Article

Generating conjectures on fundamental constants with the Ramanujan Machine

https://doi.org/10.1038/s41586-021-03229-4

Received: 30 April 2020

Accepted: 13 November 2020

Published online: 3 February 2021

Check for updates

Gal Raayoni[1,4], Shahar Gottlieb[1,4], Yahel Manor[1,2,4], George Pisha[1], Yoav Harris[1], Uri Mendlovic[3], Doron Haviv[1], Yaron Hadad[1] & Ido Kaminer[1]✉

Fundamental mathematical constants such as e and π are ubiquitous in diverse fields of science, from abstract mathematics and geometry to physics, biology and chemistry[1,2]. Nevertheless, for centuries new mathematical formulas relating fundamental constants have been scarce and usually discovered sporadically[3–6]. Such discoveries are often considered an act of mathematical ingenuity or profound intuition by great mathematicians such as Gauss and Ramanujan[7]. Here we propose a systematic approach that leverages algorithms to discover mathematical formulas for fundamental constants and helps to reveal the underlying structure of the constants. We call this approach 'the Ramanujan Machine'. Our algorithms find dozens of well known formulas as well as previously unknown ones, such as continued fraction representations of π, e, Catalan's constant, and values of the Riemann zeta function. Several conjectures found by our algorithms were (in retrospect) simple to prove, whereas others remain as yet unproved. We present two algorithms that proved useful in finding conjectures: a variant of the meet-in-the-middle algorithm and a gradient descent optimization algorithm tailored to the recurrent structure of continued fractions. Both algorithms are based on matching numerical values; consequently, they conjecture formulas without providing proofs or requiring prior knowledge of the underlying mathematical structure, making this methodology complementary to automated theorem proving[8–13]. Our approach is especially attractive when applied to discover formulas for fundamental constants for which no mathematical structure is known, because it reverses the conventional usage of sequential logic in formal proofs. Instead, our work supports a different conceptual framework for research: computer algorithms use numerical data to unveil mathematical structures, thus trying to replace the mathematical intuition of great mathematicians and providing leads to further mathematical research.

Throughout history, simple formulas of fundamental constants symbolized simplicity, aesthetics and mathematical beauty[2]. A couple of well known examples include Euler's identity $e^{i\pi}+1=0$ and the continued fraction representation of the golden ratio:

$$\varphi = 1 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\dots}}}. \tag{1}$$

We use the term regular formulas (RFs) for any mathematical expression that can be encapsulated using a computable expression[14], such as equation (1).

The act of discovering new RFs is often attributed to profound intuition, such as in the case of Gauss' ability to see meaningful patterns in numerical data that led to the famous prime number theorem and new fields of analysis such as elliptic and modular functions. He is even famous for saying: "I have the result, but I do not yet know how to get it"[15], which emphasizes the role of identifying patterns and RFs in data as enabling acts of mathematical discovery.

In a different field but a similar manner, Johannes Rydberg's discovery of his formula of hydrogen spectral lines[16] resulted from his analysis of the spectral emission by chemical elements: $\lambda^{-1} = R_{\mathrm{H}}(n_1^{-2} - n_2^{-2})$, where λ is the emission wavelength, R_{H} is the Rydberg constant, and n_1 and n_2 are the upper and lower quantum energy levels, respectively. This insight, emerging directly from identifying patterns in data, had profound implications on quantum mechanics and modern physics.

Unlike measurements in physics and all other sciences, most mathematical constants can be calculated to an arbitrary precision (number of digits) with an appropriate formula, thus providing an absolute ground truth. In this sense, mathematical constants contain an unlimited amount of data (for example, the digits in an irrational number), which we use as ground truth for finding new RFs. Since the

[1]Technion—Israel Institute of Technology, Haifa, Israel. [2]The Technion Harry and Lou Stern Family Science and Technology Youth Center, Pre-University Education, Haifa, Israel. [3]Google, Tel Aviv, Israel. [4]These authors contributed equally: Gal Raayoni, Shahar Gottlieb, Yahel Manor. ✉e-mail: kaminer@technion.ac.il

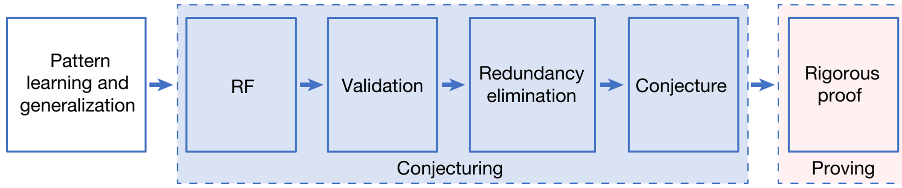


Fig. 1 | Conceptual flow of the wider concept of the Ramanujan Machine. First, using approaches of pattern learning and generalization, we can generate a space of RF conjectures, for example, PCFs. We then apply a search algorithm, validate potential conjectures, and remove redundant results. Finally, validated results form mathematical conjectures that need to be proven analytically, thus closing a complete research endeavour from pattern generation to proof, potentially yielding further mathematical insight.

fundamental constants are universal and ubiquitous in their applications, finding such patterns can reveal new mathematical structures with broad implications, for example, the Rogers–Ramanujan continued fraction (which has implications on modular forms)[17]. Consequently, having systematic methods to derive new RFs could help research in many fields of science.

In this Article, we present a concept of learning mathematical relations of fundamental constants and provide a list of conjectures found using this method. Although the concept can be leveraged for many forms of RFs, we demonstrate its potential with equations involving polynomial continued fractions (PCFs)[18]

$$a_0+\cfrac{b_1}{a_1+\cfrac{b_2}{a_2+\cfrac{b_3}{a_3+\dots}}} \tag{2}$$

where the partial numerators and denominators a_n, b_n are the evaluations (at $x=n$) of polynomials $\alpha(x), \beta(x) \in \mathbb{Z}[x]$, respectively. PCFs have been of interest to mathematicians for centuries and still are, for example, William Broucker's π representation[19] and Zudilin's work on difference equations and Catalan's constant (for example, ref. [5]).

One reason we chose to focus on PCFs is their ability to balance simplicity and broad implications. Their structure is accessible for computer-based exploration using large integer operations, making them a good testing-ground for automated conjecturing. At the same time, PCFs are related to many special functions and generalize all infinite sums. PCFs also allow us to isolate unique aspects of importance to fundamental constants such as testing irrationality and normality[6] using efficient computation methods—see Supplementary Information sections D and G. Moreover, PCFs are abundant in many areas of mathematics[20,21] because they constitute an important special case of a general mathematical object: linear recurrence relations with polynomial coefficients. Recurrences of depth 2 correspond to PCFs and appear in this form in many problems (PCFs with alternating polynomials, as shown below, correspond to recursion depths >2). The solutions to such recurrences are usually very complex and include special functions (for example, hypergeometric functions and the incomplete gamma function). For this reason, finding new PCF identities is valuable for different mathematical objects, especially when incorporated as a part of symbolic calculation programs (such as Maple and Wolfram Mathematica). More on PCFs in the Supplementary Information.

We demonstrate our approach by finding identities between a PCF and a fundamental constant substituted into a rational function. For efficient enumeration and expression aesthetics, we limit ourselves to integer polynomials on both sides of the equality. We propose two search algorithms: The first algorithm uses a meet-in-the-middle (MITM) technique, first executed to a relatively small precision to reduce the search space and eliminate mismatches. We then increase its precision with a higher number of PCF iterations on the remaining matching sequences to validate them as conjectured RFs—the algorithm is therefore called MITM-RF. The second algorithm uses an optimization-based gradient descent (GD) method, which we call Descent&Repel, converging to integer lattice points that define conjectured RFs.

Our MITM-RF algorithm was able to produce several novel conjectures that have short proofs, for example:

$$\frac{4}{3\pi-8}=3-\cfrac{1\times1}{6-\cfrac{2\times3}{9-\cfrac{3\times5}{12-\cfrac{4\times7}{\dots}}}}$$

$$\frac{2}{\pi+2}=0-\cfrac{1\times(3-2\times1)}{3-\cfrac{2\times(3-2\times2)}{6-\cfrac{3\times(3-2\times3)}{9-\cfrac{4\times(3-2\times4)}{\dots}}}}$$

$$\frac{e}{e-2}=4-\cfrac{1}{5-\cfrac{2}{6-\cfrac{3}{7-\cfrac{4}{\dots}}}}$$

$$\frac{1}{e-2}=1+\cfrac{1}{1+\cfrac{-1}{1+\cfrac{2}{1+\cfrac{-1}{1+\cfrac{3}{\dots}}}}} \tag{3}$$

These RFs are auto-generated conjectures for mathematical formulas of fundamental constants, generated by applying the MITM-RF algorithm. These conjectures were proven by contributions from the community following the first appearance of our work on the arXiv preprint server[53] (see Supplementary Information section F). Both results for π converge exponentially, and both results for e converge super-exponentially. Supplementary Information section A presents additional results (Supplementary Tables 1–3) found by our algorithms along with their convergence rates.

Our MITM-RF algorithm also produced novel conjectures that are currently still unproved:

$$\frac{8}{\pi^2}=1-\cfrac{2\times1^4-1^3}{7-\cfrac{2\times2^4-2^3}{19-\cfrac{2\times3^4-3^3}{37-\cfrac{2\times4^4-4^3}{\dots}}}}$$

$$\frac{12}{7\zeta(3)}=1\times2-\cfrac{16\times1^6}{3\times12-\cfrac{16\times2^6}{5\times32-\cfrac{16\times3^6}{7\times62-\cfrac{16\times4^6}{\dots}}}}$$

$$\frac{8}{7\zeta(3)}=1\times1-\cfrac{1^6}{3\times7-\cfrac{2^6}{5\times19-\cfrac{3^6}{7\times37-\cfrac{4^6}{\dots}}}}$$

$$\frac{2}{-1+2G}=3+0\times7-\cfrac{6\times1^3}{3+1\times10-\cfrac{8\times2^3}{3+2\times13-\cfrac{10\times3^3}{\dots}}} \tag{4}$$

To the best of our knowledge, these results are previously unknown conjectures. ζ refers to the Riemann zeta function, and G refers to the Catalan constant. The section 'Efficient computation and irrationality bounds of the Catalan constant' presents implications of our results for the computation of the Catalan constant.

One may wonder whether the conjectures discovered by this work are indeed mathematical identities or merely mathematical coincidences that break down once enough digits are calculated. However, the method employed in this work makes it fairly unlikely for the conjectures to break down. For example, the probability of finding a false positive for an enumeration space of 10^9 with accuracy of 50 digits is smaller than 10^{-40}. Our algorithms tested the conjectures for up to 2,000 digits of accuracy.

Nevertheless, high accuracy will never substitute for a formal proof, as there exist mathematical coincidences of RFs that appear to accurately represent a constant to a high degree of approximation despite being fallacies[22]. We believe and hope that proofs of new computer-generated conjectures on fundamental constants will help to create mathematical knowledge.

In contrast to the method we present, many known RFs for fundamental constants were discovered by conventional proofs, that is, sequential logical steps derived from known properties[23]. In our work, we aim to reverse this process, finding new RFs for the fundamental constants using their numerical data alone, without any prior knowledge about their mathematical structure (Fig. 1). Each RF may enable reverse-engineering of the mathematical structure that produces it. In certain cases, where the proof uses new techniques, it may also provide insight into the field. Our approach could be especially valuable when applied for empirical constants, such as the Feigenbaum constant from chaos theory (Table 1), which are derived numerically from simulations and have no analytic representation.

Given the success of our approach to finding new RFs for fundamental constants, there are additional avenues for more advanced algorithms and future research. Inspired by worldwide collaborative efforts in mathematics such as the Great Internet Mersenne Prime Search (GIMPS; https://www.mersenne.org/), we launched the initiative http://www.RamanujanMachine.com, dedicated to finding new RFs for fundamental constants. The general community can donate computational time to find RFs, propose mathematical proofs for conjectured RFs, or suggest new algorithms for finding them (Supplementary Information section B). Since its inception[53], the Ramanujan Machine initiative has already yielded fruit, and several of the conjectures posed by our algorithms have already been proved (Supplementary Information section F).

Related work

The process of mathematical research is complex, nonlinear and often leverages abstract mathematical intuition, all of which are difficult to express and study thoroughly. Respecting this fact, one may think in an oversimplified manner about mathematical research as being separated into two main steps: conjecturing and proving (as in Fig. 1).

Although both steps have received some attention in the literature, it is the second step that has been studied more extensively in the computer science literature and is known as automated theorem proving (ATP)[24], which focuses on proving existing conjectures. In ATP, algorithms have already proved many theorems such as the Four Colour Theorem[8], the Robbins' problem[10], the Kepler Conjecture on the density of sphere packing[11], a conjectured identity for $\zeta(4)$ (see ref. [25]), and various combinatorial identities[9]. There are also recent machine learning applications for ATP, such as graph neural networks[12,13].

Our work focuses on automating the first step of the process, automated conjecture generation. Early work on automated conjecture generation appeared 60 years ago[26] and included substantial contributions such as the Automated Mathematician and EURISKO[27–29], which envisioned the use of computers for the entire process of scientific discovery. Notable work by Fajtlowicz (called GRAFFITI) has found new conjectures in graph theory and matrix theory[30] by analysing properties such as chromatic index and independence number on a large number of graphs and deducing general rules. Recent work applied machine learning techniques to analyse millions of elliptic curves[31] and explore their characteristics. Automated conjecture generation has also been used as part of a combined approach with ATP[32,33], for example, on the irrationality measure of π (ref. [4]).

A particularly noteworthy algorithm in this context is PSLQ[34], which was employed to study the Riemann zeta function, finding formulas "by a combination of inspired guessing and extensive searching"[35]. PSLQ numerically discovered a new formula for π (which was later proved)[36],

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n}\left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6}\right).$$

With further manipulation and analysis, this formula gave rise to an algorithm that computes strings of binary or base-16 digits of π starting at a given position, without needing to know the preceding digits.

The general approach of using algorithmic and computational tools to explore the mathematical universe and discover conjectures worthy of further examination is known in the mathematical literature as experimental mathematics. A famous example is the work of Wolfram, who has championed experimental-computational methods to investigate the properties of cellular automata[37].

A similar approach for using computer algorithms in mathematical research is now developed and applied in physics. Specifically, supervised and unsupervised machine learning have been applied to discover physical laws from measured data (for example, refs. [38–43]).

Our work differs from all of the above in several respects. We present an end-to-end automated conjecture generation that can validate conjectures to arbitrary precision using numerical data as ground truth and allowing for a fully-automatic process that removes redundancy and false positives without user input. Our conjectures focus on formulas for fundamental constants.

Proposing conjectures is sometimes more important than proving them. For this reason, some of the most original mathematicians and scientists are known for their famous unsolved conjectures rather than for their solutions to other problems, such as Fermat's last theorem,

Table 1 | A sample of fundamental constants that are relevant targets for our method

Field	Name	Decimal expansion
Related to continued fractions	Lévy's constant	$\gamma = 3.275822...$
	Khinchin's constant	$K_0 = 2.685452...$
Chaos theory	First Feigenbaum constant	$\delta = 4.669201...$
	Second Feigenbaum constant	$\alpha = 2.502907...$
	Laplace limit	$r^* = 0.662743...$
Number theory	Twin prime constant	$\Pi_2 = 0.660161...$
	Meissel–Mertens constant	$M = 0.261497...$
	Landau–Ramanujan constant	$\Lambda = 0.764223...$
Combinatorics	Euler–Mascheroni constant	$\gamma = 0.577215...$
	Golomb–Dickman constant	$\lambda = 0.624329...$

There are thousands of additional constants for which enough numerical data exist, and our method is applicable. For all of these, new RF conjectures will point to deep underlying connections. With further improvement in our approach, along with new algorithms provided by the community, we expect that more expressions will be found. Note that some constants in the table, such as the Feigenbaum constants, have no analytical expression whatsoever, and so far can only be computed using numerical simulations. Therefore, having a RF for them will reveal a hidden truth not only about the constant but also about the entire field to which it relates. A wider list of constants is available in ref. [1].

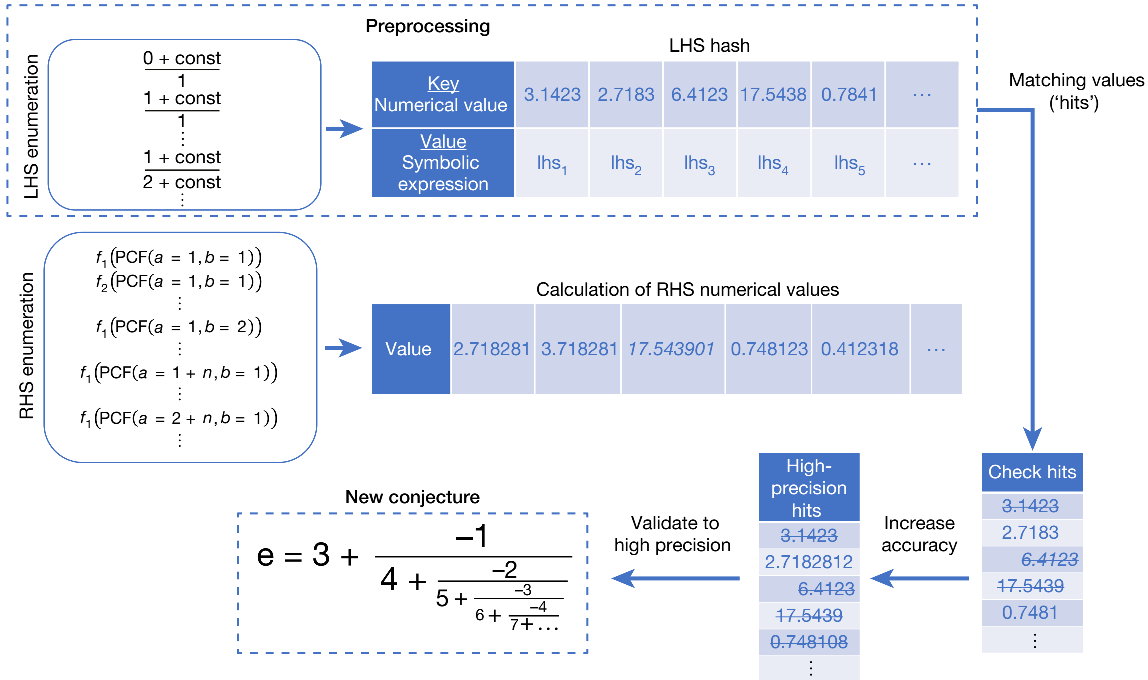


Fig. 2 | The Meet-In-The-Middle Regular Formula algorithm. The figure describes the MITM-RF algorithm that finds PCFs for fundamental constants. First, we enumerate the LHS to a low precision (for example, 10 digits) and store the results in a hash table. Second, we enumerate over the RHS at low precision and search for matches. Finally, the matches are re-evaluated to higher precision and compared again, thus eliminating false positives. The final results are then presented as new conjectures.

Hilbert's problems, Landau's problems, and of course the Riemann Hypothesis[44,45]. Maybe the most famous example is Ramanujan, who posed dozens of conjectures involving fundamental constants and considered them to be revelations from his family's goddess[7]. Our work aims to automate the process of conjecture generation and demonstrate it by providing new conjectures for fundamental constants. By analysing mathematical relationships of fundamental constants that are aesthetic and concise, the Ramanujan Machine can eventually extend the work of great mathematicians such as Gauss, Riemann and Ramanujan.

The MITM-RF algorithm

The first algorithm we present searches for a PCF of a given fundamental constant c (for example, $c = \pi$) of the following form:

$$\frac{\gamma(c)}{\delta(c)} = f_i\,(\mathrm{PCF}(\alpha, \beta)), \qquad (5)$$

for a set of four integer-coefficient polynomials (α, β, γ and δ), and a given set of functions $\{f_i\}$ (for example, $f_1(x) = x, f_2(x) = \frac{1}{x}, \ldots$). PCF($\alpha$, β) means the PCF with the partial numerator $a_n = \alpha(n)$ and denominator $b_n = \beta(n)$ defined in equation (2).

As showcased in Fig. 2, we start by enumerating over the two sides of equation (5) and successively generate integer polynomials for α, β, γ and δ. We calculate the left-hand side (LHS) of each instance up to limited precision and store the results in a hash table. We continue by evaluating the right-hand side (RHS) and attempt to match each result in the hash table, where successful attempts are considered candidate solutions. The RHS is calculated with arbitrary-size integers, directly using the recurrence formula for the numerators p_n and the denominators q_n of the rational approximation of the PCF:

$$\begin{aligned} q_{-1} &= 0, & p_{-1} &= 1, \\ q_0 &= 1, & p_0 &= a_0, \\ q_{n+1} &= a_{n+1}q_n + b_{n+1}q_{n-1}, & p_{n+1} &= a_{n+1}p_n + b_{n+1}p_{n-1}. \end{aligned} \qquad (6)$$

Since the LHS and RHS calculations are performed up to a limited precision, some of the candidate solutions are typically false positives, eliminated by calculating the RHS and LHS to higher precision in the last stage (Fig. 2). See Methods for the algorithm complexity and implementation details (see code at http://www.RamanujanMachine.com).

Our proposed MITM algorithm discovered previously known PCFs and new PCF conjectures for mathematical constants such as $\zeta(3)$ (that is, the Apéry constant) and the Catalan constant, presented in equation (4). (Supplementary Information section A provides details of the constants for which we ran searches, successful or otherwise). After discovering dozens of PCFs, we empirically observed (and later proved, Supplementary Information section D) a relationship between the ratio of the polynomial order of a_n and b_n, and the formula's convergence rate (Extended Data Fig. 1). Supplementary Information section C provides a wider outlook on PCFs.

The Descent&Repel algorithm

We propose a GD optimization method and demonstrate its success in finding RFs. Although proved successful, the MITM-RF method is not trivially scalable. This issue can be targeted by either a more sophisticated variant or by switching to an optimization-based method, as is done by the following algorithm (Fig. 3).

To find integer solutions to equation (5), we write the following constrained optimization problem with the loss function $\mathcal{L}$:

$$\min_{\alpha,\beta,\gamma,\delta} \mathcal{L} = \left\| \frac{\gamma(\pi)}{\delta(\pi)} - \mathrm{PCF}(\alpha, \beta) \right\| \text{ where } \{\alpha, \beta, \gamma, \delta\} \subset \mathbb{Z}[x]. \qquad (7)$$

Solving this optimization problem with GD seems implausible because we are only satisfied with exact $\mathcal{L} = 0$ for integer parameters. Non-zero $\mathcal{L}$ solutions are usually meaningless as mathematical conjectures, as they are only approximations.

Nevertheless, we found a feature of $\mathcal{L}$ that helped us develop a slightly modified GD, which we name Descent&Repel (Fig. 3). Examples of the

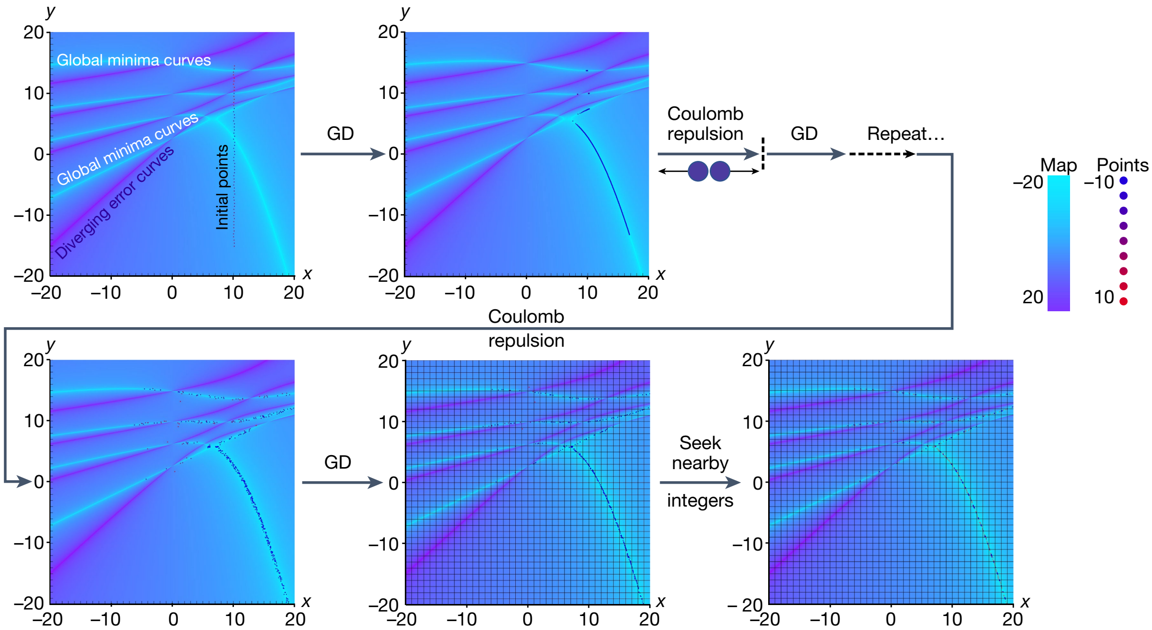


Fig. 3 | The Descent&Repel algorithm. The figure describes the Descent&Repel algorithm that finds RFs for fundamental constants by relying on GD optimization. The x and y axes are parameters defining the polynomials of the continued fraction (in this case $\alpha(n) = n$, $\beta(n) = n^2 + yn + x$; see Supplementary Information section E, Supplementary Table 4, and Supplementary Fig. 1). The key observation that enables this method is that almost all minima have zero loss ($\mathcal{L} = 0$) and appear as $(d-1)$-dimensional manifolds, where d is the number of optimization variables. Starting with our initial conditions (in this example, consisting of 600 points on a vertical line), we perform ordinary GD alternated with 'Coulomb' repulsion between all the points. Finally, we alternate two GD optimizations to reach grid points: towards integer points and the minimum curves. Lastly, we check whether any point satisfies the equation. The colours indicate the loss $\mathcal{L}$ (logarithmic scale): for the background, purple indicates larger loss and white indicates zero loss; for the points, red indicates larger loss and dark blue indicates zero loss.

results appear in Extended Data Table 1. Without the restriction of being integers, the zero $\mathcal{L}$ minima are not 0-dimensional points but rather $(d-1)$-dimensional manifolds with d being the number of optimization variables. Specifically, in the case plotted in Fig. 3, there are $d = 2$ optimization variables, and therefore a 1-dimensional manifold of minima, appearing as bright curves in the maps. This dimensionality of the minima is expected given the definition of the loss function $\mathcal{L}$, which poses only a single constraint. Consequently, the GD process is expected to result in a solution with $\mathcal{L} = 0$. The high dimension of the manifold of minima motivates our approach of adding the repel step to the algorithm since most minima have a neighbourhood that contains additional minima. See Methods for the algorithm initialization and stages.

We ran the algorithm on several different search spaces (mostly with $d = 2$, Supplementary Information section E). The current implementation of the algorithm serves as a proof of concept and as a testing environment for GD variants. As such, it had not yet been executed on large search spaces. The success we had in finding conjectures in these limited runs shows the prospects of using this algorithm on larger search spaces with different parameter choices.

Irrationality bounds of the Catalan constant

Finding RFs for fundamental constants can have important prospects for proving their intrinsic properties. An example is Apéry's proof that $\zeta(3)$ is irrational, which uses a PCF representation[3], and led to similar proofs for other constants[46]. Finding fast-converging RFs could also provide more efficient ways of computing fundamental constants; for example, one of the most efficient historical methods of computing π was based on a formula by Ramanujan[47]. Similarly, the fastest-converging expression for the Catalan constant was a PCF by Zudilin[5] until a relatively recent contribution[48]. The latter was recently used in the y-cruncher algorithm for calculating the record number of digits of the Catalan constant. Efficient formulas for calculating fundamental constants to high precision are used for checking their statistical consistencies and properties, such as normality (the distribution of digits in different integer bases)[49].

As a consequence of the MITM-RF results for the Catalan constant, we found an infinite family of PCFs for the Catalan constant (see Methods). Part of these PCFs have faster convergence rates than the current best formula[48]. Figure 4a summarizes the convergence rates alongside the computational effort per term, conveying the comparative advantage of the new PCFs we found.

Another important implication for such expressions is their potential to help prove the irrationality of the Catalan constant. Each PCF provides a Diophantine approximation sequence that can be characterized by an effective irrationality exponent that quantifies how 'efficiently' it approximates the constant (see Methods).

A paper from 2003[5] found the state-of-the-art exponent of the Catalan constant to be approximately 0.524. A paper from 2016[50] proved this value and presented the better value of about 0.554 as a conjecture. These values are now the best exponents available in the literature. One of the PCFs we found here has an exponent of around 0.567, which surpasses all the previous values in the literature, as shown in Fig. 4b. Finding an explicit sequence for which the exponent is larger than 1 will directly prove irrationality. However, it is not trivial to find such a sequence explicitly (see, for example, ref. [51]), and thus, it is of interest to try to find sequences for which the exponent is as large as possible.

Figure 4b summarizes the convergence of the approximation exponent as a function of the number of computed terms. This comparison includes the best values in the literature and several of our PCFs (detailed in Supplementary Information section G). We write the numerical value of approximation exponent for each of the results in Supplementary Tables 5 and 6. Looking forward, it may well be that the automated exploration of PCF Diophantine approximation sequences

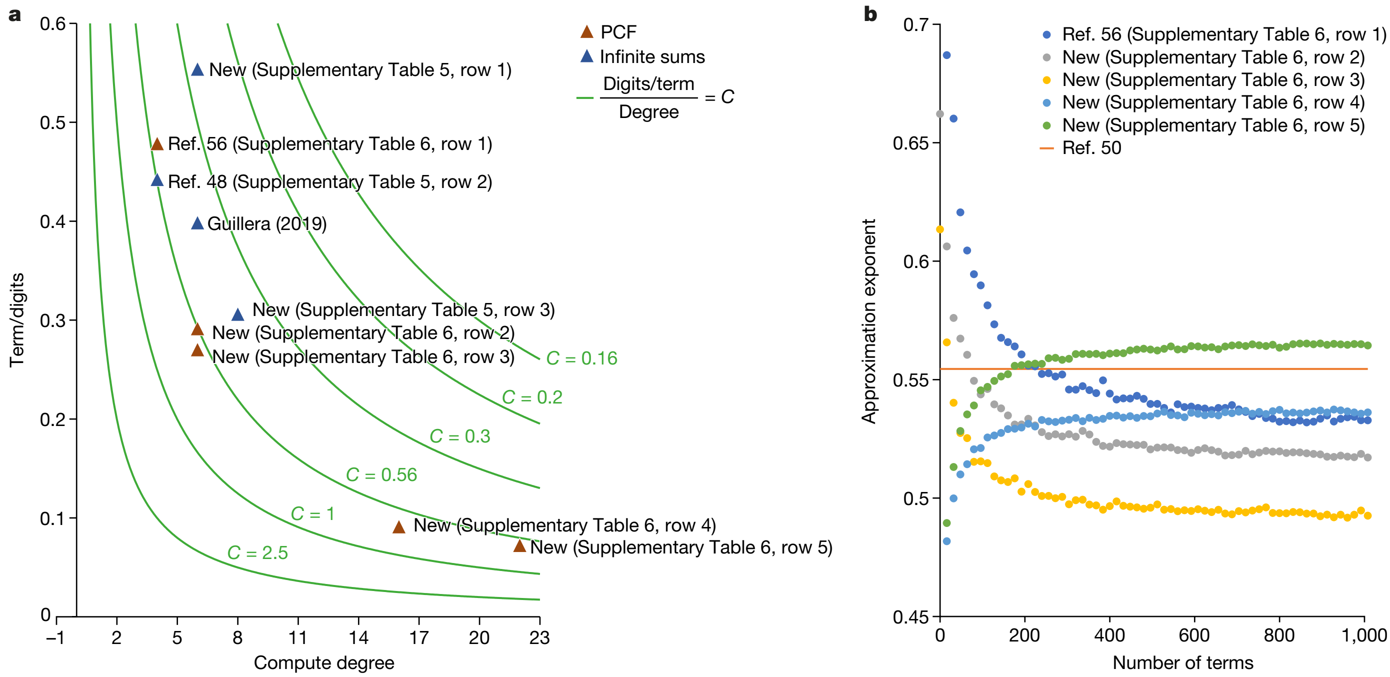


Fig. 4 | Efficient computation of the Catalan constant with new PCFs. Comparison of computational metrics with previous results. **a**, For each formula computing the Catalan constant, the scatter plot shows the asymptotic number of terms required per digit of accuracy, relative to the computational effort (compute degree: defined as the smallest possible polynomial degree that can be used in the calculation, found after transforming the PCF into a matrix of balanced degrees). Green hyperbolas mark the relative efficiencies. Readers should search 'Guillera (2019)' within the page http://www.numberworld.org/y-cruncher/internals/formulas.html to see this result. **b**, The convergence of the effective irrationality exponent (lower bound on the Liouville–Roth irrationality measure, see Methods) as a function of the number of computed terms. The previous best result, first found in ref. [5], is presented in dark blue. A conjecture for a better value, from ref. [50], is presented by a horizontal orange line. The new PCF marked in green surpasses both previous values and yields the new best value for the Catalan constant's approximation exponent. See Supplementary Information section G (specifically, Supplementary Tables 5 and 6).

will eventually provide a higher approximation exponent that can lead to proving the irrationality of the Catalan constant.

The same approach can also be used with other constants. More generally, we expect further explorations of PCFs based on the Ramanujan Machine to lead to additional advances in Diophantine approximations and irrationality measures. For example, it could be intriguing to look for PCFs for values of the Riemann zeta function at odd integers, and specifically $\zeta(5)$ (ref. [52]), because such PCFs may help prove their irrationality and provide more efficient ways of calculating ζ values.

Correspondence with the community

Following the appearance of the initial version of our work on arXiv in 2019[53], numerous people ran our algorithms, some found new conjectures, and a few provided proofs for the new formulas. Over the span of a few months, proofs for all the original manuscript formulas were presented. This led us to expand our search with the MITM-RF algorithm and find more intriguing results such as PCFs for $\zeta(3)$, π^2, and Catalan's constant, most of which are still unproved.

This back-and-forth dynamics between algorithms and mathematicians is the essence of what we believe can be achieved with automatically generated conjectures of fundamental constants. A recent example of this successful correspondence is the work of Zeilberger's group[54], generalizing and proving part of the conjectures that appeared in the earlier arXiv version of our work[53] (Supplementary Information section F.3). An example from their paper is the elegant formula

$$1+k+\cfrac{a\times 1}{2+k+\cfrac{a\times 2}{3+k+\dots}}=\frac{a^{a+k+1}}{(a+k)!\left(e^{a}-\sum_{s=0}^{a+k}\frac{a^{s}}{s!}\right)},$$
$$a>-k,\ a\in\mathbb{N},\ k\in\mathbb{Z},$$

Their method combines the proof as an inherent part of the discovery and thus can be viewed as a successful case study of algorithms that combine automated conjecture generation and ATP.

A wide range of such identities is likely to be useful in future approaches for different math problems, especially in adjacent fields (for example, proving the irrationality of Riemann zeta function values[20]). More generally, automatically discovered formulas can assist further research efforts by enriching the modern 'integral books', which are software and computing environments such as Maple or Wolfram Mathematica. This process provides an elegant example of the symbiosis between computer-generated mathematics and human-generated mathematics.

Although our work focuses on PCFs, we think that it can be systematically extended to other space of candidate RF conjectures. We envision harvesting the scientific literature (for example, over 1.5 million papers on http://arXiv.org) to generalize known formulas and identify new RFs using machine learning algorithms such as clustering methods (see, for example, ref. [55]). The scientific literature provides a strong ground truth for candidate RFs, and this method may discover mathematical conjectures that go far beyond PCFs.

Outlook on the universality of fundamental constants

Our work provides the groundwork for a more comprehensive study into fundamental constants and their underlying mathematical structure. Our proposed algorithms found PCFs for the constants π, e, Catalan's constant and $\zeta(3)$. Table 1 presents a selection of additional fundamental constants of particular interest to our approach. For some of them, such as the Feigenbaum constants, no PCF (or any RF) is known. Potentially the most interesting constants for further research are from fields like number theory (not so ironically, some of them are also named after Ramanujan) and various fields of physics. There, any new RF can point to a hidden connection between fields of science. We believe it would be particularly interesting to extend our work to test RFs that involve several different constants. With such algorithms applied to the thousands of fundamental constants in the literature, we expect many new RFs to be found.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03229-4.

1. Finch, S. *Mathematical Constants* (Cambridge Univ. Press, 2003).
2. Bailey, D., Plouffe, S. M., Borwein, P. & Borwein, J. The quest for pi. *Math. Intell*. **19**, 50–56 (1997).
3. Apéry, R. Irrationalité de ζ(2) et ζ(3). *Asterisque* **61**, 11–13 (1979).
4. Zeilberger, D. & Zudilin, W. The irrationality measure of pi is at most 7.103205334137…. *Moscow J. Combin. Number. Theory* **9**, 407–419 (2019).
5. Zudilin, W. An Apéry-like difference equation for Catalan's constant. *J. Combin*. **10**, R14 (2003).
6. Hardy, G. H. & Wright, E. M. *An Introduction to the Theory of Numbers* 5th edn (Oxford Univ. Press, 1980).
7. Berndt, B. C. *Ramanujan's Notebooks* (Springer Science & Business Media, 2012).
8. Appel, K. I. & Haken, W. *Every Planar Map Is Four Colorable* Vol. 98 (American Mathematical Society, 1989).
9. Wilf, H. S. & Zeilberger, D. Rational functions certify combinatorial identities. *J. Am. Math. Soc*. **3**, 147–158 (1990).
10. McCune, W. Solution of the Robbins problem. *J. Autom. Reason*. **19**, 263–276 (1997).
11. Hales, T. C. A proof of the Kepler conjecture. *Ann. Math*. **162**, 1065–1185 (2005).
12. Lample, G. & Charton, F. Deep learning for symbolic mathematics. In *ICLR Conf*. https://openreview.net/forum?id=S1eZYeHFDS (2020).
13. Cranmer, M. et al. Discovering symbolic models from deep learning with inductive biases. Preprint at https://arxiv.org/abs/2006.11287 (2020).
14. Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc*. **s2-42**, 230–265 (1937).
15. Asimov, I. & Shulman, J. A. *Isaac Asimov's Book of Science and Nature Quotations* (Weidenfeld & Nicolson, 1988).
16. Bohr, N. *Rydberg's Discovery Of The Spectral Laws* (C.W.K. Gleerup, 1954).
17. Shimura, G. Modular forms of half integral weight. In *Modular Functions of One Variable I* 57–74 (Springer, 1973).
18. Cuyt, A. A., Petersen, V., Verdonk, B., Waadeland, H. & Jones, W. B. *Handbook Of Continued Fractions For Special Functions* (Springer Science & Business Media, 2008).
19. Scott, J. F. *The Mathematical Work Of John Wallis (1616–1703)* (Taylor and Francis, 1938).
20. Bowman, D. & Laughlin, J. M. Polynomial continued fractions. *Acta Arith*. **103**, 329–342 (2002).
21. McLaughlin, J. M. & Wyshinski, N. J. Real numbers with polynomial continued fraction expansions. *Acta Arith*. **116**, 63–79 (2005).
22. Press, W. H. *Seemingly Remarkable Mathematical Coincidences Are Easy To Generate* (Univ. Texas, 2009).
23. Euler, L. *Introductio In Analysin Infinitorum* Vol. 2 (MM Bousquet, 1748).
24. Petkovšek, M., Wilf, H. S. & Zeilberger, D. *A = B* (A. K. Peters Ltd., 1996).
25. Bailey, D., Borwein, J. & Girgensohn, R. Experimental evaluation of Euler sums. *Exp. Math*. **3**, 17–30 (1994).
26. Wang, H. Toward mechanical mathematics. *IBM J. Res. Develop*. **4**, 2–22 (1960).
27. Lenat, D. B. & Brown, J. S. Why AM and EURISKO appear to work. *Artif. Intell*. **23**, 269–294 (1984).
28. Lenat, D. B. The nature of heuristics. *Artif. Intell*. **19**, 189–249 (1982).
29. Davis, R. & Lenat, D. B. *Knowledge-Based Systems In Artificial Intelligence* (McGraw-Hill, 1982).
30. Fajtlowicz, S. On conjectures of Graffiti. In *Annals of Discrete Mathematics* Vol. 38, 113–118 (Elsevier, 1988).
31. Alessandretti, L., Baronchelli, A. & He, Y.-H. Machine learning meets number theory: the data science of Birch–Swinnerton-Dyer. Preprint at https://arxiv.org/abs/1911.02008 (2019).
32. Chen, W. Y., Hou, Q. H. & Zeilberger, D. Automated discovery and proof of congruence theorems for partial sums of combinatorial sequences. *J. Diff. Equ. Appl*. **22**, 780–788 (2016).
33. Buchberger, B. et al. *Theorema*: towards computer-aided mathematical theory exploration. *J. Appl. Log*. **4**, 470–504 (2006).
34. Ferguson, H., Bailey, D. & Arno, S. Analysis of PSLQ, an integer relation finding algorithm. *Math. Comput. Am. Math. Soc*. **68**, 351–369 (1999).
35. Bailey, D., Borwein, P. & Plouffe, S. On the rapid computation of various polylogarithmic constants. *Math. Comput. Am. Math. Soc*. **66**, 903–913 (1997).
36. Bailey, D. & Broadhurst, D. J. Parallel integer relation detection: techniques and applications. *Math. Comput*. **70**, 1719–1737 (2000).
37. Wolfram, S. *A New Kind Of Science* Vol. 5 (Wolfram Media, 2002).
38. Schmidt, M. & Lipson, H. Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).
39. He, Y.-H. Deep-learning the landscape. Preprint at https://arxiv.org/abs/1706.02714 (2017).
40. Wu, T. & Tegmark, M. Toward an artificial intelligence physicist for unsupervised learning. *Phys. Rev. E* **100**, 033311 (2019).
41. Greydanus, S., Dzamba, M. & Yosinski, J. Hamiltonian neural networks. In *Advances in Neural Information Processing Systems (NEURIPS2019)* Vol. 32, 15379–15389 (2019).
42. Iten, R., Metger, T., Wilming, H., del Rio, L. & Renner, R. Discovering physical concepts with neural networks. *Phys. Rev. Lett*. **124**, 010508 (2020).
43. Udrescu, S. & Tegmark, M. AI Feynman: a physics-inspired method for symbolic regression. *Sci. Adv*. **6**, eaay2631 (2020).
44. Wiles, A. Modular elliptic curves and Fermat's last theorem. *Ann. Math*. **141**, 443–551 (1995).
45. Smale, S. Mathematical problems for the next century. *Math. Intell*. **20**, 7–15 (1998).
46. Van der Poorten, A. & Apéry, R. A proof that Euler missed…. *Math. Intell*. **1**, 195–203 (1979).
47. Borwein, J., Borwein, P. & Bailey, D. Ramanujan, modular equations, and approximations to pi or how to compute one billion digits of pi. *Am. Math. Mon*. **96**, 201–219 (1989).
48. Pilehrood, K. H. & Pilehrood, T. H. Series acceleration formulas for beta values. *Discret. Math. Theor. Comput. Sci*. **12**, 223–236 (2010).
49. Kim, S. Normality analysis of current world record computations for Catalan's constant and arc length of a lemniscate with $a = 1$. Preprint at https://arxiv.org/abs/1908.08925 (2019).
50. Nesterenko, Y. V. On Catalan's constant. *Proc. Steklov Inst. Math*. **292**, 153–170 (2016).
51. Zudilin, W. Well-poised hypergeometric service for diophantine problems of zeta values. *J. Théor. Nomb. Bordeaux* **15**, 593–626 (2003).
52. Zudilin, W. One of the odd zeta values from ζ(5) to ζ(25) is irrational. By elementary means. *Symmetry Integr. Geom*. **14**, 028 (2018).
53. Raayoni, G. et al. The Ramanujan machine: automatically generated conjectures on fundamental constants. Preprint at https://arxiv.org/abs/1907.00205 (2019).
54. Dougherty-Bliss, R. & Zeilberger, D. Automatic conjecturing and proving of exact values of some infinite families of infinite continued fractions. Preprint at https://arxiv.org/abs/2004.00090 (2020).
55. Tshitoyan, V. et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Complexity of the MITM-RF algorithm

A naïve enumeration is very computationally intensive with time complexity of $O(MN)$, where M and N are the LHS and RHS space size, respectively, and space complexity of $O(1)$. In our algorithm, we store the LHS in the hash table in order to substantially reduce computation time at the expense of space. This makes the algorithm's time complexity $O(M+N)$ and its space complexity $O(M)$. We also implemented another version of the algorithm in which the hash table stores the RHS (the PCF results). In both cases, the hash table can be saved and reused to reduce the duration of future enumerations. The main computational bottleneck is the enumeration and calculation of the RHS terms ($O(N)$ time complexity). By parallelizing the process on C central processing unit (CPU) cores, we are able to speed up the process, and the time complexity drops to $O(N/C)$. Moreover, we decrease the space needed in the memory by using a Bloom filter to store the LHS hash-table keys (instead of keeping the whole hash table in memory). Using a Bloom filter decreases the space by a factor of about 100 during the RHS enumeration.

Our code handles edge cases, like discarding PCFs that provide representations of rational numbers by skipping β polynomials with roots at natural numbers. For the full implementation of our MITM-RF algorithm, see the code on http://www.RamanujanMachine.com.

Generalizations of the MITM-RF algorithm

We also generalized the algorithm to allow for α and β to be integer sequences generated by any countable parametric function. For example, α and β can be interlaced sequences, that is, they may consist of multiple (alternating) integer polynomials. For example, in the case of just two interlaced sequences, odd values of n are equal to one polynomial, and even values of n are equal to a different polynomial.

Seeing how successful our algorithm was despite its relative simplicity, we believe there is still ample room for new results. By leveraging more sophisticated algorithms, other results will follow, thus discovering hidden truths about even more fundamental constants, perhaps with formulas that are more complex than the PCFs used in this work.

Stages of the Descent&Repel algorithm

We chose the optimization problem's variables as the coefficients of the $\alpha, \beta, \gamma, \delta$ polynomials in equation (7). The algorithm is initialized with a large set of points. In the specific examples we present, all initial conditions were set on a line, as shown in Fig. 3.

The algorithm is then constructed of three main stages: GD, 'Repel', and Lattice GD. We iterate between the first two stages and then perform the third stage once to converge to a possible solution.

(1) GD. We perform a standard GD separately for each point x_i, which is a d-dimensional vector. The loss function $\mathcal{L}$ is defined in equation (7), and thus, for each point x_i, we define its next iteration $t+1$ as $x_i^{(t+1)} = x_i^{(t)} - \mu \nabla \mathcal{L}|_{x_i^{(t)}}$, where μ is some small enough step size.

(2) 'Repel'. We update the values of all the points so that they 'push off' one another via a Coulomb-like repulsion proportional to $\frac{1}{\| x_i - x_j \|^2}$. Namely, we define the 'repel' iterations as

$$x_i^{(t+1)} = x_i^{(t)} + \nu \sum_j \frac{x_i^{(t)} - x_j^{(t)}}{\left\| x_i^{(t)} - x_j^{(t)} \right\|^3},$$

with another small step size ν that accounts for the strength of the repulsion. The 'repel' mechanism is used to increase the search space to more effectively cover the space of integer parameters and thus increase the probability of finding a match. We tune the repulsion strength heuristically.

(3) Lattice GD. We enforce the constraint of integer results by alternating the GD optimization between the original loss $\mathcal{L}$ of equation (7) and a different loss function $\mathcal{L}_I$ that scales like the square of the difference between the value x_i and its closest integer (round), $\mathcal{L}_I = \|\text{round}(x_i) - x_i\|$. In cases where this stage converges (being a heuristic algorithm, this is not guaranteed), the method can find points that satisfy $\mathcal{L} = \mathcal{L}_I = 0$ for both losses, meaning an integer solution to our optimization problem.

An infinite family of PCFs for the Catalan constant

The PCF results for the Catalan constant in Supplementary Table 3 can be generalized to an infinite family of PCFs. This generalization revealed an underlying mathematical structure related to the Catalan constant (there is now active research regarding additional algebraic properties of this mathematical structure, to be presented in a separate publication). We produce eight examples of formulas resulting from this generalization and present them in Supplementary Information section G, in Supplementary Tables 5 and 6. Interestingly, part of the PCF results can be expressed as infinite sums (Supplementary Table 5). However, not every PCF can be written as a sum, as is the case in the expressions in Supplementary Table 6, which we found to have a faster convergence rate than the state of the art[48]. Importantly, the complexity of these expressions may help to demonstrate how the approach proposed in this work can handle complexity that may be difficult to address without computer algorithms (we show here specific examples of polynomials of order >20 with coefficients that have >30 digits).

The irrationality measure of a constant and its lower bound

The irrationality measure of x, sometimes called the approximation exponent or the Liouville–Roth constant[6], is defined as the largest $\mu = \mu(x)$ for which there exists a sequence of rational numbers p_n/q_n that satisfy $0 < |x - p_n/q_n| < q^{-\mu}$. For every x, $\mu(x)$ is always either exactly 1 when x is rational or ≥2 when x is irrational.

We can define the effective irrationality exponent of a sequence as the largest (supremum) that satisfies the inequality. Sequences of this kind are called Diophantine approximations[6]. Every PCF we find is such a sequence of rational numbers and it has an effective irrationality exponent μ'. Generally, each explicit Diophantine approximation sequence gives a μ' that can be calculated by $\mu' = \liminf_{n \to +\infty} \frac{\log(|x - p_n/q_n|)}{\log(q_n/\gcd(p_n, q_n))}$, where gcd indicates the greatest common divisor. Each μ' provides a lower bound for the irrationality measure $\mu(x)$ of the value x to which the sequence converges.

However, finding an explicit sequence from which the value of $\mu(x)$ can be extracted is a challenge. This challenge motivated the search for such sequences for important fundamental constants, with the goal of extracting bounds on their value of $\mu(x)$. When a constant is not known to be rational, the sequences all still have $\mu' \le 1$, as in the case of the Catalan constant. Then, finding an explicit sequence for which $\mu' > 1$ will directly prove irrationality. In principle, there must be a sequence with $\mu' = 1$ or $\mu' \ge 2$. However, it is not trivial to find such a sequence explicitly[51,56,57], and thus, it is of interest to try to find sequences p_n/q_n for which μ' is as large as possible.

An infinite family of PCFs with complex variables

Example outcomes of the mathematics–algorithm correspondence in our work are aesthetic generalizations that we found based on results of the Ramanujan Machine algorithms. One example is the following PCF with a complex variable:

$$\forall z \in \mathbb{C}: \quad 1 + \cfrac{1 \times (2 \times z - 1)}{4 + \cfrac{2 \times (2 \times z - 3)}{7 + \cfrac{3 \times (2 \times z - 5)}{10 + \cfrac{4 \times (2 \times z - 7)}{13 + \dots}}}} = \frac{2^{2 \times z + 1}}{\pi \binom{2 \times z}{z}}$$

This PCF was found as a conjecture—by generalizing several automatically generated conjectures (specific integer values for z), generated by the MITM-RF algorithm. Like many other results involving π, it can be proved using generalized hypergeometric functions. The proof is quite straightforward, provided one finds certain identities involving ratios of generalized hypergeometric functions, presented in Supplementary Information section F.2.1 along with other proofs and related information. It remains to be seen whether related methods would be able to prove the unproved conjectures in Supplementary Tables 1–3 of Supplementary Information section A. The above family of PCFs is brought here as an example of how automatically generated conjectures can be generalized to a wider conjecture and later a proof. We believe that this process could be used more widely with future results of the Ramanujan Machine, so that automatically generated conjectures on fundamental constants become a catalyst for mathematical research. For an extended discussion, see Supplementary Information section B.

Data availability

All the results of the Ramanujan Machine project are shared in the paper, with newer updates appearing periodically on the project website.

Code availability

Code is available at: http://www.ramanujanmachine.com/ and the GitHub links therein.

56. Zudilin, W. A third-order Apéry-like recursion for ζ(5). *Mathematical Notes [Mat. Zametki]* **72**, 733–737 [796–800] (2002).
57. Rivoal, T. Rational approximations for values of derivatives of the Gamma function. *Trans. Am. Math. Soc.* **361**, 6115–6149 (2009).

Acknowledgements We thank M. Soljačić, B. Weiss, D. Soudry and D. Carmon for helpful discussions. I.K. is grateful for the support of R. Magid and B. Magid and for the support of the Azrieli Faculty Fellowship. Y.M. acknowledges the support and guidance of the Israeli Alpha Program for Excellent High-School Students.

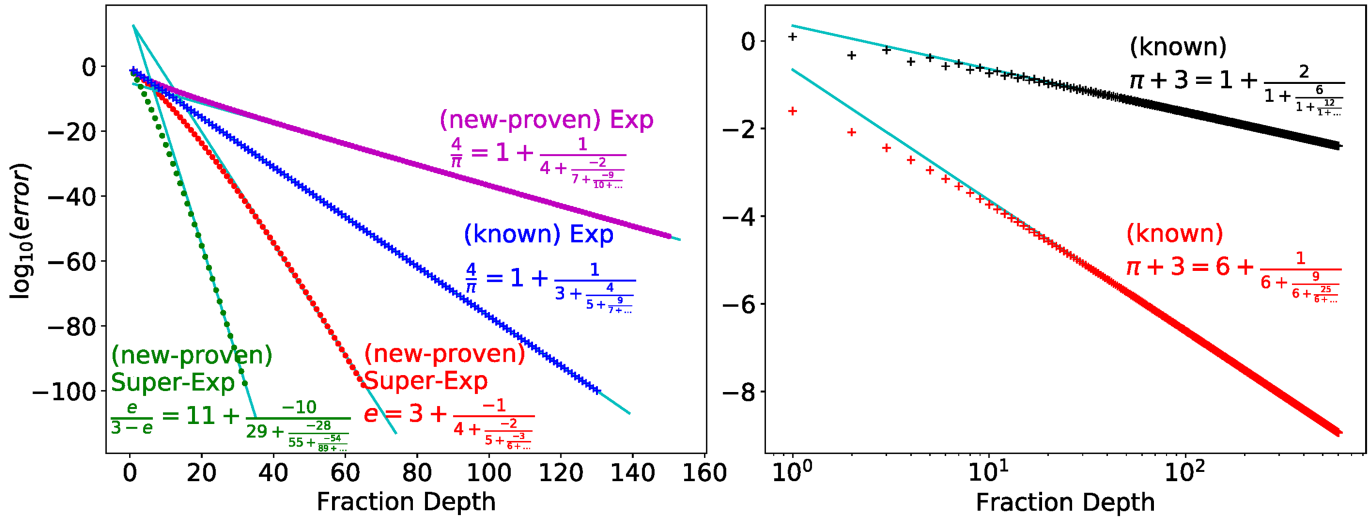
Author contributions G.R., G.P. and I.K. implemented the first proof-of-concept algorithms. G.R. implemented the first generation MITM-RF algorithm. S.G. and Y. Harris implemented the state-of-the-art MITM-RF algorithm. S.G. made the developments that led to the discovery of the ζ(3) and Catalan PCFs. Y.M. implemented the Descent&Repel algorithm. Y.M., S.G., U.M. and I.K found how to convert the Catalan PCFs into expressions with record approximation exponents and fast convergence rates. U.M., Y.M., G.R., S.G., Y. Harris and I.K. proposed parts of the algorithms and developed proofs for some of the conjectures. D.H. and Y. Hadad developed the online community. Y. Hadad, G.P. and I.K. came up with the conceptual flow of the wider concept. I.K. conceived the idea and led the research. All authors provided substantial input to all aspects of the project and to the writing of the paper.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03229-4.

Correspondence and requests for materials should be addressed to I.K.

Peer review information *Nature* thanks Yang-Hui He, Doron Zeilberger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Convergence rates of the PCFs. The plots present the absolute difference between the PCF value and the corresponding fundamental constant (that is, the error) versus the number of terms calculated in the PCF. On the left are PCFs with exponential/super-exponential convergence rates, and on the right are PCFs that converge polynomially. The majority of previously known PCFs for π converge polynomially, whereas all of our newly found results converge exponentially.

Extended Data Table 1 | RFs for π and e found in a proof-of-concept run of the Descent&Repel algorithm

Convergence	Known / New	Formula	Polynomials
Exponential	known	$\frac{4}{\pi} = 1 + \cfrac{1}{3 + \cfrac{4}{5 + \cfrac{9}{7 + \cdots}}}$	$a_n = 1 + 2n,\ b_n = n^2$
Super-Exponential	proven	$e = 3 + \cfrac{-1}{4 + \cfrac{-2}{5 + \cfrac{-3}{6 + \cdots}}}$	$a_n = 3 + n,\ b_n = -n$