Appendix 13A: Other anatomies

There are features of the anatomies of certain other mathematical objects, when broken up into their indecomposable components, that are very similar to the integers. We explore that briefly here; for more information, presented in a rather different format, see the graphic novel [**2**].

13.5. The anatomy of polynomials in finite fields

Monic polynomials (over $\mathbb{C}$ or in $\mathbb{F}_p$) can be factored in a unique way (up to order) into monic irreducible polynomials. There are p^n monic polynomials of degree n in $\mathbb{F}_p$, and in (4.12.3) of appendix 4C, we showed the number of monic irreducible polynomials of degree n in $\mathbb{F}_p$ is $\frac{1}{n}\sum_{d|n}\mu(d)p^{n/d}$. This is close to p^n/n; that is, roughly 1 out of every n polynomials of degree n is irreducible. We "calibrate" this with the proportion $1/\log x$ of integers up to x that are prime to compare the anatomies of polynomials in finite fields with those of integers.

In appendix 4D we showed that, on average, integers $\le x$ have about $\log x$ divisors. For the analogous result, note that if a given monic polynomial $f(x)$ of degree m divides a monic polynomial of degree n then it can be written as $f(x)g(x)$ for some monic polynomial $g(x)$ of degree $n-m$. Therefore, the average number of monic polynomials dividing a monic polynomial of degree n is

$$\frac{1}{p^n}\sum_{\substack{h(x)\text{ monic}\\ \text{of degree } n}}\sum_{m=0}^{n}\sum_{\substack{f(x)\text{ monic}\\ \text{of degree } m\\ f(x)\text{ divides }h(x)}} 1 = \frac{1}{p^n}\sum_{m=0}^{n}\sum_{\substack{f(x)\text{ monic}\\ \text{of degree } m}}\sum_{\substack{g(x)\text{ monic}\\ \text{of degree } n-m}} 1$$
$$= \frac{1}{p^n}\sum_{m=0}^{n} p^m\cdot p^{n-m} = n+1,$$

which was calibrated with $\log x$.

The number of monic polynomials of degree n with exactly two irreducible monic polynomial factors is

$$\frac{1}{2}\sum_{d=1}^{n-1} \sum_{\substack{f(x) \text{ monic irreducible} \\ \text{of degree } d}} 1 \cdot \sum_{\substack{g(x) \text{ monic irreducible} \\ \text{of degree } n-d}} 1$$

less the cases where $f = g$. Now the formula for the number of monic irreducibles is complicated so let's just work with main terms, so we see that the above is roughly

$$\frac{1}{2}\sum_{d=1}^{n-1} \frac{p^d}{d} \cdot \frac{p^{n-d}}{n-d} - \frac{p^{n/2}}{n/2} \approx \frac{p^n}{2}\sum_{d=1}^{n-1} \frac{1}{d(n-d)} = \frac{p^n}{n}\sum_{d=1}^{n-1} \frac{1}{d} \approx \frac{p^n}{n}\log n.$$

This can be compared with the number of integers $\le x$ with exactly two prime factors, $\approx \frac{x}{\log x}\log\log x$, which is the same, replacing n by $\log x$. A similar argument yields that the number of monic polynomials of degree n in $\mathbb{F}_p$ with exactly k irreducible monic polynomial factors is roughly $\frac{p^n}{n} \cdot \frac{(\log n)^{k-1}}{(k-1)!}$, at least if k is not too large (in terms of n).

Let $\omega(F)$ denote the number of distinct monic irreducible factors of F. The mean value of $\omega(F)$, over monic F of degree n, is

$$= \frac{1}{p^n} \sum_{\substack{F(x) \text{ monic} \\ \text{of degree } n}} \sum_{m=1}^{n} \sum_{\substack{g(x) \text{ monic irreducible} \\ \text{of degree } m \\ g \text{ divides } F}} 1 = \frac{1}{p^n}\sum_{m=1}^{n} \sum_{\substack{g(x) \text{ monic irreducible} \\ \text{of degree } m}} \sum_{\substack{F(x) \text{ monic} \\ \text{of degree } n \\ g \text{ divides } F}} 1$$

$$= \frac{1}{p^n}\sum_{m=1}^{n} \sum_{\substack{g(x) \text{ monic irreducible} \\ \text{of degree } m}} p^{n-m} = \sum_{m=1}^{n} \frac{1}{p^m} \cdot \frac{1}{m}\sum_{d|m} \mu(d)p^{m/d} \approx \sum_{m=1}^{n} \frac{1}{m} \approx \log n,$$

taking only the $d = 1$ terms.

One can then prove, in one of several ways (analogous to how we approached the prime factors of integers), that the variance is also about $\log n$, and so almost all polynomials of degree n in $\mathbb{F}_p$ have about $\log n$ distinct monic irreducible factors.

Exercise 13.5.1. Sketch a proof that almost all polynomials in $\mathbb{F}_p$ of degree $2d$ are *not* the product of two polynomials of degree d, as d gets large.

13.6. The anatomy of permutations

Permutations can be represented as a product of cycles in a unique way; and a given set of cycles defines a permutation. A cycle is an irreducible permutation. Let S_N be the set of permutations on N letters. The number of permutations on N letters is $|S_N| = N!$. There are $(N-1)!$ cycles on N letters, since the first letter can be sent to any of the other $N-1$ letters, that letter to any of the $N-2$ remaining letters, etc. The cycles form a proportion $(N-1)!/|S_N| = (N-1)!/N! = 1/N$ of all the permutations in S_N. We "calibrate" this with the proportion $1/\log x$ of integers up to x that are prime to compare the anatomies of permutations with those of integers.

If $n = p_1 \cdots p_k$, the factorization of squarefree n into primes, then each divisor can be written as $p_{j(1)} \cdots p_{j(\ell)}$ for some $\{j(1), \dots, j(\ell)\}$ of $\{1, \dots, k\}$ (and each such

product gives a divisor of n). In this language, the analogy for permutations would therefore be: If $\sigma = C_1 \cdots C_k$, the factorization of σ into cycles, then each divisor can be written as $C_{j(1)} \cdots C_{j(\ell)}$ for some subset $\{j(1), \dots, j(\ell)\}$ of $\{1, \dots, k\}$. This set of cycles acts on some subset S of the N letters, permuting the elements of S (and of the complementary set, T). That is, we can partition the N letters into $S \cup T$ and σ fixes S and T. If σ is a cycle then the only subsets it fixes are $\emptyset$ and itself, very much in analogy with how we define primes. The average number of divisors of a permutation of a set Λ of N letters, is

$$\frac{1}{N!} \sum_{\sigma \in S_N} \sum_{\substack{S \cup T = \Lambda \\ \sigma \text{ fixes } S \text{ and } T}} 1.$$

If $|S| = k$ then there are $k! \cdot (N-k)!$ permutations of S and T, and so this is the number of $\sigma \in \Lambda$ which fix S and T. Therefore the above equals

$$\frac{1}{N!} \sum_{k=0}^{N} \sum_{\substack{S \subset \Lambda \\ |S|=k}} k! \cdot (N-k)! = \frac{1}{N!} \sum_{k=0}^{N} \binom{N}{k} \cdot k! \cdot (N-k)! = \sum_{k=0}^{N} 1 = N+1.$$

The number of permutations with exactly two cycles is

$$= \frac{1}{2} \sum_{k=1}^{N-1} (k-1)!(N-k-1)! \sum_{\substack{S \cup T = \Lambda \\ |S|=k}} 1 = \frac{1}{2} \sum_{k=1}^{N-1} (k-1)!(N-k-1)! \binom{N}{k}$$

$$= \frac{N!}{2} \sum_{k=1}^{N-1} \frac{1}{k(N-k)} = \frac{N!}{2N} \sum_{k=1}^{N-1} \left(\frac{1}{k} + \frac{1}{N-k} \right) = \frac{N!}{N} \sum_{k=1}^{N-1} \frac{1}{k} \approx \frac{N!}{N} \cdot \log N.$$

A similar argument yields that the number of permutations with exactly k cycles is

$$\frac{N!}{N} \frac{1}{(k-1)!} \sum_{\substack{a_1, \dots, a_{k-1} \geq 1 \\ a_1 + \dots + a_{k-1} \leq N-1}} \frac{1}{a_1 \cdots a_{k-1}} \approx \frac{N!}{N} \cdot \frac{(\log N)^{k-1}}{(k-1)!}.$$

at least if k is not too large (in terms of N), as we prove in the following exercise:

Exercise 13.6.1. (a) Prove that

$$\left(\sum_{a \leq A/m} \frac{1}{a} \right)^m \leq \sum_{\substack{a_1, \dots, a_m \geq 1 \\ a_1 + \dots + a_m \leq A}} \frac{1}{a_1 \cdots a_{k-1}} \leq \left(\sum_{a \leq A} \frac{1}{a} \right)^m$$

(b)‡ Prove that if $m \leq \frac{\log A}{(\log \log A)^2}$ then the two terms at either end of the inequalities in (a) differ by a multiplicative factor which gets arbitrarily close to 1 as A grows.

We will now determine the average number of cycles in a permutation. First note that the number of permutations containing a given cycle C of length k is $(N-k)!$, since one determines all the ways that σ can act on the letters not acted on by C. The number of cycles of length k is $\binom{N}{k}(k-1)! = \frac{N!}{(N-k)!\,k}$. Therefore the

average number of cycles per permutation of S_N is

$$\frac{1}{N!}\sum_{\sigma\in S_N}\sum_{\substack{C \text{ a cycle}\\ C\in\sigma}} 1 = \frac{1}{N!}\sum_{\sigma\in S_N}\sum_{k=1}^{N}\sum_{\substack{C \text{ a cycle}\\ |C|=k\\ C\in\sigma}} 1 = \frac{1}{N!}\sum_{k=1}^{N}\sum_{|C|=k}\sum_{\substack{\sigma\in S_N\\ C\in\sigma}} 1$$

$$= \frac{1}{N!}\sum_{k=1}^{N}\frac{N!}{(N-k)!\,k}\cdot(N-k)! = \sum_{k=1}^{N}\frac{1}{k} \approx \log N.$$

To determine the variance, we calculate

$$\frac{1}{N!}\sum_{\sigma\in S_N}\left(\sum_{\substack{C \text{ a cycle}\\ C\in\sigma}} 1\right)^2 = \frac{1}{N!}\sum_{\sigma\in S_N}\sum_{\substack{C \text{ a cycle}\\ C\in\sigma}} 1 + \frac{1}{N!}\sum_{\sigma\in S_N}\sum_{\substack{C\cup D \text{ disjoint cycles}\\ C\cup D\in\sigma}} 1.$$

We just calculated the first term. For the second we note that given $C\cup D$ with $|C|=k,\ |D|=\ell$, the number of $\sigma\in S_N$ with $C\cup D\in\sigma$ in $(N-(k+\ell))!$. The number of pairs of disjoint cycles $C\cup D$ with $|C|=k,\ |D|=\ell$, is $\frac{N!}{(N-k)!\,k}\cdot\frac{(N-k)!}{(N-k-\ell)!\,\ell} = \frac{N!}{(N-k-\ell)!\,k\ell}$. Therefore the second term in the last displayed equation equals

$$\frac{1}{N!}\sum_{\substack{k,\ell\ge 1\\ k+\ell\le N}}\frac{N!}{(N-k-\ell)!\,k\ell}\cdot(N-(k+\ell))! = \sum_{\substack{k,\ell\ge 1\\ k+\ell\le N}}\frac{1}{k\ell}.$$

Therefore the variance equals

$$\sum_{k=1}^{N}\frac{1}{k} + \sum_{\substack{k,\ell\ge 1\\ k+\ell\le N}}\frac{1}{k\ell} - \left(\sum_{k=1}^{N}\frac{1}{k}\right)^2 < \sum_{k=1}^{N}\frac{1}{k} \approx \log N$$

by the following exercise. We deduce that almost all permutations on N letters have about $\log N$ cycles.

Exercise 13.6.2. Prove, by taking $m=k+\ell$, that

$$0 < \left(\sum_{k=1}^{N}\frac{1}{k}\right)^2 - \sum_{\substack{k,\ell\ge 1\\ k+\ell\le N}}\frac{1}{k\ell} = \sum_{\substack{1\le k,\ell\le N\\ k+\ell> N}}\frac{1}{k\ell} = 2\sum_{m=N+1}^{2N}\frac{1}{m}\sum_{k=m-N}^{N}\frac{1}{k} = 2\sum_{k=1}^{N}\frac{1}{k}\sum_{m=N+1}^{N+k}\frac{1}{m} < 2.$$

There are many other aspects of the anatomies of polynomials in finite fields, and of permutations, that mirror the anatomy of integers.

More on mathematical anatomies

[1] Richard Arratia, A.D. Barbour, and Simon Tavaré, *Random combinatorial structures and prime factorizations*, Notices Amer. Math. Soc. **44** (1997), 903-910.

[2] Andrew Granville, Jennifer Granville and Robert J. Lewis, *Prime suspects: The anatomy of integers and permutations*, Princeton University Press (2019).

[3] Anatoly M. Vershik, *Asymptotic combinatorics and algebraic analysis*, Proc ICM Zurich (1994), 1384–1394.