovius remained in steady contact. In our letter Schwarz refers to the paper "Bestimmung zweier speciellen (sic!) periodischen (sic!) Minimalflächen, auf welchen unendlich viele gerade Linien und unendlich viele ebene geodätische Linien liegen" (Determination of two special periodic minimal surfaces on which there lie infinitely many straight lines and infinitely many plane geodetic lines) (see [N]). In this paper, which is his "Habilitationsschrift" at the University in Helsingfors, Neovius constructed a minimal surface that today bears his name. (For an image of the surface see [WW], a reproduction of the original image in [N].)

Neovius visited Schwarz several times when the latter taught in Göttingen, and Schwarz sometimes took summer vacations in Helsinki. During a visit to Göttingen in 1885, Neovius noticed that Schwarz was completely overworked and longed for some quiet time to recover. Schwarz thus joined him on the trip back to Helsinki, where he stayed in his home. It was during this stay with Neovius that he completed his famous paper *Über ein die Flächen kleinsten Flächeninhaltes betreffendes Problem der Variationsrechnung. Festschrift zum Jubelgeburtstag des Herrn Karl Weierstrass* (On a problem of the calculus of variation concerning surfaces of minimal area) (see [Sch2]), a Festschrift marking the occasion of Weierstrass's 70th birthday. Many experts regard this paper as the most important of all of Schwarz's work.

In 1897 Neovius accepted an appointment as president of an important bank, and soon afterward became a member of its board. In 1900 he resigned from the university; he became a member of the Finnish cabinet as head of the country's finance department. The first years of the new century proved to be a politically difficult time for Finland. Following a pan-Slavic course, Tsar Nicholas II of Russia sought to gain more influence over internal Finnish matters, hoping to absorb the autonomous Grand Duchy of Finland into the Russian Empire. Finland was faced with the delicate question of how to react to the unilateral measures of its powerful Russian neighbour. Some Finns favoured a cautious policy, others preferred passive or even active resistance. The clash between these two attitudes grew more and more bitter, eventually culminating with the murder of the Russian governor-general Bobrikov in 1904, followed by a general strike in the autumn of 1905. Only a few months before this, Neovius left his cabinet post in bitterness, harshly criticized by those who favoured a stronger policy toward Russia. Unable to regain his chair at the university – it was now occupied by Ernst Lindelöf – he decided to leave Finland and spend the rest of his life in Denmark, his wife's native country. There he continued to follow his mathematical interests, pursuing research on minimal surfaces, and occasionally giving courses at the university of Copenhagen.

Edvard Rudolf was born into a mathematical family: his father and his uncle were both mathematicians, and two of his brothers became mathematicians as well. Moreover, he was the nephew of Leonard Lorenz Lindelöf, who was professor of mathematics at the University in Helsingfors. The latter's son, Ernst Lindelöf, became more famous than his father as a professor at this university. It is interesting to note that the mathematical genes in the wider family of Edvard Rudolf Neovius were passed to the next generation. Around 1906 some members of the family changed their name to Nevanlinna. A brother of Edvard Rudolf, Otto Wilhelm Neovius-Nevanlinna – one of the two brothers who became mathematicians – had two sons, Rolf and Frithiof Nevanlinna. Both went on to become world-famous mathematicians.

In his letter, Schwarz also mentions a text by Friedrich Edler, probably referring to the paper [E2]. Edler gives a simplified and rigorous proof of the result that for every polygon in the plane there exists a circle with smaller circumference but larger area. The paper is a sequel to his earlier [E1], in which he gave a rigorous proof of the isoperimetric property of the circle. One of the numerous lemmas leading up to the final result states that among the polygons with $2n$ edges and given circumference, the regular polygon has the largest area. In [E2] Edler presents a simplified proof.

Friedrich Edler, born in 1855 in Mühlhausen (Thüringen), began studies of mathematics in 1877 at the University in Halle, where he received his doctorate in 1882.[5] We were unable to find any information about his later life.

Schwarz's reference to *inaccuracies* in Steiner's work on maxima and minima mentions the criticism of Friedrich Otto Rudolf Sturm. Sturm received his doctorate 1863 in Breslau, taught in Darmstadt and Münster for the next twenty years, and then returned to the University of Breslau. He published a number of papers on synthetic geometry in the spirit of Steiner and wrote several textbooks on descriptive and synthetic geometry. In his paper [Stu1] he explicitly lists several gaps in Steiner's proofs. Later he published a little booklet, *Maxima und Minima in der elementaren Geometrie* (see [Stu2]), in which he gives a complete and correct presentation of Steiner's results.

Finally, at the close of his letter, Schwarz explains his own criticism of Steiner's proof of his "first fundamental theorem" for the equilateral triangle in [St2], which reads as follows:

[5] We thank Ms Karin Keller from the *Archiv der Universität Halle* for this kind information.

De deux de ces triangles [i.e., triangles with the same baseline], *celui qui aura l'angle le plus petit ou le plus grand à la base, ou bien dont l'un des côtés sera le plus petit ou le plus grand, sera le plus petit lui-même, et réciproquement.* Steiner's proof takes about half a printed page and requires a rather complicated figure. By contrast, Schwarz sketches a direct and extremely simple proof, so simple that Schwarz found it strange that Steiner could have overlooked this argument. As noted previously, Schwarz's critique does not concern a real mistake, it merely points to sloppiness on Steiner's part.

The *Berlin Collection*[6]

A document like this long-lost letter obviously sheds new light on the history of early work on isoperimetric problems. But how did the booklet containing Schwarz's paper and letter reach the shores of Lake Michigan? Here, again, we have managed to trace a line back to the year 1891 when the University of Chicago was first founded. In that same year its president, William Rainey Harper, made a trip to Germany. In Berlin, he stopped to visit the bookshop *S. Cavaltry* at Unter den Linden 17, where he discovered a large collection of books for sale. This bookshop was owned by a certain G. Heinrich Simon, a man of advanced age who wanted to retire. According to Simon, the collection consisted of about 300,000 books and 150,000 smaller booklets, including dissertations and the like. Among them were rare and distinguished works going back to the 15th and 16th centuries, but also more recent scholarly studies in philosophy, classical philology, Greek and Roman archaeology, and the sciences.

Harper realized that this collection would make an ideal acquisition for the new library at the University of Chicago. He decided to purchase the entire collection, even though this meant finding sponsors to pay for it when he returned to the United States. The contract he signed was for 180,000 marks, or about U.S. $ 45,000, a considerable sum of money. Yet it seemed to Harper a good buy. He had contacted several specialists who confirmed that the market value of the collection was significantly higher. After his return to Chicago, he was gratified that members of the Board of Trustees of the University of Chicago privately pledged large sums in order to make the purchase possible.

Later, once the transport of the books had begun, certain difficulties arose. It seems that the contract had not specified with sufficient clarity what was meant by a "volume". In fact, the number had to be revised downward: instead of 300,000 there were actually only 120,000 books and about 80,000 booklets. It is not clear how many books were actually delivered in the end; Harper's Presidential Reports of the year 1897 to 1898 mention 175,000 volumes, but this number could not be verified later. Of course, the fact that a smaller number of books was delivered resulted in a reduction of the price: only part of the pledged 180,000 marks had finally to be paid.

The purchase of this large and precious collection of books was widely publicized all over the United States, and it received considerable attention. It was seen as a very important step in the building of a university system. The books of the so called "Berlin Collection" thus became part of the University Library in Chicago; they were marked as such with a special label. It does not seem possible to determine when the booklet containing Schwarz's publication and enclosed letter was removed from the library stacks. Yet, somehow it surfaced at a used book sale in Chicago more than 100 years after President Harper purchased it in Berlin. The handwritten letter, written in old-fashioned German script, aroused the interest of an open-eyed mathematician, a remarkable and fortunate accident for historians of mathematics.

LITERATURE

[B] Blaschke, Wilhelm: *Kreis und Kugel*. Leipzig, 1916.

[CR] Courant, Richard and Robbins, Herbert: *What is Mathematics?*, Oxford University Press, New York, 1941.

[E1] Edler, Friedrich: Ueber Maxima und Minima bei ebenen Figuren. Hoffmann's Zeitschrift X (1879), pp. 245–259.

[E2] Edler, Friedrich: Vervollständigung der Steiner'schen elementargeometrischen Beweise für den Satz, dass der Kreis grösseren Flächeninhalt besitzt, als jede andere ebene Figur gleich grossen Umfanges. Göttinger Nachrichten, 1882, pp. 73–80.

[Elf] Elfving, Gustav: *The History of Mathematics in Finland*, 1828–1918. Helsinki, 1981.

[FS] Frei, Günther and Stammbach, Urs: *Mathematicians and Mathematics in Zurich, at the University and the ETH*. Schriftenreihe A der ETH-Bibliothek. Wissenschaftsgeschichte Band 8, Zürich, 2007.

[L] Lehto, Olli: *Erhabene Welten: das Leben Rolf Nevanlinnas*. Birkhäuser Verlag, Basel, 2008.

[N] Neovius, Edvard Rudolf: *Bestimmung zweier speciellen (sic!) periodischen (sic!) Minimalflächen*. Akad. Abhandlungen, Helsingfors, 1883.

[RR] Rosenthal, Robert: The Berlin Collection. A History. http://www.lib.uchicago.edu/e/spcl/excat/berlin/history.html, last visited on January 27, 2010.

[Sch1] Schwarz, Hermann Amandus: Beweis des Satzes, dass die Kugel kleinere Oberfäche besitzt, als jeder andere Körper gleichen Volumens. Göttinger Nachrichten 1884, pp. 1–13.

[Sch2] Schwarz, Hermann Amandus: Über ein die Flächen kleinsten Flächeninhaltes betreffendes Problem der Variationsrechnung. Festschrift zum Jubelgeburtstage des Herrn Karl Weierstrass. Helsingfors 1885. Erschienen in Acta Societatis Scientiarum Fennicae 15 (1888), pp. 315–362.

[St1] Weierstrass, Karl (Ed.): Jacob Steiner's Gesammelte Abhandlungen. Two Volumes, Reimer Berlin, 1882.

[St2] Steiner, Jacob: Sur le maximum et le minimum des figures dans le plan, sur la sphère et dans l'espace en général. Premier mémoire. J. Math. Pures et Appl., VI (1841), pp. 105–170, and J. reine und angew. Math. XXIV (1842), pp. 93–152. – Second mémoire. J. reine und angew. Math. XXIV (1842), pp. 189–250. – See also the reprint following the original German manuscript: Ueber Maxima und Minima bei den Figuren der Ebene, auf der

[6]The information about the history of the *Berlin Collection* stems from the detailed account by Robert Rosenthal: *The Berlin Collection. A History*, see [RR].

Kugelfläche und im Raum überhaupt. In [St1], vol. II, pp. 177–242, 243–308.

[Stu1] Sturm, Friedrich Otto Rudolf: Bemerkungen und Zusätze zu Steiners Aufsätzen über Maxima und Minima. J. reine und angew. Math. 96 (1884), 36–77.

[Stu2] Sturm, Friedrich Otto Rudolf: *Maxima und Minima in der elementaren Geometrie*, Teubner Verlag, Leipzig und Berlin, 1910.

[WW] Weber, Matthias and Wolf, Michael: *Early images of minimal surfaces*. Bull. Amer. Math. Soc. 48 (2011), 457–460.

Appendix: German transcription of the letter

Göttingen, den 28ten Januar 1884
(Weender Chaussée 17 A.)

Hochgeehrter Herr Director!
Ihr Heft der Steiner'schen Vorlesung über Maxima und Minima, welches Sie so gütig gewesen sind, mir so lange Zeit zu leihen, sende ich Ihnen heute mit bestem Danke zurück. Bei dieser Gelegenheit erlaube ich mir, Ihnen einige kleine Arbeiten von mir und die Dissertationen zweier von meinen früheren Zuhörern, die akademische Abhandlung meines ehemaligen Zuhörers Dr. E. Neovius, jetzt Professor an der Universität zu Helsingfors, und zwei Exemplare des von mir unserer Gesellschaft der Wissenschaften vorgelegten Beweises für den Satz, dass die Kugel unter allen Körpern gleichen Volumens die kleinste Oberfläche besitzt, zu überreichen. Ich habe auch zwei Exemplare des Aufsatzes des Herrn Edler beigefügt. Das zweite Exemplar bitte ich Sie für die Bibliothek Ihrer Anstalt freundlich von mir annehmen zu wollen. Der die Kugel betreffende Beweis ist wohl der erste Beweis des bezüglichen Satzes, der überhaupt bis jetzt geführt worden ist, d.h. der erste wirklich strenge Beweis, da alle anderen Beweisversuche als mißlungen angesehen sehen werden müssen, da sie auf unbewiesenen Voraussetzungen beruhen. Als den Verfasser der irrthümlich Steiner zugeschriebenen, im 2-ten Bande der gesammelten Werke auf Seite 728 unten u[nd] Seite 729 oben abgedruckten Mittheilung erlaube ich mir, Ihnen den Schreiber dieser Zeilen vorzustellen. Herr Prof. Geiser hat den von ihm begangenen Irrthum bereits zugestanden. Die im Übrigen so bewunderungswürdigen Abhandlungen Steiners über Maxima, und Minima, enthalten leider, außer den von Herrn Sturm angemerkten Ungenauigkeiten noch recht viele andere, auf die ich bei meinen wiederholten über diesen Gegenstand an hiesiger Universität auf Grundlage der Steinerschen Veröffentlichungen gehaltenen öffentlichen Vorlesungen aufmerksam geworden bin.

Der Beweis für den Satz, daß unter allen Dreiecken gleichen Umfanges das gleichseitige ein Maximum des Inhalts habe, läßt sich trotz der Mühe, welche Steiner zweifellos auf diesen Beweis verwendet hat, dennoch nicht unerheblich vereinfachen. Es ist grundsätzlich nicht zu billigen dort, wo man mit einem Zwischenglied ausreicht, davon 2 zu gebrauchen, wie Steiner angibt.

Die höchst einfache Sache, welche Steiner nicht bemerkt zu haben scheint, – denn sonst hätte er es gewiß angegeben, – ist folgende:

Es sei gegeben ein ungleichseitiges Dreieck. Man wähle die mittlere der drei Seiten, d.h. die der Länge nach mittlere Seite, zur Grundlinie, construire auf dieser Grundlinie ein dem gegebenen isoperimetrisches Dreieck, welches mit dem gleichseitigen isoperimetrischen eine Seite gemein hat. Dieses Dreieck, welches stets existiert, wähle man als Zwischendreieck, dass es sowohl mit dem gegebenen, als auch mit dem isoperimetrischen gleichseitigen je eine Seite gemein hat und nach dem zweiten Theile des ersten Fundamentalsatzes direkt mit beiden bezüglich des Flächeninhalts vergleichbar ist, so erhält man den Beweis sofort, vollständig und auf die einfachste Weise. Steiners Beweis ist viel künstlicher. Es gibt übrigens noch drei Beweise, welche mit dem vorhergehenden so ziemlich auf gleiche Linie gestellt werden können, nach folgendem Schema: Zwischendreieck hat gemeinsam mit

	dem ungleichseitigen Dreieck,	dem gleichseitigen Δ [Dreieck]
I.	die mittlere Seite	eine Seite
II.	den mittleren Winkel	eine Seite
III.	die mittlere Seite	einen Winkel a[n] der Grundl[inie]
IV.	den mittleren Winkel	einen Winkel a[n] d[er] Grundl[inie]

Es ist mir geradezu unerklärlich, daß Steiner jene Vereinfachung, die der Beweis für den Fall I vor jedem anderen Beweis hat, auch vor dem für den Fall II, auf den der Steinersche durch Auswahl des "mittleren" Winkels, als des gemeinsamen, zurückgeführt werden kann, wobei dann die vorhergehende Construction eines gleichschenkligen Hülfsdreiecks erspart wird, nicht gefunden hat. Nochmals herzlichen Dank für Ihre große Güte!
Mit vorzüglicher Hochachtung

Ihr
ergebenster
H. A. Schwarz

Felix Hausdorff in Bonn

ROBERT JONES

Does your hometown have any mathematical tourist attractions such as statues, plaques, graves, the café where the famous conjecture was made, the desk where the famous initials are scratched, birthplaces, houses, or memorials? Have you encountered a mathematical sight on your travels? If so, we invite you to submit an essay to this column. Be sure to include a picture, a description of its mathematical significance, and either a map or directions so that others may follow in your tracks.

➢ Please send all submissions to Mathematical Tourist Editor, **Dirk Huylebrouck**, Aartshertogstraat 42, 8400 Oostende, Belgium
e-mail: huylebrouck@gmail.com

"Hausdorff remembered and forgotten," reads the cover of *The Mathematical Intelligencer*, vol. 30 (2008), no. 4. Hausdorff remembered: the brilliant mathematician and author of the groundbreaking *Grundzüge der Mengenlehre*. Hausdorff forgotten: the once well-known writer, pseudonymously Paul Mongré.

In the "Years Ago" column for that issue [1], Walter Purkert, the director of Hausdorff's *Gesammelte Werke*, reunited Hausdorff and Mongré in one remarkable man and described his astonishingly rich and varied life. Purkert also described his tragic end: Hausdorff, his wife, and her sister chose suicide in the face of internment in Bonn-Endenich and then Nazi death camps in the east. Mongré, *mon gré*, my choice.

Hausdorff's choice is still largely forgotten. I came across it recently, in the Poppelsdorf Cemetery of Bonn, Germany. It is covered with dust and green moss and small stones placed there by visitors who do remember (Fig. 1).

Hausdorff's Grave

The Poppelsdorf Cemetery in Bonn was opened early in the 19th century. Today it is one of several tended by the city department of green surface management (shrubs, grass, trees, etc.). Before the time of Nazi tyranny, the area of Bonn was noted for tolerance, and people of many faiths were interred in Poppelsdorf [2].

There are two paths that a tourist might take through the Poppelsdorf Cemetery to the grave of Felix Hausdorff. One leads through the main entrance, extends past the administration building (Fig. 2a), stops at the grave of another mathematician, Rudolf Lipschitz (Fig. 2b), and then follows an upward path to the grave of Felix Hausdorff.

Figure 1. The grave of Felix Hausdorff. Placing a stone on a tombstone is an ancient Jewish custom.

DOI 10.1007/s00283-011-9257-9

Figure 2. (a) The administration building of the Poppelsdorf Cemetery; (b) the grave of Rudolf Lipschitz.

Figure 3. (a) The path from the Baroque church to the grave of Felix Hausdorff; (b) the viewing bench with a view over Bonn.

The other route leads past a small Baroque church (Fig. 3a), then enters the cemetery through a side entrance, and pauses to look out across the city of Bonn from a viewing bench (Fig. 3b). On a clear day, one can see the twin towers of the Cathedral of Cologne. The path leads directly to Hausdorff's grave.

ROBERT JONES is a student of mathematics and a retired computer consultant who enjoys listening to Leonard Bernstein discuss musical scales. He acquired a bachelor's degree in mathematics in 1955 from Michigan State University, and he then worked on the experimental use of an early hard-disc drive, the IBM 305 RAMAC. In 1965 he began working in Germany, where he developed an early virtual memory system. He especially enjoys listening to Leonard Bernstein's "Young People's Concerts."

Universitaet Duesseldorf
Rurweg 3, D-41844 Wegberg,
NRW
Germany
e-mail: jones@uni-duesseldorf.de

The partially legible inscription on the tombstone reads:

Felix Hausdorff
Prof. der Mathematik
** 8. November 1868 + 26. Januar 1942*
Charlotte Hausdorff
geb. Goldschmidt
** 7. September 1873 + 26. Januar 1942*
Edith Pappenheim geb. Goldschmidt
** 21. März 1883 + 29. Januar 1942*
Lenore König
** 1. Februar 1900 + 8. September 1991*
Prof. Dr. Arthur König
** 13. Oktober 1896 + 24. April 1969.*

Charlotte: Hausdorff's wife. Edith Pappenheim: her sister. Felix and Charlotte's daughter Lenore and her husband Arthur König were buried here later.

In a farewell letter to a friend, Hans Wollstein, Hausdorff wrote, "Auch Endenich ist noch vielleicht das Ende nich!" (Hausdorff intentionally omitted the final "t" that the reader expects) and enclosed statements from himself, his wife, and her sister requesting cremation. This explains why their grave is in the part of the Poppelsdorf Cemetery reserved for urns.

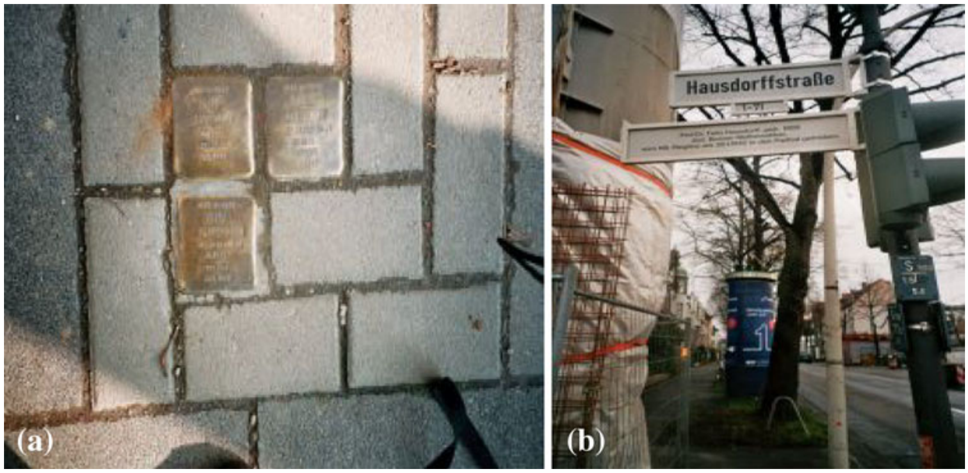


Figure 4. (a) The flagstones in Hausdorff Street; (b) the Hausdorff story on a street sign.

Figure 5. Hausdorffstrasse 61, Bonn.

Figure 7. Entrance of the Hausdorff Center for Mathematics.

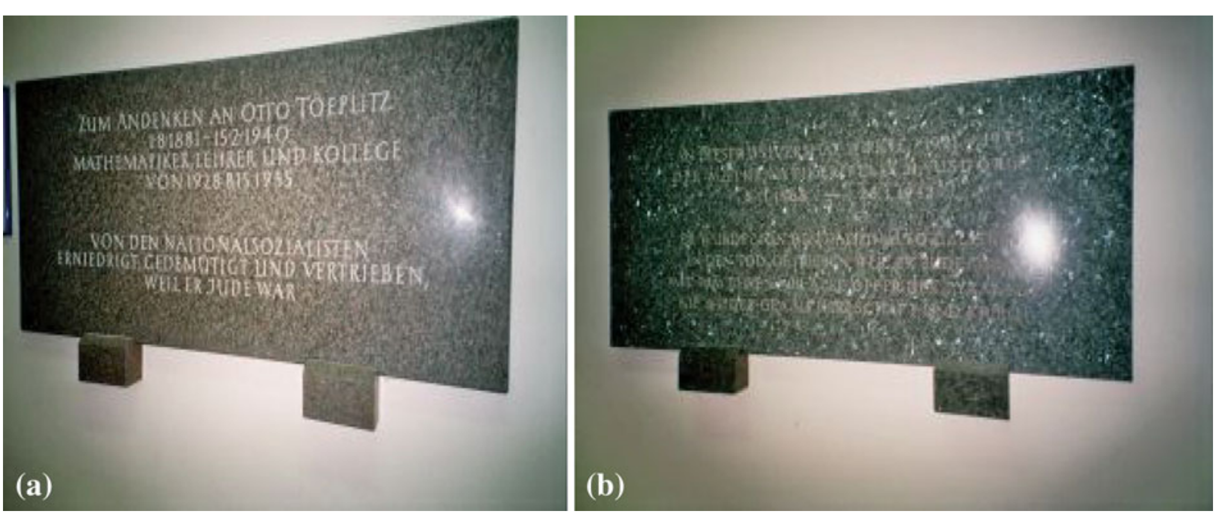


Figure 6. (a) Plaque in memory of Otto Toeplitz; (b) Plaque in memory of Felix Hausdorff.

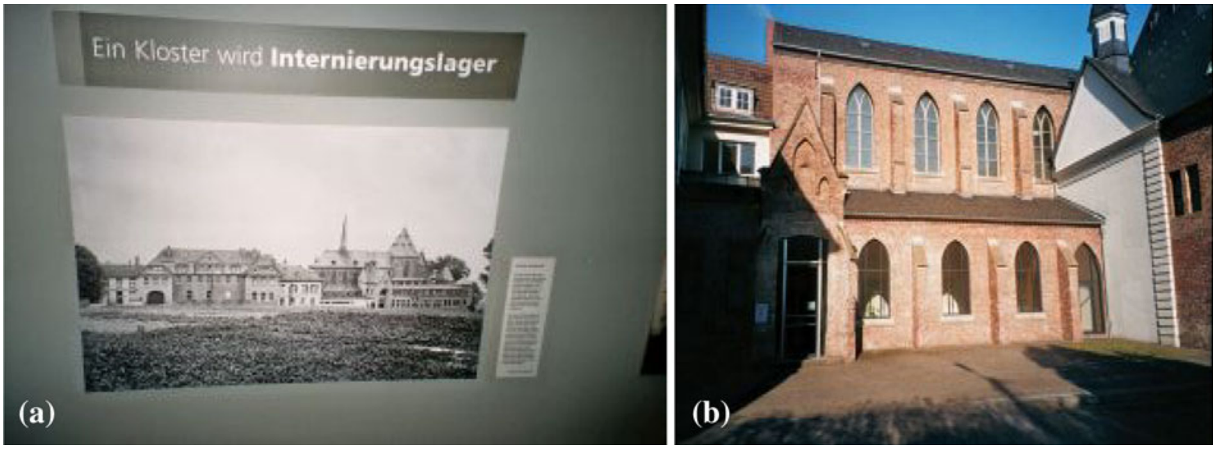


Figure 8. (a) The Cloister Endenich in the Bonn City Museum; (b) the entrance to the monastery as it is today.

Hausdorff's House

In Bonn, Felix Hausdorff lived in what was then Hindenburgstrasse 61. The street was later renovated and renamed in his honor, with the numbering retained. Today the story of his life and death, in abbreviated form, is related in brass flagstones (Fig. 4a) embedded in the sidewalk in front of his house. It is also told, in slightly less abbreviated form, on the street sign on the block in which his house still stands (Fig. 4b):

> *Professor Doctor Felix Hausdorff, born 1868, Jewish Mathematician in Bonn, driven to suicide on the 26th of January, 1942 by the Nazi regime.*

Among his many talents, Hausdorff was an accomplished pianist. Today his house (Fig. 5) is alive with music once again. The present owners, Georg Brinkmann and Vanessa Vromans, experts on klezmer music, argue that music is mathematics, writ by scales.

The Former Mathematical Institute

The former Mathematical Institute building is on Wegeler Street. In the entrance hallway there are two plaques engraved in granite (Figs. 6a,b).

The plaque on the left reads (in translation), "To the memory of Otto Toeplitz, 1.8.1881 – 15.2.1940, mathematician, teacher and colleague, by the Nazis demeaned, humiliated and exiled, because he was a Jew." On the other side of the eerily silent hallway, we read, "At this university, 1921 – 1935, the mathematician Felix Hausdorff, 8.11.1868 – 26.1.1942, was active. He was driven by the Nazis to his death, because he was a Jew. With him we honor all victims of tyranny. Never again the rule of violence and war!"

The imposing Hausdorff Center for Mathematics, at Endenicher Allee 62 (Fig. 7), is an easy stroll from the Hausdorff Institute of Mathematics. The Hausdorff Edition [1] has its offices in this building, and there is a treasure trove from the legacy and collected writings of Felix Hausdorff.

The Cloister of Endenich

In the Bonn City Museum you will find a photograph of a Benedictine monastery in Endenich, the "cloister [that] became a concentration camp" (Fig. 8a). The order to proceed there prompted Hausdorff's suicide. Endenich is a part of the city of Bonn, and it is not far from the center. After the war, the "Internierungslager" again became a cloister. It can be visited today (Fig. 8b).

The Hausdorff Family Today

There is none. Walter Purkert answered my inquiry,

> *There are unfortunately no longer any descendants of Felix Hausdorff, and we also have been unable to find any living relatives. His two sisters married and lived with their families in Prague. As far as we are able to determine, from Czechoslovakian archives, both families fell victim to the Nazis. We do not know whether some of the children of these families may have succeeded in fleeing. One son of the Brandeis family was definitely murdered.*
>
> *Mrs. König had two sons, Felix born in 1927, and Hermann born in 1929. Hermann suffered from a metabolic disorder which leads, with normal diet, to mental retardation (today all babies are tested and it is possible, with dietary restrictions, to prevent the development of the consequences of this disorder). Hermann survived the Nazi era in a children's home. He lived for several years in a medical care home in Wiesloch (near Heidelberg). The nurses there told us that he had the mentality of a four-year-old. We do not know whether he is still alive. Felix died with his wife in an auto accident. They did not have children.*
>
> *That is the entire sad story.*

ACKNOWLEDGMENTS

I would like to thank Georg Brinkmann and Vanessa Vromans for their interest, for their assistance, and for giving the Hausdorff home new life. I am also grateful to Walter Purkert for patiently answering my many questions.

Web Links

Map image showing the baroque church, called the Kreuzbergkirche, at Stationsweg 21, near the grave of Felix Hausdorff:

http://wiki.worldflicks.org/kreuzbergkirche.html#coords=%2850.71524128335446,7.081359028816223%29&z=18. Shifting this map upward, near the oval-shaped path near the center of the picture, one can see the location of Felix Hausdorff's grave.

Map picture centered on the grave of Felix Hausdorff: http://wiki.worldflicks.org/kreuzbergkirche.html#coords=%2850.71642329060522,7.082592844963074%29&z=18

Germany / North-Rhine-Westphalia / Poppelsdorf (today Bonn-Poppelsdorf): http://www.porcelainmarksandmore.com/northrhine/poppelsdorf_1/00.php

http://www.nu-klezmer.de

http://www.people.umass.edu/gmhwww/382/pdf/12-music%20scales.pdf

REFERENCES

[1] Purkert, Walter, ''The Double Life of Felix Hausdorff/Paul Mongré,'' *The Mathematical Intelligencer*, 30(4), Winter 2008, 36-50.

[2] Bajor, Stefan, *Archives in Stone*, *Jewish Life and Jewish Cemeteries in North-Rhine-Westphalia* [Archiv aus Stein, Jüdisches Leben und jüdische Friedhöfe in Nordrhein-Westfalen], Asso Verlag, Oberhausen, Germany, 2005.

Feel like writing a review for The Mathematical Intelligencer? You are welcome to submit an unsolicited review of a book of your choice; or, if you would welcome being assigned a book to review, please write us, telling us your expertise and your predilections.

➢ Column Editor: **Osmo Pekonen**,
Agora Centre, 40014
University of Jyväskylä, Finland
e-mail: osmo.pekonen@jyu.fi

Erhabene Welten: Das Leben Rolf Nevanlinnas

by Olli Lehto

BASEL, BOSTON, BERLIN: BIRKHÄUSER VERLAG, 2008, 300 PP., 92.52 EUR, ISBN 978-3-7643-7701-4

REVIEWED BY OSMO PEKONEN

Academician Olli Lehto, a Finnish complex-analyst born in 1925 and a former secretary of the International Mathematical Union, has written a biography of his own *Doktorvater*, the iconic Finnish mathematician Rolf Nevanlinna. Given the controversial aspects of the life of Nevanlinna to whom the author owes much of his own career, it could have been feared that biography turns into hagiography, but having read both the Finnish [1] and the present German version of the book, I can assure readers that this is not the case. The dark sides of Rolf Nevanlinna – if there were any – have been covered with remarkable honesty in this gripping book that should be made available to an English-speaking audience, as well. The political legacy of Rolf Nevanlinna, especially his behavior during WWII, has always been very much a hot potato for the Finnish mathematical community. Being a Finn myself, let me reveal my own cards: I have also published, in Finnish, a minor alternative biography of Nevanlinna [2] and have created a documentary movie about him, which was broadcast by Finnish television in 1995 on the occasion of Nevanlinna's centennial. This means that I have read through essentially the same source material as the present book's author. Even if some of my conclusions were more radical, I can but admire Lehto's capacity to take critical distance from a life-long close partnership with his *monstre sacré* and to lay bare the facts, both pleasant and unpleasant, which he exposes with elegance and sometimes with a healthy dose of irony. Nevanlinna is said to have exercised over his entourage a charm, which, according to his student Lars Ahlfors, "worked equally well on men and women". As for me, I am just old enough to have seen Nevanlinna once before he passed away, but I never knew him personally. Nevanlinna's colorful destiny amounts to a retelling of Finland's political history in the 20th century. Therefore, we must ask for the reader's indulgence to accept that this review contains a heavy dose of politics even if Nevanlinna was primarily a mathematician – but politics is very much present also in the book under review.

DOI 10.1007/s00283-011-9251-2

Rolf Nevanlinna was born in Joensuu, Finland, on 22 October 1895. The Grand Duchy of Finland was then a part of the Russian Empire. The extraordinary talent of the Nevanlinna family (earlier also called Neovius) has abounded through centuries. First a priestly and military family, the Nevanlinna lineage has also produced dozens of mathematicians. MathSciNet lists publications by four professors named Nevanlinna, but there have been many more, taking into account secondary-school teachers, textbook authors, and professors in other fields. Indeed, the family name has become eponymous: a "Neovian" approach to academic life, characterized by a "lofty view" (*erhabene Anschauung*), has been coined as an expression in Finland for a certain cultural ideal supposedly incarnated by the Nevanlinna family.

Rolf Nevanlinna's grandfather, Major-General Edvard Engelbert Neovius, taught mathematics and topography at a Russian military academy and conceived a pioneering approach to SETI, Search for Extraterrestrial Intelligence. In a book that appeared in Swedish, Russian, and French, he proposed a method for contacting the inhabitants of Mars, using light signals projected to the Red planet with huge beacons [3]. Rolf's father, Otto Neovius (from 1906 Nevanlinna), was a physicist and a mathematician. He married Margarete Romberg, daughter of Herman Romberg, the German chief astronomer of the Pulkovo observatory in Saint Petersburg, so that Rolf was of German, and indeed Prussian, descent on his mother's side. Rolf's elder brother Frithiof also became a well-known mathematician with whom he collaborated. Moreover, Otto Nevanlinna's cousin was Ernst Lindelöf, an eminent professor of mathematics at the University of Helsinki and the future mentor of both Rolf and Frithiof.

Figure. 1. Young Rolf Nevanlinna in 1922, at the peak of his mathematical creativity.

The young doctor Otto started his career as a school teacher of physics and mathematics at Joensuu, a rural town far away from the capital. In 1903, the family, with four children, returned to Helsinki where Rolf entered a grammar school that emphasized Greek and Latin. He was hesitant at first as to whether he should pursue classical studies at the university, but Ernst Lindelöf's example soon convinced him of his true calling, which was to be Analysis.

In December 1917, in the aftermath of the Russian revolution, Finland declared independence. The following spring, the country was torn by a bloody Civil War, which opposed Red and White Guards. The Reds were supported by revolutionary Russia, whereas the Whites received military assistance from imperial Germany. The young Rolf joined the White Guards but did not see military action. Helsinki was first held by the Red Guards, but they were swiftly expelled by the arrival in April 1918 of a German expeditionary force that probably saved the Finnish capital from massive destruction experienced elsewhere in the country. In the bourgeois circles of Helsinki, the mood was enthusiastically Germanophile. Independent Finland went as far as to elect as her first king, Prince Frederick Charles of Hesse, a brother-in-law of the Kaiser, but Germany collapsed before he could arrive to claim his throne, and Finland quickly switched into a republican constitution. A family such as Nevanlinna's certainly regarded the first German *Niederlage* with disappointment.

Frithiof Nevanlinna received his doctorate in 1918, and his brother Rolf received his in 1919, both in the field of Complex Analysis. Immediately after completing his dissertation, Rolf married his cousin Mary Selin, but in the 1950s this marriage ended in divorce. In 1958, he married his second wife Sinikka Kallio. He was to have five children.

Nevanlinna established his worldwide fame as a leading complex-analyst during a three-year period of intensive creation from 1922 through 1925. Together with Frithiof, he created the "Nevanlinna Theory", or the value distribution theory of meromorphic functions. The mathematical technicalities of the theory are not discussed in Lehto's book. Despite their outstanding research contributions, neither of the Nevanlinna brothers then held an academic position. Rolf taught mathematics at a school for about twenty hours a week, whereas Frithiof was employed by an insurance company where he later hired Rolf as well. Nevanlinna recalled that his school teaching took up most of his time, but the evenings, weekends, and holidays were devoted to research (Figure 1).

In 1926 Rolf Nevanlinna, then 30, finally received a professorship at the University of Helsinki. His best student Lars Ahlfors created a sensation in the mathematical world in 1928 when, at the age of 21, he proved the Denjoy Hypothesis. Ahlfors was awarded the Fields medal in 1936. Nevanlinna visited Zürich and Paris together with Ahlfors. During the term 1936–1937, Nevanlinna was a visiting professor at Göttingen, where Oswald Teichmüller attended his lectures. He successively became *Ehrendoktor* of the University of Heidelberg, *Ehrenbürger* of Göttingen, and member of the *Preussische Akademie der Wissenschaften*. Before lavishing such honours on him, Nazi Germany investigated Nevanlinna's racial antecedents, which were excellent, Nevanlinna's mother

Margarete being of the best Prussian stock as confirmed in a document signed by Rudolf Hess. Indeed, some male members of the Romberg family have been photographed wearing the *Pickelhaube*, the characteristic spiked Prussian military helmet. Moreover, Margarete was a member of the *Arbeitsgemeinschaft der deutschen Frau*.

The political development of Nevanlinna in the 1930s parallels that of his student Teichmüller. He viewed Adolf Hitler as a harbinger of German greatness, comparable with his other heroes Frederick the Great and Otto von Bismarck, and ended his correspondence with like-minded German colleagues, Helmut Hasse, for instance, with an enthusiastic *Heil Hitler!* We have to remember that in the interwar period it was the Soviet Union that many ordinary middle-class people in many European countries, not only in Finland, viewed as the potential aggressor, whereas Hitler, for many, appeared to be what he promised: a bulwark against Bolshevism. André Weil who stayed with Nevanlinna in Finland in the summer of 1939 reports [4] that he first read *Mein Kampf* in his household. He could have found other similar readings on Nevanlinna's bookshelves, because his host nurtured relationships with several Nazi organizations. Hermann Weyl labelled Nevanlinna as a "Finnish Nazi", but technically he wasn't one because Finland did not have a full-blown Nazi party. Nevanlinna went as far as to try to create a Nazi movement of his own, but nothing much came of it. "Merely a handful of school-boys attended the founding meeting", a Finnish counterintelligence agency report dryly notes. Nevanlinna made no secret of his Nazi sympathies, and the government viewed him with increasing suspicion. The various far-right-wing factions where Nevanlinna had found his political home belonged to the margins of the Finnish political scene. Having largely recovered from the wounds of the Civil War, the Finland of the 1930s was a blossoming Scandinavian democracy, which was preparing not for war but to host the Olympic Games of 1940.

Nevanlinna's German sympathies suffered a blow at the news of the Molotov-Ribbent7rop pact in August 1939. A consequence of the treaty was the Soviet invasion of Finland, which Germany supported by blocking the Baltic Sea. This was the famous "Winter War" (30 November 1939–13 March 1940) that was to become Finland's hour of glory; astonishing the world, the tiny nation was able to defend herself. Nevanlinna's mathematical abilities and international networks were put into use. Together with Ernst Lindelöf, he did ballistic research for the improvement of field artillery firing tables and was awarded the Cross of Liberty, Second Class, for his services. His Swedish colleague, the master codebreaker Arne Beurling [5], came to Helsinki to join the Finnish war effort as a volunteer, as did many other Swedes.

A strange episode of the Winter War is the often-told story of André Weil's brief detention (from 30 November until 12 December 1939) in Finland. Having knocked at the door of the Soviet embassy – to seek a visa for a lecturing tour – totally oblivious of the war, he was suspected of being a Soviet spy. In his autobiography [4], Weil claims that he was going to be executed and that Nevanlinna saved his life. As I have shown elsewhere in a joint article with Weil himself [6], based on my research in the archives of the Finnish counterintelligence agency, this is a hoax. The espionage suspicion was initially real, but it soon evaporated, and Weil was simply expelled to Sweden. Nevanlinna may have concocted the dramatic aspects of the story to improve his public image after the war: he possibly wanted to have on record that he had saved the life of at least one Jew. For that matter, Weil's Jewishness was no issue at all in the Finland of 1939, whereas Nevanlinna's pro-Nazi fervour, remarkably enough, never included anti-Semitism, as is testified by Weil and also confirmed by my personal research. Finland, whose Jewish community is not very large, does not have a tradition of rampant anti-Semitism. Almost all the Finnish Jews, actively protected by their government, survived WWII; there was essentially no Finnish Holocaust [7].

Figure 2. Rector Magnificus Rolf Nevanlinna in the ruins of Helsinki University after a Soviet air raid. He contemplates a snow-covered sculpture symbolizing Finland and the goddess of Freedom.

After the Winter War, a precarious truce followed (13 March 1940–25 June 1941). As a price of peace, Finland had ceded to the Soviet Union some territories often loosely referred to as "Karelia" (in fact, the territorial issue is more complicated) but maintained her national independence. However, Finland's involvement in a new conflict appeared inevitable. In the summer of 1941, there were both Soviet and Nazi troops within Finland's former national borders, both with government approval. The Soviets had established a naval base in Hanko, not far from the capital, whereas the Nazis regularly transferred troops to occupied Norway through Finnish Lapland. When the titans clashed, the Finnish national territory was doomed to become a battlefield anyway. No such option as neutrality – which Sweden bought by supplying iron ore to the Axis – existed for Finland. The traditional Germanophilia of the Finnish upper classes flourished again, and the German troops were as welcome as in 1918. The German SS General Felix Steiner, who paid a courtesy call to Nevanlinna's home, reported to Berlin that

Figure 3. A dangerous friendship. Rolf Nevanlinna together with SS General Felix Steiner at Nevanlinna's summer cottage in 1943.

the entire family "wept for joy" at the news of Operation Barbarossa on 22 June 1941. Finland remained nominally neutral for three days. Conveniently enough, the Soviet Union opened hostilities by bombing Finland's cities on 25 June 1941, which provided Finland an excuse to join in the German attack. In the same summer, Nevanlinna was elected Rector Magnificus of the University of Helsinki, an eminent public position that implied many new duties and marked the end of his most fruitful research period. He performed well in his administrative role despite the dismal wartime conditions. The university's main building itself was badly damaged by a Soviet bomb (Figure 2).

The Finnish-German *Waffenbrüderschaft* initially worked very well. The Germans were not perceived as occupiers but as actual allies, and they behaved remarkably well. Finnish troops rapidly advanced deep into Karelia, grabbing back the territories lost in 1940, and a bit more (Figure 2). Rolf Nevanlinna's ballistic skills were needed again. He was personally invited to the front to open fire when the Finnish artillery started shelling the Soviet base of Hanko: a strange duty for a university rector without any military rank. He also toured the Karelian front, and brought home a curious trophy of war: a Soviet pirate edition of 1941 of his main work *Eindeutige analytische Funktionen*. Despite the hostilities, he always remained a well-respected mathematician among the Russian colleagues.

There were ultimately about 200,000 German troops in Finland. They were mainly concentrated in Lapland, trying to break their way to Murmansk. Politically, the Finnish-German alliance remained shaky because on the diplomatic front Finland did everything possible to distance herself from full collaboration with the Axis and frantically promoted the argument of a separate war with the Soviet Union, not with the Western Allies. The United States bought the claim and never declared war on Finland whereas the United Kingdom did. SS- Reichsführer Heinrich Himmler personally arrived in Finland to demand closer cooperation. As a token of political loyalty, Finland had to recruit a volunteer corps for the German armed forces, which, moreover, was to be incorporated into the SS and not into the Wehrmacht, as the Finns would have preferred. To handle the diplomatically burdensome issue, the recruitment was organized independently of the Army as an ostensibly private initiative. The Finnish government decided to put Rolf Nevanlinna's well-known Germanophilia to good use in the national interest. In 1942, at the request of the Foreign Minister, Nevanlinna made himself available as chairman of the Committee for the Finnish Volunteer Battalion of the Waffen-SS. In this role, he rubbed shoulders with Heinrich Himmler and Alfred Rosenberg, and was awarded the German Eagle with diamonds. The committee recruited a highly academic battalion of about 1400 Finnish volunteers that was attached to the Scandinavian SS Division Wiking under the command of General Felix Steiner (Figure 3). Two nephews of Rolf Nevanlinna saw action in the battalion that was deployed on the Caucasian *Ostfront*. The battalion was never accused of war crimes. Also, the charges against General Steiner were dropped at the Nuremberg Trials. Indeed, the Germans reserved most of the *Ostfront* mass killings for themselves, not to politically alienate the foreign troops under their banners. Even so, the existence of a Finnish SS unit became such a diplomatic burden that the Finnish government unilaterally called home and disbanded it as early as July 1943. This was the beginning of the end of *Waffenbrüderschaft*. Nevanlinna gave a pretty farewell speech to the disbanded Finnish SS, duly making the Nazi salute that was forbidden in the Finnish army. German footage of the ceremony exists, and I incorporated it into my TV documentary of 1995.

Nevanlinna's participation in SS activities labelled him forever as a major Nazi collaborator. As a matter of fact, he was surely the most visible pro-Nazi figure not only among Finnish mathematicians, but in Finland at large. On the other hand, chairing the recruitment committee was the only political position that he ever had. Outside Academia, he was not a decision-maker. He had no sway on Finland's political or military leadership, and his pro-Nazi stances did not influence the course of major events. His political statements appeared in small pro-Nazi magazines of the lunatic fringe. Having read them all – a passionate praise of Hermann Göring [8], for instance – I am led to the conclusion that the man was a brilliant mathematician but a political fool (not the only one in our profession, for that matter). Olli Lehto, who lived through the dramatic years of WWII as an ordinary frontline soldier, finds a more tactful way of expressing this.

Finland's fate was discussed by the Big Three in Tehran in November 1943. President Roosevelt urged Finland to step out of the Axis before it would be too late, but this was militarily impossible, Germany still being strong enough to occupy the country (as it occupied Hungary in March 1944). The Nazis certainly had a plan B for Finland, in case of non-compliance. In the event of occupation and installation of a puppet regime, who did the Nazis have in mind as a potential

Führer of Finland? Perhaps Nevanlinna? Allied press floated such rumours in the spring of 1944 (e.g., *Daily Times*, 16 May 1944). Happily enough, Nevanlinna's dictatorship remains pure science fiction, for history went differently. After a massive, and successful, defensive campaign against the Red Army on the Karelian Isthmus in the summer of 1944, a window of opportunity for a separate peace with Stalin opened for Finland. A new truce with the Soviets was signed on 19 September 1944. Finland lost again the same territories as in the previous truce and now faced the task of chasing 200,000 German troops out of her national borders. This was achieved in the "War of Lapland", which dragged on until 27 April 1945. Olli Lehto – who has also published his autobiography [9] – first served on the Eastern front against the Soviets, then on the Northern front against the Nazis. Privately, he likes to point out that in WWII he had the rare luxury to take a few shots at "both the Commies and the Nazis". Thanks to her Army, and a dose of good luck, Finland resurfaced from the turmoil with her democratic institutions intact. Among the European capitals at war only three never saw enemy occupation: London, Moscow, and Helsinki.

So just how bad was Nevanlinna? After the war, no massive denazification was necessary in Finland, but Nevanlinna was understandably branded by the Finnish Left as a key collaborator. He never apologized, and he made no secret of his wartime views in his autobiography [10]. He even kept in touch with the former SS General Steiner who had become a family friend and came to Finland for holidays. To be sure, the second German *Niederlage* and the full revelation of Nazi horrors were devastating blows for Nevanlinna, but until the end of his days he remained more Germanophile than many Germans themselves. His share in creating the German war machine had been to chair the Committee for the Finnish Volunteer Battalion of Waffen-SS, but this was done at the specific request of a democratically elected government of his own country. Surely the Finnish government exploited Nevanlinna in 1942–1943 as a "useful political idiot", but, fairly enough, official Finland did not let him down after the war. At the Prime Minister's request, he had to step down from the Rector's office in 1945, but, as a compensation, only three years later he was elected member of the newly founded Academy of Finland, which was a lofty position with a good salary. Meanwhile, he also had accepted a position at the University of Zürich, occupying a chair where he succeeded his former student Ahlfors, who had moved on to Harvard. The Soviets were quick to rehabilitate Nevanlinna; they did not try to block his election to the presidency of the International Mathematical Union for the four-year term 1959–1962. Nevanlinna was the president of the Stockholm ICM in 1962, he chaired the program committee of the Moscow ICM in 1966, and he was the honorary president of the Helsinki ICM in 1978.

Nevanlinna left his mark in academic life in Zürich where he had several students and where his memory still lives on in the form of the Finnish-Swiss Rolf Nevanlinna Colloquia. He mentored in Zürich Lê Văn Thiêm, who was to become the father of the Vietnamese school of mathematics in Hanoi. After his retirement, Nevanlinna continued to play a public role in Finland. He was, for instance, Chancellor of the University of Turku from 1965 to 1970. He played a certain political role in the massive enlargement of tertiary education in Finland since the 1960s. Very early he advocated equipping the Finnish universities with computers, thus heralding the birth of the Information Society whose successful champion Finland has become today. He received some of Finland's highest decorations. Nevanlinna passed away in Helsinki on 28 May 1980. His funeral was a major national event. In the Warsaw ICM in 1983, the International Mathematical Union established the Rolf Nevanlinna Prize to recognize work on mathematical aspects of information science.

A few final words about the "Neovian" approach to life and mathematics. Nevanlinna was a charismatic lecturer who had a special ability to convey the sense of aesthetics in his mathematical experience. In an increasingly specialized academic world, he insisted much on the unity of Western Culture and the usefulness of practising Arts and Sciences together. "There are not Two Cultures; there is only one", he famously refuted C. P. Snow. His public lectures on general relativity or the foundations of geometry were sometimes attended by cultural figures, including poets and composers. Nevanlinna often cited the view of Jacob Burckhardt, who claimed that, besides a civilized person's main profession, his life should involve at least two other interests that are more than just hobbies. In the case of Nevanlinna, music was his second passion, and German literature perhaps the third one. He was an amateur of the violin and a somewhat fanatic aficionado of Jean Sibelius's music; if he was not happy with the performance of an orchestra, he would go and tell the conductor. At one point of his life, he almost identified himself with Goethe.

The life of Rolf Nevanlinna cannot be understood apart from the destinies of Finland and Germany, the beloved home countries of his father and mother. As a conservative nationalist of a somewhat Prussian upbringing, he was not the only 20th-century intellectual tempted by the dark side of the Force. For me, he brings to mind Doktor Faustus in a pandemonic world. Appropriately enough, Thomas Mann's *Doktor Faustus*, which he liked, was translated into Finnish by Sinikka Kallio, his second wife.

REFERENCES

[1] Olli Lehto, *Korkeat maailmat. Rolf Nevanlinnan elämä.* Helsinki: Otava, 2001 [in Finnish].

[2] Osmo Pekonen, ''Rolf Nevanlinna, matemaatikko ja kulttuurivaikuttaja''. In: *Kansallisgalleria. Suuret suomalaiset*, Vol. 4. Helsinki: Weilin & Göös, 1996, pp. 66–71 [in Finnish].

[3] Raimo Lehti, ''Edvard Engelbert Neovius, an early proponent of interplanetary communication''. *Acta Astronautica* 42: 10–12 (1998), pp. 727–738.

[4] André Weil, *Souvenirs d'apprentissage*, Basel: Birkhäuser, 1991. *The Apprenticeship of a Mathematician*, Basel: Birkhäuser, 1992. *Lehr- und Wanderjahre eines Mathematikers*, Basel: Birkhäuser 1993.

[5] Bengt Beckman, *Codebreakers: Arne Beurling and the Swedish crypto program during World War II.* Providence, RI: American Mathematical Society, 2002.

[6] Osmo Pekonen, ''L'affaire Weil à Helsinki en 1939''. *Gazette des mathématiciens* 52 (avril 1992), pp. 13–20. With an epilogue by André Weil.

[7] Hannu Rautkallio, *Finland and the Holocaust: the Rescue of Finland's Jews.* New York: Holocaust Library, 1987.

[8] Rolf Nevanlinna, ''Hermann Göring''. *Kustaa Vaasa* 3 (1941), pp. 23–26 [in Finnish].

[9] Olli Lehto, *Ei yliopiston voittanutta.* Helsinki: Otava, 1999 [in Finnish].

[10] Rolf Nevanlinna, *Muisteltua*. Helsinki: Otava, 1976 [in Finnish].

Agora Center
University of Jyväskylä
PL 35, 40014 Jyväskylä, Finland
e-mail: osmo.pekonen@jyu.fi

Gösta Mittag-Leffler: A Man of Conviction

by Arild Stubhaug

BERLIN, HEIDELBERG, NEW YORK: SPRINGER-VERLAG, 2010, X, 800 PP., 65 ILLUS., 8 IN COLOR, HARDCOVER, 59.95 EUR, ISBN: 978-3-642-11671-1.

REVIEWED BY HÅKAN HEDENMALM

This is a biography of Gösta Mittag-Leffler (1846–1927), a prominent figure in Swedish mathematics. Mittag-Leffler was instrumental in bringing the mathematical developments of the great scientific centers of the time to the backwaters of Sweden, becoming one of the most influential operators in mathematics of his time, mediating between the French and German schools. The author, Arild Stubhaug, has also written biographies of the Norwegian mathematicians Niels Henrik Abel and Sophus Lie.

Stubhaug begins with a snapshot from 1899–1900, when Mittag-Leffler (with his wife Signe and his personal physician Wettervik) traveled to Egypt to calm his nerves and to try to cure, with the warm sun and the dry desert climate, the physical ailments that kept troubling him. There the Mittag-Lefflers met Selma Lagerlöf and her close friend Sophie Elkan; they even stayed at the same hotel in Egypt. Selma Lagerlöf (1858–1940) was a renowned writer of novels who received the Nobel prize for Literature in 1909.

Magnus Gustaf (Gösta) Mittag-Leffler was a son of Johan Olof (Olle) Leffler (1813–1884) and Gustava Vilhelmina Leffler (née Mittag; 1817–1903). Gösta added Mittag to his last name as a tribute to his mother's family. Olle was a successful rektor (headmaster) of a school in Stockholm and a member of Parliament before he fell mentally ill in 1870. The onset of his illness meant that the family's economic situation worsened drastically, at a time when the talented young Gösta was a student in Uppsala.

Gösta had two brothers, Leopold Fredrik Alexander (Frits) and Artur Lorens Olof Abraham (Artur), and a sister, Anna Charlotte Gustava (Anne Charlotte). Anne Charlotte was a writer of drama, short stories, and novels. She became friends with the famous Russian mathematician Sofia Kovalevskaya, and in 1887 they wrote the drama *The struggle for happiness* together. In those years, the role of the woman in society was about to be recast, and both Leffler and Kovalevskaya were pioneering the change. Anne Charlotte died of appendicitis in 1892 in Naples, leaving behind a son, Gaetano. Frits (who wrote the family name as Läffler) was a linguist, but he suffered from mental illness like his father. Artur, an engineer, later was involved in his brother Gösta's various business ventures.

Gösta Mittag-Leffler was interested in mathematics by high school. He obtained his undergraduate degree (*filosofie kandidatexamen*) and his Ph.D. degree (1872) at Uppsala University, with the grade *laudatur*, where his thesis was influenced by Hjalmar Holmgren. Uppsala was a university town laden with centuries of traditions, and not a place where novel scientific thinking could flourish. His father's illness probably propelled the young Mittag-Leffler to do his best to succeed in science, and he decided to go to the mathematical centers in France and Germany to learn about the latest developments. He successfully obtained the Byzantian travel scholarship, which enabled him to visit Paris, Göttingen, and Berlin, in the period 1873–1876. In Paris he interacted with Charles Hermite and Henri Poincaré, in Göttingen with Ernst Schering, and in Berlin with Leopold Kronecker and most notably with Karl Weierstrass and his unofficial student Sofia Kovalevskaya. The connection with Kovalevskaya led to an 1876 visit to St. Petersburg, Russia, and after that, he went to Helsinki where a position as a university professor was available. Mittag-Leffler presented his Habilitationsschrift, in Swedish, on the theory of elliptic functions, a topic of current interest in the circles around Weierstrass, and he won the competition for the position, with a positive evaluation by the appointed expert Lorenz Leonard Lindelöf. In 1877 he was appointed professor of Mathematics in Helsinki. Being a professor was a sign of distinction in those days, and Mittag-Leffler used his position to connect with the higher social circles in Helsinki. At that time, Finland was a Grand Duchy of the Russian empire, but Swedish was rather commonly spoken, especially in the wealthier strata of society.

In Helsinki, Mittag-Leffler adopted a rather extravagant lifestyle, living beyond his means, as he had been doing since the travels to continental Europe. This was made possible by friends and acquaintances from Sweden; one of these was Johan Hagströmer, his friend from high school and university days, who was from a well-off family. Mittag-Leffler's contact with the politically influential Gustaf af Ugglas, whose son Samuel he tutored while studying at Uppsala, helped him to return to Sweden from Finland. Gustaf af Ugglas had served as Swedish finance minister from 1867 to 1870. It was easy for a person such as af Ugglas to help Mittag-Leffler obtain a professorship in Stockholm, at the newly founded Stockholms Högskola (Stockholm College, now Stockholm University), and the appointment was finalized in 1881. In the meantime Mittag-Leffler courted Signe Lindfors (1861–1921), the daughter of Major General Julius af Lindfors. They married in 1882. Some unpleasantness from interaction with the Fennoman movement – a Finnish patriotic movement aimed at promoting the Finnish language and culture – probably played a role in the decision to leave Helsinki for Stockholm.

Mittag-Leffler decided to enter the world of business, and in 1882 co-founded the insurance company Victoria. Another significant event that year was his founding of the renowned mathematical journal *Acta Mathematica*. In 1884, influenced by Weierstrass, he secured a position for Sofia Kovalevskaya at Stockholms Högskola. This was no small feat given the difficulties women encountered in academia at that time. But Kovalevskaya was not happy with her job in Stockholm and tried to leave repeatedly. The year before, in 1883, her husband Vladimir Kovalevsky had committed suicide following

Original Norwegian edition published by Aschehoug, Oslo, 2007. Swedish translation published by Atlantis, Stockholm, 2007.

DOI 10.1007/s00283-011-9261-0

financial troubles. Later, she met Maxim Kovalevsky and had a tumultuous affair with him for several years. Kovalevskaya was oriented toward the nihilist movement, and her sister Anna was a devout communist. Kovalevskaya was awarded the *Prix Bordin* by the French Academy of Sciences in 1888. She died suddenly of an infection (probably pneumonia) in 1891 in Stockholm.

In 1889, Mittag-Leffler made sure that Poincaré was given the new prize for mathematics awarded by King Oscar II. The King intended the prize to be annual, but it was awarded only once. The Nobel prizes, which date from the same time, are still awarded every year. The money was donated by the industrialist Alfred Nobel. For reasons unknown, there is no Nobel prize for mathematics. It is said that Nobel excluded mathematics because Mittag-Leffler had an affair with his (Nobel's) wife. Stubhaug says there is no real substance to the story; if there was any kind of rivalry between the two eminent Swedes over a woman, she was Sofia Kovalevskaya.

As a member of the Swedish Royal Academy of Sciences (KVA) with political clout, Mittag-Leffler wielded influence over the choice of Nobel prize laureates. For a number of years he gave a private party every December for that year's Nobel laureates at his villa in Djursholm. He was instrumental in seeing to it that Marie Curie (1867–1934) was awarded the Nobel prize, not once but twice. He also made several unsuccessful attempts to give the prize for Physics to Poincaré.

Mittag-Leffler made his main mathematical contributions in the 1880s. Later he devoted most of his time to his private business ventures (mostly related to hydropower stations and associated chemical processes) and to scientific politics; he even tried his hand in ordinary politics. Nonetheless he kept a keen interest in the latest scientific advances, such as Kristian Birkeland's discovery of a technique to make a nitrous fertilizer (saltpetre) from nitrogen in the air. Stubhaug contrasts this with Ivar Bendixson who, as his own scientific prowess waned, claimed (it is said) that science did not advance much anymore. Mittag-Leffler and Bendixson also disagreed about the future of Stockholms Högskola: Mittag-Leffler wanted it to develop into a research school, whereas Bendixson wanted it to become a traditional university similar to the ones in Uppsala and Lund. Eventually, Bendixson prevailed.

Mittag-Leffler's marriage was childless. This probably explains why he and his wife in 1916 decided to donate their private villa and most of their personal wealth to the KVA to establish a mathematical research institute. After his death in 1927, the Mittag-Leffler Institute lay more or less dormant, with Torsten Carleman serving as its first director after Mittag-Leffler himself. The more active institute we know today was made possible by the efforts of Lennart Carleson, appointed director in 1969. He raised the additional financing necessary to realize the original intentions of Mittag-Leffler.

Mittag-Leffler was a celebrated mathematician, with six honorary doctorates, and memberships in forty-five scientific societies, such as KVA, the Royal Society of London, l'Académie de Sciences (Paris), and the Russian Academy of Sciences (the Imperial St. Petersburg Academy of Sciences).

Stubhaug's monumental biography relies on a wealth of primary and secondary resources, including more than 30,000 letters in the Mittag-Leffler archives. The effort needed to turn all that material into an easily readable presentation is impressive. The book will be of great interest both to mathematicians and to general readers interested in science and culture. But the reader who may hope to discover whether there was a romantic aspect in Weierstrass's and Mittag-Leffler's support for Kovalevskaya will be disappointed.

I thank Kelly Jabbusch for helpful suggestions.

Department of Mathematics
The Royal Institute of Technology
S-100 44, Stockholm
Sweden
e-mail: haakanh@math.kth.se

math.ch/100 Schweizerische Mathematische Gesellschaft Société Mathématique Suisse Swiss Mathematical Society 1910–2010

edited by Bruno Colbois, Christine Riedtmann, and Viktor Schroeder

ZURICH: EUROPEAN MATHEMATICAL SOCIETY, 2010, 526 PP., HARDCOVER, 68.00 EURO, ISBN 978-3-03719-089-0

REVIEWED BY MAX-ALBERT KNUS

This unusual book was published in 2010 by the Swiss Mathematical Society (SMS) to celebrate its 100^{th} anniversary, and a copy was presented to every member. It consists of 23 contributions covering many aspects of a century of mathematics in Switzerland. The authors evoke prominent personalities as well as institutes and institutions—even personal memories and student life find a place. There is great diversity in the contributions: French, German, and English are used, and nearly half of the articles are reprinted from older publications. Authors were encouraged to write what they liked as they liked. Many contributions are full of life, whereas others reflect some academic dryness. Authors also freely chose their illustrations, with the result that many photographs emerged from little-known private albums. Paraphrasing the title of one of the essays, I would describe the volume primarily as a collection of glimpses. It is not (and was not intended to be) an encyclopædia giving a systematic and detailed account of mathematics in Switzerland, but it is an exciting source of information. The steering committee of the SMS, Bruno Colbois, Christine Riedtmann, and Viktor Schroeder, edited the work carefully. The result is an anniversary gift with a special touch.

The first contribution is a memorial essay on the history of mathematics in Switzerland, written by Michel Plancherel on the occasion of the 50^{th} anniversary of the SMS. The second is a history of the SMS commissioned by our centenarian society. The science historian Erwin Neuenschwander has written a well-documented report on the origin and activities of the society during its first hundred years, including a useful chronicle of year-to-year events. The remaining pieces in the volume appear in alphabetic order (based on the authors' names). I will not review the essays individually nor follow their order, but try instead to provide a personal synthesis of the book along a historical thread.

The first countrywide scientific organization in Switzerland was the Swiss Society for Natural Sciences (Schweizerische naturforschende Gesellschaft, SNG), established in 1815 to encourage natural sciences in Switzerland through the organization of annual meetings and the publication of a scientific journal. (Local societies were already in existence.) At the beginning of the twentieth century, specialized scientific societies were organized as subsections of the SNG: the Swiss Chemistry Society in 1901, the Swiss Physics Society in 1908 and, last but not least, the Swiss Mathematical Society (SMS) in 1910.

The SMS founders were Rudolf Fueter, Hans Fehr, and Marcel Grossmann. Fueter, of Basel origin, left Swizerland in 1899 to study under David Hilbert. After earning his doctorate, he spent time in various European mathematical institutes before returning to Basel as a professor in 1908. He soon realized how fragmented the mathematical community was in Switzerland. Mathematicians from different Swiss universities had no contact with each other whereas mathematicians of German origin were rather looking across the Rhine. One reason Fueter created the SMS was certainly to prevent local and international isolation.

Fueter was a recognized number-theorist. His book on "Synthetische Zahlentheorie", published in 1917 soon after he moved to the University of Zurich, was very successful; a third edition appeared in 1950. Fueter was also involved in many activities outside of mathematics. It is thanks to him and his colleague, the group-theorist Andreas Speiser, that the Swiss House in the Cité universitaire of Paris was built by the famous architect Le Corbusier in 1930. Less known is his activity during the first year of World War II as a colonel of the Swiss army. Responsible for press censorship, Fueter fought courageously for the democratic rights of the Swiss press in 1940. Some of the influential newspapers were openly critical of the Nazi regime. This was viewed as a provocation by Germany, and quite a broad part of official Switzerland supported the German accusations.

Fehr began his career as a mathematics teacher in Geneva and became professor at the University of Geneva in 1900. One of his major contributions was the journal L'Enseignement Mathématique, which he founded in 1899, together with his friend Charles-Ange Laisant, as the first international journal emphasizing mathematics teaching.

Grossmann was a fellow student and friend of Albert Einstein. In 1907 he became professor of descriptive geometry at ETH. After Einstein came back to ETH as a professor of theoretical physics in 1912, Grossmann pointed out to him the relevance of the tensor calculus to general relativity. Their collaboration led to the first paper on the general theory of relativity in 1913. During World War I, Grossmann was very active in improving relations between the French- and German-speaking parts of Switzerland.

The first meeting of the SMS was held in Basel, September 4–10, 1910, with Fueter as president. Fehr succeeded him, and Grossmann was the third president. During its first year, more than one hundred people joined the society; today the SMS has around five hundred members.

Besides national solidarity, mathematicians in Switzerland united to edit the complete works of Leonhard Euler. The far-seeing decision to publish the "Leonhardi Euleri Opera omnia" was taken in 1909 by the SNG. Publication is still under way; it will extend to around one hundred volumes when complete. Many mathematicians in Switzerland contributed to the edition.

DOI 10.1007/s00283-011-9265-9

As with most scientific societies, meetings and journals are essential activities of the SMS. There are two meetings every year, the Fall meeting or general assembly, and the Spring meeting, which is dedicated to specialized topics. The SMS owns two journals. "Commentarii Helvetici Mathematicae" was founded in 1928 to make the research of mathematicians in Switzerland better known internationally. Today it is recognized worldwide. Here too Fueter played a decisive role as its main editor until his death. The second journal, "Elemente der Mathematik", was established in 1946 by Louis Locher-Ernst to convey the relevance of today's mathematics to a large circle of readers. Since 2005, both journals have been published by the European Mathematical Society Publishing House in Zurich, founded in 2001 at the initiative of Rolf Jeltsch (ETH Zurich). The decision to leave the previous publisher, Birkhäuser Verlag, was taken after a lively discussion in an extraordinary general assembly of the SMS.

One historical duty of the SMS is to represent the Swiss mathematical community in international organizations. As an example, Fueter met with J. C. Fields in the summer of 1931 and the SMS approved the idea of the Fields medal in the following general assembly. In the last century, the small Swiss mathematical community has clearly been overrepresented internationally. The first International Congress of Mathematicians was held in 1897 in Zurich, as were two later ICMs (1932 and 1994). After the rebirth of the International Mathematical Union in 1952, five IMU presidents were active in Switzerland: Heinz Hopf, Rolf Nevanlinna, Georges de Rham, Komaravolu Chandrasekharan, and Jürgen Moser. More recently Jeltsch was president of the European Mathematical Society.

Today, some nontraditional activities, such as developing links with high-schools, are emerging. However, the SMS is a small society with a very small administration. Its success depends largely on the personal commitment of its members.

The golden age of Swiss mathematics, with the Bernoullis and Euler, was the eighteenth century. As the Bernoulli family occupied the Chair of Mathematics of the University of Basel for nearly one hundred years, Euler had to leave Basel for opportunities in Saint Petersburg and Berlin. In the first half of the nineteenth century, Switzerland had only a few prominent mathematicians. Two of them were active outside Switzerland, the geometer Jakob Steiner in Berlin and the analyst Charles Sturm in Paris. A third was the geometer Ludwig Schäfli, who spent his career at the newly created University of Bern. At that time mathematics had little influence or importance in Swiss higher education. In spite of international recognition, Schläfli waited for more than thirty years to be promoted from "Privat-Dozent" to full professor. His main mathematical contributions are in the field of higher-dimensional Euclidean geometry.

In the second half of the nineteenth century, Switzerland moved from the old confederation of cantons to a modern federal state. The opening of the first federal higher teaching institution, the Zurich Polytechnikum (ETH), in 1855 was a major event of this era. One of its first aims was to train qualified engineers for new Swiss industries and good science teachers for Swiss high-schools. However, basic research was strongly supported very soon. Most of the teaching was in German but, ETH being a federal institution, some courses, particularly in mathematics, were also taught in French.

Many of the new professors came from Germany. A number of young, highly talented German mathematicians took the opportunity in Zurich to start successful careers, among them Richard Dedekind, Elwin Bruno Christoffel, Hermann Amandus Schwarz, Heinrich Weber, and Ferdinand Georg Frobenius. Ever since, it has been a tradition in Swiss universities to invite people without regard to national borders. Later examples are Ludwig Bieberbach and Erich Hecke in Basel, Matyas Lerch in Fribourg, Hermann Minkowski, Adolph Hurwitz, Hermann Weyl at ETH, and Ernst Zermelo at the University of Zurich. Other more recent cases are described in the volume. Some mathematicians remained in Switzerland for their entire careers, but many stayed for only a short time.

Weyl was active in Zurich from 1913 until 1930 and then moved for three years to Göttingen, where he assumed the chair of his former teacher David Hilbert. In 1933, he accepted a call to the Institute for Advanced Study in Princeton, but he never lost his contacts with Zurich, in part for family reasons. As Weyl once said, his seventeen years in Switzerland were the most fruitful of his life: seven books and nearly seventy papers appeared in that period. It was therefore not easy for ETH to cope with the loss of Weyl. After both Emil Artin and Rolf Nevanlinna declined the ETH position, its board followed the advice of Issai Schur and offered it to Heinz Hopf. (Nevanlinna later accepted a position at the University of Zurich).

Born near Breslau, Hopf wrote his thesis in 1925 in Berlin under Bieberbach. He spent the year after his doctorate at Göttingen where he met Paul Alexandrov and began a lifelong friendship. He then stayed for the academic year 1927 to 1928 with Alexandrov at Princeton University, on a Rockefeller fellowship. Solomon Lefschetz, Oswald Veblen, and James W. Alexander were all at Princeton at the time. Back in Berlin, Hopf began working with Alexandrov, at the suggestion of Richard Courant, on a book on topology. ETH was fortunate to find in him a very active young mathematician, a wonderful teacher, and an extremely nice person. Using original sources, Urs Stammbach provides an enthralling description of all that happened on this occasion. Heinz Hopf quickly became well integrated into life in Switzerland. In another recent publication, Stammbach recalls how Hopf and his wife were arrested by the Gestapo in 1939 while on a visit to his parents, and how they were freed thanks to the massive help of the president of ETH, Swiss friends, and colleagues. They became Swiss citizens in 1944. Nevertheless, Hopf's candidacy for the presidency of the SMS in 1947 was rejected, one of the reasons being that he had not been born in Switzerland. The first mathematician of foreign origin to serve as president of the SMS was the French algebraist Peter Gabriel (University of Zurich), for the period 1980 to 1981.

Heinz Hopf had a large number of Ph.D. students. Among them were Eduard Stiefel, Beno Eckmann, and Michel Kervaire. All three played major roles in mathematics in Switzerland, as I will mention later in this review.

To respect seniority, we next turn to two eminent members of the SMS from the French-speaking part of Switzerland: Michel Plancherel and Georges de Rham. Famous for his theorems in harmonic analysis, Plancherel completed his Ph.D. in Fribourg under a mathematician of Czech origin, Mathias Lerch. Professor from 1911 to 1920 in Fribourg, Plancherel then moved to ETH to succeed Hurwitz. His teaching was well known for its clarity but also for being delivered in very rapid French. In their essay, Norbert Hungerbühler and Martine Schmutz recall that Plancherel was also an excellent organizer and a committed person, who undertook many charges both inside academia and outside in public life. As a high officer of the Swiss army he was responsible for press censorship in Switzerland during the period 1943 to 1945, some years after Fueter. He retired in 1955 and was killed by a car in 1967 on a pedestrian crossing near ETH.

After studies in Lausanne, de Rham worked in Paris under Henri Lebesgue. Very attached to the Romandie and its mountains, he received a joint position at the universities of Geneva and Lausanne. He interacted with many mathematicians and had an important role in the development of mathematics, both in Switzerland and internationally. His work is still strongly represented in modern mathematics. A "glimpse of the de Rham era" offered by Shristi Chatterji and Manuel Ojanguren in the book illuminates many aspects of his attractive personality. Based on de Rham's correspondence and related documents, the essay also provides information about episodes that were not well known or were poorly understood. It is one of the few places in the volume where events during World War II are evoked.

At de Rham's initiative, workshops began to be organized around 1958 in the French-speaking part of Switzerland. Later an official framework, the "troisième cycle romand", financially supported by the different cantons, was created for these workshops. In 1971 Kervaire succeeded de Rham in Geneva. After completing his Swiss thesis with Hopf in 1955, Kervaire, a French citizen, submitted another thesis (on higher-dimensional knots) in Paris in 1964, under Henri Cartan. Claude Weber told me that Cartan was hoping that Kervaire would accept a position in Paris; at that time the French "Doctorat d'État" was a requirement. Nonetheless Kervaire stayed at the Courant Institute from 1959 to 1971 before moving to Geneva. Immediately de Rham asked him to take responsibility for the workshops. Weber, who very efficiently assisted Kervaire in administrative matters, gives us a lively and detailed account of the workshops, especially the famous ones in the small mountainous village of Plans-sur-Bex. The very rustic character of the residence was highly compensated for by the quality of the food, as the numerous participants from all over the world will forever remember!

Stiefel, who completed his Ph.D. thesis in 1935 as one of Hopf's first students in Switzerland, was appointed full professor at ETH in 1943. Well known for his contributions to topology and Lie groups (Stiefel manifolds, Stiefel-Whitney classes!), he switched fields completely around 1948. His new Institute of applied mathematics rapidly gained international recognition, and his leasing of the Zuse electromechanical computer remains famous. Stiefel is primarily responsible for establishing electronic scientific computing in Switzerland. Heinz Rutishauser, one of Stiefel's early collaborators, was very active in the definition of the pioneering algorithmic language Algol 60. Like Fueter and Plancherel, Stiefel was a colonel in the Swiss army.

Stiefel's most famous masters-degree student was Armand Borel, born in la Chaux-de-Fonds in Switzerland. As Stiefel was changing his field of research, Borel moved to Paris to work with Jean Leray for his Ph.D. He immediately participated in the very active Parisian mathematical life. Back to Zurich in 1955 as a full professor, he shortly thereafter left for the Institute for Advanced Study in Princeton. As André Haefliger emphasizes, Borel was not only a great mathematician, but also a man of culture and a propagator of new ideas. One example of this may be seen in the "Borel Seminars" organized every summer in Bern from 1983 to 1986, when Borel came back to Switzerland, again as a faculty member of ETH. Professors and students from all the Swiss universities as well as many foreign guests took part in these seminars. The tradition still continues, although in a different way. The 2011 seminar, for instance, had as its theme real and complex hyperbolic geometry.

Eckmann, born in Bern, studied mathematics at ETH and did his Ph.D. under Hopf in 1941. He began his career in Lausanne, where he had close contacts with de Rham, and he then returned to ETH in 1948. He also attracted many Ph.D. students in differential geometry, algebraic topology, and algebra. Eckmann was interviewed for the EMS newsletter in January 2007, as he was nearing 90; the editors include the interview in this volume. Recalling his student days, Eckmann remembers that Plancherel was a very old-fashioned teacher, but in fact not a bad one! By contrast, Hopf was very modern, teaching in the style of van der Waerden's "Moderne Algebra".

In 1962, as president of the SMS, Eckmann recommended the establishment a Swiss Mathematical Research Institute under the patronage of the SMS that would be financed by the Swiss National Science Foundation. The idea was rejected as being too centralizing. Soon afterward, Eckmann founded the Forschungsinstitut für Mathematik at ETH, today well known as FIM, to have an organization linked to the department, and to welcome visitors to collaborate with faculty members. Although many such institutes exist nowadays, this was not true then. Eckmann ran the Institute for twenty years; Jürgen Moser succeeded him in 1984. Eckmann was also involved in mathematical publishing, particularly as the initiator of the Springer Lecture Notes in Mathematics.

This volume contains contributions dedicated to many other mathematicians, their work, and their interaction with mathematics in Switzerland: Martin Eichler, Hugo Hadwiger, Heinz Huber, Jürgen Moser, Rolf Nevanlinna, Alexander Ostrowski, and Andreas Speiser. Choices had to be made by the editors, but it is a pity that personalities such as George Pólya, Paul Bernays, and Bartel Leendert van der Waerden are absent. Born in Hungary, Pólya came as a Privat-Dozent to the ETH in 1914. Promoted to full professor in 1928, he left Switzerland for the United States in 1940. Of Swiss origin, Bernays left Göttingen in 1933 for Zurich, first in a temporary position at ETH and then as faculty member. Van der Waerden succeeded Fueter at the University of Zurich in 1951 and spent the rest of his career there.

It is striking to observe that women are entirely missing. "In Switzerland mathematics is not considered as a science for women, even if women study mathematics here since more than [a] hundred years," Christine Riedtmann begins her contribution on women mathematicians in Switzerland. She provides much interesting information that was not easy to collect: about the first female students, their short biographies, women involved in the SMS. Things are changing very slowly here. Let us recall that Switzerland's women were granted full voting rights only in 1971. With the increasing number of women students, there is hope that the proportion of women in faculties will also increase.

There is no systematic presentation of mathematics institutes in the volume, although some of them (Fribourg, Neuchâtel, and the École polytechnique in Lausanne) are briefly described. The École was first an engineering school, a dependency of the University of Lausanne, until 1969 when it was taken over by the federal state as the second Federal Institute of Technology (EPFL) in Switzerland (the first being ETH in Zurich). New departments of chemistry, mathematics, and physics were created at EPFL independently of the existing ones at the University of Lausanne. Charles Blanc, who introduced applied mathematics at the École polytechnique de Lausanne, played a central role in the creation of the new Mathematics Department. To everyone's satisfaction, chemists, mathematicians, and physicists of the University joined EPFL in 2003.

An indirect consequence of the setting up of EPFL was the cessation of the courses taught in French at ETH. Today, as in other institutions, many classes are in English.

Alain Robert's presentation of mathematics in Neuchâtel from 1950 to 1990 is full of interesting details. Robert was a direct witness, first as a student and then as a faculty member, of many events that occurred there, and his report is enhanced by a number of personal memories. Besides, he makes no bones about the hopes and difficulties of the smallest mathematical institute in Switzerland.

Other personal remembrances published in the volume include Christian Blatter's nice description of his life in Basel as a student of mathematics during the late 1950s, and Hirzebruch's essay. Blatter tells us why he decided to study mathematics, how he was thrilled by most of his teachers, and how he worked under Heinz Huber for his Ph.D. on extremal lengths. Friedrich Hirzebruch remembers how warmly Hopf and his wife welcomed him, and gives us a lot of information about his time in Zurich (1948 to 1950), which he obviously greatly enjoyed.

I would like to thank Elaine McKinnon Riehm and Marjorie Wikler Senechal for their help in the preparation of this review.

Department of Mathematics
ETH
CH-8092, Zurich
Switzerland
e-mail: knus@math.ethz.ch

The Mathematics of Life

by Ian Stewart

NEW YORK: BASIC BOOKS, 2011, 358 PP., US $27.99,
ISBN 978-0-465-02238-0

Life's Other Secret: The New Mathematics of the Living World

by Ian Stewart

NEW YORK: JOHN WILEY AND SONS, 1998, 285 PP., US $16.95,
ISBN 978-0-471-29651-5

REVIEWED BY JOHN PASTOR

It is interesting to speculate that mathematics and biology both may have had their origin in a common problem faced by our ancient ancestors: is there enough of something to feed all of us? Of course, it took many thousands of years before Georg Cantor showed that this could be answered by putting elements of a set into a one-to-one correspondence (if the food is a set of discrete objects such as apples) and for Charles Elton [1] to show that this problem is the first step in the development of the ecological concept of a food web. But many biologists have, until recently, looked on their more mathematically oriented colleagues with a certain degree of suspicion. The biological implications of equations often strike experimental biologists as akin to pulling a rabbit out of a hat; most biologists prefer conclusions derived from real rabbits. But beginning with the twentieth century, the interplay of mathematics and biology has proven increasingly fruitful, to the point where Joel Cohen [2] could now say: "mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better."

Ian Stewart agrees in these two books. Molecular biologists are fond of telling the rest of us that once we understand genes, then we will understand the rest of life. But, as Stewart says in both books, "there is more to life than genes," because genes express themselves within the context of physical laws, and these laws are mathematical. This means that mathematical biologists have to think somewhat differently, more like physicists, than experimental or observational biologists. Indeed, a surprising number of today's mathematical biologists, such as John Harte, Don DeAngelis, or Robert May, as well as many others, began their careers as physicists.

Although published a dozen years later, *The Mathematics of Life* is not an update of *Life's Other Secret*, as Stewart's *From Here to Infinity* [3] was an update of his earlier *The Problems of Mathematics* [4]. In fact, *The Mathematics of Life* does not refer to *Life's Other Secret* at all, which is curious since the two books touch on many of the same topics, although often (but not always) from different points.

Stewart takes his inspiration for both books from D'Arcy Thompson's 1917 book *On Growth and Form* [5]. *On Growth and Form* is a curious book. In more than 1000 pages of very literary prose, Thompson argues that the development of organisms is as much a result of the constraints of physics as it is the result of evolution. If organismal growth and development is constrained by physics, Thompson said, then it should be describable by mathematics just like any other physical process. Despite this insight, there is very little real mathematics in *On Growth and Form*. Perhaps Thompson's deepest mathematical insight was that the shape of many organisms can be modeled as topological transformations of the shape of other, usually closely related, organisms. This was a remarkable insight for the times: Poincaré had just recently begun developing topology, and the discovery of regulatory genes, which apparently control organism shape, was to occur many decades in the future. Today, we realize that a mutation in a single gene, which regulates a host of other genes responsible for development, has the potential to distort topologically the surface area, and hence the shape, of an organism.

But other than that, when people refer to Thompson's book at all, it is usually for the literary quality of its prose. Stewart, as well as Stephan Jay Gould and some others, claim that Thompson was the founder of mathematical biology, but I do not know a single working mathematical biologist who has ever read Thompson's book or would even make such a claim. If asked, most of us would trace our discipline back to Lotka [6], Volterra [7], Hardy [8, 9], Fisher [10], and Haldane [11], all contemporaries of Thompson. (For an excellent book on the early history of mathematical biology, see [12]). As far as I am aware, none of these contemporaries, except Lotka, ever refers to Thompson, and Lotka only in passing in two footnotes in [6]. In fact, an argument can be made that Lotka was even more successful than Thompson in showing mathematically why the patterns of life are constrained by physical processes – that is why he first titled his book *Elements of Physical Biology*, although Dover later reissued it as *Elements of Mathematical Biology*. It is curious that Stewart never refers to these contemporaries of Thompson at all, especially as their lives are sources of the sort of enlightening anecdote that Stewart is a master of using, as he did so well in [4] and [3].

In the past twenty years, there is hardly a major area of mathematics that has not proved useful to understanding biological problems. Game theory is applied to evolution and animal behavior, and knot theory to the way DNA is tangled up in a cell and to the very difficult problem of how a particular section of a DNA strand, which constitutes a gene, is untangled to translate the information onto a strand of messenger RNA. Dynamical systems theory is used to understand the dynamics of ecosystems and food webs and how small changes in parameters lead to large and abrupt changes in behaviors through various bifurcations. Turing's reaction-diffusion theory is applied to understanding the origin of patterns at many levels, from the stripes on a shell to the patterns of plankton in the sea. Geometry and topology are used to understand the shapes of viruses.

Stewart is at his best when he clearly conveys mathematical models for biological processes and theories. In several

DOI 10.1007/s00283-011-9262-z

problems, he shows how thinking about dynamics on a surface in a phase space can convey biological insight. One example is Sewall Wright's theory of the Adaptive Landscape [13] described in *Life's Other Secret*. The biological problem here is that both mutation and natural selection cause a population to move in a space of adaptive traits, but how do we understand this movement? Wright proposed that we think of a landscape, or surface, of reproductive fitness in this space with numerous peaks and valleys and saddle points between them. The peaks are combinations of traits with high reproductive fitness, and the valleys are combinations with low reproductive fitness. Individuals in a population, as well as their descendants, have unique combinations of traits that place each of them at a particular point; the entire population is a set of points dispersed on this surface. Both mutation and natural selection can move populations across this adaptive surface. Reproductive fitness increases as populations climb adaptive peaks from adjacent valleys. Stewart suggests that we think of mutations moving the distributions of populations randomly across the adaptive surface, because descendants with a new value of a trait occupy new points dispersed around those occupied by their parents. But because mutations are random, these new points are just as likely to be lower in a valley as they are to be higher on a peak or even across a saddle point. In contrast, natural selection moves them in a biased way up the peaks, because individuals with combinations of traits with greater reproductive fitness leave more descendants than those with lower reproductive fitness. If a population is split in two, or if a mutation moves an individual and its descendants across a saddle point, each subpopulation may eventually occupy different peaks. They then become new species, because any interbreeding of one population with another moves the descendants off the peaks, thereby decreasing their fitness. Thus, the whole population wanders on this surface, because mutation maintains variation whereas natural selection moves that variation up the adaptive peaks by differential reproductive survival of individuals and their descendants.

Dynamics on surfaces in phase spaces also appear in *The Mathematics of Life* to help us understand how proteins fold. Proteins are long chains of amino acids, the sequence of amino acids being determined by the analogous sequence of base pairs in DNA. Proteins, known as enzymes, catalyze reactions between two other molecules by allowing the two molecules to come into contact with each other when they temporarily bind to adjacent sites on the protein's surface. Thus, the shape of the enzyme's surface is crucial to its proper functioning. But if the DNA determines the sequence of amino acids in the chain, what determines the shape? The shape comes about because the long chain of amino acids folds to minimize its free energy: a chain of amino acids has very high free energy, and that energy is minimized by folding into different shapes. Thus, we can think of the free energy of a protein as a surface with very steep valleys and saddle points in a phase space whose dimensions are other chemical properties of the protein that change during the process of folding. The protein folds in a sequence that decreases the free energy, sending the configuration down a valley on the free energy surface. Eventually the bottom of the valley is reached, free energy is minimized, and the protein shape is stable. However, some proteins don't fold properly because random fluctuations cause them to pass over a saddle point and enter the wrong valley. When such proteins, known as prions, are released in the environment, they can cause diseases such as Mad Cow Disease in cattle or Chronic Wasting Disease in deer.

Clearly, the dynamics of protein folding and evolution on surfaces in phase (and parameter) spaces is a fundamental mathematical model for biological processes. But such mathematical ideas are not always so carefully conveyed in these books as the dynamics of biological systems on surfaces in phase space. To give an example, Chapter 13 of *The Mathematics of Life* is about pattern formation. It begins with Turing's discovery of spatial instability and hence pattern formation in reaction-diffusion equations. These are two coupled differential equations, one for an "activator" that is autocatalytic (that is, whose abundance increases because of a self-promoting positive feedback) and one for an "inhibitor" that consumes or otherwise inhibits the activator. Both diffuse spatially. In biochemistry, activators can be pigments and inhibitors can be enzymes that destroy pigments [14]. In ecology, the activator can be a prey species and the inhibitor can be a predator [15]. Unfortunately, we are never shown even a simple example of these equations. Intuitively, diffusion would seem to destroy patterns, not create them. So how does the "reaction" between activator and inhibitor produce patterns when coupled with diffusion? The answer Stewart gives (p. 200) is: "…local nonlinearity plus global diffusion creates striking and often complex patterns (see Figure 54)." Although this is a succinct mathematical insight, I doubt that any biologist not already familiar with this body of work will understand the biological meaning of "local nonlinearity" or "global diffusion" or why the two together create "striking patterns." In fact, they don't always, only when the diffusion of the inhibitor is much greater than that of the activator. Otherwise, the diffusion of the two destroys any spatial patterns. John Maynard Smith gives a very good graphical explanation for pattern formation by reaction-diffusion mechanisms in Figure 38 of his book *Mathematical Ideas in Biology* [16]. Although Maynard Smith's book is written at a more advanced level than Stewart's, his graphical explanation makes transparent the reason for pattern formation, and the explanation could easily have been incorporated into *The Mathematics of Life*. We are also given no historical background of Turing's life and tragic end. Although such biographical details are not essential for understanding either the biology or the mathematics, the details enrich both and allow the reader to place them in a larger context. Stewart did this masterfully in *The Problems of Mathematics*, which, along with its superb exposition of important mathematical ideas, makes that earlier book such a joy to read.

So what are the mathematical ideas that underlie biology? Stewart's list includes phase space, continuity, connectivity, feedback, information, order/disorder, bifurcation, and symmetry, and he thinks we need even more. Perhaps. But it is equally interesting to consider a minimum set of mathematical ideas from which most of biology can be derived. My list, which partly overlaps Stewart's, would include: 1. Discreteness (of individuals) versus continuity (of concentrations of limiting resources along gradients, such as the gradient of light down through a canopy); 2. Combinations (how many ways can nucleotides, genes, or species be combined, and which produce viable organisms or food webs?); 3. Distributions (of traits within populations and across species); 4. Coupled systems and their dynamics (via feedbacks and bifurcations); 5. Graphs (of food webs or neurons). Is either list complete? It is too early in the game to know. But Stewart would say, and I agree, that the next several decades will be exciting times for both biology and mathematics. We can only hope that new mathematical ideas will come out of the synergy between these two disciplines that have been too long apart.

REFERENCES

[1] C. S. Elton, *Animal Ecology*, Sidgwick and Jackson, London, 1927.

[2] J. Cohen, Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better, *PLoS Biology* 2 (2004): e439. doi:10.1371/journal.pbio.0020439.

[3] I. Stewart, *From Here to Infinity*, Oxford University Press, Oxford, 1996.

[4] I. Stewart, *The Problems of Mathematics*, Oxford University Press, Oxford, 1987.

[5] D'Arcy W. Thompson, *On Growth and Form*, Cambridge University Press, Cambridge, 1917.

[6] A. J. Lotka, *Elements of Physical Biology*, Williams and Wilkins, Baltimore, 1925.

[7] V. Volterra, Fluctuation in the abundance of a species considered mathematically, *Nature* 118 (1926), 558-560.

[8] G. H. Hardy, Mendelian proportions in a mixed population, *Science* 28 (1908), 49-50.

[9] I. Rajapaske, L. Muir, and P. Martin, Hardy's ''small discovery'' remembered, *Notices of the American Mathematical Society* 55 (2008), 384-386.

[10] R. A. Fisher, *The Genetic Theory of Natural Selection*, Oxford University Press, Oxford, 1930.

[11] J. B. S. Haldane, A mathematical theory of natural and artificial selection, *Transactions of the Cambridge Philosophical Society* 23 (1924), 19-41.

[12] S. E. Kingsland, *Modeling Nature*, University of Chicago Press, Chicago, 1995.

[13] S. E. Wright, The roles of mutation, inbreeding, crossbreeding, and selection in evolution, *Proceedings of the Sixth International Congress on Genetics* (1932), 355–366.

[14] H. Meinhardt, *The Algorithmic Beauty of Sea Shells*, Springer, Heidelberg, 1995.

[15] A. Okubo and S. A. Levin, *Diffusion and Ecological Problems*, Springer, Heidelberg, 1980.

[16] J. Maynard Smith, *Mathematical Ideas in Biology*, Cambridge University Press, Cambridge, 1968.

Department of Biology
University of Minnesota Duluth
Duluth, MN 55812
USA
e-mail: jpastor@d.umn.edu

Figuring It Out: Entertaining Encounters with Everyday Math

by Nuno Crato

BERLIN, HEIDELBERG: SPRINGER-VERLAG, 2010, 237 PP., ISBN 978-3-642-04832-6, E-ISBN 978-3-642-04833-3

REVIEWED BY PAMELA GORKIN

Can you think of ten mathematical tales that most people would find interesting? Let's try: There's the RSA code, the birthday problem, the Monty Hall problem, voting paradoxes, and the Möbius band is pretty cool. Then there are problems and theorems that appear in books and films. *The DaVinci Code* has Fibonacci sequences, the golden ratio and Leonardo's Vitruvian Man making star appearances. Stieg Larsson's heroine, Lisbeth Salander, can't stop thinking about Fermat's last theorem. In the movie *The Mirror Has Two Faces* we get a detailed description of the twin prime conjecture. Ten stories weren't that hard to think of. But can you think of twenty? Nuno Crato can. In fact, in *Figuring it Out*, Crato thinks of more than fifty really fascinating stories. How do I know they are fascinating? Because I checked.

For example, Crato talks about "The Most Beautiful of All," and I know that according to readers of *Physics World*, the equation he is describing really is the most beautiful equation of all. Kenneth Chang described it this way in the *New York Times* (October 24, 2004): "Euler's equation [is] a purely mathematical construct that finds wide use in theoretical physics." Crato describes it better: "Using only seven symbols, it includes three basic operations and states the relationship between the five most important numbers in mathematics. Its beauty derives from its simultaneous simplicity, and its profundity."

A survey seemed like a good way to gauge interest, so I asked my students what they thought of Crato's story on the number of ways to tie a shoelace. Every single student thought it was at least a 3.5 out of 5, and most thought it was a 4 or 5. (There are a lot of ways to tie a shoelace. Thirty-six of these are described on "Ian's Shoelace Site": http://www.fieggen.com/shoelace/lacingmethods.htm.)

Since June 2011, Crato has been the Minister of Education, Higher Education and Science, of his country, Portugal. Since (in Crato's words) "even the well-educated often demonstrate a surprising ignorance of the history of mathematics and its advancements," his goal in this book is to bring mathematics to a general and probably well-educated audience. It is a collection of interesting mathematical tales arranged neatly into sections, each containing several short stories. The section "Everyday Matters" is the first of these.

What's a mathematical everyday matter? You know: things like computational complexity, fair division, sphere-packing problems, orthogonal squares, and increasing intelligence scores. Now these topics don't sound like "everyday matters," but that's what I find most impressive about Crato's writing. Each story begins with an example that captures your attention, followed by an explanation almost any interested reader would find understandable, and then there's just a little bit of mathematics that might be out of the reader's grasp. Overall, I think it's a good model for the intended audience, and I think it works.

The first story, for example, begins with the following situation: You have five friends who have some pretty intense conflicts with each other. How do you arrange for a "hassle-free" evening? We learn that there is an algorithm for this, and that this is known as a satisfiability or an SAT problem. Each restriction has two possibilities (Should we invite Bob or Barbara?), so this is a 2-SAT problem. We're already halfway into the (very short) chapter when we learn about Richard Karp's task "to find an automatic process for designing circuits with as few transistors as possible." Toward the end of the article, the reader learns that there are problems "that quickly become impossible to solve because their complexity increases exponentially with the number of variables and restrictions." This may challenge the general reader, whereas an expert may balk at some of the descriptions. Consider the following: "At present, a distinction is made between the 'type P' problems in which the complexity increases in polynomial time with the rise in the number of variables, and the 'non-P' type problems in which this does not happen." As a colleague of mine (who found this somewhat "murky") said, "'non-P' problems, as described here, are not non-P necessarily. That's the whole point of the *P* versus *NP* problem; are they distinct?"

Following "Everyday Matters" is a section entitled "The Earth is Round," in which Crato tackles how things work, such as your GPS or celebrating birthdays in the event that your birthday falls on February 29. Incidentally, you might be interested to know that if you were born on February 29, you might have to wait eight years between birthdays to celebrate. If you already knew that, did you know why? If not, you can find that out by reading the story aptly titled "February 29." This section is followed by "Secret Affairs," beginning with Alice and Bob, Inviolate Cybersecrets, Quantum Cryptography, the FBI Wavelet, and the Enigma Machine. Although this section is probably familiar to most readers of *The Mathematical Intelligencer*, for most non-mathematical readers, it's a nice introduction. If the reader doesn't like it, it's over pretty quickly – Alice and Bob's story takes only four pages. If the reader does like it, you can recommend Simon Singh's *The Code Book* as a way to follow up.

The next section, "Art and Geometry" includes The Golden Number, Escher and the Möbius Strip, Pollack's Fractals, Voronoi Diagrams, and the Most Beautiful of All. The opening story in this section is again a good example of Crato's ability to generate interest. In "The Vitruvian Man," Crato points out that "*The DaVinci Code* is a mixture of facts and fiction, which is perfectly acceptable as it is a novel. But as a reader you are entitled to know which of the book's facts have not been embellished."

"Art and Geometry" is followed by "Mathematical Objects" including The Power of Math, The Difficulty of Chance, Conjectures and Proofs, π Day, and The Best Job in the World. I'll focus on just two stories here. (To find out which

DOI 10.1007/s00283-011-9264-x

job is the best in the world, you'll need to purchase the book.) "Conjectures and Proofs" is a mathematical tale that focuses on the $3n + 1$ conjecture or "Collatz's conjecture," as it appears here. It's a conjecture that is easily understood, and it's easy for a general reader to check cases to see whether or not the conjecture has any chance at all of being true. In short, it's a good example of a conjecture. The second story, "The Difficulty of Chance," is both a mathematical and a psychological tale. You may remember that Benny Hill once said, "The odds against there being a bomb on a plane are a million to one, and against two bombs a million times a million to one. Next time you fly, cut the odds and take a bomb." It's a joke that most people get, but most people don't get. Crato drives this point home through a discussion of coin flipping. He tells us that "In one famous experiment psychologists asked people to write sequences of zeros and ones by imagining they were tossing a coin and writing 0 every time heads appeared and 1 every time tails appeared." If you haven't seen this before, some questions are natural: To what experiment is Crato referring? Who did it? Where? When? After discussing this with a psychologist, I learned that, in fact, these studies are well known and there are many that followed, such as the hot hand in basketball by T. Galovich. I don't mean to be too critical here. After all, the point is that I was so interested in the story that I wanted to know more.

The back cover tells us that this book contains "mathematical stories – funny and puzzling mathematical stories. They tell of villains who try to steal secrets, heroes who encode their messages, and mathematicians who spend years on end searching for the best way to pile oranges." For the mathematician who does study orange packing, this set of stories probably isn't puzzling. But the stories are clever and funny, the writing is clear and interesting, and the author is remarkably creative. Stories such as these are what draw people to mathematics, and they help non-mathematicians to see the mystery, power, and beauty of mathematics.

Department of Mathematics
Bucknell University
Lewisburg, PA 17837
USA
e-mail: pgorkin@bucknell.edu

Recent Mathematical Stamps: 2008–2009

Aristotle (384–322 BC)

Aristotle became a student at Plato's Academy at the age of 17, and he stayed there for 20 years. Fascinated by logical questions, he systematized the study of logic and deductive reasoning, discussing such syllogisms as: "All men are mortal; Socrates is a man; therefore Socrates is mortal." This unlikely depiction of him is from a set of stamps celebrating Europe's Year of Astronomy, 2009.

Johannes Kepler (1571–1630)

Kepler is remembered for the three laws of elliptical planetary motion from his *New Astronomy* (1609) and his *Harmony of the World* (1619). He was fascinated by conics and introduced the word *focus* into mathematics. He was also interested in polyhedra, discovering the cuboctahedron and the antiprisms, and by summing thin disks he calculated the volumes of over 90 solids of revolution, thereby foreshadowing the development of the integral calculus.

Damodar Dharmananda Kosambi (1907–1966)

Kosambi was an Indian mathematician and statistician, a Sanskrit scholar, and a numismatist. His main areas of mathematical interest were differential geometry and statistics, though he also contributed to many other fields. He is remembered in statistics as the originator of the technique of *proper orthogonal decomposition*, often named after its later discoverers Karhunen and Loève. In 1944 he introduced the *Kosambi map function*, shown on the stamp.

Isaac Newton (1642–1727)

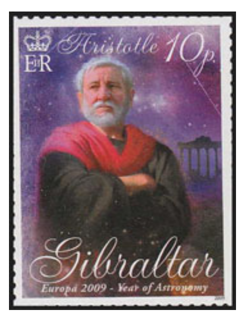
Several stamps have featured Sir Isaac Newton and his contributions to gravitation. This stamp, issued by the Congo, features his work on optics. Before Newton, telescopes contained lenses that were subject to chromatic aberration (colouring around the edges). Here we see Newton's reflecting telescope, which cured the problem and led to his election to Britain's Royal Society in 1672.

Blaise Pascal (1623–1662)

At the age of 16 Pascal introduced his "hexagon theorem" involving six points on a conic. Later he wrote a treatise on binomial coefficients (Pascal's triangle), and he investigated the theory of probability, atmospheric pressure (Pascal's principle in hydrodynamics), and the properties of cycloids and other curves. Here we see his calculating machine (the *Pascaline*), which could be used to add and subtract.

Polygon

Many stamps have assumed nontraditional shapes (scalene triangles, pentagons, parallelograms, ellipses, and others), but this irregular hexagonal stamp, issued by Macau for the 2008 Olympic Games in Beijing and featuring the "bird's nest stadium," is particularly unusual.



Aristotle

Kepler

Kosambi

Newton

Pascal

Polygon

➢ Please send all submissions to the Stamp Corner Editor,
Robin Wilson, Faculty of Mathematics,
Computing and Technology,
The Open University, Milton Keynes, MK7 6AA, UK
e-mail: r.j.wilson@open.ac.uk

DOI 10.1007/s00283-011-9269-5