The HOLONOMIC ANSATZ
I. FOUNDATIONS and Applications to Lattice Path Counting

Doron ZEILBERGER [1]

The Ansatz Ansatz

Thomas Kuhn *famously* believed that Science is *paradigm-based*. His approach to science could be thus dubbed the **Paradigm Paradigm**. Doron Zeilberger, *not-yet so famously*, believes that Mathematics, in the future, will be *ansatz-based*, so my approach to mathematical research could be called the **Ansatz Ansatz**.

What is an Ansatz?

According to Eric Weisstein's `mathworld.com` wonderful website, in an entry contributed by Mark D. Carrara[CaW],

"An ansatz is an assumed form for a mathematical statement that is not based on any underlying theory or principle."

In other words, you make a *wild guess* that the desired solution has a certain *form*, featuring some *undetermined* coefficients, "plug" that form into the conditions of the problems, and try to solve for the coefficients. If *in luck*, you find a solution, and then, since *the proof of the pudding is in the eating*, you have an **a posteriori** justification for choosing that ansatz, and more importantly for your short-term goals, you have **solved** the problem! In addition, your present success will give you more confidence that this ansatz might possibly work for similar problems in the future.

A More Relaxed Definition of Ansatz

Let's not insist that our conjectured form *not* be "based on any underlying theory or principle". If it is, all the better, but in that case it might be better to forget this fact.

In the applications to *lattice walk counting* described later in this first installment, WZ theory indeed presents such a justification([A]), but this fact will not be used here, since it will detract from our *general methodology*, that is purely empirical and abhors 'theory'.

On the other hand in the applications to be described in [Z1], we don't have to 'forget' anything. There is no iota of (a priori) theoretical justification (at present) for choosing the holonomic ansatz. The miracle is that it works (in many cases)!

---

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg at math dot rutgers dot edu` , `http://www.math.rutgers.edu/~zeilberg/`.
First version: Jan. 30, 2006. Accompanied by Maple packages GuessHolo2 and GuessHolo3 available from Zeilberger's website. Supported in part by the NSF.

Some Examples of Famous Ansatzes

As far as I know, the first *named* ansatz was the celebrated *Bethe Ansatz* from statistical physics, but physicists have been speaking the prose of ansatz from time immemorial. In fact, every time they fit data into a curve (usually transformed so that it would be a straight line), they implicitly use an ansatz, and are looking for the parameters describing the curve (the slope and intercept, in case of a line). Of course, these poor experimenters have to worry about *experimental error*, a problem that we mathematicians (usually) don't have.

The Constant Ansatz

Physical scientists have their *physical constants*, e.g. c, h, G, and to determine their values they perform lots of careful experiments and hope to get the same answer all the time, within the error bars of experimental accuracy, that they try to make as small as possible. But even in mathematics the seemingly trivial *constant ansatz* has a place.

An old puzzle of Martin Garden

The 7th problem of Chapter 7 of [G] has the following information.

10am: Mr. and Mrs. Smith leave their Connecticut home.

11am: Mrs. Smith asks Mr. Smith: "How far have we gone dear?", Mr. Smith replies: "half of the distance from here to Patricia Murphy's Restaurant".

12noon: They arrive at the restaurant, and eat for an unspecified time and then continue.

5pm: They are 200 miles from where they were at 11am. Mrs. Smith ventures another question: "How much further do we have to go dear?" Mr. Smith replies: "Half as far as the distance from here to Patricia Murphy".

7pm: Arrive at the Pennsylvania home of Mr. Smith's in-laws.

Question: How far did the Smiths travel?

Solution: Obviously the timing is irrelevant, and using the meta-information that the problem is well-posed and has a *unique* solution, it follows that the answer is *independent of the distance traveled between 10am and 11am*, hence the answer belongs to the **constant ansatz**. Plugging-in any value for that distance will give you the right answer, so why not make it zero! When that distance is zero, the locations arrived at 10am, 11am, 12noon are **identical** (in particular Patricia Murphy's restaurant is located at the Smith home). At 5pm they were 200 miles from their home, and still had $(1/2) \cdot 200 = 100$ miles to go, making the total distance $200 + 100 = 300$ miles. **Ans.**:300 miles.

Of course if we weren't so trustworthy of the question, then we would have had to use algebra and calling x the distance traveled between 10am to the 11am, the distance traveled between 11am and 12noon would be $2x$, the distance between the 12am and 5pm locations would be $200-2x$ and the remaining distance would be another $100-x$ miles, making the total distance $x+2x+(200-2x)+(100-x)=300$ miles.

The Polynomial Ansatz

This is often used, implicitly, by takers of IQ tests, when they have to continue a sequence such as $1, 3, 6, 10, 15, 21, \ldots$. A very striking example can be found in [Z3]. Since it is so snappy, let me quote it in its entirety.

For a permutation π, let $inv\ \pi$ be the number of (i, j), such that $1 \leq i < j \leq n$ and $\pi[i] > \pi[j]$, and $maj\ \pi$ be the sum of i, such that $1 \leq i < n$ and $\pi[i] > \pi[i+1]$. Svante Janson asked Don Knuth, who asked me, about the covariance of inv and maj. The answer is $\binom{n}{2}/4$. To prove it, I asked Shalosh to compute the average of the quantity $(inv\ \pi - E(inv))(maj\ \pi - E(maj))$ over all permutations of a given length n, and it gave me, for $n = 1, 2, 3, 4, 5$, the values $0, 1/4, 3/4, 3/2, 5/2$, respectively. Since we know a priori[2] that this is a polynomial of degree ≤ 4, this must be it! $\square$

The Schützenberger Ansatz

A sequence $a(n)$ belongs to that ansatz whenever its generating function is a solution of an algebraic equation whose coefficients are polynomials in x, the paradigmatic example being the Catalan numbers. See my Maple package SCHUTZENBERGER available from my website.

My favorite: The Holonomic Ansatz

Recall that a sequence $a(n)$ (of a single discrete variable n) is called *holonomic* if it satisfies a (homogeneous) *linear recurrence equation with polynomial coefficients*, i.e. there exists a positive integer L (the order), and $L+1$ polynomials $p_0(n), p_1(n), \ldots, p_L(n)$ such that

$$\sum_{i=0}^{L} p_i(n)a(n+i) = 0 \quad .$$

This concept was implicit for a long time, but was first explicated in Richard Stanley's seminal paper [S]. Stanley called such sequences *P-recursive.*

In my seminal paper [Z4], I show that almost everything in sight in enumerative combinatorics, and a lot elsewhere, is holonomic. I also gave a "slow" algorithm that was later made much faster by Frederic Chyzak and others, and in later developments I found much faster algorithms for important special cases. But this is not the point of the present article. I claim that many *specific*

[2] This is the old trick to compute moments of combinatorial 'statistics', described nicely in Graham, Knuth, and Patashnik's 'Concrete Math', section 8.2, by changing the order of summation. It applies equally well to covariance. Rather than actually carrying out the gory details, we observe that this is always a polynomial whose degree is trivial to bound.

results discoverable and provable via WZ theory, can be found **without it**, using the present **purely empirical** approach. Often the present approach is much slower, and sometimes intractable, but in other cases it is faster. But 'slow vs. fast' is not the main point here. I want to illustrate the **ansatz ansatz** using the "Holonomic ansatz" as a **case study**, and here lies the **main significance** of the present endeavor.

Discrete Holonomic Functions of Several Variables

A function on Z^d or N^d, $a(n_1, \ldots, n_d)$ is *holonomic*, if for *each* of its variables, n_j, there exist an integer L_j and polynomials $p_i^{(j)}(n_1, \ldots, n_d)$ such that

$$\sum_{i=0}^{L_j} p_i^{(j)}(n_1, \ldots, n_d) a(n_1, \ldots, n_{j-1}, n_j + i, n_{j+1}, \ldots, n_d) = 0 \quad .$$

Of course, the polynomial coefficients $p_i^{(j)}$ over-determine the sequence, and hence must satisfy lots of compatibility conditions.

The METHOD of Guessing-The-Answer-And-Then-Proving-It-By-Induction and Long Live "Essentially Verification"

G.H. Hardy lamented that all the then known proofs of the Rogers-Ramanujan identities were either complicated or 'essentially verification'. Indeed, most mathematicians, at least until recently, looked down on guessing, with the notable exception of my hero George Polya. Locally, of course, mathematicians are guessing all the time, this is what research is all about. However, many times they *don't even know that they are guessing*, and when they do, they downplay it, and cover their traces after they discover a proof.

But with the mighty computer, guessing can be carried to new heights! And not only for making conjectures, but also for proving them! Because what is a proof? It is just another mathematical object, and as such we can guess it and search for it! But we can't make the haystack too large, that's why we need anstazes. For any good ansatz, we need to teach the computer how to guess results (and, whenever possible, proofs) using that ansatz. In this series of articles we do it for the *holonomic ansatz.*

An Example of the Guess-And-Prove Approach that Even Humans can do

Suppose that we want to solve the following

Problem: Find a formula for $F(m, n)$, the number of ways of walking in the positive quadrant of the 2D square lattice, from the origin $(0, 0)$ to the point (m, n), using unit fundamental steps.

Solution: By the *obvious* combinatorics, $F(m, n)$ is characterized by the *linear partial recurrence equation* (with constant coefficients)

$$F(m, n) = F(m-1, n) + F(m, n-1) \quad , \qquad (Pascal)$$

subject to the **boundary conditions** $F(m,0)=1$ and $F(0,n)=1$.

In a way, this is already an answer! We can use it to write a double do-loop that iteratively prints out a table of the values of F for $0 \leq m,n \leq K$ for any desired K. In particular, to compute $F(100000,100000)$ all we need is compute 10^{10} intermediate values. But we can do better! We can try and **guess** a **nice** formula for $F(m,n)$.

So let's, interactively, ask the computer (in this case paper-and-pencil and perhaps just mental math suffices) to crank out the first few values of $F(m,1)$, say $m=0,1,2,\ldots,10$. Using (*Pascal*) with $n=1$ we get, starting from $m=0$,

$$F(0,1)=1 \quad , \quad F(1,1)=2 \quad , \quad F(2,1)=3 \quad , \quad F(3,1)=4 \quad , \quad F(4,1)=5 \quad , \quad F(5,1)=6 \quad ,$$

and anyone with IQ ≥ 90 would guess that

$$F(m,1)=m+1 \quad .$$

Now we do the same thing for $n=2$ and get

$$F(0,2)=1 \quad , \quad F(1,2)=3 \quad , \quad F(2,2)=6 \quad , \quad F(3,2)=10 \quad , \quad F(4,2)=15 \quad , \quad F(5,2)=21 \quad ,$$

and anyone with IQ ≥ 100 would guess that

$$F(m,2)=\frac{(m+2)(m+1)}{2} \quad .$$

Now we do the same thing for $n=3$ and get

$$F(0,3)=1 \quad , \quad F(1,3)=4 \quad , \quad F(2,3)=10 \quad , \quad F(3,3)=20 \quad , \quad F(4,3)=35 \quad , \quad F(5,3)=56 \quad ,$$

and anyone with IQ ≥ 110 would guess that

$$F(m,3)=\frac{(m+3)(m+2)(m+1)}{6} \quad .$$

Similarly

$$F(m,4)=\frac{(m+4)(m+3)(m+2)(m+1)}{24} \quad .$$

Now anyone with meta-IQ ≥ 100 would guess that

$$F(m,n)=\frac{(m+n)(m+n-1)\cdots(m+1)}{n!} \quad .$$

And Hurray!, we have a gorgeous **conjecture**. But how do we prove it? We **plug-it in** into the *defining relation* of $F(m,n)$, namely (*Pascal*). So let's call the conjectured right-hand-side, $(m+n)!/(m!n!)$, $G(m,n)$:

$$G(m,n):=\frac{(m+n)!}{m!n!} \quad .$$

To *prove* that indeed $F(m,n) = G(m,n)$, we first check the boundary conditions, and get $G(0,n) = (0+n)!/(0!n!) = 1$, and $G(m,0) = (m+0)!/(m!0!) = 1$. Our conjecture would follow by induction (on $m+n$) once we check that

$$G(m,n) = G(m-1,n) + G(m,n-1) \quad . \qquad (Pascal')$$

Dividing this by $G(m,n)$ this is equivalent to

$$1 = \frac{G(m-1,n)}{G(m,n)} + \frac{G(m,n-1)}{G(m,n)} \quad ,$$

which in turn, using the properties of factorials, is equivalent to

$$1 = \frac{m}{m+n} + \frac{n}{m+n} \quad ,$$

which is indeed true, thanks to *high-school algebra*. Yea for us!, we have just rediscovered and reproved

The Levi Ben Gerson Theorem

$$F(m,n) = \frac{(m+n)!}{m!n!} \quad .$$

A triumph to the **GuessAndProve "heuristics"**!

An Example of the Guess-And-Prove Approach that Humans can't do as easily

Suppose that we want to solve the following

Problem: Find a formula for $H(m,n)$, the number of ways that a forward-going King can walk from $(0,0)$ to (m,n), in an infinite chessboard. In other words, the number of walks in the positive quadrant of the 2D square lattice, from the origin $(0,0)$ to the point (m,n), where the set of fundamental steps is $\{(1,0),(0,1),(1,1)\}$.

Solution: By the *obvious* combinatorics, $H(m,n)$ is characterized by the *linear partial recurrence equation* (with constant coefficients)

$$H(m,n) = H(m-1,n) + H(m,n-1) + H(m-1,n-1) \quad , \qquad (ChessKing)$$

subject to the **boundary conditions** $H(m,0) = 1$ and $H(0,n) = 1$.

In a way, this is already an answer! We can use it to write a double do-loop that iteratively prints out a table of the values of H for $0 \le m,n \le K$ for any desired K. In particular to compute $H(100000,100000)$, all we need is compute 10^{10} intermediate values. But we can do better! We can try and **guess** a "formula" for $H(m,n)$, but let's not insist on it being "nice".

So let's interactively ask the computer (in this case paper-and-pencil and perhaps just mental math suffices) to crank out the first few values of $H(m,1)$, say $m = 0, 1, 2, \ldots, 10$. Using $(ChessKing)$ with $n = 1$ we get, starting from $m = 0$,

$$H(0,1) = 1 \quad , \quad H(1,1) = 3 \quad , \quad H(2,1) = 5 \quad , \quad H(3,1) = 7 \quad , \quad H(4,1) = 9 \quad , \quad H(5,1) = 11 \quad ,$$

and anyone with IQ ≥ 95 would guess that

$$H(m,1) = 2m + 1 \quad .$$

Now we do the same thing for $n = 2$ and get

$$H(0,2) = 1 \quad , \quad H(1,2) = 5 \quad , \quad H(2,2) = 13 \quad , \quad H(3,2) = 25 \quad , \quad H(4,2) = 41 \quad , \quad H(5,2) = 61 \quad ,$$

and anyone with IQ ≥ 110 would guess that

$$H(m,2) = 1 + 2m + 2m^2 \quad .$$

Now we do the same thing for $n = 3$ and get

$$H(0,3) = 1 \quad , \quad H(1,3) = 7 \quad , \quad H(2,3) = 25 \quad , \quad H(3,3) = 63 \quad , \quad H(4,3) = 129 \quad , \quad H(5,3) = 231 \quad ,$$

and anyone with IQ ≥ 125 would guess that

$$H(m,3) = 1 + \frac{8}{3}m + 2m^2 + \frac{4}{3}m^3 \quad .$$

Similarly

$$H(m,4) = 1 + \frac{8}{3}m + \frac{10}{3}m^2 + \frac{4}{3}m^3 + \frac{2}{3}m^4 \quad ,$$

and one can continue, and use the *polynomial ansatz* to conjecture, for each specific n_0, an explicit formula

$$H(m, n_0) = p_{n_0}(m) \quad ,$$

for some explicit polynomial $p_{n_0}(m)$ of degree n_0 in m.

But **no human**, not even Andrew Wiles or Arthur Benjamin, has sufficient meta-IQ to conjecture a general 'nice' formula for $H(m,n)$ for **arbitrary** n.

But perhaps we have to be more liberal and relax the meaning of *nice*. In the previous example, we conjectured, and then proved, a **closed-form** formula for $F(m,n)$

$$F(m,n) = \frac{(m+n)!}{m!n!} \quad , \qquad (EXPLICIT)$$

but this is equivalent to the following facts

$$(m+1)F(m+1,n) - (m+n+1)F(m,n) = 0 \quad , \qquad (RecurrenceM)$$

and

$$(n+1)F(m,n+1)-(m+n+1)F(m,n)=0 \quad , \qquad (RecurrenceN)$$

together with the *initial condition* $F(0,0)=1$, in other words $(EXPLICIT)$ is equivalent to $F(m,n)$ being a solution of the **system** of (homogeneous) *linear-recurrence equations with polynomial coefficients*, $(RecurrenceM)$, $(RecurrenceN)$ of the **FIRST-ORDER**. Can we do something similar for $H(m,n)$? Let me remind you that $H(m,n)$ stands for the number of ways of walking from $(0,0)$ to (m,n) where the allowed steps are one unit up $((0,1))$, one unit to the right $((1,0))$, and one unit-diagonal step up $((1,1))$. Of course the **partial** recurrence equation $(ChessKing)$ is already some kind of answer, since it enables us, using a computer, to compile a table of $H(m,n)$ for $0 \leq m,n \leq L_0$ in $3L_0^2$ operations and roughly $O(L_0^2)$ memory (with a slightly less straightforward implementation, $O(L_0)$ memory suffices), but does $H(m,n)$ satisfy **pure** recurrences, like $(RecurrenceM)$ and $(RecurrenceN)$ for $F(m,n)$?

Now, if you know WZ theory you would know that the answer to the above question is **yes**. Indeed, since

$$H(m,n)=CT\frac{1}{(1-x-y-xy)x^m y^n} \quad ,$$

where CT stands for the *"constant term of"*, and since the *constant-termand* is *holonomic* in x,y,m,n, WZ theory immediately guarantees that $H(m,n)$ is holonomic in (m,n), which is another way of saying that $H(m,n)$ satisfies *pure* recurrences both in the m variable and in n variable, just like $(RecurrenceM)$ and $(RecurrenceN)$, but **not necessarily first-order!**. Now WZ theory not only guarantees the *existence* of such recurrences, but also has algorithms to find them, but that's not the point right now. Let's forget about WZ theory, and pretend that we are, completely empirically, trying to find an 'almost nice' way of representing $H(m,n)$, but giving up on it being *as nice* as the explicit expression $((m+n)!/(m!n!))$ we found for $F(m,n)$. In other words, we will relax the condition that the pure recurrence equations be first-order.

So let's be optimistic, and try the **ANSATZ** of second-order linear recurrence in m with polynomial coefficients that are of degree 1 in (m,n), using *undetermined coefficients*:

$$(c_0+c_1m+c_2n)H(m+2,n)+(b_0+b_1m+b_2n)H(m+1,n)+(a_0+a_1m+a_2n)H(m,n)=0 \quad , \qquad (Hopeful)$$

where $a_0,a_1,a_2,b_0,b_1,b_2,c_0,c_1,c_2$ are nine numbers yet to be determined. We want it to be true for *all* $m,n \geq 0$, so by plugging-in all possible pairs, we get a system of $\infty^2=\infty$ linear equations with 9 unknowns. In practice, of course, it is enough to pick 10 random values, and to be on the safe side, let's pick 20 such pairs (m_0,n_0), plug them into (Hopeful), and solve the system of 20 linear equations for the 9 unknown numbers $a_0,a_1,a_2,b_0,b_1,b_2,c_0,c_1,c_2$. Of course, we get the $H(m_0,n_0)$ for specific (m_0,n_0) by using the partial recurrence $(ChessKing)$.

If you have a random system of linear homogeneous equations with more equations than unknowns, it is extremely unlikely to get a *non-zero* solution. If you do, it means that you discovered a conjecture!. In this case, it turns out that $a_0=-1, a_1=-1, a_2=0, b_0=-1, b_1=0, b_2=-2, c_0=2, c_1=1, c_2=0$ is the unique solution (up to a constant multiple, of course) yielding the *pure* recurrence

$$(m+2)H(m+2,n) - (2n+1)H(m+1,n) - (m+1)H(m,n) = 0 \quad . \qquad (RecurrenceMking)$$

Similarly, (and in our case, by symmetry),

$$(n+2)H(m,n+2) - (2m+1)H(m,n+1) - (n+1)H(m,n) = 0 \quad , \qquad (RecurrenceNking) \quad ,$$

which together with the *initial condition* $H(0,0) = 1$, and the convention that $H(m,-1) \equiv 0$ and $H(-1,n) \equiv 0$ uniquely define $H(m,n)$.

Why is the Pair of Pure Recurrences Better than the Defining Partial-Recurrence?

While the pair of pure recurrences $(RecurrenceMking)$ and $(RecurrenceNking)$ satisfied by $H(m,n)$ is not quite as 'nice' as the pair $(RecurrenceM)$ and $(RecurrenceN)$ satisfied by $F(m,n)$, computationally it is almost as good. For example, to compute $H(1000,2000)$ you no longer need $3 \cdot 1000 \cdot 2000$ operations and to remember $1000 \cdot 2000$ facts, but only $2(1000+2000)$ operations and to remember that many facts. In fact, of course, in the former case we only need to remember up to $2(1000+2000)$ values *at a time*, but in the latter case we only need **2** previous values at a time.

The diagonal recurrence

If you are only interested in the *diagonal* $H(n,n)$, then you can try to **conjecture** a recurrence of the form

$$p_0(n)H(n,n) + p_1(n)H(n+1,n+1) + p_2(n)H(n+2,n+2) = 0 \quad , \qquad (MainDiagonalRecurrenceAnsatz)$$

for polynomials in n, $p_0(n), p_1(n), p_2(n)$. Trying out generic polynomials of degree 1 for $p_0(n), p_1(n), p_2(n)$, and plugging into $(MainDiagonalRecurrenceAnsatz)$, $n = 0, 1, \ldots 10$, (using the numerical values $H(0,0), \ldots, H(12,12)$ obtained from $(ChessKing)$), yields 11 linear equations for the $6 = 3 \cdot (1+1)$ unknown coefficients, that result in the following **conjecture**

$$nH(n,n) - 3(2n+1)H(n+1,n+1) + (n+1)H(n+2,n+2) = 0 \quad . \; (MainDiagonalRecurrence)$$

What about a general diagonal? In this case the recurrence is still second-order but the polynomials $q_0(m,n), q_1(m,n), q_2(m,n)$, that feature as coefficients of the recurrence

$$q_0(m,n)H(m,n) + q_1(m,n)H(m+1,n+1) + q_2(m,n)H(m+2,n+2) = 0 \quad , \qquad (GeneralDiagonalRecurrenceAnsatz)$$

turn out to be of degree 3, and doing the analogous linear algebra yields the following diagonal recurrence

$$(n+1)(m+1)(m+n+4)H(m,n) - (m+n+3)(2n^2+2nm+2m^2+9n+9m+12)H(m+1,n+1)$$

$$+(m+2)(n+2)(m+n+2)H(m+2,n+2) = 0 \quad . \qquad (GeneralDiagonalRecurrence)$$

Asymptotics for H(n,n)

Once you have a linear recurrence with polynomial coefficients, the Birkhoff-Trijinski method (see [WimZ]) tells you how to extract the asymptotics to any desired order. Applying this method to ($MainDiagonalRecurrence$), yields

$$H(n,n) \asymp C(3+2\sqrt{2})^n n^{-1/2}(1+\frac{3\sqrt{2}-8}{32n}+O(1/n^2)) \quad ,$$

where C is a constant that is approximately 0.571.

How to prove the conjectured recurrences?

If you are skeptical, then you can check these conjectured recurrences not just for the above 20 pairs (m_0, n_0), used to generate the conjecture, but for the 10000 pairs (m_0, n_0) for $0 \leq m_0, n_0 \leq 100$, and this would be *overwhelming* empirical evidence for their veracity. But it would still not be a *proof.* We want to know that ($RecurrenceMking$) is true for *all* (m,n) not just for the first billion values. Recall that in order to prove the 'nice' formula $(m+n)!/(m!n!)$ for $F(m,n)$ we plugged $G(m,n) := (m+n)!/(m!n!)$ into ($Pascal$) and verified that $G(m,n)$ satisfies the analogous recurrence $G(m,n) = G(m-1,n)+G(m,n-1)$. In the more general case, of higher-order recurrences, this is no longer possible, since the conjectured 'nice' expression is not closed-form, but is defined implicitly by ($RecurrenceMking$) and ($RecurrenceNking$).

What we need is rephrase all our recurrences in *operator notation.* Introducing the shift-operators

$$Mf(m,n) := f(m+1,n) \quad , \quad Nf(m,n) := f(m,n+1) \quad ,$$

the *defining* partial-recurrence for $H(m,n)$, Eq. ($ChessKing$) can be written

$$(MN - M - N - 1)H(m,n) \equiv 0 \quad , \quad (m \geq 0, n \geq 0) \quad , \qquad (ChessKingOpe)$$

[plus we need the initial conditions $H(m,0) = 1$, $(m \geq 0)$, $H(0,n) = 1$, $(n \geq 0)$]. This is our **given.** We **have to prove**

$$[(m+2)M^2 - (2n+1)M - (m+1)]H(m,n) \equiv 0 \quad (m \geq 0, n \geq 0) \quad .$$

Let's give the operators names:

$$\mathcal{P} := MN - M - N - 1 \quad ,$$

$$\mathcal{Q} := (m+2)M^2 - (2n+1)M - (m+1) \quad .$$

We have to prove, under the initial conditions, that

$$\mathcal{P}H \equiv 0 \quad \Rightarrow \quad \mathcal{Q}H \equiv 0 \quad .$$

Let's compute the *commutator*

$$\mathcal{Q}_1 := \mathcal{P}\mathcal{Q} - \mathcal{Q}\mathcal{P} \quad .$$

If it is indeed true that $\mathcal{Q}H \equiv 0$ then of course, since $\mathcal{P}H \equiv 0$, we would have $\mathcal{Q}_1 H \equiv 0$. But also vice-versa! If it is true that $\mathcal{Q}_1 H \equiv 0$, then it would follow that $\mathcal{P}[\mathcal{Q}H] \equiv 0$. So $\mathcal{Q}H$ is annihilated by $\mathcal{P}$. Now being annihilated by $\mathcal{P}$ is not enough to make a discrete function identically zero, after all, our original function $H(m,n)$ is annihilated by $\mathcal{P}$, and is not zero. But now we have that $\mathcal{Q}H(0,0)$ is 0, so we have different initial conditions. In fact we can directly see that $\mathcal{Q}H(m,n)$ vanishes at $m=0$ and $n=0$. Since the initial conditions vanish, *now* being annihilated by $\mathcal{P}$ entails being identically 0. Hence $\mathcal{Q}H(m,n)$ is identically 0.

It remains to prove that $\mathcal{Q}_1$ annihilates $H(m,n)$. But Maple (and in this simple case you can do it by hand) shows that

$$\mathcal{Q}_1 = (M-1)M(MN-M-N-1) = (M-1)M\mathcal{P} \quad ,$$

so the operator $\mathcal{Q}_1$ turned out to be a multiple of the operator $\mathcal{P}$, and since, by assumption, $\mathcal{P}H \equiv 0$, it follows that

$$\mathcal{Q}_1 H(m,n) = (M-1)M\mathcal{P}H(m,n) \equiv 0 \quad .$$

QED!

How to prove Recurrence (*MainDiagonalRecurrence*)**?**

You can't prove it directly! We don't have enough elbow room to apply the inductive approach of the previous section. As is often the case in mathematics, the more general a statement is, the easier it is to prove.

What we can do is prove (*GeneralDiagonalRecurrence*). Let's use the letter $\mathcal{Q}$ once again to denote this operator:

$$\mathcal{Q} := (n+1)(m+1)(m+n+4)-(m+n+3)(2n^2+2nm+2m^2+9n+9m+12)MN+(m+2)(n+2)(m+n+2)M^2N^2 \quad .$$

Now we need a *tower of operators*: $\mathcal{Q}_1 := [\mathcal{P}, \mathcal{Q}]$, $\mathcal{Q}_2 := [\mathcal{P}, \mathcal{Q}_1]$, $\mathcal{Q}_3 := [\mathcal{P}, \mathcal{Q}_2]$. Since $\mathcal{P}$ is a partial recurrence operator with *constant* coefficients, its commutator with any linear recurrence operator with polynomial coefficients whose coefficients are polynomials of degree d, say, is another such operator but with the coefficients being polynomials of degree $d-1$. Since the degree of $\mathcal{Q}$ (in (m,n)) is 3, the degree of $\mathcal{Q}_1$ is 2, the degree of $\mathcal{Q}_2$ is 1, and the degree of $\mathcal{Q}_3$ is 0, i.e. $\mathcal{Q}_3$ turns out to be *constant-coefficients* and in fact is a left-multiple of $\mathcal{P}$. Hence $\mathcal{Q}_3 H = 0$ which entails that $\mathcal{Q}_2 H = 0$ which entails that $\mathcal{Q}_1 H = 0$, which finally entails that $\mathcal{Q}H = 0$. As before at each stage we have to check that the initial conditions vanish, but this is routine and capable of automation.

The General Case

All that we said above for the King's walks applies to the function counting the number of lattice walks using *any* finite set of allowed steps, in *any* number of dimensions (i.e. not just the plane) and more generally for the discrete function describing the power-series coefficients of *any* rational function. It also applies to "ballot-style" walks (i.e. walks that are restricted to stay in $m_1 \geq$

$m_2 \geq \ldots \geq m_k \geq 0$ (or even $m_1 \geq m_2 \geq \ldots \geq m_k \geq -c$ for any positive integer c). It can also do probabilities (if each step is assigned a certain probability), and find recurrences along diagonal directions.

As above we first use the 'obvious', or if you wish, *defining* partial-recurrence equation with **constant** coefficients, together with the obvious inital conditions, to crank-out enough values of the discrete function under investigation. Then we use the **holonomic ansatz** to guess **pure** recurrences in each discrete variable. Of course if our discrete function is symmetric, we need consider only one direction.

Also the outline of the proof of the validity of the conjectured recurrences can be done in general, as well as deducing the asymptotics of the diagonal terms.

Clever Guessing

The main workhorse of our method is guessing a recurrence like

$$\sum_{i=0}^{L_j} p_i^{(j)}(n_1, \ldots, n_d) a(n_1, \ldots, n_{j-1}, n_j + i, n_{j+1}, \ldots, n_d) = 0 \quad ,$$

where $p_i^{(j)}(n_1, \ldots, n_d)$ is a polynomial in its variables.

A naive approach would be to guess an upper bound for the order of the recurrence L_j and for the degrees of the $p_i^{(j)}$, write everything generically and solve for the undetermined coefficients. But for higher dimensions, sooner than later, the number of equations would explode. Also it results in a lot of wasted effort, because we start with the order being 1, and the degree being 0 and work our way up.

A more efficient approach would be to first investigate the probable order L_j and the degrees of the $p_i^{(j)}$ by freezing $n_1, n_2, \ldots, n_{j-1}, n_{j+1}, \ldots n_d$, and only consider the sequence of single variable n_j. We then get an idea for the expected order L_j and the expected degree in n_j. To be on the safe side, we can freeze them at several places and look for a consensus. Then we can unfreeze each of the other variables, *one at a time* and get an idea about the degrees in the other variables. Also it turned out to be more efficient (at least for us working in Maple), to find the operators with such specializations in many cases, and then "reconstruct" the real thing using linear algebra once again.

The Maple package GuessHolo2

All of the above, for the two-variable case, is implemented in the Maple package `GuessHolo2` accompanying this article. It is available from

`http://www.math.rutgers.edu/~zeilberg/tokhniot/GuessHolo2` .

After you downloaded it to your directory, saving it as `GuessHolo2`, go into Maple, by typing `maple [Enter]`, and then type `read GuessHolo2:` and follow the instructions given there. In particular to get the list of functions type `ezra();`, and to get help with a specific function type `ezra(FunctionName);`. For example, to get help with procedure `GH2`, type `ezra(GH2);`.

The Maple package GuessHolo3

The analogous package for three variables, in particular to counting lattice "flights", i.e. walks in the positive three-dimensional cubic lattice, download

`http://www.math.rutgers.edu/~zeilberg/tokhniot/GuessHolo3` .

It also contains most of the functions of `GuessHolo2`.

What about GuessHolo4, GuessHolo5, ...?

In principle I should have written *just* one package, that can handle *any* dimension, but at this time of writing, at least with my computer, the general-purpose approach is too slow, and one needs lots of *dirty tricks*, specific for small dimensions to make it run in *real* time. Of course, with more effort I could have combined them and written a general program, but life is too short to get stuck on one project. Of course, *you* are welcome to improve and generalize!

Sample input and output

The webpage of this article

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzI.html`,

contains numerous sample input and output files illustrating the power of the packages `GuessHolo2` and `GuessHolo3`.

REFERENCES

[A] Moa Apagodu, *Applications of the Wilf-Zeilberger method to Lattice Path Counting*, in preparation.

[CaW] Eric W. Weisstein et al. "Ansatz." From MathWorld–A Wolfram Web Resource.
`http://mathworld.wolfram.com/Ansatz.html` [contributed by Mark. D. Carrara]

[G] Martin Gardner, *"The Unexpected Hanging and Other Mathematical Diversions"*, Simon and Schuster, 1969.

[S] Richard Stanley *Differentiably finite power series*, European J. Combin. **1** (1980), 175-188.

[WimZ] Jet Wimp and Doron Zeilberger, *Resurrecting the asymptotics of linear recurrences*, J. Math. Anal. Appl. **111** (1985), 162-177.

[Z1] Doron Zeilberger, *The Holonomic Ansatz II: Automatic DISCOVERY(!) and PROOF(!!) of Holonomic Determinant Evaluations*, in preparation.

[Z2] Doron Zeilberger, *The Holonomic Ansatz III: Automatic Proofs of Positivity Theorems for Power-Series Coefficients of Rational Functions of Several Variables*, in preparation.

[Z3] Doron Zeilberger, *The Joy of Brute Force: The Covariance of The number of Inversions and the Major Index*, Personal Journal of Ekhad and and Zeilbeger,
`http://www.math.rutgers.edu/~zeilberg/pj.html` .

[Z4] Doron Zeilberger, *A Holonomic systems approach to special functions identities*, J. of Computational and Applied Math. **32**, 321-368 (1990).

[Z5] Doron, Zeilberger, *Theorems for a price: Tomorrow's semi-rigorous mathematical culture*, Notices of the Amer. Math. Soc. **40 # 8**, 978-981 (Oct. 1993). Reprinted: Math. Intell. **16**, no. 4, 11-14 (Fall 1994).