Combinatorics Notes

Trevor Gunn

2024-03-08

Table of contents

Preface 3

I Inclusion-Exclusion 4

1 Inclusion/Exclusion 5
- 1.1 For two sets . . . 5
- 1.2 For three sets . . . 5
- 1.3 Binomial coefficients and I/E . . . 6
- 1.4 Avoiding properties . . . 7
- 1.5 Simplifying notation. . . . 7
- 1.6 Application 1: Surjections . . . 9
- 1.7 Application 2: Derangements . . . 10

2 Euler's Totient Function 11
- 2.1 The totient function . . . 12
- 2.2 Applying Inclusion/Exclusion . . . 13
- 2.3 Wrapping up . . . 14

II Permutations 15

3 Permutations 16
- 3.1 As a shuffle of the symbols . . . 16
- 3.2 As a function from X to itself . . . 16
  - 3.2.1 Composition of functions . . . 17

4 Visual representations of permutations 19
- 4.1 As a collection of cycles . . . 19
  - 4.1.1 Notation . . . 19
- 4.2 Braids . . . 20
  - 4.2.1 Transpositions . . . 21
  - 4.2.2 Braids versus permutations . . . 22

5 Generators 24
- 5.1 Squaring relation . . . 24

5.2 Commutating 24
5.3 ABA = BAB 25
5.4 Other relations? 25

6 Sign of a Permutation **27**
6.1 Step 1: reducing to the identity 27
6.2 Step 2: setting up the induction hypothesis 28
6.3 Step 3: reorganizing the product 28
6.3.1 Case 1: disjoint 29
6.3.2 Case 2: two numbers in common 29
6.3.3 Case 3: larger number in common 29
6.3.4 Case 4: smaller number in common 30
6.4 Summary 31
6.5 Notation/Computation 31

7 Inversions **32**
7.1 Parity of the number of inversions 32
7.2 Analysis 33

8 Exercises **34**

Preface

Combinatorics notes extending the book Applied Combinatorics by Keller and Trotter.

Part I

Inclusion-Exclusion

1 Inclusion/Exclusion

1.1 For two sets

Example 1.1. Suppose there are 100 CS students an a school. Of these, 50 know Java and 60 know Python (everyone knows at least one of these two languages). If we add $50 + 60$ we get the 100 total students plus 10 students who know both.

$$50 + 60 = 100 + 10.$$

Given two sets A and B, we can compute $|A \cup B|$ by first adding everything in A and adding everything in B. This counts all of $A \cup B$ except the items in $A \cap B$ were counted twice. Thus

$$|A| + |B| = |A \cup B| + |A \cap B|.$$

1.2 For three sets

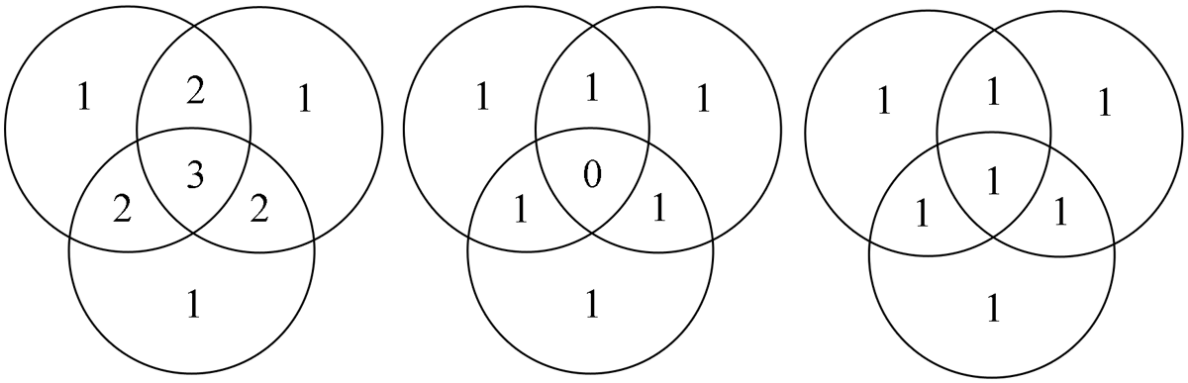


Figure 1.1: Inclusion/Exclusion for 3 sets

We can do a similar process for three sets. Start by adding up $|A| + |B| + |C|$ then subtract all the intersections of pairs then add back in $|A \cap B \cap C|$. For items in each region of the Venn diagram, think about how many times an item is added/subtracted overall.

E.g. items in A and B but not C will be added in twice in the first step ($|A| + |B|$) and subtracted once in the second step ($|A \cap B|$) so they are counted once.

Items in $A \cap B \cap C$ are added in $3 - 3 + 1 = 1$ times.

1.3 Binomial coefficients and I/E

Suppose we are looking at a union of n sets: $A_1 \cup \cdots \cup A_n$. Consider an element x belonging exactly to $A_1, \ldots, A_m$ and no other sets. If we do the same procedure of adding $|A_1| + \cdots + |A_n|$ then we are counting x a total of m times. Then, when we subtract all the pairwise intersections, $|A_i \cap A_j|$, we are subtracting from the count of x a total of $\binom{m}{2}$ because there are $\binom{m}{2}$ ways to choose two indices $\{i, j\} \subset \{1, \ldots, m\}$.

Thus, if we alternate: adding in single sets, subtracting pairs of intersections, adding tripples, subtracting quadruples, etc. we are counting x a total of

$$\binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \binom{m}{4} + \cdots \tag{1.1}$$

times.

Let's have a closer look at this. We know the Binomial Theorem says that

$$(1 + x)^m = \binom{m}{0}x^0 + \binom{m}{1}x^1 + \binom{m}{2}x^2 + \cdots + \binom{m}{m}x^m.$$

Substituting $x = -1$, we get

$$0 = \binom{m}{0} - \binom{m}{1} + \binom{m}{2} - \cdots + \binom{m}{m}(-1)^m.$$

And since $\binom{m}{0} = 1$, if we move all the other terms to the other side of the equation, we see that Equation 1.1 evaluates to 1.

> **Note 1**
>
> For this problem we are counting the size of $A_1 \cup \cdots \cup A_n$ so each element is in at least one set ($m \geq 1$). This is important so that we can say that $0^m = 0$. In the context of the binomial theorem, $(1 - 1)^0 = \binom{0}{0} = 1$. This will be important later when we want to count elements not belonging to any set.

The general I/E formula is then

$$\begin{aligned}|A_1 \cup \cdots \cup A_n| &= \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \dots \\ &= \sum_{t=1}^{n} (-1)^{t+1} \sum_{i_1<\cdots<i_t} |A_{i_1} \cap \cdots \cap A_{i_t}|.\end{aligned}$$

Note: $(-1)^{t+1}$ takes the values $1, -1, 1, -1, \dots$ starting at $t = 1$.

1.4 Avoiding properties

Inclusion/Exclusion appears most frequently in combinatorics not as a means to count $|A_1 \cup \cdots \cup A_n|$ directly but rather as a means to count everything *not* in $A_1 \cup \cdots \cup A_n$.

For example, let \$[m] = {1,...,m} and suppose we want to count the number of functions $f : [m] \to [n]$ which *don't* miss anything in the codomain. I.e. if A_i is the set of functions where $f(x)$ never equals i, then we want to count every function *not* in any set A_i.

Let X be the total set of functions. Then to count the functions avoiding $A_1 \cup \cdots \cup A_n$ we do

$$\begin{aligned}|X \setminus (A_1 \cup \dots A_n)| &= |X| - |A_1 \cup \dots A_n| \\ &= |X| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \cdots\end{aligned}$$

> **Caution**
>
> The formula $|A \setminus B| = |A| - |B|$ works only when B is contained entirely inside A.

1.5 Simplifying notation.

Let $\mathscr{P} = \{P_1, \dots, P_n\}$ be a collection of sets representing negative properties—conditions we want to avoid. For a subset $S \subseteq \{1, \dots, n\}$, let $N_{\geq}(S)$ be the number of items satisfying at least those properties in S. I.e. $\bigcap_{i \in S} P_i$. We will say "x satisfies S" for short. Also, if $S = \{a, b, c\}$, let us write for example, $N_{\geq}(a, b, c)$ instead of $N_{\geq}(\{a, b, c\})$ for simplicity.

Let $N_{=}(S)$ be the number of x satisfying exactly those properties in S and no properties not in S. When $\mathscr{P}$ represents properties we wish to avoid, then $N_{=}(\varnothing)$ is the number of elements satisfying none of the properties.

Theorem 1.1 (Inclusion/Exclusion)**.**

$$\begin{aligned} N_{=}(\varnothing) &= N_{\geq}(\varnothing) - \sum_i N_{\geq}(i) + \sum_{i<j} N_{\geq}(i,j) - \cdots \\ &= \sum_S (-1)^{|S|} N_{\geq}(S). \end{aligned}$$

Proof. As in Section 1.3, we will break up the sum focusing on each element. So first, we will write

$$\sum_S (-1)^{|S|} N_{\geq}(S) = \sum_S (-1)^{|S|} \sum_{\substack{x \\ x \text{ satisfies } S}} 1$$

Here we replace the number $N_{\geq}(S)$ by a count of 1 for each element x which satisfies S. This introduces a second sum into the picture and that will allow us to swap the order of summations. Currently, the second sum is over all x with respect to the relation "x satisfies S." When we put the sum over x outside, we still have the relation "x satisfies S" but rather than summing over all x with this property, we sum over all S:

$$\sum_S (-1)^{|S|} \sum_{\substack{x \\ x \text{ satisfies } S}} 1 = \sum_x \sum_{\substack{S \\ x \text{ satisfies } S}} (-1)^{|S|}.$$

We also move the $(-1)^|S|$ inside the second sum because that quantity depends on $|S|$.

Next, just as we did in Section 1.3, let us say that x satisfies exactly $P_{i_1}, \ldots, P_{i_m}$ (m will depend on x) and let $S_x = \{i_1, \ldots, i_m\}$. This set is the largest set that x satisfies. Every other set that x satisfies will be a subset of S_x.

We now break up our sum based on the size of those subsets $S \subseteq S_x$, using the binomial coefficients to count the number of such S:

$$\sum_x \sum_{\substack{S \\ x \text{ satisfies } S}} (-1)^{|S|} = \sum_x \sum_{k=0}^{m} \underbrace{\binom{m}{k}(-1)^k}_{\substack{|S|=k \text{ and } x \text{ satisfies } S \\ \iff |S|=k \text{ and } S \subseteq S_x}}.$$

This, by the Binomial Theorem, is the same as

$$\sum_x (1-1)^m.$$

Remember here that m depends on x.

As discussed in Note 1, $0^m = 0$ if $m \geq 1$ but if $m = 0$ then $0^0 = 1$. Saying $m = 0$ means that x satisfies exactly 0 properties—which is what we are looking for. So the final simplification looks like

$$\sum_x 0^m = \sum_{\substack{x \\ m=0}} 1 = N_=(\varnothing).$$

□

1.6 Application 1: Surjections

As in Section 1.4, let X be the set of all functions from $[m]$ to $[n]$ and let P_i be the set of functions where i is never an output: $f(x) \neq i$ for any input x.

Lemma 1.1.

a) The number of functions from an m element set to an n element set is n^m.
b) The number of functions from an m element set to an n element set that avoid k outputs is $(n-k)^m$.

Proof.

a) There are n choices for $f(1)$ and n choices for $f(2)$, etc. So there are n^m choices in total for $f(1), \ldots, f(m)$.
b) If we are avoiding k outputs then there are only $n - k$ choices for each of $f(1), \ldots, f(m)$ so $(n-k)^m$.

□

With this in mind, we have $N_{\geq}(S) = (n-k)^m$ if $|S| = k$ (we avoid at least the specified k outputs). So by inclusion exclusion, the number of functions which avoid *no* outputs (i.e. surjections) is

$$\begin{aligned} &N_{\geq}(\varnothing) - \sum_i N_{\geq}(i) + \sum_{i<j} N_{\geq}(i,j) - \cdots \\ &\qquad = n^m - \sum_i (n-1)^m + \sum_{i<j} (n-2)^m - \cdots . \end{aligned}$$

And since there are $\binom{n}{k}$ ways to choose k outputs to avoid, we can also write this as

$$n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \cdots = \sum_k \binom{n}{k}(-1)^k (n-k)^m.$$

The textbook calls this number $S(m, n)$.

> **Note**
>
> In the above application, $N_{\geq}(S) = (n-k)^m$ only depended on the size k of S. This is true in many but not all examples.

1.7 Application 2: Derangements

Let X be the set of all permutations of $[n]$. We wish to count the permutations with no fixed points. So let P_i be the property that i is a fixed point. Then $N_{=}(\varnothing)$ is the number of permutations with zero fixed points. This number is called d_n in the textbook.

Lemma 1.2.

a) The number of permutations of $[n]$ is $n!$
b) The number of permutations with at least k specified fixed points is $(n-k)!$

Proof.

a) Discussed in Chapter 2 of the book.
b) To say that there are k specified fixed points means were are permuting the other $n-k$ items. Similar to (a), there are $(n-k)!$ ways to permute $n-k$ items.

□

Applying Inclusion/Exclusion, we thus have

$$d_n = n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots = \sum_k \binom{n}{k}(-1)^k(n-k)!.$$

Again, there are $\binom{n}{k}$ ways to choose a set S of k fixed points and each of these has the same number $N_{\geq}(S) = (n-k)!$.

2 Euler's Totient Function

Consider the number system of the integers mod n. This means we group the numbers based on whether they have the same remainder when divided by n. For example, $3 \cdot 7 \equiv 1 \pmod{10}$ because both 21 and 1 have the same remainder when divided by 10.

This number system works a lot like the inteegers. We have addition, subtraction and multiplication. The question is: when can we divide? For example, is it possible to solve $2x \equiv 3 \pmod{10}$ for x? I.e. is there a way to move the 2 to the other side of the equation? The answer is no, because $2x$ will never have a remainder of 3.

What about $7x \equiv 3 \pmod{10}$? This is trickier. Let's look at our multiples of 7:

$$0, 7, 14, 21, 28, 35, 42, 49, 56, 63$$

Ok so we see that $7 \cdot 9 \equiv 3 \pmod{10}$.

The next theorem gives the criterion for when we can "divide both sides by a."

Theorem 2.1. *We can solve the equation* $ax \equiv b \pmod{n}$ *for any* b *if and only if* $\gcd(a, n) = 1$.

Proof. First, suppose $\gcd(a, n) \neq 1$ and call this gcd d. Consider the equation $ax \equiv 1 \pmod{n}$. This means that ax has a remainder of 1 when divided by n or in other words ax is 1 more than a multiple of n: $ax = qn + 1$. Now, d is a divisor of both a and n by definition, which means $ax - qn$ should be a multiple of d. But $ax - qn = 1$ and we said $d \neq 1$, that is impossible. So $ax \equiv 1 \pmod{n}$ has no solution.

Next, suppose $\gcd(a, n) = 1$. By Theorem 3.9 of the Keller/Trotter book, also known as Bézout's theorem, we can find integers u and v such that $au + nv = \gcd(a, n) = 1$. If we rearrange, we get $au = 1 - nv$. So au and 1 differ by a multiple of n.

Let's apply this to our equation: take the equation $ax \equiv b \pmod{n}$ and multiply by u to get

$$aux = (1 - nv)x \equiv bu \pmod{n}$$

Since $(1 - nv)x$ and x differ by a multiple of n, they share the same remainder. Thus our solution is $x \equiv bu \pmod{n}$. □

Example 2.1. Let's follow the steps to solve $7x \equiv 3 \pmod{1}0$. First, we do Euclid's algorithm to find the gcd of 7 and 10.

$$\begin{aligned} 10 &= 7 + 3 \\ 7 &= 2 \cdot 3 + 1 \end{aligned}$$

To get Bézout's identity: we solve each equation for the remainder term and then substitute the bigger remainder into the next equation:

$$\begin{aligned} 3 &= 10 - 7 \\ 1 &= 7 - 2 \cdot 3 = 7 - 2 \cdot (10 - 7) = 3 \cdot 7 - 2 \cdot 10. \end{aligned}$$

So 3 is our u. Multiply both sides of $7x \equiv 3 \pmod{10}$ by 3 and we get

$$\begin{aligned} 3 \cdot 7x &\equiv 3 \cdot 3 \pmod{10} \\ 1 \cdot x &\equiv 9 \pmod{10} \end{aligned}$$

2.1 The totient function

Numbers where $\gcd(a, n) = 1$ are called "coprime to n." They show up often when talking about modular arithmetic. We have a function, $\phi(n)$ called "Euler's totient function" which counts the number of integers less than n which are coprime to n. For example, $\phi(10) = 4$ counting $1, 3, 7, 9$.

Euler gave the following formula for $\phi(n)$ in terms of the prime factors p of n:

$$\phi(n) = n \cdot \prod_{\substack{p \mid n \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right)$$

Here $p \mid n$ means p divides n or p is a factor of n.

This is the form we will prove in a minute, but let us also rewrite this formula in maybe a more helpful way. Say $n = p_1^{k_1} \cdots p_r^{k_r}$ is the prime factorization of n. Then

$$\phi(n) = p_1^{k_1} \cdots p_r^{k_r} \cdot \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right).$$

Now look at each pair of factors involving p_i:

$$p_i^{k_i} \left(1 - \frac{1}{p_i}\right) = p_i^{k_i} \left(\frac{p_i - 1}{p_i}\right) = p_i^{k_i - 1}(p_i - 1).$$

Putting that together, we get

$$\phi(n) = p_1^{k_1-1}(p_1-1)\cdots p_r^{k_r-1}(p_r-1).$$

Example 2.2. For $n = 18 = 2 \cdot 3^2$ we have

$$\phi(12) = 2^0(2-1)3^1(3-1) = 6.$$

You can double check this by listing all the numbers coprime to 18.

2.2 Applying Inclusion/Exclusion

For each of the primes p_i which are factors of n. Consider the property P_i which says that a is divisible by p_i. As we've been doing, for a subset of these primes, we want to count the number of a in $0, \ldots, n-1$ which are divisible by at least those primes in S.

Lemma 2.1. *Let m be a divisor of n. The number of integers $a < n$ which are a multiple of m is n/m.*

Proof. Let's look at an example first. Suppose $n = 100$ and $m = 5$. Then we list the numbers from $0, \ldots, 99$ and write the list mod 5. That is:

$$0, 1, 2, 3, 4, 0, 1, 2, 3, 4, 0, 1, 2, 3, 4, \ldots, 0, 1, 2, 3, 4.$$

Because $n-1$ is one less than n, our last number has a remainder of one less than 5. So we see every number $0, 1, 2, 3, 4$ exactly the same number of times. Since there are 5 possible remainders, we much see every remainder $100/5 = 20$ times. This includes a remainder of 0.

This pattern works generally: we repeat remainders of $0, \ldots, m-1$ and we end on $m-1$ because $n-1$ is one less than a multiple of m. □

So putting this into our inclusion-exclusion notation, if $S = \{i_1, \ldots, i_k\}$, then

$$N_{\geq}(S) = \frac{n}{p_{i_1} \cdots p_{i_k}}.$$

2.3 Wrapping up

By Inclusion/Exclusion, the number of whole numbers $a < n$ not divisible by any p_i is

$$N_{\geq}(\varnothing) - \sum_i N_{\geq}(i) + \sum_{i<j} N_{\geq}(i,j) - \cdots$$

and that is

$$n - \sum_i \frac{n}{p_i} + \sum_{i<j} \frac{n}{p_i p_j} - \cdots = n\left(1 - \sum_i \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \cdots\right).$$

We should be able to factor a $(1 - 1/p_1)$ out of this so let's work at giving that a go by separating terms including p_1 from those that don't:

$$\begin{aligned}
&1 - \sum_i \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \cdots \\
&\quad = 1 - \frac{1}{p_1} - \sum_{1<i} \frac{1}{p_i} - \frac{1}{p_1}\sum_{1<j} \frac{1}{p_j} + \sum_{1<i<j} \frac{1}{p_i p_j} - \cdots \\
&\quad = \left(1 - \frac{1}{p_1}\right) - \left(\sum_{1<i} \frac{1}{p_i} - \frac{1}{p_1}\sum_{1<j} \frac{1}{p_j}\right) + \left(\sum_{1<i<j} \frac{1}{p_i p_j} - \frac{1}{p_1}\sum_{1<i<j} \frac{1}{p_i p_j}\right) - \cdots \\
&\quad = \left(1 - \frac{1}{p_1}\right)\left(1 - \sum_{1<i} \frac{1}{p_i} + \sum_{1<i<j} \frac{1}{p_i p_j} - \cdots\right).
\end{aligned}$$

Note: in the second term in the second to last line, the sum over all $j > 2$ is the same as the sum over all $i > 2$ just with a different name given to the variable.

So we are able to factor out $(1 - 1/p_1)$ to get a similar expression with one fewer primes. Similarly, we can factor out $(1 - 1/p_2)$ and so on until there are no more primes. This gives us Euler's product representation for the totient function.

Part II

Permutations

3 Permutations

A permutation of a finite set X can be thought of in a few ways. The set of all permutations of $\{1, \ldots, n\}$ will be denoted S_n. This is also called *the symmetric group*.

3.1 As a shuffle of the symbols

E.g. 53214 is a shuffle of $X = \{1, 2, 3, 4, 5\}$.

This gives us one way of counting the number of permutations. For the first position in the shuffle, we may write any of the n numbers. For the second position, we have $n - 1$ numbers to choose from—we cannot repeat the first. Likewise the third has $n - 2$ to choose from, not repeating the first or second. In this way, the number of permutations is

$$n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1, \text{ denoted } n!.$$

3.2 As a function from X to itself

E.g. The shuffle 53214 may be thought of as a function $\pi : X \to X$ where $\pi(i)$ equals the i-th number in the shuffle:

$$\pi(1) = 5, \pi(2) = 3, \pi(3) = 2, \pi(4) = 1, \pi(5) = 4.$$

These functions are often represented by a table:

i	1	2	3	4	5
$\pi(i)$	5	3	2	1	4

Something new added by the table/function representation is that we can talk about the inverse function which undoes the shuffle:

$j = \pi(i)$	5	3	2	1	4
$i = \pi^{-1}(j)$	1	2	3	4	5

or reordering the columns:

j	1	2	3	4	5
$\pi^{-1}(j)$	4	3	2	5	1

3.2.1 Composition of functions

Given two permutations, $\pi_1, \pi_2 : X \to X$, we can compose them to get a new function $(\pi_1 \circ \pi_2)$. One way to compute this is via the following procedure:

1. Write the two functions as tables.
2. Reorder the columns of the outermost function in the composition to align with the output of the innermost function.
3. Stack the tables on top of eachother.

E.g. Take π_2 to be the function represented by the shuffle 53214 and take π_1 to be represented by the shuffle 13254. So as a table

i	1	2	3	4	5
$\pi_1(i)$	1	3	2	5	4

Now we shuffle the columns so the top is 53214:

i	5	3	2	1	4
$\pi_1(i)$	4	2	3	1	5

Then stack this with π_2:

i	1	2	3	4	5
$j = \pi_2(i)$	5	3	2	1	4
$\pi_1(j)$	4	2	3	1	5

The bottom row is $\pi_1 \circ \pi_2$.

Exercise 3.1. Compute this in the other order: $\pi_2 \circ \pi_1$. *Observe that the order matters.*

Note: in some sources (e.g. books, software), compositions are witten as a product in the reverse order. Here are computation in the software SageMath.

```
π1 = Permutation([1,3,2,5,4])
π2 = Permutation([5,3,2,1,4])
π2 * π1
```

```
[4, 2, 3, 1, 5]
```

```
π1 * π2
```

```
[5, 2, 3, 4, 1]
```

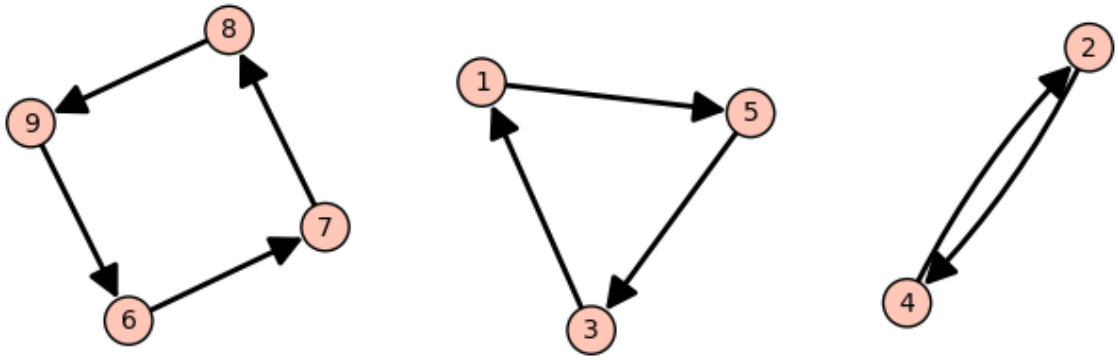
4 Visual representations of permutations

4.1 As a collection of cycles

Let's look at a longer permutation now: 541237896. As in Section 3.2, we can view this as a function where 1 goes to 5, 2 goes to 4 etc. Following the path of a single number we get a cycle: 1 goes to 5 goes to 3 goes to 1. Likewise 2 goes to 4 goes to 2. And so on.

The entire permutation can be broken up into cycles:

```
π = Permutation([5,4,1,2,3,7,8,9,6])
π.to_digraph().plot(vertex_size=1500)
```



4.1.1 Notation

The cycle $1 \to 5 \to 3 \to 1$ can be written as (153). In general, $(a_1, \ldots, a_n)$ represents the cycle $a_1 \to a_2 \to \cdots \to a_n \to a_1$. We often ommit the commas when every number is a single digit.

Here is how to compute the cycles in SageMath:

```
π.cycle_string()
```

```
'(1,5,3)(2,4)(6,7,8,9)'
```

The answer is given as a product (i.e. composition) of cycles.

Elements which do not go anywhere—also called *fixed-points*—are represented by cycles of length 1 E.g. $(123)(4)$ represents the two cycles $1 \to 2 \to 3 \to 1$ and $4 \to 4$. We often ommit length 1 cycles from the notation so $(123)(4) = (123)$ as a permutation of $1, 2, 3, 4$.

Exercise 4.1. Compute the cycle decomposition for 341859672. You can verify your answer in SageMath.

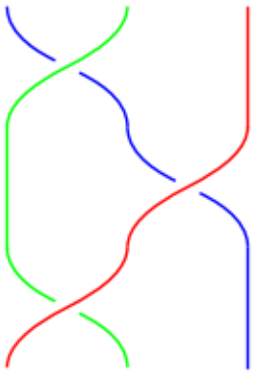
```
π = Permutation([3,4,1,8,5,9,6,7,2])
π.cycle_string()
```

Note: as mentioned earlier: SageMath ommits the cycle (5) representing the fixed point $\pi(5) = 5$.

4.2 Braids

Another visual representation used frequently in the mathematical study of knots is that of crossing lines. Let's look at an example:

```
B.<a,b> = BraidGroup(3)
plot(a * b * a)
```



This represents a permutation where the first end ends up in the third position ($1 \to 3$), the second end ends up in the second position ($2 \to 2$) and the third end ends up in the first position ($3 \to 1$). Overall, 1 and 3 are swapped, so this permutation is (13).

4.2.1 Transpositions

Cycles of length 2 are called *transpositions*. E.g. (14) is a transposition which switches 1 and 4. The picture above represents a composition of 3 transpositions (13) = (12)(23)(12).

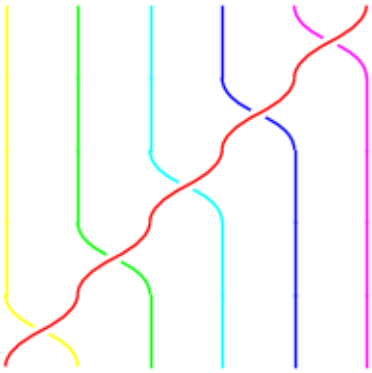
Theorem 4.1 (Product of transpositions)**.**

a) Every permutation can be written as a product of transpositions
b) Every permutation can be written as a product of transpositions of adjacent elements.

Another way to say this is that by swapping pairs of elements at a time, we can obtain any possible shuffle.

Proof. We know that a permutation may be written in terms of cycles. So if we can show that any cycle can be written as a successive sequence of swaps we are good. We will give a visual demonstration of this fact:

```
B.<t1,t2,t3,t4,t5> = BraidGroup(6)
plot(t1 * t2 * t3 * t4 * t5)
```



So the cycle (123456) = (12)(23)(34)(45)(56). Remember that compositions are read from right to left: e.g. $(f \circ g)(i) = f(g(i))$ means first do g then do f.

This kind of decomposition generalizes: you can replace (123456) with any cycle of any length. E.g. (1456) = (14)(45)(56) although one more reminder that products in SageMath are backwards from the function composition standpoint:

```
S = SymmetricGroup(6)
S((5,6)) * S((4,5)) * S((1,4))
```

```
(1,4,5,6)
```

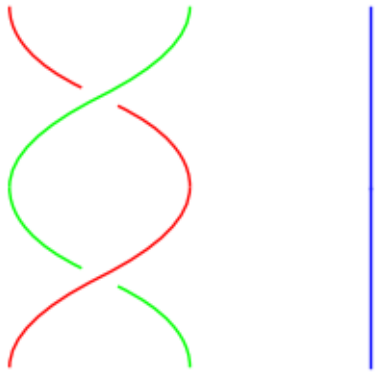
□

For an alternative proof, this decomposition into a sequence of adjacent swaps is exactly how the sorting algorithm BubbleSort works. We can sort any list using BubbleSort which does only adjacent swaps. So the shuffle is obtained by reversing those swaps to go from sorted to shuffled.

4.2.2 Braids versus permutations

You may have noticed in what we did, the word "braid" was used. Braids are similar to permutation except that we keep track of which strand goes above and which strand goes below. E.g. the transposition (12) applied twice looks like

```
B.<a,b> = BraidGroup(3)
plot(a * a)
```



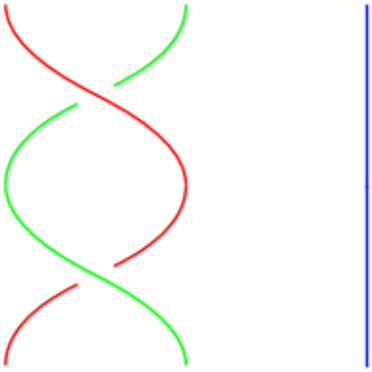
As a permutation this is the trivial shuffle 123. But as a braid it is still twisted.

While we are only focused on permutations rather than braids here, we still make use of braids because SageMath is able to create diagrams for us.

4.2.2.1 The Braid Group in SageMath

The generators of the braid group are adjacent swaps. So the line `B.<a, b> = BraidGroup(3)` sets `a` to a swap of $1, 2$ and sets `b` to a swap of $2, 3$. We can obtain the reverse swap with `a^-1` or `b^-1`

```
plot(a^-1 * a^-1)
```



When converting from braids to permutations, we ignore whether a strand goes over or under and just focus on the swapping.

5 Generators

In Theorem 4.1 we saw that every cycle and hence every permutation can be written as a product of adjacent swaps: $(12), (23), (34), \ldots, (n-1, n)$. Here we will identify some relations that hold among these generating elements.

Let us call $\tau_i = (i, i+1)$ the swap of i and $i+1$.

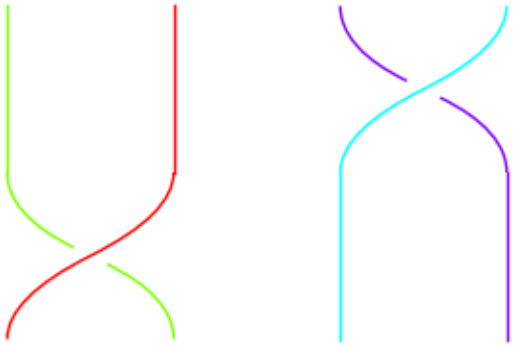
5.1 Squaring relation

We have $\tau_i^2 = 1$ where 1 represents the identity permutation (no shuffling). This says that if we swap i and $i+1$ and then swap again, we get back to a sorted list.

5.2 Commutating

From Exercise 3.1, we saw that in general $\pi_1\pi_2 \neq \pi_2\pi_1$. Nonetheless, if we are swapping disjoint sets of pairs like (12) followed by (34) then there is no interaction between the swaps. So the order doesn't matter: $(12)(34) = (34)(12)$. Specifically, τ_i and τ_j commute ($\tau_i\tau_j = \tau_j\tau_i$) provided i and j are at least 2 apart.

```
B.<a,b,c> = BraidGroup(4)
plot(a * c)
```



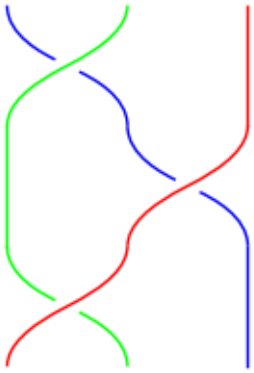
```
a * c == c * a
```

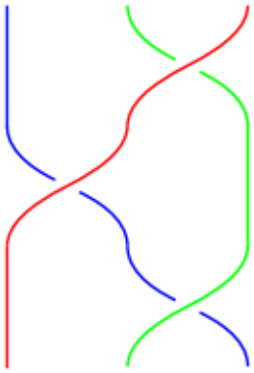
```
True
```

5.3 ABA = BAB

At the top of Section 4.2 we saw that $(12)(23)(12) = (13)$. We also have $(23)(12)(23) = (13)$. Compare the following pictures.

```
B.<a,b> = BraidGroup(3)
plot(a * b * a)
plot(b * a * b)
```





We can see visually that the middle green strand is sliding from one side of the blue/red crossing to the other.

5.4 Other relations?

It turns out, every way to simplify or manipulate products of transpositions can be reduced to exactly these three rules:

1. $\tau_i^2 = 1$
2. $\tau_i \tau_j = \tau_j \tau_i$ for $|i - j| \geq 2$ (at least two apart)
3. $\tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1}$

The proof of this has two stages. One which we have already seen. First, you show that every permutation can be written as a product of transpositions of adjacent elements (Theorem 4.1). This shows that you can use $\tau_1, \ldots, \tau_n$ and these rules to write every element of S_n. I.e. the number of objects generated by these rules is at least $n!$.

The second step is showing that the number of objects generated by these rules is no more than $n!$ so that it is exactly the same as S_n. This requires more tools than we have available to us (i.e. group theory). A proof for those in-the-know may be found here for instance.

We will take it as a fact that these rules descrie S_n exactly.

6 Sign of a Permutation

We established in Theorem 4.1 that every permutation can be written as a product of transpositions. But more than that, in all our examples, the number of transpositions required to perform a shuffle was always consistently odd or consistently even. For instance, $(13) = (12)(23)(12)$ has an odd number of transpositions on both sides.

If we look at Section 5.4, we said that every possible way to manipulate a permutation can be reduced to three rules, each of which preserves the parity (odd or even) of the number of transpositions.

In case you were unsatisfied with relying on the fact that every relation reduces to the three given relations, we will present two different proofs that the parity is always consistent: one here, one in the next section.

6.1 Step 1: reducing to the identity

Suppose we have two different ways to write a given permutation as a product of transpositions:

$$\pi = \tau_1\tau_2\ldots\tau_m = \tau_1'\tau_2'\ldots\tau_n'.$$

(Here the letter τ means any transposition, not just for adjacent elements.)

Our goal is to show that m and n have the same pairty (both are odd or both are even). Notice that if we multiply both sides by τ_1, we get

$$\tau_1\cdot\tau_1\tau_2\ldots\tau_m = \tau_1\cdot\tau_1'\tau_2'\ldots\tau_n'$$

but $\tau_1\cdot\tau_1$ means swap the same pair twice, and doing this twice is the same as not doing it at all. Therefore the $\tau_1\cdot\tau_1$ cancels and we are left with

$$\tau_2\ldots\tau_m = \tau_1\cdot\tau_1'\tau_2'\ldots\tau_n'.$$

Keep doing this from left to right until we have

$$1 = \tau_m \dots \tau_2 \tau_1 \tau_1' \tau_2' \dots \tau_n'.$$

So saying "m and n have the same parity" is equivalent to saying "1 cannot be written as an odd number of transpositions."

6.2 Step 2: setting up the induction hypothesis

Let's introduce some letters now for our transpositions. Say

$$1 = (a_1 b_1)(a_2 b_2) \cdots (a_k b_k) \tag{6.1}$$

where $a_i \neq b_i$ for all i (we're always swapping two different things).

We know that k isn't 1 since a single transposition is not the same as the identity. So look now at $k \geq 2$. We will show that we can always rewrite this product using $k-2$ transpositions.

There are a couple ways to phrase this next part of the proof. First, we could say: take as an inductive hypothesis that for all $k' < k$ if the identity is written as a product of k' transpositions then k' is even. In this way, going from k to $k' = k-2$ we know that $k-2$ is even and hence so is k.

Another way we could say this: we are always reducing by 2 every time so $k-2-2-2-\dots$ will either end up at 0 or 1 and we know it can't end up at 1.

6.3 Step 3: reorganizing the product

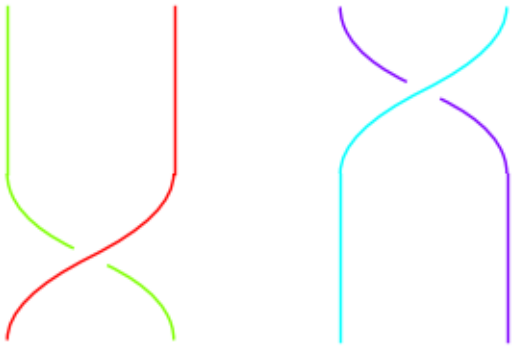
Look back at Equation 6.1. Since $(ab) = (ba)$, to keep things consistent, we will assume that $a_i < b_i$ always. Let a be the smallest number appearing in any transposition. E.g. for $(23)(34)(23)(24)$ we have $a = 2$.

We will slide all the the transpositions with an a in them to the right (towards the beginning of the composition). Now we can't just move things without changing the product so we have to be strategic.

Let's say we have in the middle of our expression $(uv)(xy)$ where $u = a$ and $x \neq a$ and so we want to move (uv) to the right of (xy).

6.3.1 Case 1: disjoint

If u, v, x, y are all distinct, then the two swaps can be done in either order and $(uv)(xy) = (xy)(uv)$.

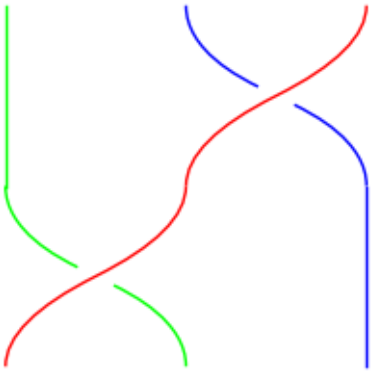


6.3.2 Case 2: two numbers in common

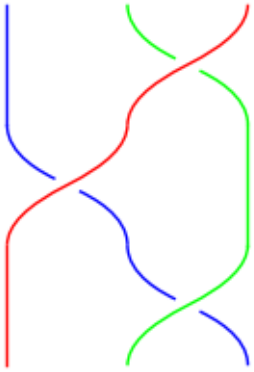
If $(uv) = (xy)$ then $(uv)(xy) = (uv)^2 = 1$ and we have reduced the number of transpositions by 2 as we said we would.

6.3.3 Case 3: larger number in common

Consider the product $(12)(23) = (123)$ (swap $2, 3$ first then $1, 2$)



Compare this with the identity $(23)(12)(23) = (13)$ that we worked out in Chapter 5:



Multiply both sides on the left by (23) and simplify using $(23)(23) = 1$ to get $(12)(23) = (23)(13)$.

In this way, the smallest number is always moving to the right. The rule is $(uv)(xy) = (xy)(uv')$ where v' is the number in (xy) which wasn't in common.

Exercise 6.1. Work out the diagrams to show that $(12)(23) = (23)(13)$. (I would show you in SageMath, but the software doesn't draw (13) very well.)

6.3.4 Case 4: smaller number in common

Using the previous case, we move all the terms having $u = a$ to the right. So now we have a bunch of terms $(av_1)(av_2)\dots(av_r)$ in the right of our product. We know we have at least two pairs because if we swap a out of position, something later on has to swap it back into position.

Also, if any of the adjacent pairs are equal, we can cancel them as in Case 2. Otherwise, we do something similar to Case 3 where

$$(12)(23)(12) = (13) \implies (12)(23) = (13)(12)$$

by multiplying on the right by (12)

The feature of this identity is that we have one fewer 1's, or more generally, one fewer a's. And we can keep doing this until we are left with either a single a (impossible) or we eventually find a pair that cancels.

Exercise 6.2. Draw the diagram for this identity.

6.4 Summary

We showed that we can push all the transpositions containing a 1 to the right and then moving those transpositions past each other until we had fewer and fewer 1's and eventuall there must be a pair with both numbers in common because we can't just move 1 a single time in our sequence of transpositions.

Example 6.1. Start with (12)(23)(12)(13). Use Case 3 to swap the first and second: (23)(13)(12)(13). Now use Case 4 to swap the second and third: (23)(23)(13)(13). Now cancel using Case 2.

6.5 Notation/Computation

If π is a permutation, common notations for its parity or sign are: $\operatorname{sgn}\pi$ and $(-1)^\pi$. We say the sign is $+1$ if it is an even length product and the sign is -1 if it is an odd length product.

We have the identity $\operatorname{sgn}(\pi_1\pi_2) = \operatorname{sgn}\pi_1 \operatorname{sgn}\pi_2$ because, for example, multiplying two odd length products creates an even length product $((-1)\cdot(-1) = +1)$

One way to compute this is by decomposing π into cycles. We saw in Theorem 4.1 how to write cycles in terms of transpositions: (12345) = (12)(23)(34)(45). The observation here is that odd length cycles become an even length product of transpositions and vice versa.

So the algorithm is:

1. Convert the permutation into a product of cycles.
2. Write a $+1$ if the cycle has an odd length and a -1 if it has an even length.
3. Multiply those numbers together to find the sign.

Example 6.2. Consider the shuffle 376819254. We can write this as (1369485)(27) this is a product of a cycle of length 7 and a cycle of length 2. Therefore the sign is $(+1)(-1) = -1$.

This computation in SageMath:

```
π = Permutation([3,7,6,8,1,9,2,5,4])
sign(π)
```

```
-1
```

7 Inversions

We present a second (or third depending if you count the proof we didn't prove) proof that the number of transpositions is always ever even or always odd.

Let π be a permutation. An **inversion** for π is a pair of outputs which are not in sorted order. I.e. we have $i < j$ to begin with and $\pi(i) > \pi(j)$ to end with.

In counting inversions, we look at where $\pi(i)$ is greater than $\pi(j)$ for some j that comes after i.

Example 7.1. Consider the permutation represented by 31254.

We have $\pi(1) = 3$ and this is greater than both $\pi(2) = 1, \pi(3) = 2$.

We have $\pi(2) = 1$ but this is not greater than anything that comes after.

We have $\pi(3) = 2$ but this is not greater than anything that comes after.

We have $\pi(4) = 5$ which is greater than $\pi(5) = 4$.

So this permutation has 3 inversions: $(1 < 2) \to (\pi(1) > \pi(2))$ and $(1 < 3) \to (\pi(1) > \pi(3))$ and $(4 < 5) \to (\pi(4) > \pi(5))$.

We can use SageMath to check our work:

```
π = Permutation([3,1,2,5,4])
π.inversions()
```

```
[(1, 2), (1, 3), (4, 5)]
```

7.1 Parity of the number of inversions

Inversions as a concept have a few uses in combinatorics. Relavent to us now is that the number of inversions has the same parity as the permutation. To show this, we will consider how a single transposition affects this parity. But since the total number can go up or down, let us define a quantity which we can analyze to describe the parity.

For the identity permutation, define

$$V(1) = \prod_{i<j}(j-i)$$

E.g. for $n = 4$ this is $(2-1)(3-1)(4-1)(3-2)(4-2)(4-3)$. We're not interested in the absolute value here but rather the sign—which for the identity permutation is $+1$.

More generally, define

$$V(\pi) = \prod_{i<j}(\pi(j) - \pi(i)).$$

Note that we will have a factor of $V(\pi)$ which is negative whenever $i < j$ and $\pi(i) > \pi(j)$. So the sign of $V(\pi)$ tells us the parity of the number of inversions.

What's useful here is that factoring out the various -1's and reordering, we can write $V(\pi) = \pm V(1)$ where it is a $+1$ if we have an even number of inversions and a -1 if we have an odd number.

7.2 Analysis

We want to show that a single transposition changes $V(1)$ to $-V(1)$. Then a sequence of odd length will have a sign of -1 and one of even length will have a sign of $+1$, matching what we did in Chapter 6.

Suppose we apply the transposition (ij) with $i < j$. Right away, we get one inversion because now $\pi(i) > \pi(j)$. Let's look at the other numbers. Suppose $x < i < j$. Then after swapping, we still have x on the left of $i, j\$$ so no inversions here. Likewise, if $i < j < x$ then x will still be on the right afterwards.

The last case is that $i < x < j$. Then we get two inversions going from i, x, j to j, x, i. One between i and x and one between x and j since both these pairs are now out of order.

To summarize, this single transposition gives us:

- 0 inversions for anything left of i or right of j,
- 2 inversions for each number in between,
- 1 inversion for the pair i, j

So overall we get an odd number of inversions for each swap. This shows that the number of inversions is odd if and only if the number of swaps is odd.

8 Exercises

1. Convert $2, 12, 3, 5, 4, 9, 8, 7, 6, 10, 1, 11$ to a product of cycles and draw the associated digraph.
2. Convert $(1, 8, 12)(2, 3, 6, 7, 9)(4, 10, 11)$ to a shuffle of $1, \ldots, 12$
3. Show that $(12)(23)(34)(23)(12) = (14)$ by drawing the braid diagram.
4. Compute the sign of the permutations in 1. and 2.
5. Write (182635) as a product of transpositions.