In How Many Ways Can You Reassemble Several Russian Dolls?

Doron ZEILBERGER[1]

Dedicated to my favorite identical twins: Thotsaporn and Thotsaphon THANATIPANONDA

Preface

In the Fall of 1981, I had the pleasure and honor of living next-door to the eminent combinatorialist Joel Spencer, who was then a visiting professor at the Weizmann Institute of Science in Israel. Joel was always glad to talk to me, provided that I spoke in Hebrew. Joel is also a great punner, and was very proud when he made his first pun in Hebrew, during a volleyball game with other math faculty and students, when he said "kadur sheli" that could mean "my ball" but also "Lee [Segal]'s ball". One day I got as a present a set of "Russian Dolls" (a.k.a. as *Matryoshkas* or *Babushkas*), which is a nested set of dolls. Trying to test Joel's knowledge of enumeration (after all he is an expert in "Hungarian", rather than enumerative, combinatorics), I asked him:

In how many ways can one reassemble an n*-nested Russian Doll?*,

and he immediately replied: *This is a Stirling question, it rings a Bell.*

Several Russian Dolls

Almost thirty years later, my brilliant student, Thotsaporn "Aek" Thanatipanonda asked me what happens if you have *several*, say r, identical Russian Dolls. Thanks to *Sloane*, *MathSciNet*, and *Google scholar*, we quickly found out that the case $r=2$ goes back to Comtet[C] and is featured in Sloane's **A020554**. See also [Ba] and [R], and for an insightful *species* treatment see [L] and [P]. The case $r=3$ was nicely handled by Ed Bender[Be], while the general case was given its *coup de grâce* by John Devitt and David Jackson[DJ], who used a very ingenious generatingfunctionology approach.

As Herb Wilf[W1] famously said, an "answer" to an enumeration question is an *efficient algorithm* to generate many terms in the enumerating sequence. Explicit formulas are just *one* such way, and often not the most efficient one! In this article, I will describe a "calculus" approach, using *differential operators*, that has the following advantages.

1. It is somewhat faster than using the Devitt-Jackson[DJ] complicated exponential generating function.

---

[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` . Sept. 16, 2009. Revised: Dec. 16, 2009. Accompanied by Maple package `BABUSHKAS` downloadable from `http://www.math.rutgers.edu/~zeilberg/tokhniot/BABUSHKAS` . Exclusively published in the Personal Journal of Ekhad and Zeilberger `http://www.math.rutgers.edu/~zeilberg/pj.html` and `http://arxiv.org` . Supported in part by the USA National Science Foundation.

2. It is so flexible that it can handle *non-identical* Russian Dolls. Given a_1 single-births, a_2 pairs of (identical) twins, a_3 sets of (identical) triplets, a_4 sets of (identical) quadruplets, $\dots$, a_k sets of (identical) k-tuplets, $\dots$, (and you can't tell identical twins etc. apart form each other), in how many ways can you partition them into (not necessarily distinct) *sets*? This is useful if the school principal has to assign them into (non-ordered) classes such that no identical-looking children would be in the same class (or else their teacher won't be able to tell them apart, and the children can play tricks on her).

3. It can be used in conjunction with Wilf's [Wi2] celebrated methodology for random selection of combinatorial objects to design quick algorithms for selecting, *uniformly at random*, a mutli-set set-partition of any given multiset.

4. It beautifully illustrates MacMahon's lovely method of "differential operators" that transcribes combinatorial operations into differential operators. MacMahon was very fond of it, and he nicely described it in the entry *Combinatorial Analysis* of the eleventh edition of *Encyclopeadia Britannica* [M].

5. It beautifully illustrates my favorite methodology of *rigorous experimental mathematics*. You teach the computer how to do the combinatorics, it derives, *all by itself*, the (symbolic) differential operator (that for $r > 3$ would be too complicated to derive by hand, let-alone apply, even for MacMahon), and then the computer goes on and uses it to crank-out as many terms as desired in the enumerating sequence.

6. It beautifully illustrates the notion of *catalytic variables*. These are variables corresponding to quantities that we may not care about, but are nevertheless needed in order to facilitate the enumeration. At the end of the day, we set them all equal to 1.

7. Last but not least, it enables me to contribute six new sequences to *Sloane*, the (beginnings of) the enumerating sequence for the number of ways of reassembling r identical n-nested Russian Dolls for $3 \leq r \leq 8$. So far, only $r = 1$ (the Bell numbers, **A000110**) and $r = 2$ (Comtet's sequence **A020554**) are present there.

The Evolution Differential operator

A Baby Example: $r = 1$

For the sake of pedagogy, let's first treat the classical case of one Russian Doll.

Suppose that you have a *set partition* of $\{1, 2, \dots, n-1\}$ with k sets. How can we accommodate a new-comer n? Either she is shy (or anti-social), and decides to form her own new set $\{n\}$, or she is outgoing and would like to join an existing set, for which she has k choices. If we call the "number of sets" the "state", then in the former case the new comer, n, caused the set-partition to move to state $k+1$, while in each of the k latter cases, it stayed in state k. If we give state k the *weight* z^k, then each and every set-partition of state k (with weight z^k) gives rise to the "evolution"

$$z^k \to z^{k+1} + kz^k \quad .$$

In other words

$$z^k \to (z + z\frac{d}{dz})z^k \quad .$$

This is true for each and every monomial z^k, and by linearity for each polynomial. So if $P_n(z)$ is the sum of the weights of all set-partitions of $\{1, \dots, n\}$, then we have the *differential-recurrence* equation:

$$P_n(z) = \mathcal{D}_1 P_{n-1}(z) \quad ,$$

where $\mathcal{D}_1$ is the *differential operator*

$$\mathcal{D}_1 f(z) := (z + z\frac{d}{dz})f(z) \quad .$$

The initial condition is $P_0(z) = 1$. If we are only interested in the *total number* of set partitions of $\{1, 2, \dots, n\}$, then at the end of the day we plug-in $z = 1$, getting

$$B_n = P_n(1) \quad .$$

This gives a quick way to crank-out a table of the first one thousand (or whatever) Bell numbers, that is memory efficient. Once we are at day n, and know $P_n(z)$, we can let the computer forget about $P_{n-1}(z)$ (and even about B_{n-1}, once it is printed out). Surprisingly, this turned out to be much more efficient (for large n and using Maple) than using

$$B_n = \sum_{i=0}^{n-1} \binom{n-1}{i} B_i \quad ,$$

or

$$\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n = e^{e^z - 1} \quad .$$

A Toddler Example: $r = 2$

Suppose that we have $n-1$ pairs of identical twins, already arranged into sets. The only restriction is that no two identical twins can be in the same set, but sets can be repeated. So we have a *multiset of sets*, whose union, as a multiset, is $1^2 2^2 \dots (n-1)^2$.

Of course, no set can be repeated more than twice (why?). Let there be a sets, $A_1, \dots, A_a$ that show up *once*, and let there be b sets $B_1, \dots, B_b$ that show up twice. We say that the *state* of this arrangement is (a, b), and its *weight* is $z_1^a z_2^b$.

The school principal has to place the newly-arrived pair of twins n and n, but he can't put them in the same set. There are **seven** cases.

Case 1: Create two new singleton sets $\{n\}^2$. There is only *one* way of doing it, and the new state is $(a, b+1)$, with weight $z_1^a z_2^{b+1}$. The corresponding operation on monomials is

$$f(z_1, z_2) \to z_2 f(z_1, z_2) \quad .$$

Case 2: Create one new singleton set $\{n\}$, and place the other n into one of the existing $A_i's$. There are a ways of doing it, and the new state is $(a+1, b)$, with weight $z_1^{a+1} z_2^b$. The corresponding operation on monomials is

$$f(z_1, z_2) \to z_1 \left(z_1 \frac{d}{dz_1} \right) f(z_1, z_2) \quad .$$

Case 3: Create one new singleton set $\{n\}$, and place the other n into one of the existing $B_i's$. There are b ways of doing it, and the new state is $(a+3, b-1)$, since one of the B's became two As, with the help of n. The weight of the new state is $z_1^{a+3} z_2^{b-1}$. The corresponding operation on monomials is

$$f(z_1, z_2) \to z_1^3 z_2^{-1} \left(z_2 \frac{d}{dz_2} \right) f(z_1, z_2) \quad .$$

Case 4: Place the two twins n and n into two (different, of course) A_i's. There are $\binom{a}{2}$ ways of doing it, and the new state remains (a, b), with weight $z_1^a z_2^b$. The corresponding operation on monomials is

$$f(z_1, z_2) \to \frac{1}{2} \left(z_1^2 \frac{d^2}{d^2 z_1} \right) f(z_1, z_2) \quad .$$

Case 5: Place one of the new twins (n or n) into one of the A_i's, and the other one into one of the B_i's. There are ab ways of doing it, and the new state is $(a+2, b-1)$, since one of the doubletons B's, let's call it B_i, was lost, and it became the two distinct sets B_i and $B_i \cup \{n\}$. The new weight is $z_1^{a+2} z_2^{b-1}$. The corresponding operation on monomials is

$$f(z_1, z_2) \to z_1^2 z_2^{-1} \left(z_1 \frac{d}{z_1} \right) \left(z_2 \frac{d}{dz_2} \right) f(z_1, z_2) \quad .$$

Case 6: Place the twins (n and n) into two different B_i's. There are $\binom{b}{2}$ ways of doing it, and the new state is $(a+4, b-2)$, since two of the doubletons B's, let's call them B_i and B_j, were lost, and they became the four distinct new sets B_i, $B_i \cup \{n\}$, B_j and $B_j \cup \{n\}$. The new weight is $z_1^{a+4} z_2^{b-2}$. The corresponding operation on monomials is

$$f(z_1, z_2) \to z_1^4 z_2^{-2} \frac{1}{2} \left(z_2^2 \frac{d^2}{dz_2^2} \right) f(z_1, z_2) \quad .$$

Case 7: Place both of the new twins (n and n) into each of the two copies of the same B_i. There are b ways of doing it, and the new state is the same, (a, b), since single sets stay single and doubletons stay doubletons. The new weight is $z_1^a z_2^b$. The corresponding operation on monomials is

$$f(z_1, z_2) \to \left(z_2 \frac{d}{dz_2} \right) f(z_1, z_2) \quad .$$

Combing, we have just proved:

Fact: Let $P_n^{(2)}(z_1,z_2)$ be the sum of the weights of all multiset set-partitions of the multiset $\{1^2 \dots n^2\}$, with the weight being z_1 to the power the number of sets that show up once times z_2 to the power the number of sets that show up twice. Let $\mathcal{D}_2$ be the partial-differential operator (where $D_1 := \frac{d}{dz_1}$, $D_2 := \frac{d}{dz_2}$),

$$\mathcal{D}_2 := z_2 D_2 + 1/2\, {z_1}^4 {D_2}^2 + {z_1}^3 D_1 D_2 + 1/2\, {z_1}^2 {D_1}^2 + {z_1}^3 D_2 + {z_1}^2 D_1 + z_2 \quad ,$$

then

$$P_n^{(2)}(z_1,z_2) = \mathcal{D}_2 P_{n-1}^{(2)}(z_1,z_2) \quad .$$

The Comtet numbers are $P_n^{(2)}(1,1)$.

The general case

Suppose that we already have a multiset set-partition of $\{1^r \dots (n-1)^r\}$, with a_1 sets that show-up once, a_2 sets that show-up twice, $\dots$, a_r sets that show-up r times. The state of this particular multiset set-partion is $(a_1,a_2,\dots,a_r)$, and its weight is $z_1^{a_1} \dots z_r^{a_r}$.

We have to place the r identical new comers n^r.

We must make the following decisions

1. How many of them would start their own singleton sets, say, c_0

2. For the remaining $r-c_0$ new members n, for $i=1,2,\dots,r$, how many of them, let's call it c_i, would be placed in sets that show up i times.

After these decisions we have a vector of non-negative integers $[c_0,c_1,\dots,c_r]$, such that $c_0+c_1+\dots+c_r=r$.

Once we decided that c_i of the new n's would go to sets that show up i times, we have to decide, amongst those c_i siblings which sets should be asked to invite them. These sets can be all different, but they could all be different copies of the same set (if $c_i \leq i$). This naturally leads to an *integer partition* $\lambda_i = 1^{m_1} 2^{m_2} 3^{m_3} \dots i^{m_i}$ (written in multiplicity notation).

We have to place m_1 of these siblings such that each of them goes to different sets. These make m_1 of the formerly a_i-repeated sets become (a_i-1)-repeated sets and creates m_1 new singleton sets. So $a_1 \to a_1+m_1$, and $a_{i-1} \to a_{i-1}+m_1$ and $a_i \to a_i - m_1$. In terms of differential operators it is

$$f(z_1,\dots,z_r) \to (1/m_1!)(z_1 z_{i-1} D_i)^{m_1} f(z_1,\dots,z_r) \quad .$$

Similarly for 2^{m_2} we have

$$f(z_1,\dots,z_r) \to (1/m_2!)(z_2 z_{i-2} D_i)^{m_2} f(z_1,\dots,z_r) \quad ,$$

and so on.

So every possible scenario of placing the new identical r siblings of the n family corresponds to an r+1 tuple:

$$T = [c_0, \lambda_1, \ldots, \lambda_r] \quad ,$$

where c_0 is an integer, $\lambda_1, \lambda_2, \ldots, \lambda_r$ are integer partitions such that the largest part of λ_i is $\leq i$, and

$$c_0 + |\lambda_1| + |\lambda_2| + \ldots + |\lambda_r| = r \quad .$$

For each such scenario corresponds the "monomial" operator

$$\mathcal{P}[T] := z_{c_0} \prod_{i=1}^{r} \mathcal{Q}[\lambda_i] \quad ,$$

where, writing $\lambda_i = 1^{m_1} 2^{m_2} 3^{m_3} \ldots i^{m_i}$ (m_1, m_2 are now *local variables*, i.e. they are different, of course, for each λ_i),

$$\mathcal{Q}[\lambda_i] = \prod_{j=1}^{i} \frac{1}{m_j!} (z_j z_{i-j} D_i)^{m_j} \quad .$$

Here $D_i := \frac{d}{dz_i}$ and $z_0 := 1$.

Finally, we can write down the *evolution operator* $\mathcal{D}_r$:

$$\mathcal{D}_r := \sum_{T \ scenario} \mathcal{P}[T] \quad ,$$

and we have the

Theorem: Let $P_n^{(r)}(z_1, z_2, \ldots, z_r)$ be the sum of the weights of all multiset set-partions of the multiset $\{1^r \ldots n^r\}$, with the weight being z_1 to the power the number of sets that show up once times z_2 to the power the number of sets that show up twice times … times z_r to the power the number of sets that show up r times, then

$$P_n^{(r)}(z_1, z_2, \ldots, z_r) = \mathcal{D}_r P_{n-1}^{(r)}(z_1, z_2, \ldots, z_r) \quad .$$

The *number* of such multi-set set partitions is of course $P_n^{(r)}(1, 1, \ldots, 1)$.

Non-Identical Russian Dolls

The operators $\mathcal{D}_r$ can be combined to yield the

Main Theorem: The number of ways of partitioning into sets a multiset consisting of m_1 elements that appear once, m_2 elements that appear twice, …, m_r elements that appear r times, or equivalently, the multiset

$$1 \ldots m_1 (m_1+1)^2 \ldots (m_1+m_2)^2 \ldots (m_1+\ldots+m_{r-1}+1)^r \ldots (m_1+\ldots+m_{r-1}+m_r)^r$$

is computed as follows. First compute the polynomial in $z_1, \dots, z_r$:

$$P(z_1, z_2, \dots, z_r) = \left(\prod_{i=1}^{r} \mathcal{D}_i^{m_i} \right) (1) \quad ,$$

and then plug-in $z_1 = 1, \dots, z_r = 1$.

Random Generation

Going back to the Stirling-Bell case, $r = 1$, the "differential recurrence" $P_n(z) = (z + z\frac{d}{dz})P_{n-1}(z)$ is equivalent to the famous recurrence

$$S(n,k) = S(n-1, k-1) + kS(n-1, k) \quad .$$

This can be used, according to Wilf[W2], to generate *uniformly at random*, a set-partion with k sets as follows.

First pre-compute a table of $S(n,k)$ using the recurrence. Now roll a loaded coin with probability of Heads being $S(n-1,k-1)/S(n,k)$ and probability of Tails being $kS(n-1,k)/S(n,k)$. If it lands Heads, recursively generate a random set-partitions of $\{1, 2, \dots, n-1\}$ with $k-1$ sets, and adjoin the singleton $\{n\}$ to it, otherwise generate recursively a random set-partition of $\{1, 2, \dots, n-1\}$ with k sets, and then roll a fair k-sided die, and accordingly decide which of the k members of the set-partition should invite n to join it.

If we want a (uniformly) random set partition, then decide on the number of sets k, by rolling a loaded n-faced die with probabilities of it landing k equalling $S(n,k)/B_n$, and then proceed as before.

The differential-recurrence of the Theorem yields to a partial recurrence for the quantity, let's call it $S^{(r)}(n; a_1, \dots, a_r)$ for the number of multiset set-partitions of $1^r \dots n^r$ with a_1 sets that show up once, $\dots$, a_r sets that show up r times. Using this the computer (all by itself!) can use the Wilf Methodology to create a random-generation algorithm. The programming details are a bit daunting, so we leave it as a challenge to the reader.

The Maple package BABUSHKAS

Everything here (except for the random-generation, for which we only have the simple $r = 1$ case) is implemented in the Maple package `BABUSHKAS` available, via a link, from the webpage of this article: `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/babushkas.html`
or directly from: `http://www.math.rutgers.edu/~zeilberg/tokhniot/BABUSHKAS` . That webpage also contains sample input and output, including the sequences for $1 \leq r \leq 8$.

The main procedure is `SeqBrn(r,n)` that uses the present approach to generate the first n terms of the enumerating sequence for the number of ways of reassembling r identical Russian Dolls. `SeqBrnDJ(r,n)` does the same thing using the Devitt-Jackson approach. We are glad to report that they agree! As yet another check, we have the program `SSP`, that actually constructs the set of *all* multi-set set partitions of any given multiset, and emables checking, for small values, with the naive count. Procedure `B(L)` handles the case of non-identical Russian Dolls, or equivalently, an arbitrary multiset, using the Main Theorem.

`SeqCrn` and `SeqCrnDJ` handle set-partitions of the multiset $1^r \dots n^r$, in other words, each set can only show up once. This is simply $P_n^{(r)}(1,0,\dots,0)$, in the above notation.

Full details are available on-line by typing `ezra();`.

The sequences

Even though much more data is available in the above-mentioned webpage, and these sequences will soon be submitted to *Sloane*, let us cite the first ten terms for $r = 1, 2, 3, 4$.

$r = 1 : 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975$ (the Bell Numbers).

$r = 2 : 1, 3, 16, 139, 1750, 29388, 624889, 16255738, 504717929, 18353177160$ (the Comtet numbers)

$r = 3 : 1, 4, 39, 862, 35775, 2406208, 238773109, 32867762616, 6009498859909, 1412846181645855$

$r = 4 : 1, 5, 81, 4079, 507549, 127126912, 55643064708, 38715666455777, 40095856807088486,$
58901884724160709571.

References

[Ba] G. Baroti, *Calcul des nombres de bicouvrements et de birevêtements d'un ensemble fini, employant la méthode fonctionnelle de Rota*, in: Combinatorial theory and its Applications I (ed. P. Erdős et al.), North Holland, Amsterdam, 1970), 93-103.

[Be] E.A. Bender, *Partitions of multisets*, Discrete Mathematics **9**(1974), 301-312.

[C] L. Comtet, *Birecouvrements et birevetements d'un ensemble fini*, Studia Sci. Math. Hungar. **3**(1968), 137-152.

[DJ] J.S. Devitt and D.M. Jackson, *The enumeration of covers of a finite set*, J. London Math. Soc.(2) **25** (1982), 1-6.

[L] G. Labelle, *Counting enriched multigraphs*, Discrete Math. **217**(2000), 237-248.

[M] P.A. MacMahon, *Combinatory Analysis*, in: *Encyclopeadia Britantica, 11th edition* (available on-line) vol. 6, 752-758 (1910).

[P] G. Paquin, *Dénombrement de multigraphes enrichis*, Mémoire, Math. Dept., Univ. Québec à Montréal, 2004.

[R] J. Reilly, *"Bicoverings of a Finite Set" by Generating Function Methods*, J. Comb. Theory, Series A **28** (1980), 219-225.

[W1] H.S. Wilf, *What is an answer?*, Amer. Math. Monthly **89** (1982), 289-292.

[W2] H. S. Wilf, *A unified setting for sequencing, ranking, and selection algorithms for combinatorial objects*, Advances in Mathematics **24**(1977), 281-291.