Integer Partitions

Stuyvesant Junior Math Team

Aditya Pahuja

May 17, 2024

Contents

1 Introduction

One of the fundamental counting problems is that of splitting a collection of n things into groups. There are several different flavors of this question: are the things distinct or identical? Do we care about the ordering of the groups? Are the groupings restricted in any way?

This sequence of lessons addresses the instance of n identical things being split into groups whose order doesn't matter. This is more concretely phrased as expressing an integer as the sum of a nonincreasing sequence of nonnegative integers.

Definition 1.1 — A **partition** of an integer n is a decomposition of n into a sum of positive integers: namely, we say that a sequence $(a_1, a_2, \dots, a_k)$ is a partition of n if $a_1 \geq a_2 \geq \cdots \geq a_k$ and $a_1 + a_2 + \cdots + a_k = n$. Moreover, the a_i are called the **parts** of this partition.

Where convenient, we will abbreviate $(a_1, a_2, \dots, a_k)$ by just writing $a_1 + a_2 + \cdots + a_k$ for a given partition of n. For example, the seven partitions of 5 are

$$5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1.$$

We generally refer to the number of partitions of n as $p(n)$, so $p(5) = 7$.

Counting partitions is (very, very) hard, and as such, the main flavor of question that appears in the theory of integer partitions is that of enumerating various sets of partitions. For example:

Theorem 1.2. The number of partitions of n with k parts is the same as the number of partitions of n whose largest part is k.

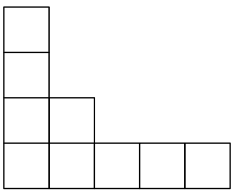
Much of our work will be dedicated to understanding different ways to prove such theorems.

2 Bijections galore

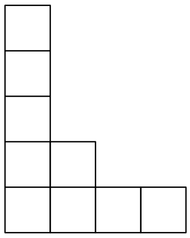
Statements like that of Theorem 1.2 often beg for bijective proofs. We thus introduce a tool for visually manipulating partitions, for pictures help us find combinatorial proofs.

Definition 2.1 — The **Ferrers diagram** of a partition $(a_1, a_2, \dots, a_k)$ of n is a set of k columns of boxes aligned at their bottom edge in which column j contains a_j boxes.

It is easier to describe these by just showing a picture; here is the Ferrers diagram of $(4, 2, 1, 1, 1)$.



Obviously, Ferrers diagrams with n boxes are in bijection with partitions of n. What's nice is that we can manipulate these diagrams so as to generate new partitions. One common construct of this flavor is the **conjugate partition**, obtained by reflecting a Ferrers diagram so that columns become rows and rows become columns (i.e. read the diagram horizontally instead of vertically). For example, the conjugate of $(4, 2, 1, 1, 1)$ is $(5, 2, 1, 1)$, as shown below.



We also have the notion of a **self-conjugate** partition, which, as you may infer, is a partition that is invariant under conjugation.

This gives us the machinery to prove Theorem 1.2 very quickly, as well as several other neat facts.

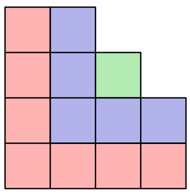
Proof of Theorem 1.2. Let $\mathcal{S}$ be the set of partitions of n with k parts, and let $\mathcal{T}$ be the set of partitions of n with largest part k. Then, we have a natural mapping $S \to T$ given by conjugation, and it is obvious that this map is bijective because it is its own inverse (in the jargon, we say that the map is an **involution**). □

> **Exercise.** If the bijectivity of the conjugation map is not obvious to you, then prove that indeed any function that is its own inverse must be bijective.

Some more partition facts:

Theorem 2.2. The number of partitions of n into distinct odd parts is the same as the number of self-conjugate partitions of n.

Proof. I'll just provide a picture for this one.



Do you see the bijection? □

Theorem 2.3. The number of partitions of n in which each part is at least two is $p(n) - p(n-1)$.

Proof. We just need to show that the number of partitions of n containing a 1 is $p(n-1)$. The corresponding bijection is easy: take a partition of n containing 1, and delete the 1! □

Theorem 2.4. The number of partitions of n whose parts are all odd is the same as the number of partitions of n whose parts are all distinct.

We will show two proofs of this fact (and a third one via generating functions later!).

Proof 1. Consider a partition into odd parts, say

$$n = a_1 \cdot 1 + a_3 \cdot 3 + a_5 \cdot 5 + \cdots .$$

Write each a_i in binary, so we have

$$a_i = b_{0,i} + 2b_{1,i} + 4b_{2,i} + \cdots$$

where $b_{x,y} \in \{0, 1\}$ for all nonnegative integers x and odd positive integers y. Then,

$$n = \sum_{x \geq 0} \sum_{y \text{ odd}} 2^x \cdot y \cdot b_{x,y},$$

which is a partition of n into distinct integers because each integer can be uniquely expressed as the product of an odd number and a power of two. This is thus an injective function mapping the partitions of n into odd integers to the partitions of n into distinct integers. The function is also invertible as, given a partition with distinct integers, we just split each part into 2^k smaller odd parts. More precisely, given n in the form

$$n = \sum_{x \geq 0} \sum_{y \text{ odd}} 2^x \cdot y \cdot b_{x,y},$$

we swap the order of summation to write

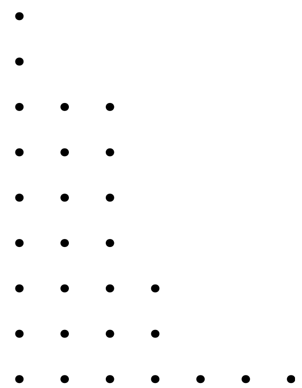
$$n = \sum_{y \text{ odd}} \left(\sum_{x \geq 0} 2^x \cdot b_{x,y} \right) y,$$

where the inner sum exactly gives a_y.

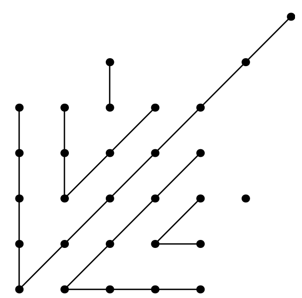
This means we have described a bijective function between the two sets of partitions, which implies the sets have the same size. $\square$

> **Remark 2.5.** This proof can also be phrased as a greedy algorithm, where the odd to distinct map is given by repeatedly pairing off identical numbers until you can't anymore.

Proof 2. This is a pretty visual proof, so again I will draw a picture and let you interpret the algorithm. It will help to draw the Ferrers diagram with dots here:



Here are the same dots, but moved around a bit (a lot):



Verifying that this is actually a bijection between the sets of partitions is left as an exercise. $\square$

Theorem 2.6. Let $p(a,b,n)$ be the number of partitions of n with at most a parts, each of which is at most b. Then,

$$p(a,b,0) + p(a,b,1) + p(a,b,2) + \cdots + p(a,b,ab) = \binom{a+b}{a}.$$

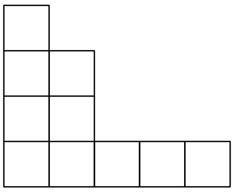
Proof. Consider an up-right lattice path in the coordinate plane from $(0,0)$ to (a,b). This path, together with the lines $x = 0$ and $y = b$, bounds a region that can be read off as the Ferrers diagram of some partition. As the path is guaranteed to have at most a parts, with each each part being at most b, we see that the number of lattice paths is

$$\binom{a+b}{b} = \sum_{n \geq 0} p(a,b,n).$$

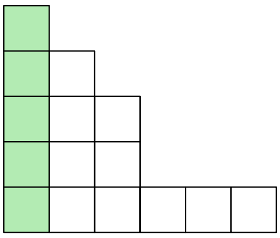
However, $p(a,b,n) = 0$ for $n > ab$ because such n are simply too large, so we can cut the sum off at $n = ab$ to get the desired result. $\square$

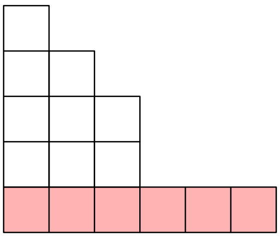
Theorem 2.7. The number of partitions of $a-c$ into $b-1$ parts, each of which is at most c, is the same as the number of partitions of $a-b$ into $c-1$ parts, each of which is at most b.

Proof. Yet again a sequence of diagrams! Here, we are taking $a = 15$, $b = 6$, and $c = 5$. One partition of $a - c = 10$ into $b - 1 = 5$ parts of size at most $c = 5$ is

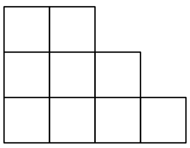


Add a part (column) of size c and then subtract one from each part (delete the bottom row):





Then, take the conjugate of the result. It should be clear that this is always going to be a partition of $a - b$ into $c - 1$ parts of size at most b.



Verifying that this is a bijection between the two sets of partitions is (you guessed it!) an exercise for you, dear reader. $\square$

Theorem 2.8. The number of partitions of n into parts not divisible by d is the same as the number of partitions of n whose parts are each repeated less than d times.

Sketch. This is a generalization of Theorem 2.4. Indeed, we can use a similar argument to the first proof: for each partition of the former type, you can write the number of occurrences it has in base d and then conclude, since each integer can uniquely be expressed as $d^x y$ where $d \nmid y$; again, reversing the map is fairly simple. ☐

Remark 2.9. Do let me know if you find a more visual argument for this one.

The following theorem leads to a beautiful proof of what's known as "Euler's pentagonal number theorem," which will be stated in its full depth in the next section.

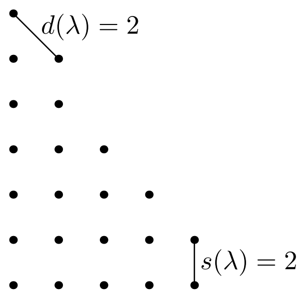
Theorem 2.10. Let $p_e(n)$ be the number of partitions of n into an even number of distinct parts, and let $p_o(n)$ be the number of partitions of n into an odd number of distinct parts. Then,

$$p_e(n) - p_o(n) = \begin{cases} (-1)^m & \text{if } n = \frac{1}{2}m(3m \pm 1) \\ 0 & \text{otherwise} \end{cases}$$

Numbers of the form $\frac{1}{2}m(3m \pm 1)$, where m is an integer, are called **pentagonal numbers**.

Proof. Let A and B respectively be the two sets of partitions in question. We will try to pair off elements of A and B, and our pairing will succeed in all but two cases.

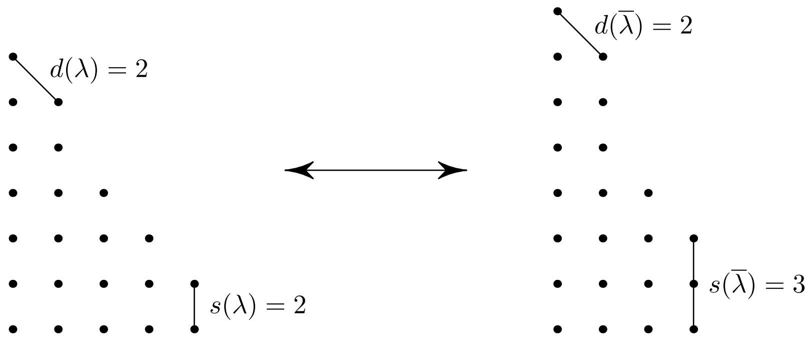
First, given a partition λ of n, we define the following two functions: $s(\lambda)$ is the smallest part of λ and $d(\lambda)$ is the greatest j with the property that the j largest parts of λ are consecutive. They are quite easy to read off of a given Ferrers diagram; here is an example with $\lambda = (7, 6, 4, 3, 2)$.



Given λ, we then define $\overline{\lambda}$ as follows:

- If $s(\lambda) > d(\lambda)$, then subtract one from each of the $d(\lambda)$ largest parts, and add in a new (smallest) part of size $d(\lambda)$.
- If $s(\lambda) \le d(\lambda)$, then delete the smallest part and add one to each of the $s(\lambda)$ largest parts.

Here is what this looks like for $\lambda = (7, 6, 4, 3, 2)$ and its counterpart $\overline{\lambda} = (8, 7, 4, 3)$:



Most of the time, this will pair an element of A with an element of B. However, there are two points of failure.

- In the first case, our transformation tries to create a new smallest part. This can fail if $\lambda = (r+1, r+2, \ldots, 2r)$, in which case $\overline{\lambda} = (r, r, r+1, \ldots, 2r-1)$ no longer has distinct parts. Here, λ is a partition of $\frac{r(3r+1)}{2}$.
- In the second case, our transformation tries to get rid of the smallest part. This can fail if $\lambda = (r, r+1, \ldots, 2r-1)$, in which case $\overline{\lambda} = (1, r+2, r+3, \ldots, 2r)$ still has the same number of parts. Here, λ is a partition of $\frac{r(3r-1)}{2}$.

Exercise. Check that, for all other partitions of n into distinct parts, our transformation works as intended (i.e. pairs off an element of A with an element of B).

Thus, if $n = \frac{r(3r\pm1)}{2}$, then, based on the parity of n, $|A| - |B| = p_e(n) - p_o(n)$ is either $+1$ or -1. Indeed, the exact value is just $(-1)^r$ since the bad λs enumerated above both have r parts. Otherwise (if n isn't a pentagonal number) then the transformation works seamlessly, and $|A| = |B|$. $\square$

The theory of partition bijections can be taken very far; a sequence of papers in the 1980s went so far as to create a *bijection machine* for generating bijections between certain kinds of restricted partitions. An overview of this sequence of papers can be found in [3].

3 Heavy artillery

3.1 Brute forcing identities

Bijections are hard to find and can require much ingenuity. We can instead approach partitions from a more computational perspective via generating functions.

Theorem 3.1. Let $\mathcal{P}(x)$ be the (ordinary) generating function for $p(n)$. Then,

$$\mathcal{P}(x) = \sum_{n \geq 0} p(n)x^n = \prod_{n \geq 1} \frac{1}{1 - x^n}.$$

Remark 3.2 (If "exponential generating function" means something to you)**.** All the generating functions we will work with are going to be ordinary generating functions, so I will omit the word "ordinary" from here on out.

To be more explicit, the product should read as

$$\prod_{n \geq 1} (1 + x^n + x^{2n} + x^{3n} + \cdots).$$

When multiplying this out, then, each term in the expansion is produced by picking exactly one $(x^n)^j$ monomial from each term of the product. The result is that each term looks like

$$x^{a_1 \cdot 1 + a_2 \cdot 2 + a_3 \cdot 3 + \cdots},$$

so (assuming the exponent is finite) this term enumerates one of the partitions of $n = a_1 \cdot 1 + a_2 \cdot 2 + a_3 \cdot 3 + \cdots$. Summing across all such terms, we do indeed end up with $\sum_n p(n)x^n$.

Remark 3.3 (Technical note on the convergence of the product)**.** At this point you may be concerned about the caveat "assuming the exponent is finite" in the previous paragraph; that was a bit of handwaving. More precisely, we can say

$$\lim_{m \to \infty} \prod_{k=1}^{m} \frac{1}{1 - x^k} = \mathcal{P}(x)$$

in that, for each k, the coefficients of x^0, x^1, ..., x^k will stabilize (at $p(0)$, $p(1)$, ..., $p(k)$ respectively) once m is big enough (in fact $m > k$ is sufficient). This is necessary even when multiplying just two of the factors in the product, actually, since the factors are really infinite sums.

If you want to be extra extra rigorous, $\mathbb{R}[[x]]$ (the set of power series in x with real coefficients) can be made a metric space with metric

$$d(P(x), Q(x)) = c^{-\nu_x(P(x) - Q(x))},$$

where c can be any constant greater than 1 of your choosing and $\nu_x(F(x))$ is the smallest integer k such that x^k has a nonzero coefficient (in particular $\nu_x(0) = +\infty$, and in this case you should take $c^{-\infty} = 0$). Convergence then works verbatim using your ε-δ definition from calculus class, with d being your distance function instead of the absolute value (hence the name "metric").

You can trust me that all infinite products I write in this document are convergent.

Naturally, we don't have to take our product over all integers and we don't need each term to be infinite — imposing such restrictions lets us work with different flavors of partitions, like the kind we saw in the previous section. For example, let's prove Theorem 2.4 with our new machinery.

Proof of Theorem 2.4. We construct two generating functions, $\mathcal{O}(x)$ for partitions with odd parts and $\mathcal{D}(x)$ for partitions with distinct parts.

For $\mathcal{O}(x)$, we only take the factors corresponding to odd parts, which gives

$$\mathcal{O}(x) = \prod_{k \text{ odd}} \frac{1}{1-x^k}.$$

On the other hand, for $\mathcal{D}(x)$, we truncate each factor after the second term:

$$\mathcal{D}(x) = \prod_{k \geq 1} (1+x^k).$$

We thus need to verify that these generating functions are identical, which we do by writing

$$\mathcal{D}(x) = (1+x)(1+x^2)(1+x^3)\cdots = \frac{1-x^2}{1-x} \cdot \frac{1-x^4}{1-x^2} \cdot \frac{1-x^6}{1-x^3} \cdots$$

and observing that all the even degree terms cancel out to yield $\mathcal{O}(x)$. □

More involved example problem:

Problem 3.4 (USAMO 1986/5)

For any partition π, define $A(\pi)$ to be the number of 1's which appear in π, and define $B(\pi)$ to be the number of distinct integers which appear in π. For example, if $n = 13$ and π is the partition $1 + 1 + 2 + 2 + 2 + 5$, then $A(\pi) = 2$ and $B(\pi) = 3$.

Prove that, for any fixed n, the sum of $A(\pi)$ over all partitions π of n is equal to the sum of $B(\pi)$ over all partitions π of n.

Solution. Let $\alpha(n)$ be the sum of A over all partitions of n and let $\beta(n)$ be the sum of B over all partitions of n. Our goal will be to find generating functions for α and β.

Generating function for α. Let $\alpha_k(n)$ be the number of partitions of n with k ones, so that we want to compute

$$\alpha(n) = \sum_{k=1}^{n} k\alpha_k(n).$$

The key observation is the following:

Claim — $\alpha_k(n)$ is the number of partitions of $n-k$ with no ones.

Proof. Obvious; just delete the ones from a partition counted by $\alpha_k(n)$. □

To this end, consider the generating function for partitions of n with no ones, namely

$$(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)(1 + x^4 + x^8 + \cdots) = \prod_{k \geq 2} \frac{1}{1-x^k} = \sum_{k \geq 0} c_k x^k.$$

We want to compute the sum $c_{n-1} + 2c_{n-2} + \cdots + (n-1)c_1 + nc_0$. Here, we employ the deep algebraic strategy of multiplying the expression by $x + 2x^2 + 3x^3 + \cdots$ to get

$$(x + 2x^2 + 3x^3 + \cdots) \prod_{k \geq 2} \frac{1}{1-x^k} = \frac{x}{(1-x)^2} \prod_{k \geq 2} \frac{1}{1-x^k} = \frac{x}{1-x} \mathcal{P}(x).$$

One can easily check that at this point the coefficients are what we want; indeed, the x^n term is a combination of terms that look like $kx^k \cdot c_{n-k} x^{n-k}$ as desired. Thus, the expression above is the generating function for α.

Generating function for β. The main idea is that we can write

$$\beta(n) = \sum_{\pi} B(\pi) = \sum_{k=1}^{n} (\text{\# of partitions containing } k) = \sum_{k=1}^{n} p(n-k),$$

where the second equality follows from an argument identical to that from Theorem 2.3. At this point, we immediately see that

$$\sum_{n\geq 1} \beta(n)x^n = (x + x^2 + x^3 + \cdots)\mathcal{P}(x) = \frac{x}{1-x}\mathcal{P}(x).$$

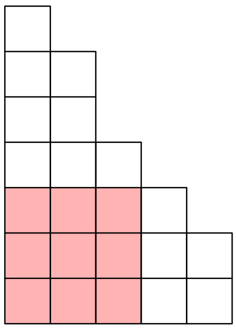
Since the generating functions for α and β are identical, we can conclude $\alpha(n) = \beta(n)$ for all n. $\square$

For more practice, you can of course go back and try to prove the theorems in the previous section with generating function machinery.

3.2 Durfee squares

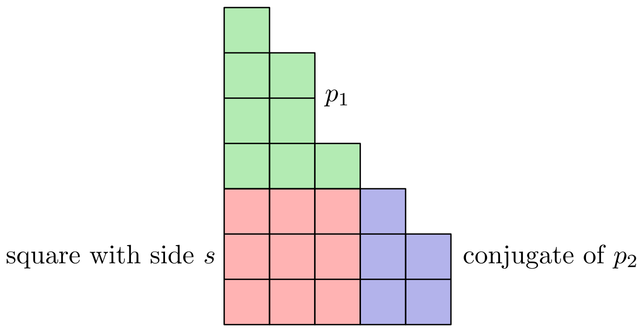
Definition 3.5 — The **Durfee square** of a partition is the largest square that can be drawn along the bottom and left edges of the partition's Ferrers diagram.

For example, the following partition has a Durfee square of side length 3.



Exercise. Describe a way to determine the side length of the Durfee square of a given partition without drawing its Ferrers diagram.

The interesting thing is that drawing the Durfee square splits a partition into three pieces: the square itself (say of side length s), a partition with at most s parts, and a partition whose conjugate has at most s parts.



In fact, the set of triples (s, p_1, p_2), where $s \geq 0$ is an integer and p_1 and p_2 are partitions with at most s parts, is in bijection with the set of all partitions, with the bijective mapping given exactly by the above picture.

This leads us very quickly to the following identity:

Theorem 3.6.

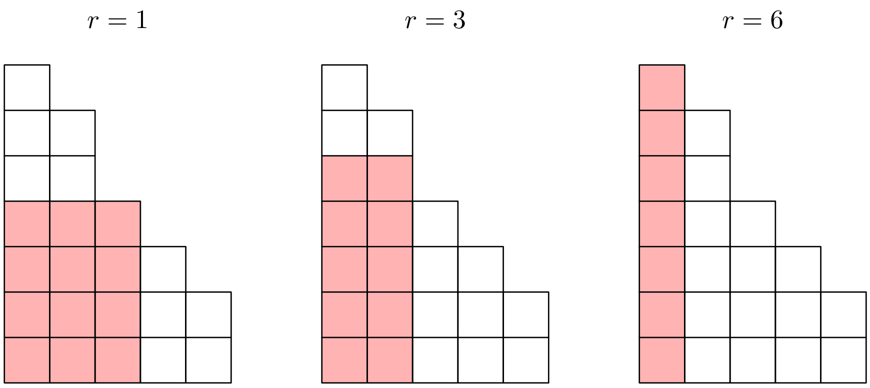
$$\sum_{s\geq 0}\left(x^{s^2}\cdot\prod_{i=1}^{s}\frac{1}{(1-x^i)^2}\right)=\mathcal{P}(x).$$

We can do a bit better, though.

Definition 3.7 (Not standard terminology) — For an integer $r \geq 0$, the **r-Durfee rectangle** of a given partition is the largest bottom-edge-aligned and left-edge-aligned rectangle whose height is r more than its width that can be drawn within the Ferrers diagram. In particular, the 0-Durfee rectangle is the Durfee square.

Exercise (Just like the previous exercise)**.** How do you determine the length of the r-Durfee rectangle of a given partition without looking at its Ferrers diagram?

Here are a couple of examples.



We can then get a more general version of Theorem 3.6:

Theorem 3.8.

$$\sum_{s\geq 0}\left(x^{s(s+r)}\cdot\prod_{i=1}^{s+r}(1-x^i)^{-1}\cdot\prod_{i=1}^{s}(1-x^i)^{-1}\right)=\mathcal{P}(x).$$

The proof goes in a very similar manner: fixing r, we get a bijection between arbitrary partitions and triples (s, p_1, p_2), where p_1 has at most s parts and p_2 has at most $s + r$ parts. The bijective mapping is directly analogous to the one used in Theorem 3.6: take an $s \times (s + r)$ rectangle and glue p_1 and p_2 into the sides with length s and $s + r$ respectively.

4 Euler's pentagonal number theorem

4.1 The combinatorial approach

The idea here is to find a generating function for $p_e(n) - p_o(n)$, as defined in Theorem 2.9. Of course, we already know the value of this function at every n, so one way to express this power series is

$$\sum_{m\in\mathbb{Z}} (-1)^m x^{m(3m-1)/2}.$$

To make this more manageable, what we can do is give each partition a weight based on the number of parts it has: in other words, consider the following sum.

$$\sum_{n\geq 0} x^n \cdot \sum_{\substack{\sum a_i = n \\ a_1 > \cdots > a_k}} (-1)^k.$$

(To be explicit, the inner sum ranges over all partitions of n with distinct parts.) Clearly, the x^n coefficient here is $p_e(n) - p_o(n)$. On the other hand, we can rewrite the sum as

$$\sum_{n\geq 0} \sum_{\substack{\sum a_i = n \\ a_1 > \cdots > a_k}} (-x^{a_1})(-x^{a_2})\cdots(-x^{a_k}),$$

which in fact is

$$(1-x)(1-x^2)(1-x^3)\cdots = \prod_{k\geq 1}(1-x^k).$$

(Observe the similarity between this and

$$\sum_{n\geq 0} \sum_{\substack{\sum a_i = n \\ a_1 > \cdots > a_k}} x^{a_1}x^{a_2}\cdots x^{a_k} = \prod_{k\geq 1}(1+x^k),$$

which is $\mathcal{D}(x)$.) Thus we obtain the titular theorem:

Theorem 4.1 (Euler's pentagonal number theorem).

$$\prod_{k\geq 1}(1-x^k) = \sum_{m\in\mathbb{Z}} (-1)^m x^{m(3m-1)/2} = 1 + \sum_{m\geq 1}(-1)^m\left(x^{m(3m-1)/1} + x^{m(3m+1)/2}\right).$$

This is an incredible fact, since it means we have exactly computed the power series $\mathcal{Q}(x)$ for which $\mathcal{P}(x)\mathcal{Q}(x) = 1$. Let's elaborate more on this fact: for $n \geq 1$, the x^n coefficient in $\mathcal{P}(x)\mathcal{Q}(x)$ is exactly

$$\sum_{m\in\mathbb{Z}} (-1)^m p\left(n - \frac{m(3m-1)}{2}\right) = p(n) - p(n-1) - p(n-2) + p(n-5) + p(n-7) - \cdots.$$

This, however, is the x^n coefficient of 1, which is obviously 0. Thus, we have a recursion for $p(n)$ in the form

$$\boxed{p(n) = \sum_{m\geq 1}(-1)^m\left(p\left(n - \frac{m(3m-1)}{2}\right) + p\left(n - \frac{m(3m+1)}{2}\right)\right)}.$$

This is pretty neat, since we have only about $\frac{2}{3}\sqrt{6n}$ nonzero terms on the right-hand side. It turns out that this lets us compute $p(n)$ fairly efficiently: in the language of computer science, one can compute all $p(n)$ up to a certain large N in $O(N\sqrt{N})$ operations.

Exercise (If you're bored)**.** Compute $p(1)$, $p(2)$, $\ldots$, $p(15)$.

Remark 4.2. In the way of computation, the following asymptotic formula for $p(n)$ is also cool (but way out of scope of this lesson):

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

where $\exp(x) = e^x$. That is, the ratio of the two expressions approaches 1 as n grows arbitrarily large.

The best known method for computing $p(n)$ for large n is the Hardy-Ramanujan-Rademacher formula:

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{k=1}^{\infty} A_k(n) k^{1/2} \left[\frac{d}{dx} \frac{\sinh((\pi/k)(\frac{2}{3}(x - \frac{1}{24})^{1/2}))}{(x - \frac{1}{24})^{1/2}}\right]_{x=n}$$

where

$$A_k(n) = \sum_{\substack{h \bmod k \\ \gcd(h,k)=1}} \omega_{h,k} e^{-2\pi i n h/k}$$

where $\omega_{h,k}$ is a particular 24th root of unity (I will spare you from the precise definition). It is a monstrosity of a formula and a truly breathtaking achievement, giving an exact formula for $p(n)$ that converges extremely rapidly (and actually gives the asymptotic formula above as a corollary).

4.2 The generating function approach

There is also a proof of the pentagonal number theorem with a more generating function based approach. In particular, we will prove the following identity, from which the equality of generating functions above will fall out nicely.

Theorem 4.3 (Jacobi's triple product identity)**.**

$$\prod_{n\geq 1}(1 + x^{2n-1}z)(1 + x^{2n-1}z^{-1}) = \prod_{n\geq 1}(1 - x^{2n})^{-1} \sum_{r\in\mathbb{Z}} x^{r^2} z^r.$$

(Normally, the three products are placed on the left side, but we will use this form.)

Remark 4.4 (More technical plumbing)**.** We now have two variables in our big product instead of one. One way to read this is "for each z^r, the coefficient (which is an element of $\mathbb{R}[[x]]$) converges as we take the limit to infinity." This obviously extends to more than two variables: as a result, you can inductively define a topology on $\mathbb{R}[[x_1, x_2, \ldots, x_n]]$ using this general idea. Again, you can trust that this holds in this case.

Proof. We follow the second proof outlined in [4].

First, fixing r, let's describe what the $x^q z^r$ coefficient counts on the left-hand side. Since the LHS is symmetric in z^1 and z^{-1}, we can assume without loss of generality that r is positive. Such a term arises from picking r more $x^{2n-1}z$ terms than $x^{2n-1}z^{-1}$ terms when expanding the product — thus, the coefficient of $x^q z^r$ counts the number of ways to express q as

$$q = \sum_{i=1}^{s+r}(2a_i - 1) + \sum_{i=1}^{s}(2b_i - 1),$$

where $s \geq 0$ and the a_i and b_i are increasing sequences of positive integers.

Thus, if we write the LHS as

$$\sum_{r \in \mathbb{Z}} f_r(x) z^r,$$

we know that the coefficients of $f_r(x)$ count the number of ways to choose the a_i and b_i. Our goal then is to show that $f_r(x) = x^{r^2}\mathcal{P}(x^2)$.

Right now, we have expressed q as a sum of two partitions with odd, distinct parts. This should be vaguely reminiscent of Theorem 3.8, which dealt with sums of partitions, albeit unrestricted ones. To start with, we can write

$$\frac{q+2s+r}{2} = \sum_{i=1}^{s+r} a_i + \sum_{i=1}^{s} b_i.$$

This gets rid of the parity restriction, but it doesn't change the fact that the a_i are distinct and the b_i are distinct. The goal, then, is to convert an increasing sequence of positive integers to a nondecreasing sequence of nonnegative integers, which is the setting that Theorem 3.8 takes place in. The standard way that this is done is to consider $a_i - i$ and $b_i - i$ (you might see this pattern in a stars-and-bars context, too): after some algebra, this gives

$$\frac{q-r^2}{2} = s(s+r) + \sum_{i=1}^{s+r}(a_i - i) + \sum_{i=1}^{s}(b_i - i)$$

To be explicit, what we have done here is reparameterized $\frac{q-r^2}{2}$ as $s(s+r)$ plus a sum of unrestricted partitions with at most $s+r$ and s parts respectively

In other words, the number of ways to express q as a sum of two increasing sequences of $s+r$ and s odd numbers is the same as the number of ways to express $\frac{q-r^2}{2}$ as $s(s+r)$ plus two arbitrary partitions with at most $s+r$ and s parts, across all nonnegative integers s. This latter object is exactly what Theorem 3.8 counts — indeed, the quantity we have been counting is in bijection with the partitions of $(q-r^2)/2$, so the quantity we have been counting is the $x^{(q-r^2)/2}$ coefficient of $\mathcal{P}(x)$.

The rest is purely algebraic manipulation to extract

$$f_r(x) = \sum_{q \geq 0} p\left(\frac{q-r^2}{2}\right) x^q$$

in terms of $\mathcal{P}(x)$. We can double all the exponents of $\mathcal{P}(x)$ by substituting $x \mapsto x^2$, and then add r^2 to all the exponents by multiplying it with x^{r^2}:

$$\begin{aligned}
\mathcal{P}(x) &= p(0) + p(1)x + \cdots + p\left(\frac{q-r^2}{2}\right) x^{(q-r^2)/2} + \cdots \\
\mathcal{P}(x^2) &= p(0) + p(1)x^2 + \cdots + p\left(\frac{q-r^2}{2}\right) x^{q-r^2} + \cdots \\
x^{r^2}\mathcal{P}(x) &= p(0)x^{r^2} + p(1)x^{2+r^2} + \cdots + p\left(\frac{q-r^2}{2}\right) x^q + \cdots
\end{aligned}$$

This implies that $f_r(x)$ agrees with $x^{r^2}\mathcal{P}(x)$ on all its coefficients, so the power series are equal. $\square$

From here, a proof of Euler's theorem takes one line: set $(x, z) = (q^{3/2}, -q^{-1/2})$ to get

$$\sum_{r \in \mathbb{Z}} (-1)^r q^{(3r^2-r)/2} = \prod_{n \geq 1}(1-q^{3n-2})(1-q^{3n-1})(1-q^{3n}) = \prod_{n \geq 1}(1-q^n).$$

5 May 14

This is a summary of what we did on the day that half the class was gone for AP English.

Theorem 5.1. For a nonnegative integer n, let b_n be the number of ways to partition n into powers of two such that each power of two appears at most once (note that $b_0 = 1$). Then, the sequence

$$\frac{b_1}{b_0}, \frac{b_2}{b_1}, \frac{b_3}{b_2}, \frac{b_4}{b_3}, \frac{b_5}{b_4}, \ldots$$

contains exactly one copy of each positive rational number.

Remark 5.2. The b_i are known as the **hyperbinary numbers**, and so we will call the partitions counted by b_i **hyperbinary partitions**.

To me, at least, this is a very pretty theorem: it gives us an explicit enumeration of all the rational numbers! As an immediate consequence, we can conclude that $\mathbb{Q}$ is countable.

The theorem itself is somewhat artificial; we will focus on determining the structure of the b_i and the b_{i+1}/b_i sequences, and then the desired result will fall out with a little bit of prodding. In this vein, we wish to find a recursion for b_n.

Lemma 5.3. For nonnegative integers n,

$$b_{2n+1} = b_n, \quad b_{2n+2} = b_n + b_{n+1}.$$

Proof 1. In fact, both equations are proved in the same way. The main idea is to create a pairing of hyperbinary partitions by doing the following:

Given a hyperbinary partition, delete all its 1s and then divide each part by two.

Clearly, this operation is reversible.

Since hyperbinary partitions of $2n+1$ always have one 1, the operation pairs each of these partitions with one of n and vice versa, proving the first equation.

Similarly, hyperbinary partitions of $2n+2$ have either zero or two 1s, so the operation pairs each one with either a partition of n or $n+1$, which gives the second equation. □

Proof 2. Let $B(x) = \sum_{n\geq 0} b_n x^n$ be the generating function for b_n. We then leverage the self-similar structure of $B(x)$ by writing

$$\sum_{n\geq 0} b_n x^n = \prod_{k\geq 0}(1 + x^{2^k} + x^{2^{k+1}}) = (1+x+x^2)B(x^2) = (1+x+x^2)\sum_{m\geq 0} b_m x^{2m}.$$

Matching coefficients, we see that x^{2n+1} can only be generated on the RHS via $b_n x^{2n} \cdot x$, so $b_n = b_{2n+1}$; and x^{2n+2} can only be generated by $b_n x^{2n} \cdot x^2$ or $b_{n+1}x^{2n+2} \cdot 1$, so $b_{2n+2} = b_n + b_{n+1}$. □

Remark 5.4. Convince yourself that the generating function argument is really just the original combinatorial argument in disguise.

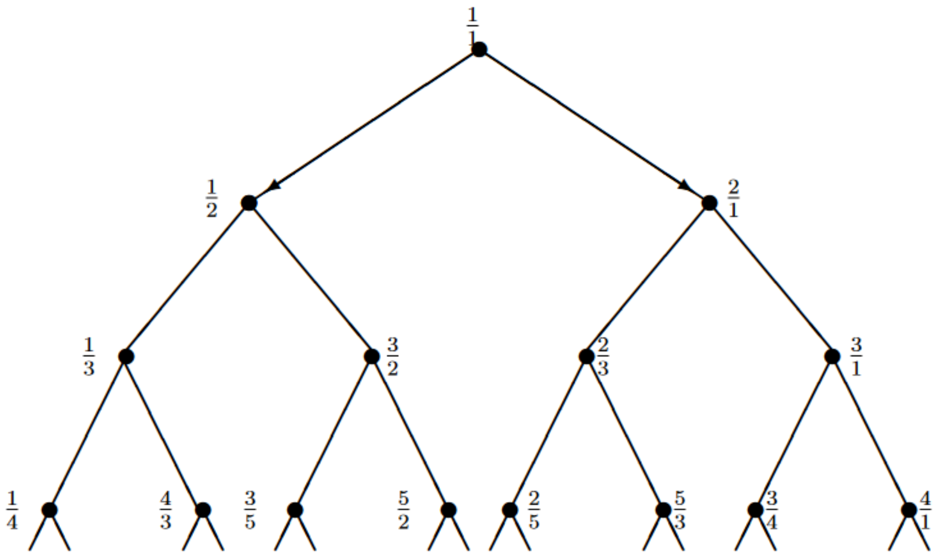
This means

$$\frac{b_{2n+2}}{b_{2n+1}} = \frac{b_n + b_{n+1}}{b_n}, \quad \frac{b_{2n+3}}{b_{2n+2}} = \frac{b_{n+1}}{b_n + b_{n+1}}.$$

In particular, if we know b_n and b_{n+1}, we can compute the two values above. This motivates putting our fractions in a rooted binary tree, defined more precisely as follows:

- The root node is $\frac{1}{1}$.
- The two children of $\frac{i}{j}$ are $\frac{i+j}{j}$ and $\frac{i}{i+j}$ — we say the former is the **right child**. and the latter is the **left child** (this is very obviously reflected pictorially).

Here's a picture:



Now, we want to show that each $\frac{a}{b}$ (where a, b are relatively prime positive integers) appears exactly once in this tree. If $\frac{a}{b}$ were indeed the child of some node, then we know whether it is a left child or a right child based on its size, since left children are always smaller than 1 while right children are always greater than 1. Thus, we should be able to trace back the pathway that generated $\frac{a}{b}$. In particular, we generate a sequence of fractions

$$\frac{a}{b}, \frac{a_1}{b_1}, \frac{a_2}{b_2}, \frac{a_3}{b_3}, \dots$$

where

$$a_i = a_{i-1} \text{ and } b_i = b_{i-1} - a_{i-1}$$

if $b_{i-1} > a_{i-1}$ or

$$a_i = a_{i-1} - b_{i-1} \text{ and } b_i = b_{i-1}$$

if $a_{i-1} > b_{i-1}$. You should recognize this process as the Euclidean algorithm! Since we have preordained $\frac{a}{b}$ to be in simplest form, the sequence of fractions will always terminate with $\frac{1}{1}$, and hence following the fractions in the opposite order gives a pathway towards $\frac{a}{b}$. Moreover, this pathway is unique due to the structure of our tree — at no point did we have a choice in which direction to trace backwards.

As an example, if we want to locate $\frac{5}{2}$ in the tree, then the corresponding sequence of fractions is

$$\frac{5}{2}, \frac{3}{2}, \frac{1}{2}, \frac{1}{1},$$

and we see in the above picture that

$$\frac{1}{1} \to \frac{1}{2} \to \frac{3}{2} \to \frac{5}{2}$$

is indeed a path in this graph.

6 References

[1] Andrews, G. E. *The Theory of Partitions*. Cambridge University Press, 1984, pp. 1–53.

[2] Stanley, R. P. *Enumerative Combinatorics, Volume 1*, second edition. Wadsworth & Brooks/Cole, 1986, pp. 68–78.

[3] Wilf, H. S. "Lectures on Integer Partitions". 2000. https://www2.math.upenn.edu/~wilf/PIMS/PIMSLectures.pdf.

[4] Sudler, C. "Two Enumerative Proofs of an Identity of Jacobi." *Proceedings of the Edinburgh Mathematical Society* 15.1 (1966): 67–71.