The Challenge of Computing Geode Numbers

Tewodros Amdeberhan, Manuel Kauers, and Doron Zeilberger

Abstract: In a fascinating recent American Mathematical Monthly article, Norman Wildberger and Dean Rubine introduced a new kind of combinatorial number, that they aptly named the "Geode numbers". While their definition is simple, they are suprisingly hard to compute, in general. The two-dimensional kind has a nice closed-form that make them easy to compute, but already the three-dimensional kind pose major computational challenges that we do meet, combining experimental mathematics and the holonomic ansatz. Alas, things get really complicated in four and higher dimensions, and we are unable to efficiently compute, for example, the 10000-th term of the 4D Geode sequence. A donatation of 200 US dollas to the OEIS, in honor of the first person to compute this number, is offered.

Maple package This article is based on the Maple package `Geode.txt` available directly from the url

`https://sites.math.rutgers.edu/~zeilberg/tokhniot/Geode.txt` .

There are numerous input and output files in the front of this article:

`https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/Cgeode.html` .

Introduction

In the May 2025 issue of the *American Mathematical Monthly*, Dean Rubine and Norman Wilderberger, in the course of stating their beautiful "explicit" solution to the general algebraic equation (with *symbolic* coefficients c_0, c_1, c_2, c_k)

$$0 = c_0 - c_1\alpha + c_2\alpha^2 + \ldots + \alpha^k \quad ,$$

needed the following numbers defined on lists of non-negative integers $[m_1, m_2, \ldots, m_k]$, that they called the *hyper-Catalan numbers*:

$$C(m_1, \ldots, m_k) := \frac{(2m_1 + 3m_2 + \ldots + (k+1)m_k)!}{(1 + m_1 + 2m_2 + \ldots + km_k)!\, m_1! \cdots m_k!} \quad .$$

Let n and k be positive integers. Define a polynomial, let's call it $P_{n,k}(t_1, \ldots, t_k)$, in the k (continuous) variables, $t_1, \ldots, t_k$, of total degree n as follows:

$$P_{n,k}(t_1, \ldots, t_k) := \sum_{m_1+m_2+\ldots+m_k=n} C(m_1, \ldots, m_k) t_1^{m_1} \cdots t_k^{m_k} \quad .$$

Rubine and Wilderberger proved the following surprising theorem (Theorem 12 of [WR]).

The Geode Theorem (Rubine and Wilderberger): For all positive integers n and k, the polynomial $P_{n,k}(t_1, \ldots, t_k)$ is divisible by $t_1 + \ldots + t_k$.

So it makes sense to define another polynomial, of degree n in $t_1, \ldots, t_k$:

Definition (The Geode polynomial). For positive integers n and k

$$Q_{n,k}(t_1, \ldots, t_k) := \frac{F_{n+1,k}(t_1, \ldots, t_k)}{t_1 + \ldots + t_k}$$

The *geode* numbers are the coefficients of this polynomial.

Definition (The Geode numbers) for every list of non-negative inetegers, $[m_1, \ldots, m_k]$, $G(m_1, \ldots, m_k)$ is the coefficient of $t_1^{m_1} \cdots t_k^{m_k}$ in $Q_{m_1+\ldots+m_k}(t_1, \ldots, t_k)$.

Procedure `G[m]` in our Maple package uses the above definition to compute these numbers.

For example, in order to find $G(4, 7, 8)$ type `G([4,7,8]);`, getting 11258614474275030033600. This works for small arguments, but don't even try to do `G([1000,1000,1000]);` . It will take your computer for ever. Since the computer would first compute the (total) degree-3003 polynomial $P_{3003,3}(t_1, t_2, t_3)$, then divide by $t_1 + t_2 + t_3$, and then exctract the coefficient of $t_1^{1000}\, t_2^{1000}\, t_3^{1000}$. As we will soon see, and that is one of the achievements of this paper, that using experimtental mathematics and *guessing* this 3910-digit integer can be computed in 0.231 seconds. Alas, at this time of writing we can't compute $G(1000, 1000, 1000, 1000)$ and one of us (DZ) is pledging to donate \$200 dollars to `http://oeisf.org/` in honor of the first person to compute it correctly.

The Two-Dimensional Geode Numbers

It turned out that $G(m_1, m_2)$ has a nice **closed-form** expression, not that different-looking than that of $C(m_1, m_2)$. It was conjectured in [WR] and soon proved (along with some other conjectures from that paper) in [AZ] and [R].

Theorem 1 (Conjectured in [WR] and proved in [AZ] and [R])

$$G(m_1, m_2) \;=\; \frac{1}{(2m_1 + 2m_2 + 3)(m_2 + m_3 + 1)} \cdot \frac{(2m_1 + 3m_2 + 3)!}{(m_1 + 2m_2 + 2)!m_1!m_2!} \quad .$$

This is implemented in procedure `G2f(m)` in our Maple package, that inputs $m = [m_1, m_2]$ and outputs $G(m_1, m_2)$. For example, computing $G(5000, 5000)$ using the definition (i.e. executing `G([5000,5000]);`) takes on our laptop 91 seconds, while using the above explcit formula only takes 0.005 seconds!

The Three-Dimensional Geode Numbers

An equivalent way to stating the above theorem is that $G(m_1, m_2)$ satisfy the following *pure*, **first-order** linear recurrences

$$G(m_1, m_2) = \frac{(2m1 + 2 + 3m2)\,(2m1 + 3m2 + 3)\,(2m1 + 1 + 2m2)\,(m1 + m2)}{m1\,(2m1 + 2m2 + 3)\,(m1 + m2 + 1)\,(m1 + 2m2 + 2)} \cdot G(m_1 - 1, m_2) \quad .$$

$$G(m_1,m_2) = \frac{(2m1+1+3m2)\,(2m1+2+3m2)\,(2m1+3m2+3)\,(2m1+1+2m2)\,(m1+m2)}{m2\,(2m1+2m2+3)\,(m1+m2+1)\,(1+m1+2m2)\,(m1+2m2+2)} \cdot G(m_1,m_2-1) \quad .$$

Together with the *initial condition* $G(1,1)=16$, this enales one to compute any $G(m_1,m_2)$ very fast, in fact it is more efficient to use the equivalent recurrence than the "closed-form" expression for $G(m_1,m_2)$ given by Theorem 1, since you divide very large integers by other very large integers.

While there is no 'closed-form' expression for $G(m_1,m_2,m_3)$, the next-best-thing is true! $G(m_1,m_2,m_3)$ satisfies **second-order** linear recurrences with polynomial coefficients. This also enables very fast computation of $G(m_1,m_2,m_3)$ for large arguments, that would be infeasible using the definition.

So we have the next very useful theorem.

Theorem 2: There exist (explicit) rational functions in m_1,m_2,m_3,

• $f_{1,1}(m_1,m_2,m_3)$ and $f_{1,2}(m_1,m_2,m_3)$ both with numerators and denominators of (total) degree 11

• $f_{2,1}(m_1,m_2,m_3)$ and $f_{2,2}(m_1,m_2,m_3)$ both with numerators and denominators of (total) degree 14

• $f_{3,1}(m_1,m_2,m_3)$ and $f_{3,2}(m_1,m_2,m_3)$ both with numerators and denominators of (total) degree 17

such that $G(m_1,m_2,m_3)$ staisty the second-order linear recurrences

$$G(m_1,m_2,m_3) \;=\; f_{1,1}(m_1,m_2,m_3)G(m_1-1,m_2,m_3) \;+\; f_{1,2}(m_1,m_2,m_3)G(m_1-2,m_2,m_3) \quad .$$

$$G(m_1,m_2,m_3) \;=\; f_{2,1}(m_1,m_2,m_3)G(m_1,m_2-1,m_3) \;+\; f_{2,2}(m_1,m_2,m_3)G(m_1,m_2-2,m_3) \quad .$$

$$G(m_1,m_2,m_3) \;=\; f_{3,1}(m_1,m_2,m_3)G(m_1,m_2,m_3-1) \;+\; f_{3,2}(m_1,m_2,m_3)G(m_1,m_2,m_3-2) \quad .$$

These rational functions are too complicated to be presented here, but are easily viewable via our Maple package by typing

`Key3(m1,m2,m3)[1][1]`, `Key3(m1,m2,m3)[1][2]`,`Key3(m1,m2,m3)[2][1]`, `Key3(m1,m2,m3)[2][2]`, `Key3(m1,m2,m3)[3][1]`, `Key3(m1,m2,m3)[3][2]`, respectively.

Note that together with the initial conditions $G(i,j,k)$ with $1 \leq i,j,k \leq 2$, these enable very fast computation of $G(m_1,m_2,m_3)$ If you are only interested in the diagonal sequence $G(n,n,n)$, then we:

Theorem 3: There exist rational functions in n, $f_1(n)$ and $f_2(n)$, whose numerators and denominators are all polynials of degree 35 in n, such that the sequence $\{G(n,n,n)\}_1^\infty$ satisfies the second-order linear recurrence

$$G(n,n,n) \,=\, f_1(n)\, G(n-1,n-1,n-1) \,+\, f_2(n)\, G(n-2,n-2,n-2) \quad .$$

Again $f_1(n)$ and $f_2(n)$ are too complicated to be displayed here, but you can see them by typing

`Rec3AllPC(n)[1][1][1])`, `Rec3AllPC(n)[1][1][2])`, respectively.

Note that together with the initial conditions $G(1,1,1) = 319$ and $G(2,2,2) = 669123$ this recurrence allows to compute, very fast the values of $G(n,n,n)$ for large n that would he hopless using the definition.

This is implemented in procedure G3diag(n). To see the values of $G(10^i, 10^i, 10^i)$ for $1 \leq i \leq 5$, see the output file:

https://sites.math.rutgers.edu/~zeilberg/tokhniot/oGeode2.txt .

How did we find there amazing (and useful!) recurrence for the 3D Geode numbers?

We have a confession to make! the above recurrences are 'only' guesses, and we didn't bother to prove them rigorously. We cranked out sufficiently many terms using the definition of $G(m_1, m_2, m_3)$ and then used *linear algebra* and *undetermined coefficients* to guess the recurrences, just like in *machine learning*. we had the advantage that the data was *exact* without noise. Once the computer found the rerurrenecs using the *training data*, we could verify it for many other cases.

While we did not bother with rigorous proofs, we know that they exist! Let us explain. Using the *holonomic ansatz* [Z][K], we are **guaranteed** that *some* linear recurrences with polynomial coefficients exist for $G(m_1, m_2, m_3)$ and hence also for $G(n,n,n)$. Here is why.

The coefficients in the summmand of $P_{n,k}(t_1, \ldots, t_k))$ are proper-hypergeometric. Hence $P_{n,k}(t_1, \ldots, t_k)$ is continuous-holonomic in the k continuous variables $t_1, \ldots, t_k$. In other words there exist pure *differential operators*, in each of the variables. The same is true for $(t_1 + \ldots + t_k)^{-1}$, and by the closure of holonomic functions it is true for with polynomial coefficients, annihilating it. This is also true for $Q_{n,k}(t_1, \ldots, t_k)$.

Since

$$G(m_1, \ldots, m_k) = (\frac{1}{2\pi i})^k \int_{t_1=1} \cdots \int_{t_k=1} \frac{Q_{n,k}(t_1, \ldots, t_k)}{t_1^{m_1+1} \cdots t_k^{m_k+1}}\, dt_1 \cdots dt_k \quad ,$$

it follows from the general *theory* [Z][K] that $G(m_1, \ldots, m_k)$ is holonomic in the discrete variables $m_1, \ldots, m_k$. In other words, the *exist* pure linear recurrences of **finite** order in each of the discrete variables $m_1, m_2, m_3, \ldots, m_k$.

While they exist powerful tools [Ko] to find such recurrences rigorously, it would take a long time, and it is not worth the trouble, since our recurrences are **absolutely certaion**.

How about 4 Dimensions and beyond?

We were unbale to guess a linear recurrence with polynomial coefficients for $\{G(n,n,n,n)\}$, or pure recurrences for $G(m_1, m_2, m_3, m_4)$. While we know that they **exist** for sure, finding them is a different matter.

One is us (DZ) is offering to donate 100 US dollars to the OEIS for the determination of

$$G(1000, 1000, 1000, 1000) \quad ,$$

and 200 US dollars for

$$G(1000, 1000, 1000, 1000, 100) \quad .$$

Good luck!

References

[AZ] Tewodros Amdeberhan and Doron Zeilbeger, *Proofs of Three Geode Conjectures*, submitted.
https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/geode.html
https://arxiv.org/abs/2506.17862 .

[Ge] Ira Gessel, *Lattice paths and the Geode*, https://arxiv.org/abs/2507.09405 .

[Go] Fern Gossow, *Ordered trees and the Geode*, https://arxiv.org/abs/2507.18097 .

[Ka] Manuel Kauers, *" D-finite functions"*, Springer, 2023.

[Ko] Christoph Koutschan. *Holonomic Functions (a Mathematica package)*. https://risc.jku.at/sw/

[R] Dean Rubine, *Hyper-Catalan and Geode Recurrences and Three Conjectures of Wildberger*,
https://arxiv.org/abs/2507.04552 .

[WR] N. J. Wildberger and D.Rubine, *Hyper-Catalan Series Solution to Polynomial Equations, and the Geode*, Amer. Math. Monthly, **132**, (2025), 383-402.
https://www.tandfonline.com/doi/full/10.1080/00029890.2025.2460966

[Z] Doron Zeilberger, *A Holonomic Systems Approach To Special Functions*, J. Computational and Applied Math **32** (1990), 321-368.
https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/holonomic.html .

Tewodros Amdeberhan, Department of Mathematics, Tulane University, New Orleans, LA 70118, USA; tamdeber@tulane.edu .

Manuel Kaurers, Institute for Algebra, J. Kepler University Linz, Austria; manuel.kauers@jku.at .

Doron Zeilberger, Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854, USA; DoronZeil@gmail.com .