Chapter 3: Enumerative Combinatorics

Thomas Britz

The Inclusion-Exclusion Principle

Inclusion-Exclusion Principle

$$\left|\bigcup_{i\in I} A_i\right| = \sum_{J\subseteq I}(-1)^{|J|+1}\left|\bigcap_{j\in J} A_j\right|$$

For instance,

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

or equivalently,

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

and

$$|A\cup B\cup C| = |A| + |B| + |C| - |A\cap B| - |A\cap C| - |B\cap C| + |A\cap B\cap C|\,.$$

Lemma.

$$\sum_{J\subseteq I} t^{|J|} = (t+1)^{|I|}$$

Proof. If $I = \emptyset$, then the sum is just the term $t^{|\emptyset|} = 1$. Otherwise, choose $i \in I$. By induction on $|I|$,

$$\begin{aligned}\sum_{J\subseteq I} t^{|J|} &= \sum_{i\in J\subseteq I} t^{|J|} + \sum_{i\notin J\subseteq I} t^{|J|}\\ &= \sum_{J\subseteq I-\{i\}} t^{|J\cup\{i\}|} + \sum_{J\subseteq I-\{i\}} t^{|J|} = (t+1)\sum_{J\subseteq I-\{i\}} t^{|J|} = (t+1)(t+1)^{|I|-1} = (t+1)^{|I|}\,. \quad\square\end{aligned}$$

Lemma.

$$\sum_{K\subseteq J\subseteq I} t^{|J|} = (t+1)^{|I-K|}t^{|K|}$$

Proof.

$$\sum_{K\subseteq J\subseteq I} t^{|J|} = t^{|K|}\sum_{J'\subseteq I-K} t^{|J'|} = (t+1)^{|I-K|}t^{|K|} \quad\square$$

Corollary.

$$\sum_{K\subseteq J\subseteq I} (-1)^{|J|} = \begin{cases} (-1)^{|K|} & \text{if } K = I\\ 0 & \text{otherwise.}\end{cases}$$

In particular,

$$\sum_{J\subseteq I}(-1)^{|J|} = \begin{cases} 1 & \text{if } I = \emptyset\\ 0 & \text{otherwise.}\end{cases}$$

Proof.

$$\sum_{J\subseteq I}(-1)^{|J|} = \sum_{i\in J\subseteq I}(-1)^{|J|} + \sum_{i\notin J\subseteq I}(-1)^{|J|} = \sum_{J\subseteq I-\{i\}}(-1)^{|J\cup\{i\}|} + \sum_{J\subseteq I-\{i\}}(-1)^{|J|} = 0 \qquad \square$$

The first lemma generalises the observation that I contains $2^{|I|}$ subsets. It also generalises the following well-known result.

The Binomial Theorem

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$$

Proof. Apply the above lemma to $t = \frac{a}{b}$. That is,

$$(a+b)^n = b^n\left(\frac{a}{b}+1\right)^n = b^n \sum_{J\subseteq[n]}\left(\frac{a}{b}\right)^{|J|} = b^n\sum_{k=0}^{n}\binom{n}{k}\left(\frac{a}{b}\right)^k = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\,. \qquad \square$$

Inclusion-Exclusion Principle, Proof I.

$$\begin{aligned}\Big|\bigcup_{i\in I}A_i\Big| &= \sum_{a\in\bigcup_{i\in I}A_i}1 = \sum_{a\in\bigcup_{i\in I}A_i}\Big(1-\sum_{J\subseteq\{j:a\in A_j\}}(-1)^{|J|}\Big) = \sum_{a\in\bigcup_{i\in I}A_i}\ \sum_{\emptyset\neq J\subseteq\{j:a\in A_j\}}(-1)^{|J|+1}\\ &= \sum_{\emptyset\neq J\subseteq I}\ \sum_{a\in\bigcap_{j\in J}A_j}(-1)^{|J|+1} = \sum_{\emptyset\neq J\subseteq I}(-1)^{|J|+1}\Big|\bigcap_{j\in J}A_j\Big| = \sum_{J\subseteq I}(-1)^{|J|+1}\Big|\bigcap_{j\in J}A_j\Big|. \qquad \square\end{aligned}$$

Inclusion-Exclusion Principle, Proof II.
Use the simple Inclusion-Exclusion Principle $|A\cup B| = |A|+|B|-|A\cap B|$ and induction on $|I|$:

$$\begin{aligned}\Big|\bigcup_{i\in I}A_i\Big| = \Big|A_n\cup\bigcup_{i\in I-\{n\}}A_i\Big| &= |A_n| + \Big|\bigcup_{i\in I-\{n\}}A_i\Big| - \Big|A_n\cap\bigcup_{i\in I-\{n\}}A_i\Big|\\ &= |A_n| + \Big|\bigcup_{i\in I-\{n\}}A_i\Big| - \Big|\bigcup_{i\in I-\{n\}}A_i\cap A_n\Big|\\ &= |A_n| + \sum_{J\subseteq I-\{n\}}(-1)^{|J|+1}\Big|\bigcap_{j\in J}A_j\Big| - \sum_{J\subseteq I-\{n\}}(-1)^{|J|+1}\Big|\bigcap_{j\in J}A_j\cap A_n\Big|\\ &= \sum_{J\subseteq I}(-1)^{|J|+1}\Big|\bigcap_{j\in J}A_j\Big|. \qquad \square\end{aligned}$$

Let S be a set, G an abelian group, and $\{A_i\}_{i\in I}$ be a finite family of S-subsets.
A function $v:\mathcal{P}(S)\to G$ is a *valuation* if, for all $A,B\subseteq S$,

$$v(A)+v(B) = v(A\cup B)+v(A\cap B)\,.$$

The Inclusion-Exclusion Principle (valuations)

$$v\Big(\bigcup_{i\in I}A_i\Big) = \sum_{J\subseteq I}(-1)^{|J|+1}v\Big(\bigcap_{j\in J}A_j\Big)$$

Proof. Use induction on $|I|$:

$$\begin{aligned}
v\Big(\bigcup_{i\in I} A-i\Big) &= v\Big(A_n \cup \bigcup_{i\in I=\{n\}} A_i\Big) \\
&= v(A_n) + v\Big(\bigcup_{i\in I-\{n\}} A_i\Big) - v\Big(\bigcup_{i\in I-\{n\}} A_i \cap A_n\Big) \\
&= v(A_n) + \sum_{J\subseteq I-\{n\}} (-1)^{|J|+1} v\Big(\bigcap_{j\in J} A_j\Big) - \sum_{J\subseteq I-\{n\}} (-1)^{|J|+1} v\Big(\bigcap_{j\in J} A_j \cap A_n\Big) \\
&= \sum_{J\subseteq I} (-1)^{|J|+1} v\Big(\bigcap_{j\in J} A_j\Big)\,. \qquad \square
\end{aligned}$$

The Inclusion-Exclusion Principle (probabilities)
For a probability $P : \mathcal{P}(S) \to [0,1]$,

$$P\Big(\bigcup_{i\in I} A_i\Big) = \sum_{J\subseteq I} (-1)^{|J|+1} P\Big(\bigcap_{j\in J} A_j\Big)$$

Proof. $P(A) + P(B) = P(A\cup B) + P(A\cap B)$ for all $A, B \subseteq S$. $\square$

The Inclusion-Exclusion Principle (weights)
For $w : S \to G$, define $w : \mathcal{P}(S) \to G$ by $w(A) := \sum_{a\in A} w(a)$. Then

$$w\Big(\bigcup_{i\in I} A_i\Big) = \sum_{J\subseteq I} (-1)^{|J|+1} w\Big(\bigcap_{j\in J} A_j\Big)$$

Proof.

$$w(A) + w(B) = \sum_{a\in A} w(a) + \sum_{a\in B} w(a) = \sum_{a\in A\cup B} w(a) + \sum_{a\in A\cap B} w(a) = w(A\cup B) + w(A\cap B)\,. \qquad \square$$

Lemma. If A and B are disjoint and $v(\emptyset) = 0$, then $v(A\cup B) = v(A) + v(B)$.

Proof.
$$v(A\cup B) = v(A) + v(B) - v(A\cap B) = v(A) + v(B) - v(\emptyset) = v(A) + v(B)\,. \qquad \square$$

If each A_i is finite, then each valuation c on $\mathcal{P}(\bigcup A_i)$ is in fact a weight function on the elements of $\bigcup A_i$, since $v(A_i) = \sum_{a\in A_i} v(a)$. This is not true for uncountable sets.

Lemma.
$$\sum_{J\subseteq I : |J|\geq r+1} (-1)^{|J|+1} = (-1)^r \binom{|I|-1}{r}$$

Proof. Suppose that $i \in I$. Then

$$\begin{aligned}
\sum_{J\subseteq I:|J|\geq r+1} (-1)^{|J|+1} &= \sum_{i\in J\subseteq I:|J|\geq r+1} (-1)^{|J|+1} + \sum_{i\notin J\subseteq I:|J|\geq r+1} (-1)^{|J|+1} \\
&= \sum_{J\subseteq I-\{i\}:|J|\geq r} (-1)^{|J\cup\{i\}|+1} - \sum_{J\subseteq I-\{i\}:|J|\geq r+1} (-1)^{|J|} \\
&= \sum_{J\subseteq I-\{i\}:|J|=r} (-1)^{|J|} \\
&= (-1)^r \binom{|I|-1}{r}.
\end{aligned}$$

$\square$

The Bonferroni Inequalities (Dunn 1959)
For even s and odd t with $s, t \leq |I|$,

$$\sum_{J\subseteq I:|J|\leq s} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big| \leq \Big|\bigcup_{i\in I} A_i\Big| \leq \sum_{J\subseteq eqI:|J|\leq t} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big|.$$

Proof. For each $a \in \bigcup_{i\in I} A_i$, set $I_a := \{i : a \in A_i\}$ and choose $i_a \in I_a$. By the preceding lemma,

$$\begin{aligned}
\Delta := \Big|\bigcup_{i\in I} A_i\Big| - \sum_{J\subseteq I:|J|\leq r} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big| &= \sum_{J\subseteq I:|J|\geq r+1} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big| = \sum_{J\subseteq I:|J|\geq r+1} \sum_{a\in\bigcap_{j\in J} A_j} (-1)^{|J|+1} \\
&= \sum_{a\in\bigcup_{i\in I} A_i} \sum_{J\subseteq I_a:|J|\geq r+1} (-1)^{|J|+1} = (-1)^r \sum_{a\in\bigcup_{i\in I} A_i} \binom{|I_a|-1}{r}.
\end{aligned}$$

Thus, $\Delta < 0$ when r is odd, and $\Delta > 0$ when r is even. The Inequalities follow. $\square$

Definition. A *derangement* is a permutation $\pi \in \Sigma_n$ with no fixed point $\pi(i) = i$.
Here, Σ_n denotes the set of permutations on $[n]$.

Proposition. The number of derangements of n elements is

$$D_n := n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

Proof. For $i \in I := [n]$, let A_i be the set of permutations $\pi \in \Sigma_n$ with $\pi(i) = i$.
Note that $|\Sigma_n| = n!$ and that $\left|\bigcap_{j\in J} A_j\right| = (n - |J|)!$ for each $\emptyset \neq J \subseteq I$.
By the Inclusion-Exclusion Principle,

$$\begin{aligned}
D_n = \Big|\Sigma_n - \bigcup_{i\in I} A_i\Big| &= n! - \sum_{J\subseteq I} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big| \\
&= n! + \sum_{\emptyset\neq J\subseteq I} (-1)^{|J|}(n-|J|)! = n! + \sum_{i=1}^{n} \binom{n}{i} (-1)^i (n-i)! = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.
\end{aligned}$$

$\square$

Corollary. About e^{-1} of all permutations are derangements.

Lemma. $D_n = (-1)^n + nD_{n-1}$.

Proof.

$$D_n = n!\frac{(-1)^n}{n!} + n(n-1)!\sum_{i=0}^{n-1}\frac{(-1)^i}{i!} = (-1)^n + nD_{n-1}\,. \qquad \square$$

Corollary. Let M be an $n \times n$ matrix with 0's on the diagonal and ± 1 elsewhere.
If n is even, then $\det M \neq 0$.

Proof. Let Δ_n be the set of derangements on $[n]$ and write $M = (m_{ij})$. Then

$$\det M = \sum_{\pi \in \Sigma_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} m_{i\pi(i)} = \sum_{\pi \in \Delta_n} \prod_{i=1}^{n} m_{i\pi(i)} \equiv \sum_{\pi \in \Delta_n} 1 \equiv D_n \pmod 2\,.$$

By the above lemma, D_n is odd for even n, so $\det M$ must also be odd. $\square$

Corollary. The number of $2 \times n$ Latin rectangles is

$$n!\sum_{i=0}^{n}\frac{(-1)^i}{i!} \approx \frac{n!}{e}\,.$$

Theorem
The number of integer solutions to $x_1 + \cdots + x_n = k$ with $x_i \geq 0$ is $\binom{n+k-1}{n-1}$.

Proof. Draw k dots in a row and draw $n-1$ lines between them.
Dots bordered by lines represent one of the x_i values and vice versa.
Bijectively then, the number of solutions is the number of ways to draw these dots on lines. $\square$

Theorem
The number of integers solutions to $x_1 + \cdots + x_n = k$ with $0 \leq x_i < \ell$ is

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

Proof. Set $I := [n]$ and let S be the set of all non-negative integer solutions. Let A_i be the set of non-negative solutions with $x_i \geq \ell$ for $i \in I$. For $j \subseteq I$, $\bigcap_{j\in J} A_j$ is the set of integer solutions with $x_j \geq \ell$ for $j \in J$. By substituting $y_j := x_j - \ell$ for each $j \in J$ and $y_j = x_j$ otherwise, we get non-negative integer solutions to $y_1 + \cdots + y_n = k - \ell|J|$. There are $\binom{n+k-\ell|J|-1}{n-1}$ non-negative integer solutions to this new equation, and thus, bijectively, that many solutions to the initial equation.

By the Inclusion-Exclusion Principle, the number of solutions is

$$\begin{aligned}
\Big|S - \sum_{a \in \bigcup_{i\in I} A_i}\Big| &= \binom{n+k-1}{n-1} - \sum_{J \subseteq I}(-1)^{|J|+1}\Big|\bigcap_{j\in J} A_j\Big| \\
&= \binom{n+k-1}{n-1} + \sum_{\emptyset \neq J \subseteq I}(-1)^{|J|}\binom{n+k-\ell|J|-1}{n-1} \\
&= \binom{n+k-1}{n-1}\sum_{i=1}^{n}\binom{n}{i}(-1)^i\binom{n+k-\ell i-1}{n-1} \\
&= \sum_{i=0}^{n}(-1)^i\binom{n+k-\ell i-1}{n-1}\,.
\end{aligned} \qquad \square$$

Definition. *Euler's totient* $\phi(n)$ is the function given by

$$\phi(n) := \Big|\{i \in [n] : \gcd(n,i) = 1\}\Big|$$

Example.

$$\begin{aligned}\phi(8) &= \big|\{1,3,5,7\}\big| = 4\\ \phi(12) &= \big|\{1,5,7,11\}\big| = 4\\ \phi(7) &= \big|\{1,2,3,4,5,6\}\big| = 6\,.\end{aligned}$$

Theorem

$$\phi(n) = n \prod_{p|n\,,\, p \text{ prime}} \Big(1 - \frac{1}{p}\Big)$$

Proof. Write $n = p_1^{\alpha_1} \dots p_s^{\alpha_s}$, where p_i is prime and set $A_i = \{m \in [n] : p_i | m\}$ for $i \in I := [s]$.
For $J \subseteq I$, $\big|\bigcap_{j\in J} A_j\big| = \frac{n}{\prod_{j\in J} p_j}$. By the Inclusion-Exclusion Principle,

$$\begin{aligned}\phi(n) = \Big|[n] - \bigcup_{i\in I} A_i\Big| &= n - \sum_{J\subseteq I} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big|\\ &= n + \sum_{\emptyset\neq J\subseteq I} (-1)^{|J|} \frac{n}{\prod\limits_{j\in J} p_j} = n \sum_{J\subseteq I} \prod_{j\in J} \frac{-1}{p_j} = n \prod_{i=1}^{s} \Big(1 - \frac{1}{p_i}\Big)\,.\end{aligned}$$ □

Example. Using the above theorem,

$$\begin{aligned}\phi(8) &= 8(1-\tfrac{1}{2}) = 4 & \phi(7) &= 7(1-\tfrac{1}{7}) = 6\\ \phi(12) &= 12(1-\tfrac{1}{2})(1-\tfrac{1}{3}) = 4 & \phi(1000) &= 1000(1-\tfrac{1}{2})(1-\tfrac{1}{5}) = 400\end{aligned}$$

Definition. Let $G = (V,E)$ be a graph.
A colouring $\chi : V \to [r]$ is *proper* if $\chi(u) \neq \chi(v)$ whenever $\{u,v\} \in E$.
Let $c(J)$ be the number of connected components in the subgraph of G with vertices V and edges $J \subseteq E$.

Theorem (Birkhoff 1912)
The number of proper r-colourings of $V(G)$ is $P(G;r) := \sum\limits_{J\subseteq E} (-1)^{|J|} r^{c(J)}$.

Proof. Let C denote the set of all r-colourings of V. Then $|C| = r^{|V|}$. For $e = \{u,v\} \in E$, let A_e be the colourings $\chi \in C$ with $\chi(U) = \chi(V)$. For each $J \subseteq E$, each colouring in $E_J := \bigcap_{e\in J} A_e$ is monochromatic in each connected component of (V, E_J) and the component colours are independent of each other. Hence, $\big|\bigcap_{e\in J} A_e\big| = r^{c(A)}$.

By the Inclusion-Exclusion Principle, the number of proper colourings is

$$\Big|C - \bigcup_{e\in E} A_e\Big| = r^{|V|} - \sum_{J\subseteq E} (-1)^{|J|+1} \Big|\bigcap_{j\in J} A_j\Big| = r^{|V|} + \sum_{\emptyset\neq J\subseteq E} (-1)^{|J|} r^{c(J)} = \sum_{J\subseteq E} (-1)^{|J|} r^{c(J)}\,.$$ □

Example.

$$P(K_n; r) = r(r-1)\dots(r-n+1)$$

Recursive identities allow efficient calculations of $P(G;r)$.

Möbius Inversion

Recall that a *poset* P is a set with a partial order $\preceq$ satisfying, for all $x, y, z \in P$,

(R) $x \preceq x$;
(A) if $x \preceq y$ and $y \preceq x$, then $x = y$;
(T) if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

Definition. A *lattice* L is a poset with binary operators *meet* $\wedge$ and *join* $\vee$ so that

- $x \wedge y$ is the *unique* maximal element smaller than both x and y.
- $x \vee y$ is the *unique* minimal element greater than both x and y.

A lattice is *graded* if it has a well-defined rank function

$$\ell(x) := |C| - 1$$

where C is any maximal chain with $\max(C) = x$.

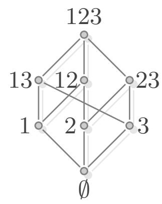
Example. $[n]$ is a graded lattice under the usual order $\leq$:

$$x \wedge y = \min\{x, y\}$$
$$x \vee y = \max\{x, y\}$$
$$\ell(x) = x - 1$$

[4]

Example. $\mathcal{B}_n$ $(= \mathcal{P}(S))$ is a graded lattice under containment $\subseteq$:

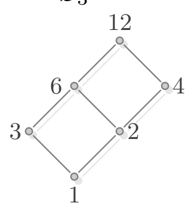
$$A \wedge B = A \cap B$$
$$A \vee B = A \cup B$$
$$\ell(A) = |A|$$



$\mathcal{B}_3$

Example. The divisors of n form a graded lattice under division $|$:

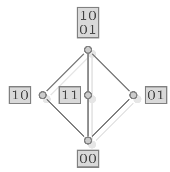
$$a \wedge b = \gcd(a, b)$$
$$a \vee b = \operatorname{lcm}(a, b)$$
$$\ell(a) = \sum_{i=1}^{s} \alpha_i, \text{ where } a = p_1^{\alpha_1} \dots p_s^{\alpha_s}$$



$(\{a : a \,|\, 12\}, |)$

Example. The subspaces of a vector space form a graded lattice under containment $\subseteq$:

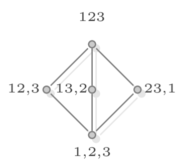
$$U \wedge V = U \cap V$$
$$U \vee V = \operatorname{span}(U \cup V)$$
$$\ell(U) = \dim(U)$$



Subspaces of $\mathbb{F}_2^2$

Example. The partitions π_n of $[n]$ form a lattice under refinement $\preceq$:

$$\ell(\pi) = n - |\pi|$$



Π_3

Definition. Let P be a finite poset. The *zeta function* of P is the matrix ζ with entries

$$\zeta(x,y) = \begin{cases} 1 & x \le y \\ 0 & \text{otherwise.} \end{cases}$$

The *Möbius function* of P is the inverse $\mu := \zeta^{-1}$. By antisymmetry, the row- and column indices of ζ may be ordered so it is upper-triangular. In that case, μ is also upper-triangular.

Theorem
Let $c_k(x,y)$ denote the number of chains in P of size k from x to y. Then

$$\mu(x,y) = \sum_{k\ge1} (-1)^{k-1} c_k(x,y)\,.$$

The proof is left as an **exercise**!

Example. For the previous example with the lattice $[n]$ under the usual order $\le$,

$$\zeta = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad \text{In general,} \quad \mu(x,y) = \begin{cases} 1 & x = y \\ -1 & x = y-1 \\ 0 & \text{otherwise.} \end{cases}$$

Example. For the previous example with the Boolean lattice $\mathcal{B}_n$ under inclusion $\subseteq$,

$$\mu(A,B) = \begin{cases} (-1)^{\ell(B)-\ell(A)} = (-1)^{|B-A|} & A \subseteq B \\ 0 & \text{otherwise.} \end{cases}$$

To prove this, we have that for $A \subseteq B$,

$$(\zeta\mu)_{AB} = \sum_C \zeta(A,C)\mu(C,B) \;=\; \sum_{A\subseteq C\subseteq B} (-1)^{|B-C|} = (-1)^{|B|} \sum_{A\subseteq C\subseteq B} (-1)^{|C|} = \begin{cases} ((-1)^{|B|})^2 & A = B \\ 0 & \text{otherwise.} \end{cases}$$

Hence, $\zeta\mu = I$. Similarly, $\mu\zeta = I$, so $\mu = \zeta^{-1}$. $\square$

Example. For the previous example with the lattice of divisors of n under division $|$,

$$\begin{aligned} \mu(a,b) &= \begin{cases} (-1)^{\ell(b)-\ell(a)} = (-1)^{\ell(b/a)} & a|b \\ 0 & \text{otherwise.} \end{cases} \\ &= \mu\Big(\frac{b}{a}\Big)\,. \end{aligned}$$

Example. For the previous example with the subspaces of $\mathbb{F}_2^2$,

$$\zeta = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} 1 & -1 & -1 & -1 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

In general, if the vector space is over $\mathbb{F}_q$, then

$$\mu(U,V) = \begin{cases} (-1)^k q^{\binom{k}{2}} & U \subseteq V \\ 0 & \text{otherwise} \end{cases}$$

where $k = \ell(V) - \ell(U) = \dim(V) - \dim(U)$.

Möbius Inversion

$$g = \zeta f \iff f = \mu g$$
$$h = f\zeta \iff f = h\mu\,.$$

Equivalently,

$$g(x) = \sum_{a : a \le x} f(a) \text{ for all } x \in P \iff f(x) = \sum_{a : a \le x} \mu(a,x) g(a) \text{ for all } x \in P$$
$$h(x) = \sum_{b : b \ge x} f(b) \text{ for all } x \in P \iff f(x) = \sum_{b : b \ge x} \mu(x,b) h(b) \text{ for all } x \in P\,.$$

Möbius Inversion is a strong generalisation of the Inclusion-Exclusion Principle. As the following proof shows, the latter is a special case of the former.

Inclusion-Exclusion Principle, Proof III.
Let $P := \mathcal{P}(I) \equiv \mathcal{B}_{|I|}$ and for each $\emptyset \neq J \in P$, define

$$f(J) := |\{a : a \in A_j \text{ if and only if } j \in J\}|$$
$$g(J) := |\{a : a \in A_j \text{ for all } j \in J\}| = \Big|\bigcap_{j \in J} A_j\Big|,$$

and $f(\emptyset) := 0$ and $g(\emptyset) := \left|\bigcup_{i \in I} A_i\right|$.
Then $g(J) = \sum_{K \supseteq J} f(K)$ for all $J \in P$. By Möbius Inversion,

$$f(J) = \sum_{K \supseteq J} \mu(J,K) g(K) = \sum_{K \supseteq J} (-1)^{|K-J|} g(K)\,.$$

Hence,

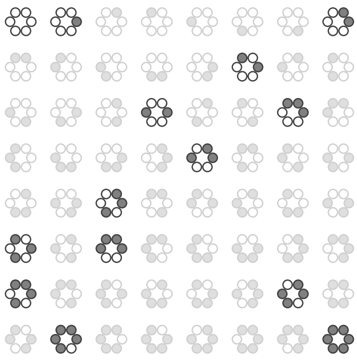
$$\begin{aligned}\Big|\bigcup_{i \in I} A_i\Big| = g(\emptyset) &= \sum_{K \subseteq I} (-1)^{|K|} g(K) - \sum_{\emptyset \neq K \subseteq I} (-1)^{|K|} g(K) \\ &= f(\emptyset) - \sum_{\emptyset \neq K \subseteq I} (-1)^{|K|} g(K) = 0 - \sum_{\emptyset \neq K \subseteq I} (-1)^{|K|} \Big|\bigcap_{j \in K} A_j\Big| = \sum_{\emptyset \neq J \subseteq I} (-1)^{|J|+1} \Big|\bigcap_{j \in J} A_j\Big|\,. \quad \square\end{aligned}$$

Pólya Counting

Pólya counting allows us to approach problems such as counting the number of configurations of a Rubik's cube and the number of different ways to tile a surface with some given tiles. These questions assume certain symmetries that we ignore and just count equivalence classes.

Example. We want to count the number of bracelets that can be made with six beads that are each either black or white. If you are wearing the bracelet, then rotating it doesn't change it. If you take it off however, then it is seen to be the same bracelet if you either rotate or turn it around (i.e., reflect it).

We can of course just count by brute force:



Of the $2^6 = 64$ strictly different bracelets, 14 are distinct up to rotation. If we ignore reflections, then the two bracelets  and  are seen as identical, and so we have only 13 distinct bracelets.

This brute force approach is not feasible in general, and we will now show to count objects under symmetries in more systematics fashion.

Burnside's Lemma (Cauchy 1845, Frobenius 1887)
The number of orbits of a permutation group G acting on a set X if

$$\frac{1}{|G|}\sum_{g\in G}\phi(g)$$

where $\phi(g) := |\{x \in X : g(x) = x\}|$ is the number of fixed points of g.

Proof. For each $x \in X$, let $O_x = \{g(x) : g \in G\}$ and $S_x = \{g \in G : g(x) = x\}$ be the orbit and stabilise of x in G. Let $\mathcal{O}$ be the set of orbits of G. By the Orbit-Stabiliser Theorem, $|G| = |S_x||O_x|$, and hence,

$$\begin{aligned}|\mathcal{O}| = \sum_{O\in\mathcal{O}} 1 = \sum_{O\in\mathcal{O}}\frac{|O|}{|O|} = \sum_{O\in\mathcal{O}}\sum_{x\in O}\frac{1}{|O|} = \sum_{x\in X}\sum_{\substack{O\in\mathcal{O}:\\ O=O_x}}\frac{1}{|O|} = \sum_{x\in X}\frac{1}{|O_x|} = \frac{1}{|G|}\sum_{x\in X}|S_x| \\ = \frac{1}{|G|}\sum_{x\in X}\sum_{\substack{g\in G:\\ g(x)=x}} 1 = \frac{1}{|G|}\sum_{g\in G}\sum_{\substack{x\in X:\\ g(x)=x}} 1 = \frac{1}{|G|}\sum_{g\in G}\phi(g)\,. \qquad \square\end{aligned}$$

Example. We can use Burnside's Lemma to count the bracelet from the example above. Here, $G = D_6$, the dihedral group of rotations and reflections on a hexagon. The orbits of G are the bracelet equivalence classes that we wish to count. The group G is generated by a $60°$ rotation r and a reflection s.

Each bracelet $x \in X$ is a fixed point of the identity $e \in G$, so $\phi(e) = 64$. For $g := r, r^{-1}$, the fixed points of g are the bracelets  and , so $\phi(g) = 2$.

We can continue to find all of the values of $\phi(g)$. They are

$$64, 2, 2, 4, 4, 8, 16, 8, 16, 8, 16, 8\,.$$

By Burnside's Lemma, the number of orbits, and thus bracelets, is

$$\frac{1}{|G|}\sum_{g\in G}\phi(g) = \frac{1}{12}(64+2+2+4+4+8+16+8+16+8+16+8) = \frac{156}{12} = 13\,.$$

Example. Consider bead strings of length two, in three colours. The group G is here just $\{e, s\}$ where e is the identity element and where s swaps the beads. Then $\phi(e) = 9$ and $\phi(s) = 3$.

By Burnside's Lemma, the number of these bead strings is

$$\frac{1}{|G|}\sum_{g\in G}\phi(g) = \frac{1}{2}(3+9) = 6\,.$$

Example. Now consider bead strings of length three, in black and white.
The group is G is again just $\{e, s\}$. Then $\phi(e) = 8$, and $\phi(s) = 4$.

By Burnside's Lemma, the number of these bead strings is

$$\frac{1}{|G|}\sum_{g\in G}\phi(g) = \frac{1}{2}(8+4) = 6\,.$$

Let A be a set and B be a set of colours, both finite.
Let G be a group of permutations acting on A.
Further, let G act on B^A, the B-colourings of A, by

$$\sigma(f) := f(\sigma^{-1}(x))\,.$$

Example. Consider again the bracelet problem. Here, we could let

- A be the set of uncoloured beads in the bracelet.
- B be the set of colours
- B^A be the bracelet colourings, i.e. the coloured bracelets.
- G be the dihedral group acting on A.

As before, the orbits of G count the different bracelets.

Definition. Let π be a permutation.
The *cycle decomposition* of π is the unique product of cycles $\pi = \sigma_1 \dots \sigma_n$.
The *type* of π is the formal product $1^{z_1} 2^{z_2} \dots$, with $z_i = |\{j : |\sigma_j| = i\}|$.

Example. The permutation
$$\pi = \binom{12345678}{34516287}$$
has cycle decomposition
$$\pi = \binom{78}{87}\binom{134}{341}\binom{256}{562}$$
and thus has type $2^1 3^2$.

Definition. The *cycle index* of a group G is the polynomial
$$Z_G(X_1, \dots, X_n) := \frac{1}{|G|} \sum_{\sigma \in G} X_1^{z_1(\sigma)} \dots X_n^{z_n(\sigma)}.$$

Example. Let G be the cyclic group C_n. For each divisor $d|n$, there are $\phi(d)$ permutations of order d, and each of these has $\frac{n}{d}$ cycles of length d. Thus, the cycle index of C_n is
$$Z_G(X_1, \dots, X_n) = \frac{1}{n} \sum_{d|n} \phi(d) X_d^{n/d}.$$

Theorem
Let $c_k(G)$ be the number of permutations with exactly k cycles. The number of orbits on B^A is
$$\frac{1}{|G|} \sum_{k=1}^{\infty} c_k(G) |B|^k.$$

Proof. A fixed point of $\sigma \in G$ is a colouring $f \in B^A$ for which $\sigma(f) = f$.
That is, $f(a) = f(\sigma^{-1}(a))$ for all $a \in A$. Hence, f is constant on each cycle σ_i of σ.

Conversely, if f is constant on each cycle σ_i, then f is a fixed point for σ. Hence, if σ has k cycles, then the number of fixed points for σ is $|B|^k$. By Burnside's Lemma, the number of orbits is thus
$$\frac{1}{|G|} \sum_{\sigma \in G} \phi(x) = \frac{1}{|G|} \sum_{k=1}^{\infty} \sum_{\substack{\sigma \in G: \\ \sigma \text{ has } k \text{ cycles}}} \phi(\sigma) = \frac{1}{|G|} \sum_{k=1}^{\infty} \sum_{\substack{\sigma \in G: \\ \sigma \text{ has } k \text{ cycles}}} |B|^k = \frac{1}{|G|} \sum_{k=1}^{\infty} c_k(G) |B|^k. \qquad \square$$

The above theorem may be re-expressed in the following way.

Pólya's Theorem (simple)
The number of orbits of G on B^A is
$$Z_G(|B|, \dots, |B|) = \frac{1}{|G|} \sum_{\sigma \in G} |B|^{z_1(\sigma) + \dots + z_n(\sigma)}.$$

This result can be generalised significantly.

Example. Consider bracelets with n beads in b colours identical under rotational symmetry but distinct under reflections. Then G is the cyclic group C_n and $|B| = b$.
By Pólya's Theorem, the number of distinct bracelets with n beads is

$$Z_G(b, \ldots, b) = \frac{1}{n} \sum_{d|n} \varphi(d) b^{n/d} .$$

For $n = 6$,

$$Z_G(b, \ldots, b) = \frac{1}{6}\big(\varphi(1)b^{6/1} + \varphi(2)b^{6/2} + \varphi(3)b^{6/3} + \varphi(6)b^{6/6}\big) = \frac{1}{6}\big(b^6 + b^3 + 2b^2 + 2b\big) .$$

For instance, the number of bracelets with 6 beads in $b = 3$ colours is

$$\frac{1}{6}\big(3^6 + 3^3 + 2 \cdot 3^2 + 2 \cdot 3\big) = 130 .$$

As we have seen, the number of bracelets with 6 beads in $b = 2$ colours is

$$\frac{1}{6}\big(2^6 + 2^3 + 2 \cdot 2^2 + 2 \cdot 2\big) = 14 .$$

Generating Functions

To appear.

References

[1] M. Aigner, *Combinatorial Theory*, Springer-Verlag, New York, 1979.

[2] R.L. Graham, M. Grötschel, and L. Lovász (eds.), *Handbook of Combinatorics. I–II*, North-Holland, Amsterdam, 1995.

[3] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992.

[4] C.L. Liu, *Introduction to Combinatorial Mathematics*, McGraw-Hill Book Co., New York, 1968.

[5] R.P. Stanley, *Enumerative Combinatorics. Vol. 1*, Cambridge University Press, 1997.