Varieties of Mathematical Explanation

Jeremy Avigad

Department of Philosophy
Department of Mathematical Sciences

Hoskinson Center for Formal Mathematics

Carnegie Mellon University

Fields Medal Symposium: Akshay Venkatesh
October 19, 2022

Conceptual understanding

It is commonly held that the goal of modern mathematics is to acquire a *conceptual* understanding.

This is often associated with:

- complicated definitions
- lots of algebraic structures
- complicated arguments and shared expertise
- de-emphasis of calculation (which either doesn't provide understanding, or is uninteresting)
- unexpected connections between apparently different domains of mathematics

The notion has origins in the nineteenth century.

Conceptual understanding

"…in our opinion truths of this kind should be drawn from the notions involved rather than from notations." [Gauss, *Disquisitiones*, 1801]

" …for these kinds of questions there exist considerations of a metaphysical nature which hover over the calculations, and often make them useless. …For, all that creates the beauty and at the same time the difficulty of this theory [i.e. Abel's theory of elliptic functions] that one has ceaselessly to indicate the path of the analysis and to anticipate the results, without ever being able to carry them out." [Galois, 1830?]

"…to seek proofs based immediately on fundamental characteristics, rather than on calculation, and indeed to construct the theory in such a way that it is able to predict the results of calculation…" [Dedekind, 1871]

Conceptual understanding

From a down-to-earth point a view, it's a matter of conceptual engineering and information management.

- We need ways to cope with complexity.
- Modularization and algebraic abstraction
  - make information salient when it is needed
  - suppress information when it isn't
  - provide clean interfaces.
- Different representations have different affordances.

The way mathematics manages to do this is fascinating.

Outline

This talk:

- conceptual understanding
- challenges: computers in mathematics
- historical comparisons:
  - Newton and *Propositio Kepleriana*
  - Dirichlet's theorem on primes in an arithmetic progression
- morals
- bigger morals

Challenges

Four challenges from discrete geometry:

- the Kepler conjecture
- packing tetrahedra
- the Keller conjecture.
- optimal sphere packings in 8 and 24 dimensions

Challenges: Hales' theorem

Theorem

The optimal density achieved by a packing of nonoverlapping equally sized spheres in Euclidean space is the one attained by the face-centered cubic packing.

Challenges: Hales' theorem

An overview:

- Johannes Kepler asserted optimality in 1611.
- Thomas Hales (with Samuel Ferguson) announced a proof in 1998.
- The proof relies on:
  - a combinatorial enumeration of "tame graphs"
  - relaxations of nonlinear constraints to linear constraints
  - checking the infeasibility of linear constraints.

  All were obtained with the help of computers.
- Frustrated by referees, Hales launched an effort to formally verify the result in 2003 and completed it (with many collaborators) in 2014.

Challenges: packing tetrahedra

Theorem

There is a packing of equally sized regular tetrahedra in Euclidean space with density $4000/4671$.

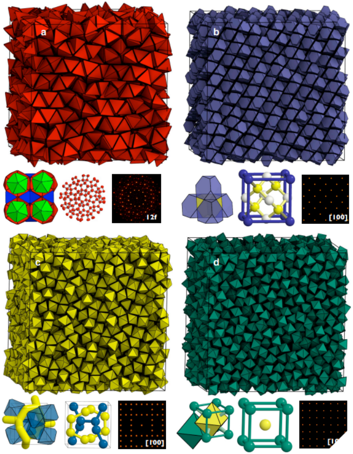


Figure 3. Structures of truncated tetrahedra self-assembled in simulation at intermediate density. In each subfigure a snapshot of the full simulation box, temporally averaged to remove thermal disorder together with a characteristic motif (bottom left),

Challenges: packing tetrahedra

Overview:

- In 2006, John Conway and Salvatore Torquato obtained a packing with a density of about 0.7175 manually.
- Other approaches parameterized families of solutions and used optimization software.
- In 2009, Sharon Glotzer and colleagues used Monte-Carlo simulations.
- The best result, due to Elizabeth Chen, Michael Engel, and Glotzer used a parameterized family, but searched for local improvements.

Challenges: higher-dimensional sphere packing

Theorem

The optimal density of an 8-dimensional sphere packing is attained by E_8*, and the optimal density of a 24-dimensional sphere packing is attained by the Leech lattice.*

Challenges: higher-dimensional sphere packing

Overview:

- Remarkably little is known about optimal sphere packings in higher dimensions.
- E_8 and the Leech lattice have long been known to provide remarkably efficient (and interesting) packings in 8 and 24 dimensions, respectively.
- In 2003, Henry Cohn and Noam Elkies showed that one could use Fourier analysis to obtain upper bounds on higher-dimensional packings.
- Using numerical methods, they found the best known bounds for dimensions 4–36. They showed that E_8 and the Leech lattice are close to optimal.
- With the right "magic functions," the methods could provide tight bounds in those two cases.

Challenges: higher-dimensional sphere packing

- Numeric calculations by Cohn and Stephen Miller gave some hints as to what these magic functions would look like.
- In 2016, Maryna Viazovska constructed a function to prove that the density achieved by the E_8 packing is optimal in eight dimensions. The construction was a *tour de force*, drawing on modular forms, properties of the Laplace transform, experimentation, and guesswork.
- Within a week of the announcement, Cohn, Abhinav Kumar, Miller, and Danylo Radchenko had joined her to extend the method to show the optimality of the packing based on the Leech lattice in twenty-four dimensions.

Challenges: the Keller conjecture

In 1930, Ott-Heinrich Keller conjectured that any tiling of n-dimensional space by unit cubes requires face-sharing.

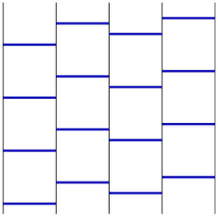


Figure 1: Two-dimensional tiling

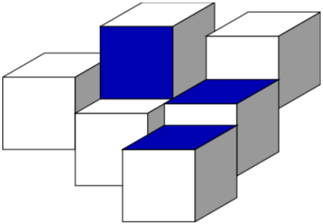


Figure 2: Three-dimensional tiling

Theorem

Keller's conjecture is true up to and including dimension 7, but fails for dimension 8 and above.

Challenges: the Keller conjecture

Overview:

- In 1940, Oskar Perron showed that the conjecture is true for $n \leq 6$.
- In 1990, Keresztély Corrádi and Sándor Szabó reduced the statement to one about finite graphs.
- In 1992, Jeffrey Lagarias and Peter Shor showed that it is false for $n \geq 10$.
- In 2002, John Mackey showed that it is false for $n \geq 8$.
- In 2019, Joshua Brakensiek, Marijn Heule, Mackey, and David Narváez showed that it is true for $n = 7$, using a SAT solver, symmetry-breaking, clever encodings, and heuristics.

Reactions

These are great results. But they don't fit the "conceptual" mold:

- The proofs are relatively short.
- They don't require a large scaffold of algebraic definitions and theorems.
- They use computers, in various ways: numeric and symbolic computation, optimization, simulation, and search.

Reactions: packing tetrahedra

4000/4671 isn't optimal.

The result is "experimental mathematics."

For some, this is a term of derision.

In any case, it doesn't feel like *real* mathematics.

Reactions: the Keller conjecture

SAT solvers are computer science, not mathematics.

"Don't expect mathematicians to be interested—this is a 'finite computation.' "

Reactions: the Kepler conjecture

"...for the purpose of this discussion, what I call a traditional mathematician is someone who has a permanent position that involves proving theorems .... What they do is try to get a conceptual and rigorous understanding of which mathematical statements are true. Both adjectives are important. To them, Hales' proof of the Kepler conjecture is nothing like solving the conjecture. And the Flyspeck project is purely computer science. That doesn't mean it is not interesting; it is something different, because at least part of the proof lacks the 'conceptual' adjective. You will often hear mathematicians, talking about their own proofs, saying things like 'there is nothing to understand here, it's only a computation.'"

Reactions: higher-dimensional sphere packing

We denote the space of weakly-holomorphic modular forms of weight k and group Γ by $M_k^!(\Gamma)$. The spaces $M_k^!(\Gamma)$ are infinite dimensional. Probably the most famous weakly-holomorphic modular form is the *elliptic j-invariant*

$$j := \frac{1728\, E_4^3}{E_4^3 - E_6^2}.$$

This function belongs to $M_0^!(\Gamma(1))$ and has the Fourier expansion

$$\begin{aligned} j(z) = q^{-1} + 744 + 196884\, q + 21493760\, q^2 \\ + 864299970\, q^3 + 20245856256\, q^4 + O(q^5), \end{aligned}$$

where $q = e^{2\pi i z}$. Using a simple computer algebra system such as PARI GP or Mathematica one can compute the first hundred terms of this Fourier expansion within a few seconds. An important question is to find an asymptotic formula for $c_j(n)$, the n-th Fourier coefficient of j. Using the Hardy–Ramanujan circle method [17, pp. 460–461] or the nonholomorphic Poincaré series [15], one can show that

Reactions: higher-dimensional sphere packing

where $Df(z) = \frac{1}{2\pi i}\frac{d}{dz}f(z)$. These identities combined with (20) and (25) give the asymptotic formula for the Fourier coefficients $c_{\phi_{-4}}(n)$, $c_{\phi_{-2}}(n)$, and $c_{\phi_0}(n)$. The first several terms of the corresponding Fourier expansions are

$$\phi_{-4}(z) = q^{-1} + 504 + 73764\,q + 2695040\,q^2 + 54755730\,q^3 + O(q^4), \tag{32}$$

$$\phi_{-2}(z) = 720 + 203040\,q + 9417600\,q^2 + 223473600\,q^3 + 3566782080\,q^4 + O(q^5), \tag{33}$$

$$\phi_0(z) = 518400\,q + 31104000\,q^2 + 870912000\,q^3 + 15697152000\,q^4 + O(q^5), \tag{34}$$

Reactions: higher-dimensional sphere packing

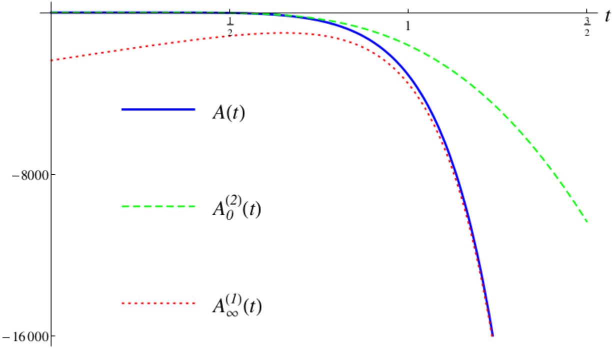


Figure 1. Plot of the functions $A(t)$, $A_0^{(2)}(t) = -\frac{368640}{\pi^2}\, t^2\, e^{-\pi/t}$, and $A_\infty^{(1)}(t) = -\frac{72}{\pi^2}\, e^{2\pi t} + \frac{8640}{\pi} t - \frac{23328}{\pi^2}$.

Our goal is to show that $A(t) < 0$ for $t \in (0, \infty)$. The function $A(t)$ is plotted in Figure 1. We observe that we can compute the values of $A(t)$ for $t \in (0, \infty)$ with any given precision. Indeed, from identities (29) and (45) we

Reactions: higher-dimensional sphere packing

Therefore, we can estimate the error terms in the asymptotic expansions (63) and (64) of $A(t)$

$$\left|A(t) - A_0^{(m)}(t)\right| \le \left(t^2 + \frac{36}{\pi^2}\right) \sum_{n=m}^{\infty} 2e^{2\sqrt{2}\pi\sqrt{n}}\, e^{-\pi n/t},$$

$$\left|A(t) - A_\infty^{(m)}(t)\right| \le \left(t^2 + \frac{12}{\pi} t + \frac{36}{\pi^2}\right) \sum_{n=m}^{\infty} 2e^{2\sqrt{2}\pi\sqrt{n}}\, e^{-\pi n t}.$$

For an integer $m \ge 0$, we set

$$R_0^{(m)} := \left(t^2 + \frac{36}{\pi^2}\right) \sum_{n=m}^{\infty} 2e^{2\sqrt{2}\pi\sqrt{n}}\, e^{-\pi n/t},$$

$$R_\infty^{(m)} := \left(t^2 + \frac{12}{\pi} t + \frac{36}{\pi^2}\right) \sum_{n=m}^{\infty} 2e^{2\sqrt{2}\pi\sqrt{n}}\, e^{-\pi n t}.$$

Using interval arithmetic we check that

$$\begin{aligned}
\left|R_0^{(6)}(t)\right| &\le \left|A_0^{(6)}(t)\right| && \text{for } t \in (0,1],\\
\left|R_\infty^{(6)}(t)\right| &\le \left|A_\infty^{(6)}(t)\right| && \text{for } t \in [1,\infty),\\
A_0^{(6)}(t) &< 0 && \text{for } t \in (0,1],\\
A_\infty^{(6)}(t) &< 0 && \text{for } t \in [1,\infty).
\end{aligned}$$

Reactions: higher-dimensional sphere packing

"…it's wonderful to see a relatively simple proof of a deep theorem in sphere packing."

"Her proof is thus a notable contribution to the story of E_8, and more generally the story of exceptional structures in mathematics."

"Viazovska … establishes a new connection between modular forms and discrete geometry."

"Instead of justifying sphere packing by aspects of the problem or its applications, we'll justify it by its solutions: a question is good if it has good answers. Sphere packing turns out to be a far richer and more beautiful topic than the bare problem statement suggests. From this perspective, the point of the subject is the remarkable structures that arise as dense sphere packings."

Reactions: higher-dimensional sphere packing

"…modular forms are deep and mysterious functions connected with lattices, as are the magic functions, so wouldn't it make sense for them to be related?"

"Despite our lack of understanding [of other higher dimensional sphere packing problems], the special role of eight and twenty-four dimensions aligns with our experience elsewhere in mathematics. Mathematics is full of exceptional or sporadic phenomena that occur in only finitely many cases, and the E_8 and Leech lattices are prototypical examples. These objects do not occur in isolation, but rather in constellations of remarkable structures."

Reactions: higher-dimensional sphere packing

A Conceptual Breakthrough in Sphere Packing

Henry Cohn

Maryna Viazovska solved the sphere packing problem in eight dimensions.

Reflection

Recent results in discrete geometry don't fit our image of mathematical understanding.

Should we reconsider that image?

Can we make room for computational results?

Let's get some historical perspective:

- Newton and the inverse square law of gravity.
- Dirchlet's theorem on primes in an arithmetic progression.

History: the inverse square law

In 1609, based on data from Tycho Brahe, Kepler determined that the planets travel in ellipses with the sun at the focus.

A centerpiece of Newton's *Philosophiæ Naturalis Principia Mathematica* (1687) is the *Propositio Keplieriana*:

> *Assuming the laws of motion, given that the path of a body is an ellipse and only force on it is directed to one focus, that force is inversely proportional to the distance.*

In other words, assuming gravitational force accounts for the motion, he derived the inverse square law.

History: the inverse square law

In Newton's *Principia* it was demonstrated that Kepler's first law- the planets follow elliptical orbits, with the sun at a focus- *implies* that gravity follows an 'inverse-square law', under the assumption that it is a an attractive central force directed along a line from the planet to the Sun (that is, towards the origin). What follows is Newton's argument in modern notation. Start from the polar equation of an ellipse, with the origin at a focus:

$$r = \frac{a}{1 + e\cos\theta}.$$

Differentiating twice in time and using l to eliminate θ', we find:

$$r' = \frac{ae\sin\theta}{(1 + e\cos\theta)^2}\theta' = \frac{le}{a}\sin\theta,$$

$$r'' = \frac{le}{a}(\cos\theta)\theta' = \frac{l^2 e}{ar^2}\cos\theta.$$

History: the inverse square law

Using the relation $e\cos\theta = ar^{-1} - 1$, we can express this in terms of r only:

$$r'' = \frac{l^2}{ar^2}(\frac{a}{r} - 1) = \frac{c^2}{r^3} - \frac{c^2}{ar^2}.$$

Now recall the expression for the radial component of Newton's second law of motion:

$$F_r = ma_r = m(r'' - r(\theta')^2),$$

to obtain (again using l to eliminate θ'):

$$F_r = m(\frac{l^2}{r^3} - \frac{l^2}{ar^2} - r\frac{l^2}{r^4}) = -\frac{ml^2}{a}\frac{1}{r^2},$$

showing that the magnitude of the force follows an inverse-square law.

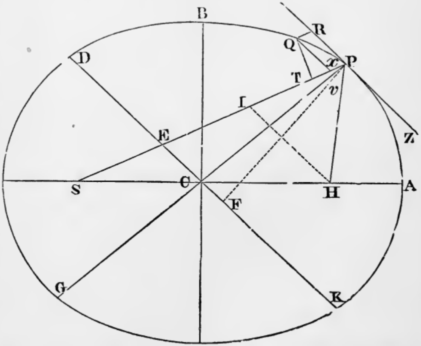
History: the inverse square law

SECTION III.

Of the motion of bodies in eccentric conic sections.

PROPOSITION XI. PROBLEM VI.

If a body revolves in an ellipsis; it is required to find the law of the centripetal force tending to the focus of the ellipsis.

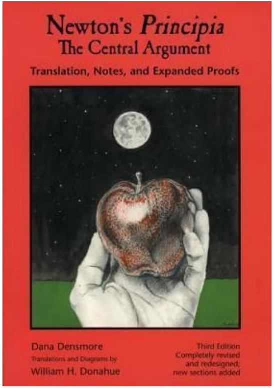


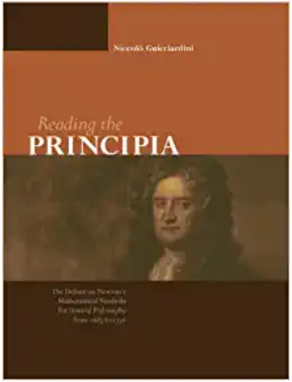
Let S be the focus of the ellipsis. Draw SP cutting the diameter DK of the ellipsis in E, and the ordinate Qv in x; and complete the parallelogram QxPR. It is evident that EP is equal to the greater semi-axis AC: for drawing HI from the other focus H of the ellipsis parallel to EC, because CS, CH are equal, ES, EI will be also equal; so that EP is the half sum of PS, PI, that is (because of the parallels HI, PR, and the equal angles IPR, HPZ), of PS, PH, which taken together are equal to the whole axis 2AC. Draw QT perpendicular to SP, and putting L for the princi al latus rectum of the ellipsis (or for

History: the inverse square law

$\frac{2BC^2}{AC}$), we shall have L × QR to L × Pv as QR to Pv, that is, as PE or AC to PC; and L × Pv to GvP as L to Gv; and GvP to Qv^2 as PC^2 to CD^2; and by (Corol. 2, Lem. VII) the points Q and P coinciding, Qv^2 is to Qx^2 in the ratio of equality; and Qx^2 or Qv^2 is to QT^2 as EP^2 to PF^2, that is, as CA^2 to PF^2, or (by Lem. XII) as CD^2 to CB^2. And compounding all those ratios together, we shall have L × QR to QT^2 as AC × L × PC^2 × CD^2, or 2CB^2 × PC^2 × CD^2 to PC × Gv × CD^2 × CB^2, or as 2PC to Gv. But the points Q and P coinciding, 2PC and Gv are equal. And therefore the quantities L × QR and QT^2, proportional to these, will be also equal. Let those equals be drawn into $\frac{SP^2}{QR}$, and L × SP^2 will become equal to $\frac{SP^2 \times QT^2}{QR}$. And therefore (by Corol. 1 and 5, Prop. VI) the centripetal force is reciprocally as L × SP^2, that is, reciprocally in the duplicate ratio of the distance SP. Q.E.I.

History: the inverse square law





History: the inverse square law

Algebra and geometry:

- Early algebra was fundamentally *about* geometric magnitudes and constructions.
- In the 1630s, Descartes showed that symbolic algebraic methods could be used to solve difficult geometric problems.
- In 1666, Newton developed a symbolic version of calculus using algebraic methods.
- He soon became very critical of Cartesian methods for not providing proper understanding.
- The *Principia*, published in 1687, is explicitly geometric.

History: the inverse square law

Christiaan Huygens, 14 years senior to Newton, was a towering figure in 17th century science.

He was critical of Newton for straying from a pure geometric interpretation.

The theory of proportions dealt with sameness of ratios: $A/B : C/D$.

You can only "compound" ratios with the same middle term, $A/B \times B/C$ is A/C.

To "compound" A/B and C/D, one has to construct an F such that $B/F : C/D$, in which case, the compound ratio is A/F.

History: Dirichlet's theorem

Theorem

Let m *and* k *be relatively prime. Then the arithmetic progression* $m, m+k, m+2k, \ldots$ *contains infinitely many primes.*

For example, there are no primes in the sequence

$$6, 15, 24, 33, 42, 51, \ldots.$$

There are infinitely many primes in the sequence

$$5, 14, 23, 32, 41, 50, \ldots$$

History: Dirichlet's theorem

Legendre assumed this in 1798, in giving a purported proof of the law of quadratic reciprocity.

Gauss pointed out this gap, and presented two proofs of quadratic reciprocity in his *Disquisitiones Arithmeticae* of 1801.

He ultimately published six proofs of quadratic reciprocity, and left two more proofs in his *Nachlass*. But he never proved the theorem on primes in an arithmetic progression.

Dirichlet's 1837 proof is notable for the sophisticated use of analytic methods to prove a number-theoretic statement.

History: Dirichlet's theorem

If G is a finite abelian group, χ is a *character on* G if it is a homomorphism from G to the nonzero complex numbers, i.e.

$$\chi(g_1 g_2) = \chi(g_1)\chi(g_2)$$

for every g_1 and g_2 in G.

The following two "orthogonality" relations hold:

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi = \chi_0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G| & \text{if } g = 1 \\ 0 & \text{otherwise} \end{cases}$$

History: Dirichlet's theorem

Modern presentations of Dirichlet's theorem use characters on $(\mathbb{Z}/k\mathbb{Z})^*$.

They define the Dirichlet L-functions:

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Euler product expansion:

$$L(s,\chi) = \prod_{q} \left(1 - \frac{\chi(q)}{q^s}\right)^{-1} = \prod_{q \nmid k} \left(1 - \frac{\chi(q)}{q^s}\right)^{-1}$$

This converges when $Re(s) > 1$.

History: Dirichlet's theorem

Taking logarithms of both sides:

$$\log L(s, \chi) = \sum_{q \nmid k} \frac{\chi(q)}{q^s} + O(1).$$

Multiply both sides by $\overline{\chi(m)}$ and sum over χ.

$$\sum_{\chi} \overline{\chi(m)} \log L(s, \chi) = \sum_{\chi} \sum_{q \nmid k} \overline{\chi(m)} \frac{\chi(q)}{q^s} + O(1).$$

Interestingly, they do not require knowing how to name, describe, or calculate any particular characters.

It can be done, though.

History: Dirichlet's theorem

In the case where the common difference is a prime, p:

- Let c be a primitive element modulo p.
- For every n coprime to p, let γ_n be such that $c^{\gamma_n} \equiv n \bmod p$.
- Characters χ correspond to $p-1$st roots of unity ω, where $\chi(n) = \omega^{\gamma_n}$.

Pick a generator Ω of the $p-1$st roots of unity, $\{\Omega^0, \ldots, \Omega^{p-2}\}$.

Instead of $L(s, \chi)$, we can write $L_m(s)$ for the L-series corresponding to the root Ω^m.

History: Dirichlet's theorem

In the case where the modulus k is not prime:

- Decompose $(\mathbb{Z}/k\mathbb{Z})^*$ into a product of cyclic groups.
- Choose generators for each cyclic group.
- A unit n modulo k has indices α_n, β_n, γ_n, $\gamma'_n, \ldots$
- Each character $\chi(n)$ corresponds to a choice of roots of unity, $\theta, \varphi, \omega, \omega', \ldots$ by the equation $\chi(n) = \theta^\alpha \varphi^\beta \omega^\gamma \omega'^{\gamma'}$.

I left the dependence on n is left implicit.

If we choose appropriate primitive roots of unity, each character is given by a list of indices $a, b, c, c', \ldots$.

Instead of $L(s, \chi)$, we can write $L_{a,b,c,c',\ldots}$.

History: Dirichlet's theorem

Man hat daher die Gleichung:

$$\Sigma \frac{1}{q^{1+\varrho}} + \tfrac{1}{2}\Sigma \frac{1}{q^{2+2\varrho}} + \tfrac{1}{3}\Sigma \frac{1}{q^{3+3\varrho}} + \cdots$$
$$= \frac{1}{p-1}\left(\log L_0 + \Omega^{-\gamma_m}\log L_1 + \Omega^{-2\gamma_m}\log L_2 + \cdots + \Omega^{-(p-2)\gamma_m}\log L_{p-2}\right),$$

wo sich die erste Summation auf alle Primzahlen q der Form $\mu p + m$ erstreckt, die zweite auf alle Primzahlen q, deren Quadrate, die dritte auf alle Primzahlen q, deren Cuben, u. s. w. in derselben Form enthalten sind. Denkt man sich nun ϱ unendlich klein werdend, so wird die zweite Seite durch das Glied $\log L_0$

History: Dirichlet's theorem

III. Es sei nun:

$$k = 2^{\lambda} p^{\pi} p'^{\pi'} \ldots,$$

wo, wie in II. 2, $\lambda \gtreqless 3$ ist, und p, p', ... von einander verschiedene ungerade Primzahlen bezeichnen. Hat man irgend eine durch keine der Primzahlen $2, p, p', \ldots$ theilbare Zahl n, und kennt man die den Moduln:

$$4, \quad 2^{\lambda}, \quad p^{\pi}, \quad p'^{\pi'}, \quad \ldots$$

und ihren primitiven Wurzeln:

$$-1, \quad 5, \quad c, \quad c', \quad \ldots$$

entsprechenden Indices:

$$\alpha_n, \quad \beta_n, \quad \gamma_n, \quad \gamma'_n, \quad \ldots,$$

so hat man die Congruenzen:

$$(-1)^{\alpha_n} \equiv n \pmod{4}, \qquad 5^{\beta_n} \equiv \pm n \pmod{2^{\lambda}},$$
$$c^{\gamma_n} \equiv n \pmod{p^{\pi}}, \quad c'^{\gamma'_n} \equiv n \pmod{p'^{\pi'}}, \quad \ldots,$$

History: Dirichlet's theorem

so kommt, mit Berücksichtigung der oben erwähnten Eigenschaften der Indices und der Gleichungen (9):

$$(10) \qquad \Pi \frac{1}{1 - \theta^{\alpha} \varphi^{\beta} \omega^{\gamma} \omega'^{\gamma'} \cdots \frac{1}{q^s}} = \Sigma \theta^{\alpha} \varphi^{\beta} \omega^{\gamma} \omega'^{\gamma'} \cdots \frac{1}{n^s} = L,$$

wo sich das Multiplicationszeichen auf die ganze Reihe der Primzahlen, mit Ausschluss von 2, p, p', ..., und das Summenzeichen auf alle positiven ganzen

History: Dirichlet's theorem

Zwecke nehme man die Logarithmen von beiden Seiten der Gleichung (10) und entwickle; man erhält so:

$$\Sigma\,\theta^{\alpha}\,\varphi^{\beta}\,\omega^{\gamma}\,\omega'^{\gamma'}\cdots\frac{1}{q^{1+\varrho}}+\tfrac{1}{2}\,\Sigma\,\theta^{2\alpha}\,\varphi^{2\beta}\,\omega^{2\gamma}\,\omega'^{2\gamma'}\cdots\frac{1}{q^{2+2\varrho}}+\cdots=\log L,$$

wo die Indices α, β, γ, γ', ... zu q gehören, und auch das Zeichen Σ sich auf q bezieht. Stellt man die Wurzeln θ, φ, ω, ω', ... auf die im §. 8 ange-

History: Dirichlet's theorem

d. h. wenn $q^h \equiv m \pmod{k}$ ist, in welchem Falle $W = K$ wird. Unsere Gleichung wird daher:

$$(15)\quad \left\{ \begin{aligned} & \Sigma \frac{1}{q^{1+\varrho}} + \tfrac{1}{2}\Sigma \frac{1}{q^{2+2\varrho}} + \tfrac{1}{3}\Sigma \frac{1}{q^{3+3\varrho}} + \cdots \\ & \qquad = \frac{1}{K}\Sigma \Theta^{-\alpha_m \mathfrak{a}} \Phi^{-\beta_m \mathfrak{b}} \Omega^{-\gamma_m \mathfrak{c}} \Omega'^{-\gamma'_m \mathfrak{c}'} \ldots \log L_{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{c}', \ldots}, \end{aligned} \right.$$

wo sich die Summationen auf der ersten Seite resp. auf alle Primzahlen q beziehen, deren erste, zweite, dritte Potenzen in der Form $\mu k + m$ enthalten sind,

History: Dirichlet's theorem

- Dirichlet 1837: Dirichlet's original proof
- Dirichlet 1840, 1841: extensions to Gaussian integers, quadratic forms
- Dedekind 1863: presentation of Dirichlet's theorem
- Dedekind 1879, Weber 1882: characters on arbitrary abelian groups
- Hadamard 1896: presentation of Dirichlet's theorem and extensions
- de la Vallée Poussin 1897: presentation of Dirichlet's theorem and extensions
- Kronecker (1901, really 1870s and 1880s): constructive, quantitative treatment
- Landau 1909, 1927: presentation of Dirichlet's theorem and extensions

History: Dirichlet's theorem

Mathematicians quickly realized that Dirichlet's proof could be modularized to suppress the details of the representations.

But they still described the representations and made the translation explicit.

In other words, the proofs were fundamentally *about* the symbolic expressions, even though they became increasingly hidden.

It was a long time for mathematicians to dispense with computational detail entirely.

Moral 1: Views of understanding change

It was once a common view that mathematics is fundamentally about geometric magnitudes, and that geometric understanding is essential to mathematics.

Moral 1: Views of understanding change

It was once a common view that mathematics is fundamentally about geometric magnitudes, and that geometric understanding is essential to mathematics.

It was once a common view that mathematics is fundamentally about symbolic representations and calculation, and that a computational understanding is essential to mathematics.

Moral 1: Views of understanding change

It was once a common view that mathematics is fundamentally about geometric magnitudes, and that geometric understanding is essential to mathematics.

It was once a common view that mathematics is fundamentally about symbolic representations and calculation, and that a computational understanding is essential to mathematics.

It is now a common view that mathematics is fundamentally about abstract structures and the relationships between them, and that a conceptual understanding is essential to mathematics.

Moral 2: change happens for good reasons

Changes in mathematics do not happen suddenly.

They are driven by pragmatic needs, and justified by mathematical success.

Moral 2: change happens for good reasons

Changes in mathematics do not happen suddenly.

They are driven by pragmatic needs, and justified by mathematical success.

In times of change, the mathematical community gradually comes to terms with

- the benefits to using new methods,
- the drawbacks,
- what is gained,
- and what is lost.

It tries to maximize the gains while minimizing the losses.

Moral 3: mathematics incorporates its past

Geometry is still fundamental to mathematics.

We know have other ways to think about what Euclid did, and what Newton did.

Computation is also fundamental to mathematics.

We now have a better understanding of what to compute.

Outline

This talk:

- conceptual understanding
- challenges: computers in mathematics
- historical comparisons:
  - Newton and *Propositio Kepleriana*
  - Dirichlet's theorem on primes in an arithmetic progression
- morals
- bigger morals

Bigger moral 1: the value of history

Studying the history of mathematics helps us understand where we are now.

Bigger moral 1: the value of history

Studying the history of mathematics helps us understand where we are now.

It also helps us *appreciate* where we are now.

Bigger moral 1: the value of history

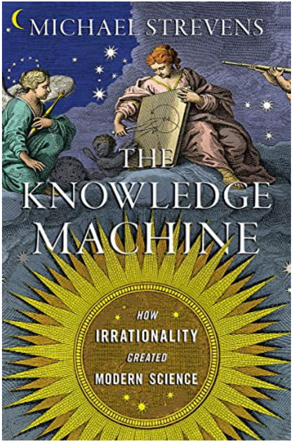
Studying the history of mathematics helps us understand where we are now.

It also helps us *appreciate* where we are now.

If we could find ways to support historical research, our students would benefit, and so would we.

Bigger moral 2: the value of philosophical reflection

According to Strevens: it is a misstep to suggest "that science would flourish if scientists knew and cared more about the rest of existence. Quite the contrary: their obliviousness is the greatest guarantee that they will follow without deviation the empirical path laid out by the iron rule."



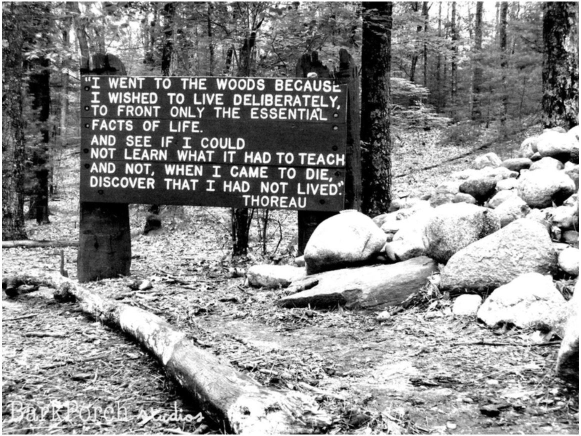
Bigger moral 2: the value of philosophical reflection

We express judgments about what makes for good mathematics every time we:

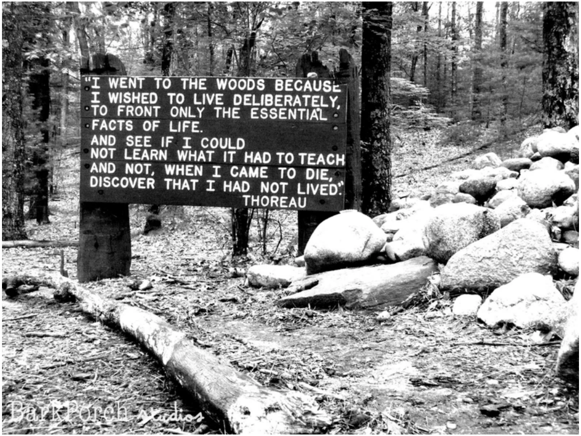
- decide what to work on,
- decide what to teach,
- write a book or article,
- review an article,
- decide who to hire or promote,
- bestow an award.

We should be better at talking about what's important to us.

Bigger moral 2: the value of philosophical reflection



Bigger moral 2: the value of philosophical reflection



We should live our mathematical lives deliberately.

Bigger moral 3: optimism

There are legitimate reasons to be concerned about how computers might change mathematics.

Bigger moral 3: optimism

There are legitimate reasons to be concerned about how computers might change mathematics.

But ultimately, we decide how to use the technology.

Bigger moral 3: optimism

There are legitimate reasons to be concerned about how computers might change mathematics.

But ultimately, we decide how to use the technology.

Mathematics knows what it's doing (even though we don't).

Bigger moral 3: optimism

There are legitimate reasons to be concerned about how computers might change mathematics.

But ultimately, we decide how to use the technology.

Mathematics knows what it's doing (even though we don't).

We should not be afraid to explore.

Bigger moral 3: optimism

These days, there are lots of things we should be worried about.

Bigger moral 3: optimism

These days, there are lots of things we should be worried about.

Mathematics isn't one of them.

Bigger moral 3: optimism

These days, there are lots of things we should be worried about.

Mathematics isn't one of them.

Future generations will understand things better than we do, and they will be thinking about things in ways we can't even imagine.

Bigger moral 3: optimism

These days, there are lots of things we should be worried about.

Mathematics isn't one of them.

Future generations will understand things better than we do, and they will be thinking about things in ways we can't even imagine.

We should be encouraged, because that's exactly why we do mathematics.