The Pentagonal Number Theorem and All That

Dick Koch

August 26, 2016

1 Introduction

When I was a sophomore, my advanced calculus midterm was given in Memorial Hall, Harvard’s grotesque memorial to students who died in the Civil War. The interior of the building is a room shaped like a cathedral, with a stained glass window at the front. But where the pews should be there are instead gigantic tables. College midterms are given in that room, with a table reserved for each class. Over there — philosophy, and there — German, and at our table — advanced calculus.

Our exam had four questions. At the end of the hour I realized that I was not going to answer any question. The saving grace was that the other students had the same problem. At the next class meeting, our professor David Widder said “You don’t know anything! You don’t even know the series expansion of $\frac{1}{1-x}$.”

Since then, I have known that series cold. Wake me from sleep and I can recite it.

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots$$

I’m giving this lecture in thanks to David Widder.

2 Euler

Newton and Leibniz invented calculus in the seventeenth century, and their immediate successors the Bernoulli’s worked at the start of the eighteenth century. But most of that century was dominated by a single man, Leonard Euler. Euler laid the foundations for mathematics of later centuries, and he is remembered with particular fondness as the master of beautiful formulas. Today I’m going to show you one of his most spectacular discoveries.
3 Partitions

A partition of a number \( n \) is a representation of \( n \) as a sum of positive integers. Order does not matter. For instance, there are 5 partitions of 4: 4, 3 + 1, 2 + 2, 2 + 1 + 1, 1 + 1 + 1 + 1.

Let \( p_n \) be the number of partitions of \( n \). Easily, \( p_1 = 1, p_2 = 2, p_3 = 3, p_4 = 5, p_5 = 7 \). Unfortunately, there is no formula for \( p_n \), and just writing the possibilities down and counting is not a good idea because \( p_n \) gets large fast: there are almost four trillion partitions of 200.

Euler discovered, however, an indirect way to compute \( p_n \). I’ll describe his method, and use it to show that \( p_{1000} = 24,061,467,864,032,622,473,692,149,727,991 \).

4 Euler’s First Formula

Euler’s technique proceeds in two steps. The first allows us to compute the \( p_n \), but slowly. The second dramatically speeds up the process. Here is his first formula:

Theorem 1

\[
(1 + x + x^2 + \ldots)(1 + x^2 + x^4 + \ldots)(1 + x^3 + x^6 + \ldots)(\ldots) = \sum p_n x^n
\]

Proof: Recall the distributive law, which I like to call the “Chinese menu formula”; to multiply \((a+b)(c+d)\), choose one of \( a \) and \( b \) from column A, and one of \( c \) and \( d \) from column B and multiply them, and then you add up all the possibilities, giving \( ac + ad + bc + bd \).

This works for more complicated products as well. To compute the product

\[
(1 + x + x^2 + \ldots)(1 + x^2 + x^4 + \ldots)(1 + x^3 + x^6 + \ldots)(\ldots)
\]

we choose \( x^{k_1} \) from column A, \( x^{2k_2} \) from column B, \( x^{3k_3} \) from column C, etc., and multiply them to obtain \( x^{k_1+2k_2+3k_3+\ldots+sk_s} \), and then add up the possibilities to get

\[
\sum_{k_1,k_2,\ldots} x^{k_1+2k_2+3k_3+\ldots+sk_s}
\]

The result is a sum of powers of \( x \),

\[
\sum_{k_1,k_2,\ldots} x^n
\]

but the term \( x^n \) will occur as many times as we can write \( n = k_1 + 2k_2 + 3k_3 + \ldots + sk_s \). However, such an expression is just a fancy way to write \( n \) as a partition, namely as a sum of \( k_1 \) 1’s, \( k_2 \) 2’s, \( k_3 \) 3’s, etc. So the final sum is \( \sum p_n x^n \). QED.
Euler’s first formula describes a way to organize a computation of \( p_n \). This method can also be described in a manner that doesn’t use algebra. Notice that the term \( 1 + x + x^2 + \ldots \) in Euler’s product counts partitions containing only 1; each integer can be written as such a sum in only one way. The product

\[
(1 + x + x^2 + \ldots)(1 + x^2 + x^4 + \ldots)
\]

counts partitions containing only 1’s and 2’s and thus equals

\[
1 + x + 2x^2 + 2x^3 + 3x^4 + 3x^5 + 4x^6 + \ldots
\]

because 0 and 1 have no extra partitions with 2’s, 2 and 3 have one additional partition with 2’s, 4 and 5 have two additional partitions with 2’s, etc. (It is useful to think of zero as having exactly one partition, the partition with no 1’s, no 2’s, no 3’s, etc. So we sometimes write \( p_0 = 1 \).)

In a similar way we can count partitions using 1’s, 2’s, and 3’s, and then partitions using 1’s, 2’s, 3’s, and 4’s, and so forth. Let’s consider one of these cases in detail. It turns out that

\[
(1 + x + x^2 + \ldots)(1 + x^2 + x^4 + \ldots)(1 + x^3 + x^6 + \ldots)(1 + x^4 + x^8 + \ldots) =
\]

\[
1 + x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + 12x^6 + 18x^7 + 24x^8 + 33x^9 + 46x^{10} + \ldots
\]

where the coefficient of \( x^n \) counts partitions of \( n \) containing only 1’s, 2’s, 3’s, and 4’s. Suppose we now want to count partitions of \( n \) containing only 1’s, 2’s, 3’s, 4’s, and 5’s. Let’s concentrate on the case \( n = 13 \). From the above product we see that there are 39 such partitions with no 5’s. A partition containing exactly one 5 will contain 1’s, 2’s, 3’s and 4’s adding up to \( 13 - 5 = 8 \), and the above product shows that there are 15 such partitions. A partition of 13 containing exactly two 5’s will contain 1’s, 2’s, 3’s, and 4’s adding up to \( 13 - 5 - 5 = 3 \) and there are 3 such partitions. There are no partitions of 13 with three or more 5’s. So the total number of partitions of 13 with 1’s, 2’s, 3’s, 4’s, and 5’s is 39 + 15 + 3 = 57.

Euler’s first formula is just a fancy way to summarize this technique. To count all partitions of \( n \), first count all partitions of all \( k \leq n \) containing only 1’s, and then count all partitions of all \( k \leq n \) containing only 1’s and 2’s, and then count all partitions of all \( k \leq n \) containing only 1’s, 2’s, and 3’s, and continue in this way up to partitions containing only 1’s, 2’s, 3’s, \ldots, \( n \)’s. If \( a_0, a_1, a_2, \ldots \) are counts of partitions with summands less than \( t \), then the \( b_0, b_1, b_2, \ldots \) counting partitions with summands less than or equal to \( t \) are given by the formula

\[
b_k = a_k + a_{k-t} + a_{k-2t} + \ldots
\]
5 For Programmers

If you are a programmer, you’ll understand this much better by writing a program to do the calculation. Almost any language will do, but you have to remember that the numbers will get large and might overflow. I’ll write a program in Mathematica because I have it handy and it can deal with arbitrarily large integers.

```mathematica
F[limit_] :=
  Block[{N, f, i, j, k, list1, list2}, (* local variables *)
    N = limit + 1; (* number of series coefficients *)
    f[s_] := 1;
    list1 = Array[f, N]; (* series coefficients *)
    list2 = Array[f, N]; (* fill initial list with 1's *)
    k = 2;
    While[k <= limit,
      Print[" "]; Print["Partitians using 1 through ", k];
      For[i = 1, i <= N, i++,
        sum = list1[[i]]; 
        For[j = i - k, j > 0, j = j - k,
          sum = sum + list1[[j]]; 
        
        list2[[i]] = sum; 
      ];
      For[i = 1, i <= N, i++, list1[[i]] = list2[[i]]];
      For[i = 1, i <= N, i++, Print[i - 1, ": ", list2[[i]]]]; 
    k++;
  ];
```

This program takes 50 seconds to compute the first fifty values of $p_n$. In particular, $p_{50} = 204,226$.

However, Euler discovered a much faster method. It takes my computer a little over one second to compute the first fifty values of $p_n$ with Euler’s second method. I’ll explain his method in the next three sections.

6 Dealing with the Analysts

You old calculus teacher is probably whispering in your ear about convergence, rigor, and all that. We’re going to tell the analysts to shut up by defining their objection away.
Definition 1 Let
\[ U = \{ 1 + a_1 x + a_2 x^2 + \ldots \mid a_i \in \mathbb{Z} \} \]

Define a product on this set by writing
\[ (1 + a_1 x + a_2 x^2 + \ldots)(1 + b_1 x + b_2 x^2 + \ldots) = 1 + c_1 x + c_2 x^2 + \ldots \]
where \( c_k = a_k + a_{k-1} b_1 + a_{k-2} b_2 + \ldots + b_k \).

Remark: Thus an element of \( U \) is a formal power series, and no convergence is required. This multiplication is well-defined because we can compute any particular element of the product in a finite amount of time. I need to warn you that \( U \) is my personal notation, not something any mathematician would recognize. To me, \( U \) stands for “units.”

Theorem 2 The set \( U \) with this product is a group.

Proof: Only inverses are unclear. At first sight, it seems ridiculous to suppose that \( U \) is a group because
\[ \frac{1}{1 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots} \]
isn’t itself a power series. But you have to remember what the axiom really says. It says the series \( 1 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots \) has an inverse
\[ 1 + b_1 x + b_2 x^2 + b_3 x^3 + \ldots \]
such that the product of the two series is the identity:
\[ (1 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots)(1 + b_1 x + b_2 x^2 + b_3 x^3 + \ldots) = 1 \]
Computing this product, we want to find \( b_i \) such that
\[ 1 + (a_1 + b_1)x + (a_2 + a_1 b_1 + b_2)x^2 + (a_3 + a_2 b_1 + a_1 b_2 + b_3) + \ldots = 1 \]
Setting each coefficient of \( x^i \) to zero, we get a series of equations for the \( b_i \), which have a unique inductive solution:
\[ b_1 = -a_1 \]
\[ b_2 = -a_1 b_1 - a_2 \]
\[ b_3 = -a_1 b_2 - a_2 b_1 - a_3 \]
\[ \ldots \]
QED.

Remark: In particular, \( \frac{1}{1-x} = 1 + x + x^2 + \ldots \), as I learned from David Widder. Indeed
\[ (1 - x)(1 + x + x^2 + x^3 + \ldots) = (1 - x) + (x - x^2) + (x^2 - x^3) + (x^3 - x^4) + \ldots = 1 \]
7 Euler’s Second Formula

Theorem 3
\[
\frac{1}{(1-x)(1-x^2)(1-x^3)\ldots} = \sum p_n x^n
\]

Proof: This follows immediately from Euler’s first formula by taking inverses of the series on the left side.

8 The Pentagonal Number Theorem

After writing this formula, Euler multiplied out the denominator by hand, hoping to find a pattern. One of my sources says he multiplied the first fifty terms, while another says he multiplied as many as one hundred terms. Amazingly, he found that

Theorem 4 (The Pentagonal Number Theorem)
\[
(1 - x)(1 - x^2)(1 - x^3)\ldots = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - x^{35} - x^{40} + \ldots
\]

Notice that pairs of terms with minus signs alternate with pairs with positive signs. Notice that the intervals between the exponents of pairs with the same sign increase by one, then two, then three, etc. Notice that the intervals between exponents of pairs with opposite signs increase by three, then five, then seven, etc.

The Pentagonal Number Theorem leads to a rapid method of computing the partition numbers. Indeed rewriting theorem 4 using theorem 3 gives

\[
(1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + \ldots)(1 + p_1 x + p_2 x^2 + p_3 x^3 + \ldots) = 1
\]

Consequently the coefficient of \(x^n\) in the product is zero, and so

\[
p_n - p_{n-1} - p_{n-2} + p_{n-5} + p_{n-7} - p_{n-12} - p_{n-15} + \ldots = 0
\]

Each of these expressions is a finite sum because \(p_0 = 1\) and \(p_k = 0\) for negative \(k\) by definition. These formulas then allow us to compute the \(p_n\) inductively starting with the value \(p_0 = 1\). Thus \(p_1 - p_0 = 0, \text{ so } p_1 = 1\). Then \(p_2 - p_1 - p_0 = 0, \text{ so } p_2 = 2\). Etc.

Using this revised formula, it takes my computer only a second to find the first 50 values of \(p_n\) and only 50 seconds to find the first 1000 values. And indeed, as promised the computer gives \(p_{1000} = 24,061467,864,032,622,473,692,149,727,991\).

For the record, I’ll show the Mathematica program I used to do this computation. You can easily rewrite this program in your favorite language, but keep in mind that the integers computed by the program will be very large.
FPentagonal[limit_] :=
Block[{N, f, P, Pinverse, k, index, n, i}, (* local variables *)
N = limit + 1; (* number of series coefficients *)
f[s_] := 0;
P = Array[f, N]; (* the p(n), initially filled with zeros *)
P[[1]] = 1; (* p(0) = p(1) = 1; arrays in Mathematica are one-based *)
P[[2]] = 1;
Pinverse = Array[f, N]; (* inverse of p(n) series *)
k = 1; (* now fill in Pinverse using the Pentagonal Number Theorem *)
index = k (3 k - 1) / 2;
While[index <= N,
   Pinverse[[index]] = (-1)^k;
   index = k (3 k + 1) / 2;
   If[index <= N, Pinverse[[index]] = (-1)^k];
   k = k + 1;
   index = k (3 k - 1) / 2;
   ];
For[n = 2, n < N, n = n + 1, (* compute p(n) as inverse of Pinverse *)
P[[n]] = 0;
For[i = 1, i < n, i = i + 1,
P[[n]] = P[[n]] - P[[n - i]] Pinverse[[i]]
   ];
P[[n]] = P[[n]] - Pinverse[[n]];
Print[n, ": ", P[[n]]];
]

I don’t know how many values of \(p_n\) were computed by Euler. In 1918, MacMahon in England computed the first 200 values of \(p_n\). This table was extended to 600 by Gupta in 1935, and to 1000 by Gupta, Gwyther and Miller in 1958. I don’t know if a computer was used for this final table.

Why is the theorem called “the pentagonal number theorem”? The reason isn’t very interesting mathematically, but here it is. Exponents in the pentagonal series with nonzero coefficients have the form \(n = \frac{k(3k-1)}{2}\) and \(n = \frac{k(3k+1)}{2}\). The numbers \(\frac{k(3k-1)}{2}\) are called “pentagonal numbers” because they count the number of dots in a pentagonal pattern, just as the numbers \(n = \frac{k(k+1)}{2}\) and \(n = k^2\) are “triangular numbers” and “square numbers” because they count dots in triangular and square patterns.
Jordan Bell wrote an interesting paper on the history of the Pentagonal Number Theorem. It can be found at http://arxiv.org/pdf/math/0510054v2.

The first mention of the theorem is in a letter from Daniel Bernoulli to Euler on January 28, 1741. Bernoulli is replying to a (lost) letter from Euler about the expansion, and he writes “The other problem, to transform \((1 - x)(1 - x^2)(\ldots)\) into \(1 - x - x^2 + x^5 + \ldots\), follows easily by induction, if one multiplied many factors. The remainder of the series I do not see. This can be shown in a most pleasant investigation, together with tranquil pastime and the endurance of pertinacious labor, all three of which I lack.”

Euler mentions the theorem many more times over the next few years, in letters we do possess to Niklaus Bernoulli, Christian Goldbach, d’Alembert, and others, and in the first publication of 1751. (This paper was written on April 6, 1741 and had no proof. Euler wrote so many papers that the publishers fell dramatically behind; they were publishing new papers many years after his death.) A typical entry, from a letter to Goldbach, reads “If these factors \((1 - n)(1 - n^2)(1 - n^3)\) etc. are multiplied out onto infinity, the following series \(1 - n - n^2 + n^5 + n^7 - \ldots\) is produced. I have however not yet found a method by which I could prove the identity of these two expressions. The Hr. Prof. Niklaus Bernoulli has also been able to prove nothing beyond induction.” Here the word “induction” means “by experiment” rather than “a proof by induction”.

Euler is not above a little trickery. Learning that d’Alembert wanted to leave mathematical research to regain his health, he wrote him “If in your spare time you should wish to do some research which does not require much effort, I will take the liberty to propose the expression \((1 - x)(1 - x^2)(1 - x^3)(1 - x^4)\) etc., which upon expansion by multiplication gives the series \(1 - x - x^2 + x^5 + x^7 - \ldots\) etc., which would seem very remarkable to me because of the law which we easily discover within it, but I do not see how his law may be deduced without induction of the proposed expression.” Eventually d’Alembert wrote back “regarding the series of which you have spoken, it is very peculiar, but I only see induction to show it. But no one is deeper and better versed on such matters than you.”
Euler finally was able to prove the theorem on June 9, 1750, in a letter to Goldbach. His proof is algebraic. The proof was first published in 1760, and Euler gives more details about points which were vague in his letter to Goldbach. You can consult Bell’s paper if you want to follow this original Euler proof.

10 Franklin’s Proof

In 1881, the American mathematician Franklin gave a proof which involves no algebra at all. Hans Rademacher called this proof “the first major achievement of American mathematics.” Here is Franklin’s proof:

Proof: The basic idea is that the series \((1 - x)(1 - x^2)(1 - x^3)(\ldots)\) can be interpreted as a sophisticated count of a certain restricted type of partitions. Let us begin with the following formula, which I’ve obtained inductively by multiplying out terms:

\[
(1+x)(1+x^2)(1+x^3)(\ldots) = 1+x+x^2+2x^3+3x^4+4x^5+5x^6+6x^7+8x^8+10x^9+12x^{10}+\ldots
\]

Looking back at the proof of theorem 1, we see that the product on the left is equal to

\[
\sum_{k_1,k_2,\ldots} x^{k_1+2k_2+3k_3+\ldots+sk_s}
\]

but this time each \(k_i\) is either zero or one. This means that in the partition \(n = k_1 + 2k_2 + 3k_3 + \ldots + sk_s\), the number of 1’s is either zero or one, the number of 2’s is either zero or one, the number of 3’s is either zero or one, etc. Thus the product is equal to

\[
\sum q_n x^n
\]

where \(q_n\) counts partitions of \(n\) as a sum of distinct positive numbers. For example, the coefficient of \(x^7\) is 5 because there are only five partitions of 7 with distinct factors, namely 7, 6 + 1, 5 + 2, 4 + 3, and 4 + 2 + 1.

We aren’t quite interested in this series, but instead in

\[
(1-x)(1-x^2)(1-x^3)(\ldots) = \sum_{k_1,k_2,\ldots} (-1)^{k_1+k_2+\ldots+sk_s} x^{k_1+2k_2+3k_3+\ldots+sk_s}
\]

So this time when we count partitions with distinct summands, we count partitions with an even number of terms positively, but partitions with an odd number of terms negatively. The coefficient of \(x^n\) is thus “the number of distinct partitions of \(n\) with an even number of terms, minus the number of distinct partitions of \(n\) with an odd number of terms.”

According to the pentagonal number theorem, this product is

\[
1-x-x^2+x^5+x^7-x^{12}-x^{15}+x^{22}+x^{26}-x^{35}-x^{40}+\ldots
\]
Notice in particular that the coefficient of $x^n$ is usually zero. Franklin concentrated on that fact, and tried to understand why the number of distinct partitions of $n$ with an even number of terms is usually exactly the same as the number of distinct partitions of $n$ with an odd number of terms. And the explanation he gave is that you can pair up each even partition with a corresponding odd partition. For example, there are 12 partitions of 11 into distinct summands, and we will see that the appropriate pairing is

\begin{align*}
10 + 1 & \leftrightarrow 11 \\
9 + 2 & \leftrightarrow 8 + 2 + 1 \\
8 + 3 & \leftrightarrow 7 + 3 + 1 \\
7 + 4 & \leftrightarrow 6 + 4 + 1 \\
6 + 5 & \leftrightarrow 5 + 4 + 2 \\
5 + 3 + 2 + 1 & \leftrightarrow 6 + 3 + 2
\end{align*}

How is this pairing defined?

Franklin’s trick is exposed in the next picture. You could finish the proof without reading further by thinking carefully about this picture. Draw a distinct partition as a pattern of rows of dots; for instance the picture below corresponds to $20 = 7 + 6 + 4 + 3$. Concentrate on the last row, and on the largest diagonal that can be drawn at the right. The idea is to move the bottom row up to form a new diagonal, or move the diagonal down to form a new row. In the picture below, the bottom row cannot be moved up because that would leave a hanging dot, but the diagonal can be moved down. Notice that moving converts a partition with an even number of terms into a partition with an odd number of terms. In the case illustrated below, it converts $7 + 6 + 4 + 3$ into $6 + 5 + 4 + 3 + 2$.

What is the rule for moving? Say the bottom row has $a$ dots and the diagonal on the right has $b$ dots. If we want to move the bottom row up without getting a hanging dot, we need $a \leq b$. If we want to move the diagonal down and get a shorter final row, we need $a > b$. 

10
So when \( a \leq b \), we can move up, but not down, and when \( a > b \) we can move down but not up.

The one legal move for a given diagram produces a new diagram. This diagram also has a unique legal move. But certainly one thing we can do is to reverse the original move and return to the original diagram, so that must be the unique legal move. It follows that our diagrams are paired: the legal move for each leads to the other. Since the number of rows increases or decreases by one, even partitions are paired with odd partitions, as promised.

The only problem with this argument is that it seems to show that all of the coefficients of \((1 - x)(1 - x^2)(1 - x^3)(\ldots)\) are zero. The truth is that there are “edge cases” where the analysis just given doesn’t quite work. These edge cases occur when the row at the bottom and the diagonal strip on the right side share a common corner, as below.

Let us again say that the bottom row has \( a \) dots and the diagonal on the right has \( b \) dots. If we want to move the bottom row up, then the diagonal length will decrease by one, and so to avoid a hanging dot we need \( a \leq (b - 1) \). If we want to move the diagonal down and get a shorter final row, then the existing final row length before the motion will decrease by one and we need \((a - 1) > b \). So there are two troublesome cases where neither motion is legal: \( a = b \) and \( a = b + 1 \). Below are two samples, the first with \( a = b \) and the second with \( a = b + 1 \). Notice that neither motion is legal for these diagrams. The left diagram corresponds to an odd partition of 12 and the right corresponds to an odd partition of 15. Note that the pentagonal number theorem expansion contains \(-x^{12} - x^{15}\).
The first list below contains diagrams with $a = b$. The second list contains diagrams with $a = b + 1$. Each list extends infinitely to the right. The number of dots in diagrams on the first list is 1, 5, 12, 22, ... and the number of dots in diagrams on the second list is 2, 7, 15, 26, ... and these are exactly the exponents which occur in the pentagonal number expansion! The reader can check that the numbers from the first list equal $k(3k - 1)/2$ and the numbers from the second list equal $k(3k + 1)/2$. QED.

11 Ramanujan

Do you know the story of Ramanujan?

The following information comes from G. H. Hardy’s book on Ramanujan. Ramanujan was born in India, where his father was a clerk in a cloth-merchant’s office; all of his relatives were very poor. His mathematical abilities became apparent in school, and by the time he was thirteen he was recognized as a quite abnormal boy. In 1903 he passed the Matriculation Examination of the University of Madras and won the Subrahmanyam scholarship. But then he concentrated so heavily on mathematics that he failed to secure promotion to the senior class and his scholarship was discontinued. In 1906 he appeared as a private student for the F. A. examination, but failed. In 1909 he married, but had no definite occupation until 1912, when he became a clerk in he Port Trust in Madras, at a salary of thirty pounds a year. In 1913 he wrote to G. H. Hardy in Cambridge, listing around one hundred formulas he had discovered, with no proofs. I have listed some of those formulas at the end of these notes. Ramanujan had written to other English mathematicians earlier that year, making no impression. But Hardy, one of the greatest English mathematicians at that time, recognized his genius, and brought him to England. Here are some of Hardy’s comments about the formulas in his letter:
I thought that, as an expert in definite integrals, I could prove 1.6, and did so, though with a good deal more trouble than I had expected. The series formulas 1.1 - 1.4 I found much more intriguing, and it soon became obvious that Ramanujan must possess much more general theorems and was keeping a good deal up his sleeve. The second is a formula well known in the theory of Legendre series, but the others are much harder than they look.

The formulae 1.10 - 1.11 are on a different level and obviously both difficult and deep. Indeed 1.10 - 1.11 defeated me completely; I had never seen anything in the least like them. A single look at them is enough to show that they could only be written down by a mathematician of the highest class. They must be true because, if they were not true, no one would have had the imagination to invent them.

The last formulae stands apart because it is not right and shows Ramanujan’s limitations, but that does not prevent it from being additional evidence of his extraordinary powers. The function in 1.14 is a genuine approximation to the coefficient, thought not at all so close as Ramanujan imagined, and Ramanujan’s false statement is one of the most fruitful he ever made, since it ended by leading us to all our joint work on partitions.

Ramanujan is often considered the greatest inventor of beautiful formulas since Euler. But his mathematical training was spotty. So Hardy was faced with the task of filling the mathematical gaps in his education without destroying his natural gifts. Ramanujan died young, in 1920.

Ramanujan was in England when Macmahon computed the first 200 values of $p_n$ using Euler’s formula. Looking at this table, Ramanujan noticed some remarkable patterns. I’ve printed the table of the first 200 values on the last pages. Before reading further, look at the table and see if you can find obvious patterns.

(As a hint, look at cases when $p_n$ is divisible by 5. There are some distracting values when a term not part of the pattern is divisible by 5, and you’ll have to ignore the distractions. If you find a pattern, congratulations, but keep looking. Ramanujan found two other patterns using other divisors, plus more subtle patterns I won’t discuss.)

Notice that $p_4 = 5$. Ramanujan noticed that every fifth entry in the table after that was also divisible by 5. In modern notation

$$p_{5m+4} \equiv 0 \pmod{5}$$

Notice that $p_5 = 7$. Ramanujan also noticed that every seventh entry after that was also divisible by 7. Similarly $p_6 = 11$ and Ramanujan noticed that every eleventh entry after that was also divisible by 11. So
\[ p_{7m+5} \equiv 0 \pmod{7} \]
\[ p_{11m+6} \equiv 0 \pmod{11} \]

But don’t get your hopes up. The generalization to 13 fails.

Eventually Ramanujan managed to prove all three observations. The proofs are not easy, and the result for 11 is decidedly harder than the results for 5 and 7. There are now many different proofs, some involving remarkable variations on the pentagonal number theorem. For example, one proof of the mod 5 congruence depends on proving the following amazing formula asserted (but not proved) by Ramanujan:

**Theorem 5** Let \[ f(x) = (1 - x)(1 - x^2)(1 - x^3)(\ldots) = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + \ldots \]

Then

\[ \sum p_{5n+4} x^n = 5 \frac{f(x^5)^5}{f(x)^6} \]

Remark: Notice that \( f(x) \in \mathcal{U} \) and \( f(x^5) \in \mathcal{U} \). So the quotient on the right is a formal series with integer coefficients. After multiplying by 5, all coefficients are divisible by 5.

Remark: Similarly the mod 7 congruence follows from the following Ramanujan formula:

\[ \sum p_{7n+5} x^n = 7 \frac{f(x^7)^3}{f(x)^4} + 49x \frac{f(x^7)^7}{f(x)^8} \]

12 The Size of \( p_n \)

Hardy and Ramanujan also provided an estimate for the size of \( p_n \). The argument, developed further by Hardy, Littlewood, and Rademacher, uses what is called the circle method and involves relating the pentagonal number series with the theory of modular forms.

**Theorem 6**

\[ p_n \sim \frac{1}{4n\sqrt{3}} e^{\sqrt{\frac{2n}{3}}} \]

*Examples:* When \( n = 200 \), this approximation is \( 4.10025 \times 10^{12} \), compared to the exact value of \( 3.97299 \times 10^{12} \). When \( n = 1000 \), the approximation is \( 2.4402 \times 10^{31} \) and the actual value is \( 2.4061 \times 10^{31} \). When \( n = 10000 \), the approximation is \( 3.6328 \times 10^{106} \), far larger than the expected number of atoms in the universe, and \( p_{10000} \) is thus a number with around 106 digits.
13 Freeman Dyson

In 1944, an Oxford undergraduate named Freeman Dyson wrote a paper on Ramanujan’s congruences. Dyson switched to physics in graduate school and became one of the most important physicists of the twentieth century. A few weeks ago, he gave a talk in Eugene.

Dyson’s undergraduate paper sketched an alternate proof of Ramanujan’s mod 5 result, although he couldn’t prove the central observation he made (it was proved ten years later by Atkin and Swinnerton-Dyer.) Suppose we have a partition $\lambda_1 + \lambda_2 + \ldots + \lambda_k$ with $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_5$. Dyson defined the rank of the partition to be $\lambda_k - k$. For a fixed $n$, Dyson divided all partitions of $5n + 4$ into five classes, namely those with rank divisible by 5, those with rank congruent to one modulo 5, those congruent to two modulo 5, ..., and those congruent to four modulo five. Purely empiracally, Dyson noticed that these five classes always have the same number of elements, and consequently the total number of partitions of $5n + 4$ must be divisible by five. The same empirical method worked for Ramanujan’s mod 7 result.

But curiously, Dyson noticed that his method did not work in the mod 11 case. So he conjectured in his paper that there was a better function from partitions to integers, replacing the rank, and working for all of 5, 7, and 11. Dyson named this unknown function the crank, and he wrote “Whatever the final verdict of posterity may be, I believe the ‘crank’ is unique among arithmetical functions in having been named before it was discovered. May it be preserved from the ignominious fate of the planet Vulcan!” Forty years later, Andrews and Garvan discovered just such a crank function.

14 Ken Ono

John Leahy is the Oregon mathematician who represented the University of Oregon at Bend when Cascade College was founded. John and George Andrews (one of the discoverers of the crank) were fellow graduate students at the University of Pennsylvania in the 1960’s. Andrews later became one of the world’s experts on Ramanujan. When Andrews visited John in Eugene, he knew fascinating stories about Ramanujan and his notebooks.

I met John at Penn, and we knew a faculty member there named Takashi Ono. Ono’s son, Ken Ono, also became a mathematician; I am unhappy to report that Ken hadn’t been born when John and I were in Philadelphia. Ken left college early to race bicycles as a member of the Pepsi-Miyata cycling team, and to become a graduate student at Chicago.

For most of the twentieth century, mathematicians working on partitions tried to understand the special significance of the primes 5, 7, and 11. Thus it came as an enormous
surprise when Ken Ono proved in 2001 that there are similar congruence relations for *every prime greater than 3*. Indeed, in 2006, Ono proved that whenever $N$ is an integer whose prime factorization does not contain 2 or 3, then there are integers $a$ and $b$ such that

$$p_{am+b} \equiv 0 \pmod{N}$$

for all $m$. This is by far the greatest advance in the theory since Ramanujan.

The status of the primes 2 and 3 is still unknown. In particular, nobody knows how to predict when $p_n$ is even, or even how to predict an infinite number of cases when $p_n$ is even.

Why didn’t Ramanujan discover these additional congruences, and why weren’t they found in the century after Ramanujan? You’ll understand as soon as I show you the next simplest cases, for the primes 13 and 17:

$$p_{17303m+237} \equiv 0 \pmod{13}$$

$$p_{48037937m+1122838} \equiv 0 \pmod{17}$$

It is just possible that you’d notice that 13 divides $p_{237}$ from the tables. But what are the chances that you’d notice that it also divides $p_{237+17303} = p_{17440}$? And if you noticed that, what are the chances that you’d leap to the conclusion that after another 17303 steps, it would divide $p_{17540+17303}$, and so on forever?
Formulas from Ramanujan’s Letter to Hardy

1.1
\[ (1 - \frac{3!}{(1!)^3} x^2 + \frac{6!}{(2!)^3} x^4 - \ldots ) (1 - \frac{x}{(1!)^3} + \frac{x^2}{(2!)^3} + \ldots ) \]

1.2
\[ 1 - 5 \left( \frac{1}{2} \right)^3 + 9 \left( \frac{1 \cdot 3}{2 \cdot 4} \right)^3 - 13 \left( \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \right)^3 + \ldots = \frac{2}{\pi} \]

1.3
\[ 1 + 9 \left( \frac{1}{4} \right)^4 + 17 \left( \frac{1 \cdot 5}{4 \cdot 8} \right)^4 + 25 \left( \frac{1 \cdot 5 \cdot 9}{4 \cdot 8 \cdot 12} \right)^4 + \ldots = \frac{2^3}{\pi^2 \left\{ \Gamma \left( \frac{3}{4} \right) \right\}^2} \]

1.4
\[ 1 - 5 \left( \frac{1}{2} \right)^5 + 9 \left( \frac{1 \cdot 3}{2 \cdot 4} \right)^5 - 13 \left( \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \right)^5 + \ldots = \frac{2}{\left\{ \Gamma \left( \frac{3}{4} \right) \right\}^4} \]

1.6
\[ \int_0^\infty \frac{dx}{(1+x^2)(1+r^2x^2)(1+r^4x^2+\ldots)} = \frac{\pi}{2(1+r+r^3+r^6+r^{10}+\ldots)} \]

1.10
If \( u = \frac{x}{1+\frac{x^5}{1+\frac{x^{10}}{1+\ldots}}} \), \( v = \frac{x}{1+\frac{x^2}{1+\frac{x^3}{1+\ldots}}} \), then \( v^5 = u \frac{1-2u+3u^2-u^3+u^4}{1+3u+6u^2+2u^3+u^4} \)

1.11
\[ \frac{1}{1+\frac{e^{-2\pi}}{1+\ldots}} = e^{-4\pi} \left\{ \sqrt{\frac{5+\sqrt{5}}{2}} - \frac{\sqrt{5}+1}{2} \right\} e^{\frac{2}{3} \pi} \]

1.14
The coefficient of \( x^n \) in \( (1 - 2x + 2x^4 - 2x^9 + \ldots)^{-1} \) is the integer nearest to
\[ \frac{1}{4n} \left( \cosh \pi \sqrt{n} - \frac{\sinh \pi \sqrt{n}}{\pi \sqrt{n}} \right) \]
### Partition Table

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26: 2436</td>
<td>51: 239943</td>
<td>76: 9289091</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>27: 3010</td>
<td>52: 281589</td>
<td>77: 10619863</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>28: 3718</td>
<td>53: 329931</td>
<td>78: 12132164</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>29: 4565</td>
<td>54: 386155</td>
<td>79: 13848650</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>30: 5604</td>
<td>55: 451276</td>
<td>80: 15796476</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>31: 6842</td>
<td>56: 526823</td>
<td>81: 18004327</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>32: 8349</td>
<td>57: 614154</td>
<td>82: 20506255</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>33: 10143</td>
<td>58: 715220</td>
<td>83: 23338469</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>34: 12310</td>
<td>59: 831820</td>
<td>84: 26543660</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>35: 14883</td>
<td>60: 966467</td>
<td>85: 30167357</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>36: 17977</td>
<td>61: 1121505</td>
<td>86: 34262962</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>37: 21637</td>
<td>62: 1300156</td>
<td>87: 38887673</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>38: 26015</td>
<td>63: 1505499</td>
<td>88: 44108109</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>39: 31185</td>
<td>64: 1741630</td>
<td>89: 49995925</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>40: 37338</td>
<td>65: 2012558</td>
<td>90: 56634173</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>41: 44583</td>
<td>66: 2323520</td>
<td>91: 64112359</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>42: 53174</td>
<td>67: 2679689</td>
<td>92: 72533807</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>43: 63261</td>
<td>68: 3087735</td>
<td>93: 82010177</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>44: 75175</td>
<td>69: 3554345</td>
<td>94: 92669720</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>45: 89134</td>
<td>70: 4087968</td>
<td>95: 104651419</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>46: 105558</td>
<td>71: 4697205</td>
<td>96: 118114304</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>47: 124754</td>
<td>72: 5392783</td>
<td>97: 133230930</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>48: 147273</td>
<td>73: 6185689</td>
<td>98: 150198136</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>49: 173525</td>
<td>74: 7089500</td>
<td>99: 169229875</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>50: 204226</td>
<td>75: 8118264</td>
<td>100: 190569292</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>125: 3163127352</td>
<td>150: 40853235313</td>
<td>175: 435157697830</td>
<td>200: 3972999029388</td>
<td></td>
</tr>
</tbody>
</table>