Discrete Mathematics lecture notes 13-2

December 8, 2013

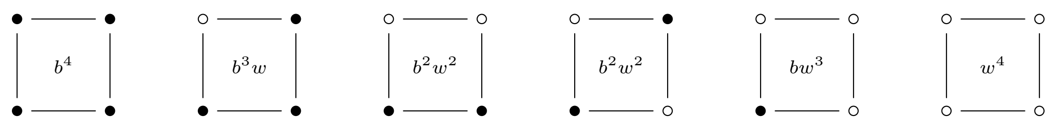
34. Pólya's Enumeration Theorem

We started examining the number of essentially distinct colorings of an object with questions of the form "How many necklaces can be made with 3 black beads and white beads?" We were able to answer these questions in a fairly *ad hoc* manner using the Lemma that is not Burnside's.

A closer examination of this method of counting led to the previous lecture's result: If the cycle index $\zeta_G(x_1, \ldots, x_n)$ of a group G of permutations of a set X is known, then X has $\zeta_G(r, r, \ldots, r)$ essentially G-distinct r-colorings.

While this more recent result is more universal, and in particular deals with all r-coloring in the same manner, it is not really a generalization of the earlier *ad hoc* counting results: We can easily compute that there are 6 essentially distinct necklaces with four beads, which can either be black or white, but this does not tell us how many necklaces are possible with a specified number of black and white beads.

Below are representatives of the 6 essentially distinct black and white colorings of the vertices of the square, which you might recognize as the first column of the example worked out in the previous lecture:



Each representative is labeled with an expression of the form b^iw^j, where i is the number of black beads and j the number of white for that coloring. We have chosen a somewhat suggestive notation: Instead of having b^3w be simply a shorthand for "there are three black beads and one white here," we'll view the expression b^3w as a monomial in the formal variables b and w. This formalism allows us to manipulate the labelings as if they were polynomials, and in particular, add all the labelings:

$$\begin{aligned} U_D(b, w) &= b^4 + b^3w + b^2w^2 + b^2w^2 + bw^3 + w^4 \\ &= b^4 + b^3w + 2b^2w^2 + bw^3 + w^4. \end{aligned}$$

Here, U_D is the *generating function for the number of distinct 2-colorings of the square*, a formal polynomial in the variables b and w.

Note that every monomial that makes up U_D is of total degree 4.[1] This makes sense, as the exponent on b is the number of black beads in a given coloring, while the exponent on w is the number of white beads; as there are a total of 4 beads to color either black or white, the observation follows.

Of greater interest is the again obvious[2] Observation that the *coefficient* of the term b^iw^j in U_D is the number of colorings with i black beads a and j white. Of course, U_D was defined by taking the sum of monomials with one terms for each essentially distinct coloring, so again this should not be surprising. However, these easy observations are actually hinting at a much deeper counting method, one that relies on the new way of packaging information that is the generating function.

Recall the notation from the previous lecture: X a finite set of size n, G a permutation group of X, $F = \{f_1, f_2, \ldots, f_r\}$ a set of r colors, and Ω the set of colorings $\omega : X \to F$. We generalize the generating function U_D above by thinking of the f_i as formal variables, not just elements of a set. Thus expressions of the form $f_1f_2^3f_3^3$, $f_1^2f_3^4 + f_2^3f_4^3$, etc., make sense.

Definition 1. Suppose that $\omega : X \to F$ is a coloring that assigns a_i elements of X to the color f_i. The *indicator* of ω is the polynomial in the formal variables $\{f_i\}$:

$$\operatorname{ind}(\omega) = f_1^{a_1} f_2^{a_2} \ldots f_r^{a_r}.$$

[1] By definition, the *total degree* of a monomial $x_1^{n_1}x_2^{n_2}\ldots x_k^{n_k}$ is $n_1 + n_2 + \ldots + n_k$. If every monomial of a polynomial $p(x_1, x_2, \ldots, x_k)$ is of the same degree d, we say p is *homogeneous of degree d*.

[2] At least in retrospect.

For example, in each of the six colorings listed above, the expression $b^i w^j$ is the indicator of the corresponding coloring.

We are interested in answering the question, "How many essentially distinct colorings of a set X are possible subject to the condition that a_i elements of X are sent to color f_i?" We'll actually introduce a more general notion[3] in order to find the general solution.

Definition 2. Let $A \subseteq \Omega$ be a set of r-colorings of X. Define the *generating function of A* to be the polynomial U_A in the variables $\{f_1, f_2, \ldots, f_n\}$, defined by

$$U_A(f_1, f_2, \ldots, f_n) = \sum_{\omega \in A} \operatorname{ind}(\omega).$$

By construction of U_A, the coefficient of the monomial $f_1^{a_1} f_2^{a_2} \ldots f_n^{a_n}$ is the number of colorings of A that send a_i elements of X to the color f_i.

The following observation is just a definition-check:

Lemma 3. *If $\omega \in \Omega$, then $\operatorname{ind}(\omega)$ is a monomial of degree $n = |X|$, and consequently U_A is a homogeneous polynomial of degree n for any $\emptyset \neq A \subseteq \Omega$.*

Example 4. Let $X = \{x_1, x_2, \ldots, x_n\}$, $F = \{b, w\}$, and $A = \Omega$. Then the coefficient of $b^i w^j$ is the number of ways i elements of X can be colored black and $j = n - i$ elements can be colored white; by definition of binomial coefficients, this number is $\binom{n}{i} = \binom{n}{j}$. Thus

$$U_\Omega(b, w) = \sum_{i=0}^{n} \binom{n}{i} b^i w^{n-i} = (b + w)^n,$$

where the last equality is of course just a restatement of the Binomial Theorem. From another point of view: Generating functions are actually not a new concept for us, and we're now just generalizing the ideas inherent in the Binomial, Trinomial, etc. Theorems.

The main structure advantage of thinking in terms of generating functions is that we can now avail ourselves of the algebraic properties of polynomials: We can add, subtract, multiply, and (sometimes) divide them, which in general can have powerful implications for whatever it is we're interested in counting. For now, we're most interested in the fact that polynomials can also be *composed*; the effect of this observation lead to the following elegant solution to all of our counting problems.

Theorem 5 (Pólya's Enumeration Theorem)**.** *Let X, G, F, Ω be as above, and let $D \subseteq \Omega$ be a set of representatives of the $\widehat{G}$-orbits on Ω.*[4] *If $\zeta_G(x_1, x_2, \ldots, x_n)$ is the cycle index of G and $\alpha_i = f_1^i + f_2^i + \ldots + f_r^i$ for $1 \leq i \leq n$, the generating function for D is*

$$U_D(f_1, f_2, \ldots, f_n) = \zeta_G(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

Example 6. In the case of the 2-colorings (with $f_1 = b$ and $f_2 = w$) of the square: We calculated last time that the cycle index of D_8 is

$$\zeta_{D_8}(x_1, x_2, x_3, x_4) = \frac{1}{8}(x_1^4 + 3x_2^2 + 2x_4 + 2x_1^2 x_2).$$

Since we only have two colors to worry about, it is relatively easy to evaluate ζ_{D_8} with $x_i = \alpha_i := b^i + w^i$:

$$\begin{aligned} \zeta_{D_8}(\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= \frac{1}{8}\big((b+w)^4 + 3(b^2+w^2)^2 + 2(b^4+w^4) + 2(b+w)^2(b^2+w^2)\big) \\ &= b^4 + b^3 w + 2b^2 w^2 + b w^3 + w^4, \end{aligned}$$

which is fortuitously the generating function $U_D(b, w)$ we calculated in the introduction.

[3] I.e., we're equipping ourselves to deal with more general sets of colorings than just "all essentially distinct" ones.

[4] Recall from the previous lecture that $\widehat{G}$ is the group of permutations of Ω induced by G. A "set of representatives of the orbits" is then a set of colorings of X, such that no two are equivalent under the induced G-action, and any other coloring is equivalent to one of our set.

One might argue that Pólya's Theorem has the potential to not be well defined: If A is one set of representatives of the $\widehat{G}$-orbits of Ω, and A' is a second, does it necessarily follow that $U_A = U_{A'}$?

Lemma 7. *For all $\omega \in \Omega$ and $g \in G$, the indicators of ω and $\widehat{g}(\omega)$ are equal:*

$$\operatorname{ind}(\omega) = \operatorname{ind}(\widehat{g}(\omega)).$$

Thus ind *is constant on the $\widehat{G}$-orbits of Ω, so U_A does not depend on the choice of representatives A.*

Proof. Both $\operatorname{ind}(\omega)$ and $\operatorname{ind}(\widehat{g}(\omega))$ are degree n monomials in the variables $f_1, f_2, \ldots, f_r$:

$$\operatorname{ind}(\omega) = f_1^{a_1} f_2^{a_2} \ldots f_r^{a_r} \qquad \text{and} \qquad \operatorname{ind}(\widehat{g}(\omega)) = f_1^{a_1'} f_2^{a_2'} \ldots f_r^{a_r'},$$

where ω (resp. $\widehat{g}(\omega)$) sends a_i (resp. a_i') elements of X to the color f_i. Recall that the coloring $\widehat{g}(\omega)$ is defined by the formula

$$(\widehat{g}(\omega))(x) = \omega(g^{-1}(x))$$

for all $x \in X$. Thus if $x_1, x_2, \ldots, x_{a_1}$ are the elements of X colored f_1 by ω, then $g(x_1), g(x_2), \ldots, g(x_{a_1})$ are precisely the elements of X colored f_1 by $\widehat{g}(\omega)$. Of course, the same is true with a_1 replaced by a_i and f_1 by f_i for any $1 \leq i \leq r$, so we see that ω and $\widehat{g}(\omega)$ send the same *number* of elements of X to each color (though the particular elements so colored can of course change). Because ind is defined simply in terms of the number of elements sent to each color, we conclude that $a_i = a_i'$ for all $1 \leq i \leq r$. The result follows. □

This Lemma suggests a generalization of the Lemma that is not Burnside's that will be useful in the proof of Pólya's Theorem:

Theorem 8 (Weighted Burnside)**.** *Let X be a set with permutation group G, R a ring,*[5] *and let $f : X \to R$ be constant on all G-orbits, i.e., $f(x) = f(g(x))$ for all $x \in X$ and $g \in G$. If $D \subseteq X$ is a set set of representatives of $G \backslash X$, then*

$$\sum_{x \in D} f(x) = \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X^g} f(x).$$

Proof. Note that if $f(x) = 1$ for all $x \in X$, this reduces to the statement of Burnside we had encountered a few lectures ago. This suggests, correctly, that we might get some mileage from the proof technique we used earlier.

Define $E := \{(x, g) \in X \times G | g(x) = x\}$. We will calculate $S := \sum_{(x,g) \in E} f(x)$ in two different ways (depending on the order of summing x and g), the equality of the two results will lead to the different sides of the desired equality.

Fix x and define $C_x := \sum_{\substack{g \in G \\ (x,g) \in E}} f(x)$. Since the summand $f(x)$ does not change with g, we see that C_x is just $f(x) \cdot |\{g \in G | (x, g) \in E\}| = f(x) \cdot |G_x|$. Since $S = \sum_{x \in X} C_x$, we have $S = \sum_{x \in X} f(x) \cdot |G_x|$. If $X_1, X_2, \ldots, X_k$ are are the G-orbits, and $x_1, \ldots, x_k$ are the representatives of X_i that live in D, then (since both $f(x)$ and $|G_x|$ are constant on G-orbits), we see

$$S = \sum_{x \in X} f(x) \cdot |G_x| = \sum_{x_i \in D} \sum_{x \in X_i} f(x) \cdot |G_x| = \sum_{x_i \in D} f(x) \cdot |G_x| \cdot |X_i| = \sum_{x_i \in D} f(x) \cdot |G| = |G| \sum_{x \in D} f(x).$$

For the second method of computing S, fix $g \in G$ and set $R_g := \sum_{\substack{x \in X \\ (x,g) \in E}} f(x)$. By definition of E, this is just $\sum_{x \in X^g}$; since $S = \sum_{g \in G} R_g$, we conclude $S = \sum_{g \in G} \sum_{x \in X^g} f(x) = |G| \sum_{x \in D} f(x)$. The result follows. □

[5] Recall that this simply means a set equipped with an addition and commutative multiplication governed by the axioms we described governing the arithmetic of $\mathbb{Z}$ at the beginning of the course. The relevant examples are $R = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and polynomial rings like $\mathbb{Z}[x]$ or $\mathbb{Z}[f_1, f_2, \ldots, f_r]$, the latter being the ring of polynomials with variables $\{f_i\}$, where addition and multiplication defined in the obvious manner.

We can now fit all the pieces together.

Proof (Pólya's Enumeration Theorem). Recall that D is a set of representatives of $\widehat{G}\backslash\Omega$. By the definition of the generating function for D, we have

$$U_D(f_1, f_2, \ldots, f_r) = \sum_{\omega \in D} \text{ind}(\omega).$$

As ind is constant on $\widehat{G}$-orbits of Ω the weighted version of the Lemma that is not Burnside's implies

$$U_D(f_1, f_2, \ldots, f_n) = \frac{1}{|\widehat{G}|} \sum_{\widehat{g} \in \widehat{G}} \sum_{x \in \Omega^{\widehat{g}}} \text{ind}(\omega) = \frac{1}{|G|} \sum_{g \in G} \left[\sum_{x \in \Omega^{\widehat{g}}} \text{ind}(\omega) \right],$$

where the second equality uses the fact that the map $G \to \widehat{G}$ is an injection. Let's focus our attention on the sum set off in square brackets.

Inside the square brackets, the element $g \in G$ is fixed, hence $\widehat{g}$ is as well. By definition of the generating function, the bracketed term is actually $U_{\Omega^{\widehat{g}}}(f_1, f_2, \ldots, f_r)$, i.e., the generating function for the colorings fixed by $\widehat{g}$. What does such a coloring look like?

For $\omega \in \Omega$, we have $\omega \in \Omega^{\widehat{g}}$ if and only if $\widehat{g}(\omega) = \omega$ if and only if for all $x \in X$

$$\widehat{g}(\omega)(x) = \omega(g^{-1}(x)) = \omega(x).$$

It follows that ω is a $\widehat{g}$-fixed coloring if and only if ω is constant on each cycle of g: If one can get from x to y by repeated application of g, ω must assign both x and y the same color.

If g has cycle type $[1^{c_1} 2^{c_2} \ldots n^{c_n}]$[6], set $k := \sum_{i=1}^{n} c_i$, so that k is the number of g-orbits of X.[7] In particular, X is partitioned $X = X_1 \amalg X_2 \amalg \ldots \amalg X_k$, where the X_i are the distinct g-cycles of X. By the above, any $\widehat{g}$-fixed coloring ω must be constant on each of the X_i, but any of the r colors of F are possible choices on each g-cycle.

Therefore, if $\ell_i := |X_i|$ and ω assigns color f_j to X_i, those ℓ_i elements contribute a total of $f_j^{\ell_i}$ to $\text{ind}(\omega)$. Taking all of the g-cycles into account, and allowing for all choices of color on each X_i, it follows that

$$U_{\Omega^{\widehat{g}}}(f_1, f_2, \ldots, f_r) = (f_1^{\ell_1} + f_2^{\ell_1} + \ldots + f_r^{\ell_1}) \cdot (f_1^{\ell_2} + f_2^{\ell_2} + \ldots + f_r^{\ell_2}) \cdot \ldots \cdot (f_1^{\ell_k} + f_2^{\ell_k} + \ldots + f_r^{\ell_k}).$$

Said differently, when we expand out this product, in each of the k parenthetical sums we have r choices of color; if we choose color f_j in the ith spot, that corresponds to the coloring that is f_j on X_i. As we range over all possible choices of how to expand the terms of the product, we get all the possible choices of how to color X_i with the r colors available, and the claim follows.

Recall that for $g \in G$, the cycle type of g gives rise to the monomial $\zeta_g(x_1, x_2, \ldots, x_n) = x_1^{c_1} x_2^{c_2} \ldots x_n^{c_n}$. If we evaluate ζ_g at $x_i = \alpha_i = f_1^i + f_2^i + \ldots + f_r^i$, we get

$$\zeta_g(\alpha_1, \alpha_2, \ldots, \alpha_n) = (f_1 + f_2 + \ldots + f_r)^{c_1} \cdot (f_1^2 + f_2^2 + \ldots + f_r^2)^{c_2} \cdot \ldots \cdot (f_1^n + f_2^n + \ldots + f_r^n)^{c_n}$$

Compare this to the expression for $U_{\Omega^{\widehat{g}}}$. In that expression, we may assume that the ℓ_i are clustered such that $\ell_1 = \ell_2 = \ldots = \ell_{c_1}$; $\ell_{c_1+1} = \ell_{c_1+2} = \ldots = \ell_{c_1+c_2}$; etc. In other words, the first c_1 terms represent the g-cycles of length 1, the next c_2 terms the g-cycles of length 2, etc. In these terms, it is clear that

$$U_{\Omega^{\widehat{g}}}(f_1, f_2, \ldots, f_r) = \zeta_g(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

[6] Recall that this means that, when written in terms of disjoint cycles, g has c_1 cycles of length 1, c_2 cycles of length 2, etc.

[7] Note the distinction between G-orbits and g-orbits: In the latter case, we simply mean those element of X that can be joined by repeated application of the permutation g; in the former, we may use any permutation of G to move around X.

Returning to the beginning, we have

$$\begin{aligned} U_D(f_1, f_2, \ldots, f_n) &= \frac{1}{|G|}\sum_{g\in G}\left[\sum_{x\in\Omega^{\bar{g}}}\mathrm{ind}(\omega)\right] \\ &= \frac{1}{|G|}\sum_{g\in G} U_{\Omega^{\bar{g}}}(f_1, f_2, \ldots, f_r) \\ &= \frac{1}{|G|}\sum_{g\in G}\zeta_g(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ &= \zeta_G(\alpha_1, \alpha_2, \ldots, \alpha_n), \end{aligned}$$

and the Theorem is proved. □

Example 9. Let's look at an example that would be a pain to work out by hand, but fairly easy with this new technology: How many essentially distinct necklaces can be made with 3 red, 4 blue, and 4 green beads? We start by noting that there are 11 beads total, so we're dealing with the group of symmetries of the regular 11-gon, known as D_{22}. By inspection,[8] we compute

$$\zeta_{D_{22}} = \frac{1}{22}(x_1^{11} + 10x_{11} + 11x_1x_2^5).$$

By Pólya's Enumeration Theorem, the generating function for the number of essentially distinct 3-colorings of the vertices of the 11-gon is given by

$$\zeta_{D_{22}}(r+b+g, r^2+b^2+g^2, \ldots r^{11}+b^{11}+g^{11}) = \frac{1}{22}\big((r+b+g)^{11}+10(r^{11}+b^{11}+g^{11})+11(r+b+g)\cdot(r^2+b^2+b^2)^5\big).$$

We could, in principal, expand this out, collect like terms, and then read off the coefficient of $r^3b^4g^4$ (which would give us our numerical answer)... or we could just read off only that coefficient directly. There are three terms that can contribute different sorts of monomials; if we figure out the contribution of $r^3b^4g^4$ from each separately, we can add them together and get the total coefficient.

$(r+b+g)^{11}$ has an $r^3b^4g^4$-coefficient of $\binom{11}{3,4,4} = \frac{11!}{3!\cdot 4!\cdot 4!}$ by the Trinomial Theorem. The middle term cannot contribute any terms of the desired form, which just leaves the term $11(r+b+g)\cdot(r^2+b^2+b^2)^5$. Ignoring the lead coefficient of 11 for a moment, we ask for the coefficient of $r^3b^4g^4$ in $(r+b+g)(r^2+b^2+g^2)^5$. Think of this as 6 terms being multiplied by one another, where we must pick a single summand from each. Clearly we must pick r from the term $r+b+g$ (so that the total r-degree will be odd), meaning that from the remaining five terms we must pick one copy of r^2, two copies of b^2, and two copies of g^2. Again by the Trinomial Theorem, there are $\binom{5}{1,2,2} = \frac{5!}{1!\cdot 2!\cdot 2!}$ such possible choices.

Putting this all together, the coefficient of $r^3b^4g^4$ in U_D is

$$\frac{1}{22}\left(\frac{11!}{3!\cdot 4!\cdot 4!} + 11\cdot\frac{5!}{1!\cdot 2!\cdot 2!}\right) = 540,$$

up to the possibility that I made a multiplication error. Thus there are 540 essentially distinct necklaces with 3 red, 4 blue, and 4 green beads.

[8] Do this!