Bell and Stirling Numbers

Recall that the partition function $p(n)$ counts the number of ways of partitioning n identical objects into nonempty piles; for example $p(4) = 5$ since (4), $(3, 1)$, $(2, 2)$, $(2, 1, 1)$, $(1, 1, 1, 1)$ are the five partitions of 4. Also $p_k(n)$ counts the number of ways of partitioning n identical objects into k non-empty piles; also, the number of partitions of n into any number of nonemtpy piles of maximum size n. For example, $p_3(7) = 4$ since $(5, 1, 1)$, $(4, 2, 1)$, $(3, 3, 1)$, $(3, 2, 2)$ are the four ways to partition 7 into three nonempty parts; also $(3, 3, 1)$, $(3, 2, 2)$, $(3, 2, 1, 1)$, $(3, 1, 1, 1, 1)$ are the four ways to partition 7 into nonempty parts of maximum size 3. What if instead we want to partition n distinct (i.e. distinguishable objects, like students or different books) instead of n indistinguishable objects? For this we must use instead the Bell numbers B_n and the Stirling numbers $\left\{ {n \atop n} \right\}$. (Actually, Stirling numbers come in two kinds, and the values $\left\{ {n \atop k} \right\}$ are known as *Stirling numbers of the second kind*; but we will not be considering the Stirling numbers of the first kind as they are less important.)

Bell Numbers

The number of ways to partition a set of n distinct objects into nonempty parts is the Bell number B_n. The sequence of Bell numbers is given by $B_n = 1, 1, 2, 5, 15, 52, 203, 877, \ldots$ for $n = 0, 1, 2, 3, 4, 5, 6, 7, \ldots$. We illustrate $B_4 = 15$ by counting partitions of a set of four students A,B,C,D. To make sure we don't miss any, we use the fact that $p(4) = 5$ and we enumerate the partitions of {A,B,C,D} based on the five partitions of 4, thus:

Shape (4): just one partition {{A,B,C,D}};

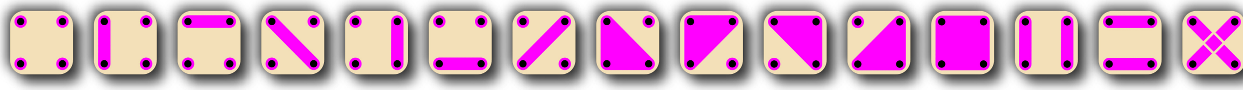
Shape (3,1): four partitions {{A,B,C}, {D}}, {{A,B,D}, {C}}, {{A,C,D}, {B}}, {{B,C,D}, {A}};

Shape (2,2): three partitions {{A,B}, {C,D}}, {{A,C}, {B,D}}, {{A,D}, {B,C}};

Shape (2,1,1): six partitions {{A,B}, {C}, {D}}, {{A,C}, {B}, {D}}, {{A,D}, {B}, {C}}, {{B,C}, {A}, {D}}, {{B,D}, {A}, {C}}, {{C,D}, {A}, {B}}; and

Shape (1,1,1,1): just one partition {{A}, {B}, {C}, {D}}.

The partitions of an n-set may be graphically depicted by arranging n dots on a circle; and for each of the B_n partitions of the dots, we color in each part (using the convex hull of each part, as it is known). As an example, we illustrate $B_4 = 15$ this way:



The last partition shown in this list is an example of a *crossing partiton* (since two parts of the partition cross each other in our picture). For larger values of $n>4$, the number of crossing partitions grows quickly; and for $n<4$, all the partitions of $[n]$ are *non-crossing partitions*. It turns out that the number of non-crossing partitions is given by the sequence of Catalan numbers $C_n=\frac{1}{n+1}\binom{2n}{n}=1,1,2,5,14,42,\ldots$. These claims explain the evident similarity between the two sequences, as well as the inequality $C_n\leqslant B_n$.

For larger values of n, the fastest way to generate the sequence by hand is recursively, using the following:

Theorem. The sequence of Bell numbers is defined recursively by $B_0=1$ and $B_{n+1}=\sum\limits_{k=0}^{n}\binom{n}{k}B_k$ for all $n\geqslant 0$.

Proof. Note that B_{n+1} is the number of partitions of $[n{+}1]=\{1,2,\ldots,n,n{+}1\}$. Any partition of $[n{+}1]$ has a part containing $n{+}1$. This part may be denoted by $A\sqcup\{n{+}1\}$ where A is an arbitrary subset of $[n]$. This part has size $k{+}1$; and the remaining parts of the partition have size adding up to $n{-}k$. Since there are $\binom{n}{k}$ ways to choose the k-subset $A\subseteq[n]$, and B_{n-k} ways to partition the remaining $n{-}k$ elements, we have

$$B_{n+1}=\sum_{k=0}^{n}\binom{n}{k}B_{n-k}=\sum_{\ell=0}^{n}\binom{n}{n-\ell}B_\ell=\sum_{\ell=0}^{n}\binom{n}{\ell}B_\ell\,.$$

Here we have substituted $\ell=n-k$ and used the identity $\binom{n}{n-k}=\binom{n}{k}$. □

For example, $B_4=\binom{3}{0}B_0+\binom{3}{1}B_1+\binom{3}{2}B_3+\binom{3}{3}B_3=1{\cdot}1+3{\cdot}1+3{\cdot}2+1{\cdot}5=15$.

Another way to generate the sequence of Bell numbers is to use its exponential generating function, which in this case is much preferable to the ordinary generating function. Just as the *ordinary generating function* of a sequence $a_0,a_1,a_2,\ldots$ is defined by $\sum_{n=0}^{\infty}a_nx^n$, the *exponential generating function* of $a_0,a_1,a_2,\ldots$ is defined by $\sum_{n=0}^{\infty}\frac{a_n}{n!}x^n$. For example, the constant sequence $1,1,1,1,\ldots$ has ordinary generating function

$$1+x+x^2+x^3+x^4+\cdots=\frac{1}{1-x}\,,$$

whereas its exponential generating function is

$$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots=e^x.$$

Theorem. The sequence of Bell numbers has exponential generating function

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

The generating function $F(x) = e^{e^x-1}$ may come as a surprise at first; but it arises in this context because it satisfies the differential equation $F'(x) = e^x F(x) = e^{e^x-1}e^x$; it is in fact the unique solution of this differential equation which also satisfies the initial condition $F(0) = 1$. Using the recurrence formula for the Bell numbers, it is not hard to show that its exponential generating function satisfies this characteristic property.

Proof. Let $B(x) = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$. Then

$$\begin{aligned} B'(x) &= \sum_{n=1}^{\infty} B_n \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} B_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} B_{n+1} \frac{x^n}{n!} \\ &= \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \binom{n}{k} B_k \Big] \frac{x^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} B_k \frac{x^n}{k!(n-k)!} \,. \end{aligned}$$

Here we are summing over all pairs of incides $\{(n,k) : n \geqslant k \geqslant 0\} = \{(k+\ell, k) : k, \ell \geqslant 0\}$, so

$$\begin{aligned} B'(x) &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} B_k \frac{x^{k+\ell}}{k!\ell!} = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} B_k \frac{x^k}{k!} \cdot \frac{x^\ell}{\ell!} \\ &= \Big[\sum_{k=0}^{\infty} B_k \frac{x^k}{k!} \Big] \Big[\sum_{\ell=0}^{\infty} \frac{x^\ell}{\ell!} \Big] = B(x) e^x. \end{aligned}$$

Thus $B(x)$ satisfies the same differential equation as $F(x) = e^{e^x-1}$. It also satisfies the initial condition $B'(0) = B_0 = 1$. So $B(x) = F(x)$. □

We demonstrate using Maple® to generate the first few terms of the Bell sequence using the recurrence formula:

Construct the first 20 (actually, 21) terms of the Bell sequence recursively

```
> B:=array(0..20):
> B[0]:=1;
```

$$B_0 := 1$$

```
> for n from 0 to 19 do
>    B[n+1]:=sum(binomial(n,k)*B[k],k=0..n): od:
> seq(B[n],n=0..20);
```

1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975, 678570, 4213597, 27644437, 190899322, 1382958545, 10480142147, 82864869804, 682076806159, 5832742205057, 51724158235372

One can instead look up the builtin command in Maple®:

There is also a black box in Maple for generating the Bell numbers, which you can find by using the Help feature.

```
> with(combinat):
> seq(bell(n),n=0..20);
```

1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975, 678570, 4213597, 27644437, 190899322, 1382958545, 10480142147, 82864869804, 682076806159, 5832742205057, 51724158235372

The Bell numbers can also be generated from the coefficients of its generating function:

Find the first 21 Bell numbers using the exponential generating function. First find the coefficients in the power series, then multiply by $n!$ to obtain B_n

```
> exp(exp(x)-1);
```

$$e^{e^x - 1}$$

```
> B:=series(%,x=0,21);
```

$$B := 1 + x + x^2 + \frac{5}{6}x^3 + \frac{5}{8}x^4 + \frac{13}{30}x^5 + \frac{203}{720}x^6 + \frac{877}{5040}x^7 + \frac{23}{224}x^8 + \frac{1007}{17280}x^9 + \frac{4639}{145152}x^{10} + \frac{22619}{1330560}x^{11} + \frac{4213597}{479001600}x^{12} + \frac{27644437}{6227020800}x^{13} + \frac{95449661}{43589145600}x^{14} + \frac{276591709}{261534873600}x^{15} + \frac{10480142147}{20922789888000}x^{16} + \frac{255755771}{1097800704000}x^{17} + \frac{97439543737}{914624815104000}x^{18} + \frac{5832742205057}{121645100408832000}x^{19} + \frac{263898766507}{12412765347840000}x^{20} + O(x^{21})$$

```
> seq(coeff(B,x,n),n=0..20);
```

$$1, 1, 1, \frac{5}{6}, \frac{5}{8}, \frac{13}{30}, \frac{203}{720}, \frac{877}{5040}, \frac{23}{224}, \frac{1007}{17280}, \frac{4639}{145152}, \frac{22619}{1330560}, \frac{4213597}{479001600}, \frac{27644437}{6227020800}, \frac{95449661}{43589145600}, \frac{276591709}{261534873600}, \frac{10480142147}{20922789888000}, \frac{255755771}{1097800704000}, \frac{97439543737}{914624815104000}, \frac{5832742205057}{121645100408832000}, \frac{263898766507}{12412765347840000}$$

```
> seq(factorial(n)*coeff(B,x,n),n=0..20);
```

1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975, 678570, 4213597, 27644437, 190899322, 1382958545, 10480142147, 82864869804, 682076806159, 5832742205057, 51724158235372

Stirling Numbers

For any set of n distinct objects, in particular the standard n-set $[n]$, the number of partitions of n into k nonempty subsets is denoted by the Stirling number $\left\{ n \atop k \right\}$. Note from the example above that

$$B_4 = \left\{ 4 \atop 0 \right\} + \left\{ 4 \atop 1 \right\} + \left\{ 4 \atop 2 \right\} + \left\{ 4 \atop 3 \right\} + \left\{ 4 \atop 4 \right\} = 0 + 1 + 7 + 6 + 1 = 15.$$

We have the obvious identity

$$B_n = \left\{ n \atop 0 \right\} + \left\{ n \atop 1 \right\} + \left\{ n \atop 2 \right\} + \cdots + \left\{ n \atop n-1 \right\} + \left\{ n \atop n \right\}$$

which is the analogue of the formula $p(n) = p_0(n) + p_1(n) + p_2(n) + \cdots + p_{n-1}(n) + p_n(n)$; and the notation $\left\{ \cdot \atop \cdot \right\}$ reminds us that this time we are counting partitions of sets rather than partitions of numbers. It is easy to check that the value of $\left\{ n \atop k \right\}$ satisfies

- $\left\{ n \atop k \right\}$ is zero unless $0 \leqslant k \leqslant n$;
- $\left\{ n \atop 0 \right\} = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \geqslant 1; \end{cases}$
- $\left\{ n \atop 1 \right\} = \left\{ n \atop n \right\} = 1$ for $n \geqslant 1$;
- $\left\{ n \atop 2 \right\} = 2^{n-1} - 2$ and $\left\{ n \atop n-1 \right\} = \binom{n}{2}$ for $n \geqslant 2$.

More general values of the Stirling numbers are best generated using the following recurrence formula (at least when one wants to generate the values by hand:

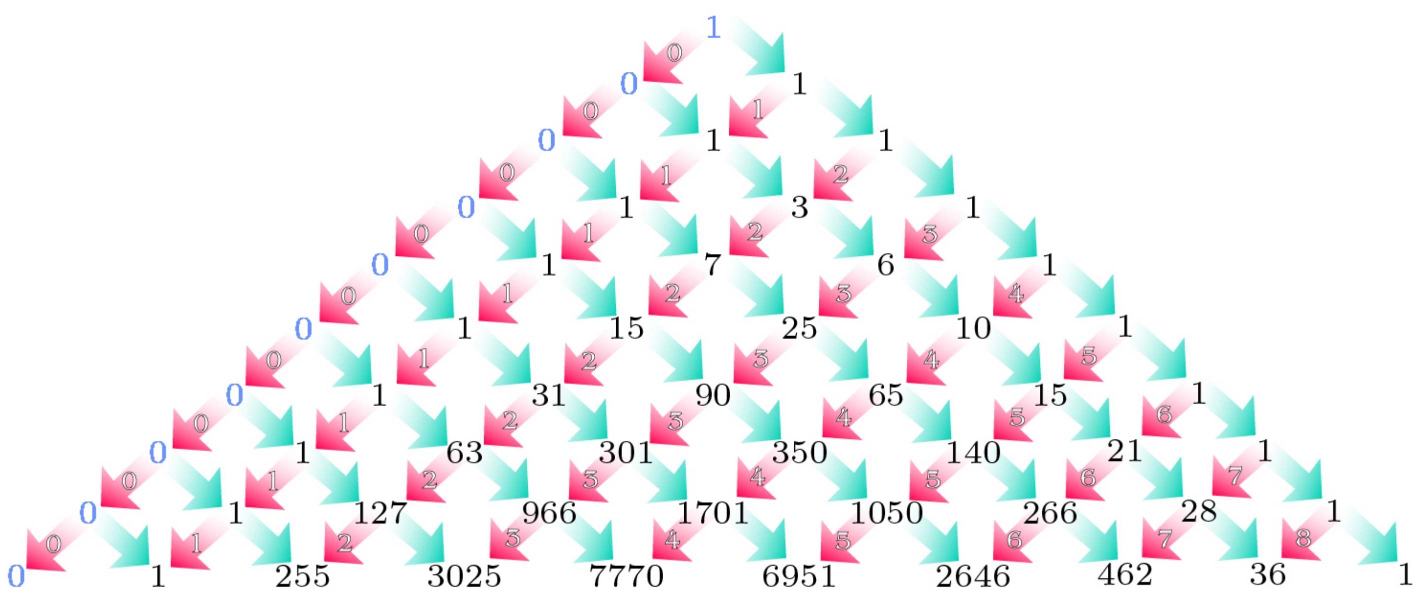
> **Theorem.** The Stirling numbers satisfy the recurrence formula
> $$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k\left\{ {n-1 \atop k} \right\}.$$

Proof. Let $n \geqslant 1$. Any partition of $[n]$ into k nonempty parts has one of two forms:

(i) We use a partition of $[n-1]$ into $k-1$ nonempty parts, and add to this one extra part $\{n\}$ of size 1. There are $\left\{ {n-1 \atop k-1} \right\}$ partitions of this form.

(ii) We use a partition of $[n-1]$ into k nonempty parts; and then add the new element n into any of the k existing parts. There are $k\left\{ {n-1 \atop k} \right\}$ ways to do this.

Adding together the number of solutions in (i) and (ii), we obtain the desired total number of partitions. □



For the purpose of quickly generating the Stirling numbers by hand, it is easiest to use an analogue of Pascal's triangle, known as *Stirling's triangle*, as shown. Entry k in row n is the Stirling number $\left\{ {n \atop k} \right\}$. The triangle is constructed recursively, very much like Pascal's triangle, each entry coming from the entries immediately above-left (green arrows) and above-right (red arrows). This differs from the Pascal triangle construction, however, in that the entry above-right (red arrow) is first multiplied by k before adding to the entry from the green arrow, as required by the recurrence relation derived above. We usually omit the leftmost edge of the triangle (which is mostly zeroes), leading to a simplified form in which the rows are now indexed $n = 1, 2, 3, \ldots$, and the entries in each row are indexed $k = 1, 2, \ldots, n$. The resulting abbreviated form of Stirling's Triangle is as shown below:

1

1 1

1 3 1

1 7 6 1

1 15 25 10 1

1 31 90 65 15 1

1 63 301 350 140 21 1

1 127 966 1701 1050 266 28 1

1 255 3025 7770 6951 2646 462 36 1

Counting Functions

Finally, we are able to count functions (including injections and surjections) between two finite sets!

> **Theorem.** The number of functions $f : [k] \to [n]$ is
>
> (a) n^k, with no further restrictions on f;
>
> (b) $P(n, k) = n(n-1)(n-2)\cdots(n-k+1)$, if f is required to be one-to-one; and
>
> (c) $\left\{ {n \atop k} \right\} k!$, if f is required to be onto.

Of course, there are no injections if $n > k$, and no surjections if $n < k$. You should check that the Theorem supplies the correct values in these cases. Also if $n = k$, then an injection is the same as a surjection; so in this case, both (b) and (c) give $n!$ as the correct answer.

Proof. Conclusion (a) is clear since there are n independent choices for each of the values $f(1), f(2), \ldots, f(k)$. If f is required to be injective, there are n choices for $f(1)$, then $n-1$ choices for $f(2)$, then $n-2$ choices for $f(3)$, . . . , and finally, $n-k+1$ choices for $f(n)$. This gives $P(n, k)$ as the answer for (b); and of course this is zero if $n > k$.

In order to construct a surjection $f : [n] \to [k]$, we must first partition the domain into k nonempty subsets (and there are $\left\{ {n \atop k} \right\}$ ways to do this); then we must match the parts of this partition to the k values in the range (and there are $k!$ ways to do this). Altogether, there are $\left\{ {n \atop k} \right\} k!$ surjections $[n] \to [k]$. □

Example: How many ways can I distribute 10 identical silver dollars to 6 students? What if each student is required to receive at least one of the silver dollars?

Solution. There are $\binom{10+5-1}{5-1} = \binom{14}{4} = 1001$ ways to distribute 10 identical silver dollars to 5 students. If each student is required to receive at least one of the silver dollars, I should give out six silver dollars first, one to each student; and then distribute the remaining 4 silver dollars in any of $\binom{4+5-1}{5-1} = \binom{8}{4} = 70$ ways.

Example: How many ways can I distribute 10 different books to 6 students? What if each student is required to receive at least one of the books?

Solution. In both cases, we are counting functions from a 10-set to a 6-set; but in the second case, the function is required to be surjective. There are $10^6 = 1{,}000{,}000$ ways to distribute the ten books. If every student is required to receive at least one book, there are $\left\{ {10 \atop 6} \right\} 6! = 22827 \cdot 720 = 16{,}435{,}440$ ways to distribute the books.

Computing Stirling Numbers

In place of a generating function, what works best for computing Stirling numbers is an approach based on polynomials. Here I will use the ring $\mathbb{Q}[x]$ of all polynomials in x with coefficients in the field of rational numbers $\mathbb{Q}$, although in place of $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ work just as well. The ring $\mathbb{Q}[x]$ is also a vector space (of infinite dimension) with basis $\mathcal{B}_1 = \{1, x, x^2, x^3, x^4, \ldots\}$. However an alternative basis $\mathcal{B}_2$ is the set of polynomials is the collection of polynomials $P(x,k)$ for $k = 0, 1, 2, 3, \ldots$. The Stirling numbers give a way of writing one basis in terms of the other. To see this, we count in two different ways the number of functions $[n] \to [k]$. On the one hand, we know this number is n^k. On the other hand, for each $i \leqslant k$ there are $\binom{k}{i}$ subsets $A \subseteq [k]$ of size $|A| = i$, and for each one there are $\left\{ {k \atop i} \right\} i!$ surjections $f : [n] \to A$. This gives the identity

$$n^k = \sum_{i=0}^{k} \binom{n}{i} \left\{ {k \atop i} \right\} i! = \sum_{i=0}^{k} \left\{ {k \atop i} \right\} P(n,i).$$

This gives the polynomial identity

$$x^k = \sum_{i=0}^{k} \left\{ {k \atop i} \right\} P(x,i)$$

expressing the basis $\mathcal{B}_1$ in terms of the basis $\mathcal{B}_2$. It is easy to express the basis $\mathcal{B}_2$ in terms of the basis $\mathcal{B}_1$, as this is a simple matter of expanding out:

$$\begin{aligned}
P(x,0) &= 1, \\
P(x,1) &= x, \\
P(x,2) &= -x + x^2, \\
P(x,3) &= 2x - 3x^2 + x^3, \\
P(x,4) &= -6x + 11x^2 - 6x^3 + x^4, \text{ etc.}
\end{aligned}$$

The coefficients in these polynomials are known as the Stirling numbers of the first kind. Solving for the basis vectors $x^k \in \mathcal{B}_1$ in terms of the basis vectors $P(x,i) \in \mathcal{B}_2$, we get

$$\begin{aligned}
1 &= P(x,0), \\
x &= P(x,1), \\
x^2 &= P(x,1) + P(x,2), \\
x^3 &= P(x,1) + 3P(x,2) + P(x,3), \\
x^4 &= P(x,1) + 7P(x,2) + 6P(x,3) + P(x,4), \text{ etc.}
\end{aligned}$$

These coefficients are the Stirling numbers of the second kind: to obtain $\left\{ {n \atop k} \right\}$, simply read off the coefficient of $P(x,k)$ in the expansion of x^n. The matrix of coefficients $\left\{ {n \atop k} \right\}$ is simply the inverse of the matrix of coefficients in the previous system. I have attached a Maple® worksheet implementing this approach, in order to compute the first ten rows of the Stirling triangle. You should check that these values are consistent with the values listed above. Also note that $\left\{ {10 \atop 4} \right\} = 34{,}105$, in agreement with the value given above.

Once again, Maple® has a builtin command for listing Stirling numbers, which you can locate using the Help feature. Simply use `with(combinat):` to load the combinatorics package, then `stirling2(10,4)` to obtain the desired value 34105.