THE ANATOMY OF INTEGERS AND PERMUTATIONS

Andrew Granville

We begin, as in any mathematical paper, with definitions:

Anatomy (`a-nat-o-my`) noun: *The scientific study of the shape and structure of an organism and the inter-relation of its various parts. The art of separating the parts of an organism in order to ascertain their position, relations, structure, and function.*

Forensic (`fo-ren-sic`) adjective: *Relating to the use of science or technology in the investigation and establishment of facts or evidence.*

If you switch on your TV in the evening then, as likely as not, you will find yourself watching an episode of a popular detective show (set in various spectacular locations) in which surprisingly dapper forensic scientists turn up evidence using careful anatomical (and other) study so as to be able to identify and prosecute a heinous criminal. Sometimes a flatfooted detective is misled by the surface evidence to suspect one person, but then the forensic team, digging deeper, turns up details that surprise not only the easily misled detective but even you, the astute viewer. For example, two seemingly unrelated corpses are found, and our hapless detective believes that the crimes are unrelated, whereas the forensic investigators turn up conclusive proof that the two corpses were in fact twins.

So what would happen if we put together a forensic team to investigate the anatomy of some of the most common mathematical objects, say of integers and permutations? Seems silly at first. Most of our training with these simple mathematical objects involves how they are used in understanding more complicated phenomena, but rarely do we look at their anatomy, the inter-relation of their constituent parts (that is, the prime factors of integers, and the cycles of permutations). So our objective is to be the forensic scientists, with the corpses of these two seemingly unrelated mathematical objects laid out before us, and it is up to us to determine whether there is more in common between the anatomies of integers and permutations than meets the eye.

This article is written as a companion piece to [0], a film-script in which we develop the connection with anatomy and forensics to create a fantasy world where forensic detectives (loosely based on famous mathematicians) prove and interpret several of the key notions exposed more precisely herein.

1. Is there a case to be made?

1.1. The basic constituent parts. When comparing the anatomies of two seemingly different organisms, the forensic scientist knows that one must calibrate their sizes otherwise one might be misled into believing that they are different, whereas they might be

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

twins organisms that had grown apart. In order to do such a calibration, one needs to find some essential feature that the organisms have in common to allow one to better compare the two objects. Often one chooses to compare the key constituents of each organism – forensic scientists consider the selection and measurement of this key constituent as much an art as a science.

So what are the key constituents of our mathematical organisms, integers[1] and permutations[2]? Each integer is a product of primes, the basic, indecomposable parts of an integer, and each permutation is a union of (disjoint) cycles, the basic, indecomposable parts of a permutation. So it would make sense to calibrate the primes with the cycles, but we have to figure out how to do so given that these are about as similar as apples and I-pods. One idea is to compare how rare primes are with how rare cycles are: We know that roughly one out of every $\log x$ integers up to x is prime, and that exactly one in every N permutations on N elements is a cycle, so we could try to calibrate by replacing N when we measure the anatomy of a permutation with $\log x$ when we measure the anatomy of an integer.

The fact that roughly 1 out of every $\log x$ integers up to x is prime is known as "the prime number theorem" and is a deep fact to prove. On the other hand the fact that there are $(N-1)! = N!/N$ cycles of length N in S_N (the set of permutations on N letters) is easy to prove: In an N-cycle, element 1 (which we will call e_1) must be mapped to some different element (there are $N-1$ possibilities), call it e_2; then e_2 must be mapped to some element other than e_1 and e_2 (there are $N-2$ possibilities), etc; and thus in total there are $(N-1) \times (N-2) \times \cdots \times 1$ possibilities.

Now we have the calibration we need to start comparing our organisms, permutations and integers, and their constituent parts, cycles and prime factors. Perhaps the most obvious things to compare are how many constituent parts a typical organism of each type has, and whether these parts are laid out in the same way.

1.2. How many constituent parts? If one selects a permutation on N letters at random then, with a probability that goes to 1 as $N \to \infty$, the permutation has about $\log N$ disjoint cycles.[3] Now, replacing the N in $\log N$ by $\log x$ (which we must do in order to calibrate this count), we expect that if one chooses an integer up to x at random then, with a probability that goes to 1 as $x \to \infty$, the integer has about $\log\log x$ distinct prime factors. And this is true (as was shown in famous work of Hardy and Ramanujan from 1917).

More compelling justification for the comparison of permutations and integers comes in asking more precise questions. For example, with what probability does one or the other organism have somewhat fewer parts, or somewhat more parts than $\log N$ (or $\log\log x$)? In 1942, Goncharov showed that, as we vary over the permutations on N letters, the probability that the number of cycles is more than or less than a given quantity is governed by a Bell curve, that is the normal distribution, with mean and variance about $\log N$. Similarly Erdős and Kac showed in 1940 that, as we vary over the integers $\leq x$, the

[1] **Integer** (`in-te-ger`) (noun): *positive or negative whole number.*

[2] **Permutation** (`per-mu-ta-tion`) (noun): *a re-arrangement of the elements of a set.*

[3] By "about $\log N$ disjoint cycles", I mean, more precisely, that the ratio of the number of disjoint cycles to $\log N$ tends to 1 as $N \to \infty$

probability that the number of distinct prime factors of an integer is more than or less than a given quantity is also governed by the normal distribution, this time with mean and variance around $\log\log x$.

1.3. The layout. For a permutation $\sigma \in S_N$ (where S_N is the set of permutations on N letters), suppose that σ decomposes into cycles as $C_1 C_2 \cdots C_\ell$ with lengths $1 \leq d_1(\sigma) \leq d_2(\sigma) \leq \cdots \leq d_\ell(\sigma) \leq N$. If we take the logarithms of these lengths then we have

$$0 \leq \log d_1(\sigma) \leq \log d_2(\sigma) \leq \cdots \leq \log d_\ell(\sigma) \leq \log N,$$

So most $\sigma \in S_N$ have about $\log N$ numbers of the form $\log d_i(\sigma)$ in the interval $[0, \log N]$. This is an interval of length $\log N$ and so there is an average distance of one between these numbers. How are these numbers distributed within the interval? Do they look random, or will they be bunched up in one part of the interval, and sparse in another part? The correct model for "random", from probability theory, is the "Poisson point process", which deals with the appearance of random variables over time. However, since we do not actually have random variables here we need to reshape this model for our situation: Suppose that we are given a sequence of finite sets $S_1, S_2, \cdots$ such that S_j is a subset of $[0, m_j]$ and contains about m_j elements, where $m_j \to \infty$ as $j \to \infty$. We say that $S_1, S_2, \cdots$ is "Poisson distributed" if, for any fixed interval length $\lambda > 0$, and given integer $k \geq 0$, the probability that a random subinterval of $[0, m_j]$ of length λ contains exactly k elements of S_j tends to $e^{-\lambda}\lambda^k/k!$ as $j \to \infty$. One can prove that the sets $\{\log d_1(\sigma), \log d_2(\sigma), \cdots \log d_\ell(\sigma)\}$ are indeed Poisson distributed, for almost all $\sigma \in S_N$, as $N \to \infty$.[4]

So how about for integers? We saw in the previous section that we should replace $\log N$ by $\log\log x$ for comparison purposes, and we might extend this to replacing the $\log d_j$ by $\log\log p_j$ for the prime factors involved. Thus we consider the sets $\{\log\log p : \ p|n\}$; we have seen that, for $n \leq x$, these sets typically have about $\log\log x$ elements inside the interval $[\log\log 2, \log\log x]$. One can prove that the sets $\{\log\log p : \ p|n\}$ are indeed Poisson distributed, for almost all $n \leq x$, as $x \to \infty$.

Our two organisms, permutations and integers, seem to be almost identical with our chosen calibration. However it is also true that the poisson and normal distributions appear in many situations in mathematics, so perhaps these successful comparisons are not too surprising – after all the cycle lengths and prime factor sizes have to be distributed somehow, so one's first guess would probably be something random, hence the poisson and normal distributions. So are there measures of permutations or integers that involve rather unusual functions, so that it would be more surprising if our two organisms calibrate so well?

1.4. The largest and smallest parts. One cannot sensibly ask how long is the longest cycle in a permutation since no particular length occurs with high probability. However, for fixed $u > 0$, the probability that the longest cycle of a randomly chosen permutation on N letters contains no more than N/u elements is about $\rho(u)$ for large N (as was also shown by Goncharov, in 1944). Evidently $\rho(u) = 1$ in the uninteresting range $0 < u \leq 1$,

[4] By "almost all" we mean that the proportion tends to 1 as $N \to \infty$.

but it does not appear to have any simple definition (that is, any closed formula) for $u > 1$. Indeed the most palatable definition that we know of is via the integral delay equation

$$\rho(u) = \frac{1}{u} \int_{u-1}^{u} \rho(t)\ \mathrm{d}t \text{ for } u > 1,$$

which certainly satisfies our requirement that the measuring function be unusual.

Evidently the analogy to longest cycle will be largest prime factor. Replacing N by $\log x$ we wish to determine how often some function of the largest prime factor is $\leq \frac{1}{u} \log x$. To be a meaningful analogy this should always be the case when $u = 1$, and since if p is the largest prime factor of $n \leq x$ then $\log p \leq \log n \leq \log x$, we guess that one can take the function of p to be $\log p$. Now $\log p \leq \frac{1}{u} \log x$ if and only if $p \leq x^{1/u}$, so our analogy would be, for fixed $u > 0$, that the probability that the largest prime factor of an integer $\leq x$ is $\leq x^{1/u}$ is about $\rho(u)$ for large x. This was proved to be true by DICKMAN in 1930.

And the same is true for the largest k cycles of a permutation or prime factors of an integer. That is for any given $1 \leq u_1 \leq u_2 \leq \cdots \leq u_k$ the probability that a randomly chosen cycle on N letters has its jth largest cycle of length $\leq N/u_j$ for $j = 1, 2, \ldots , k$ is more-or-less the same as the probability that a randomly chosen integer $\leq x$ has its jth largest prime factor $\leq x^{1/u_j}$ for $j = 1, 2, \ldots , k$, for N and x both sufficiently large.[5]

One can ask analogous question about the smallest parts: For fixed $u > 0$, one finds that the probability that the smallest cycle of a randomly chosen permutation on N letters contains at least than N/u elements is about $u\omega(u)/N$ for large N. Evidently $\omega(u) = 0$ for $0 < u < 1$, and has the simple formula $\omega(u) = 1/u$ for $1 \leq u \leq 2$. However, it does not appear to have any simple definition for $u > 2$, the most palatable being

$$\omega(u) = \frac{1}{u} \int_{0}^{u-1} \omega(t)\ \mathrm{d}t \quad \text{for all } u > 2,$$

which again satisfies our requirement that the measuring function be unusual. In 1949 BUCHSTAB showed that the probability that the smallest prime factor of an integer $\leq x$ is $\geq x^{1/u}$ is also about $u\omega(u)/\log x$ for large x. Again the correct analogy. And again the same probabilities occur when determining how often the smallest k cycles of a permutation are at least $N/u_1, N/u_2, \ldots , N/u_k$, respectively, and how often the smallest k prime factors of an integer are at least $x^{1/u_1}, x^{1/u_2}, \ldots , x^{1/u_k}$, respectively.

1.5. Smaller parts, and the constant γ.

There is no obvious way to compare the smallest cycles in a typical permutation with the smallest prime factors of a typical integer. What prime corresponds to a fixed point of permutation (that is, a cycle of length one)? Forensic scientists know that you are never going to have a perfect match, that there is some evidence which is truly irrelevant, simply superficial,[6] and in this situation it will be the cycles of bounded length, and the primes of bounded size. However we should look at cycles that are a little longer and primes a little larger, which should provide a more fundamental description of the respective anatomies. So we will estimate the probability

[5] The probability theorists call this a *Poisson-Dirichlet* distribution if it holds true for all integers k.

[6] For example, one twin may have broken her nose, the other dyed his hair.

that the shortest cycle has length $> M$, where M is some function of N, such that M and N/M tend to ∞ as N does.[7]

We will now estimate the number of elements of S_N whose shortest cycle has length $> M$, by an inclusion-exclusion argument. First we take all permutations on N letters, and then remove all those that contain a given cycle of length $\leq M$, for each possible cycle of length $\leq M$. The expected number of cycles of length j in a permutation is about $1/j$, and so our total so far is about $N! - \sum_{j=1}^{M} N!/j$, which equals $N!(1-\mu_M)$ where $\mu_M := \sum_{j=1}^{M} \frac{1}{j}$. Of course we have subtracted off too much; for a permutation that has r cycles of length $\leq M$, we have counted -1 r times. So we have to add back in the expected number of permutations that contain a pair of given cycles of length $\leq M$, for each given pair. The expected number of cycles of lengths j_1 and j_2 in a permutation is about $1/j_1 j_2$ if $j_1 \neq j_2$, and about $1/2j_1^2$ if $j_1 = j_2$. Therefore our sum is now about $N!(1 - \mu_M + \frac{\mu_M^2}{2})$. Continuing in this way, we estimate that the number of permutations in S_N that are free of cycles of length $\leq M$ is about

$$N! \sum_{r\geq 0} (-1)^r \frac{\mu_M^r}{r!} = N! e^{-\mu_M} \quad \text{which is about} \quad \frac{e^{-\gamma}}{M}\, N!.$$

Here γ, the Euler-Mascheroni constant, is defined to be $\lim_{M\to\infty} \left(\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{M} - \log M\right)$. It is not even obvious that this limit exists (but it does). Mathematicians know very little about γ, but expect that it is a transcendental number. By the way, this argument can be made into a rigorous proof so long as M and $N/M \to \infty$ as $N \to \infty$.

Now let's try estimating the number of integers $\leq x$ free of prime factors $\leq y$. We begin with the number of integers $\leq x$, and subtract the number of such integers that are divisible by p, of which there are about x/p, for all primes $p \leq y$. We then need to add the number of integers $\leq x$ that are divisible by pq back into the sum, of which there are about x/pq, for all pairs of primes $p < q \leq y$. Continuing on like this we estimate, for $P = \prod_{p\leq y} p$, that the number of integers $\leq x$ free of prime factors $\leq y$ is about

$$\sum_{\substack{d \text{ a positive integer} \\ d \text{ divides } P}} (-1)^{\nu(d)} \frac{x}{d} = x \prod_{\substack{p \text{ prime} \\ p\leq y}} \left(1 - \frac{1}{p}\right) \quad \text{which is about} \quad \frac{e^{-\gamma}}{\log y}\, x.$$

(Here $\nu(d)$ denotes the number of distinct prime factors of d.) This last approximation is known as Mertens' theorem. This argument can be made into a rigorous proof so long as $\log y$ and $(\log x)/(\log y) \to \infty$ as $x \to \infty$. Equating M with $\log y$, as in the previous section, we find yet again that the anatomies of permutations and integers are very much alike.

An academic's aside: As an analytic number theorist I find this particular piece of evidence rather special. You have to understand that the appearance of γ in Mertens' theorem is somewhat mysterious: In all of the proofs in the literature one obtains the constant in this estimate in terms of certain integrals, which evaluate to $e^{-\gamma}$ but none of these proofs

[7] The case where N/M is fixed (and equal to u) was dealt with in the previous section.

seem to give any intuition as to *why* the special constant γ should feature. However, if we believe that the anatomies of permutations and integers are indeed the same then this section gives us intuition as to why $e^{-\gamma}$ is the constant in Mertens' theorem.

2. Even more analogies

What do we make of all of this evidence as to the similarities between the anatomies of integers and permutations? A skeptic might argue that we have mostly considered the "typical" integers and permutations, whereas it might be profitable to look at the atypical, for example those permutations with exactly ℓ cycles and those integers with exactly ℓ prime factors, even for atypical values of ℓ.

2.1. The proportion with ℓ parts. We know that one in every N permutations in S_N is a cycle. We now ask what proportion have exactly two cycles, or three or more? In fact it is known that, for any fixed integer $\ell \geq 1$, the proportion of permutations with exactly ℓ cycles is about

$$\sim \frac{1}{N} \frac{(\log N)^{\ell-1}}{(\ell-1)!}$$

(as shown by Jordan in 1947); in fact this is true for all ℓ that are significantly smaller than the expected number of cycles, $\log N$.[8] Moreover if ℓ is a fixed multiple of $\log N$, then this estimate needs multiplying by a small constant, $1/\Gamma(r+1)$, depending (only) on the ratio $r := (\ell-1)/\log N$.[9]; and this remains true for all ℓ that are significantly smaller than $(\log N)^2$.

And what about integers up to x having exactly two prime factors, or three or more? Can we simply replace N by $\log x$ in the above results? It is an old theorem of Landau from 1909 that for any fixed integer $\ell \geq 1$, the proportion of integers up to x with exactly ℓ prime factors is

$$\sim \frac{1}{\log x} \frac{(\log\log x)^{\ell-1}}{(\ell-1)!},$$

and this is true for all ℓ that are significantly smaller than the expected number of prime factors, $\log\log x$. If ℓ is a fixed multiple of $\log N$, then this estimate needs multiplying by *two* small constants: First, $1/\Gamma(r+1)$, depending (only) on the ratio $r := (\ell-1)/\log\log x$, and second,

$$\prod_{p \text{ prime}} \left(1 + \frac{r}{p-1}\right)\left(1 - \frac{1}{p}\right)^r .$$

At first sight this second factor might seem to suggest a big difference between the two anatomies. However at both $r = 0$ and $r = 1$, the most usual values to consider, this product equals 1. In general this infinite product is equal, up to a small factor, to the same product but now limited to the primes $p \leq r^2$, and therefore only involves the small primes. Now we argued earlier that small primes should be regarded as an irrelevant and superficial difference, and one can prove that the contribution of this factor is always pretty

[8] That is, if $\ell/\log N \to 0$ as $N \to \infty$, thinking of ℓ as a function of N.

[9] The function Γ is the classical "Gamma function".

small, so we should feel safe in ignoring this difference. This SATHÉ-SELBERG formula (1954) remains true for all ℓ that are significantly smaller than $(\log\log x/\log\log\log x)^2$.

And how about if there are more parts, far more than the average? It seems that our formulas are getting more complicated the more parts there are, so it might not come as a surprise to hear that there are no simple functions known that describe the proportion of permutations $\sigma \in S_N$ with exactly ℓ cycles, where ℓ is larger than a fixed power of $\log N$, or the number of integers up to x with exactly ℓ prime factors, where ℓ is larger than a fixed power of $\log\log x$. In both cases we do have estimates in a wide range available but these are in terms of saddle points so that the values are implicit functions and difficult to estimate precisely. However what one can do with such estimates, which is just as useful for many applications, is to prove an accurate comparative estimate, as we will now explain.

Suppose that ℓ and $N/\ell \to \infty$ as $N \to \infty$, and let $\nu = \frac{N}{\ell}\log(\frac{N}{\ell})$. If m is a positive integer that is significantly smaller than both N and ν then

$$\frac{\text{Proportion of permutations on } N-m \text{ letters with } \ell-1 \text{ cycles}}{\text{Proportion of permutations on } N \text{ letters with } \ell \text{ cycles}} \quad \text{is about} \quad \frac{\ell}{\log\nu}.$$

This follows from deep estimates of MOSER and WYMAN (1958).[10] Similarly suppose that ℓ and $\log x/\ell(\log\log x)^A \to \infty$ as $x \to \infty$, for every fixed A, and let $\nu = \frac{\log x}{\ell}\log(\frac{\log x}{\ell})$. If d is a positive integer for which $\log d$ is significantly smaller than ν then[11]

$$\frac{\text{Proportion of integers up to } x/d \text{ with exactly } \ell-1 \text{ prime factors}}{\text{Proportion of integers up to } x \text{ with exactly } \ell \text{ prime factors}} \quad \text{is about} \quad \frac{\ell}{\log\nu}.$$

Could these results really just be a co-incidence, or is this compelling new evidence that the whole populations of permutations and integers have remarkably similar anatomy?

2.2. The layout with ℓ parts. Suppose that a permutation has exactly ℓ cycles, for a fixed integer $\ell \geq 2$. Since $\sum_{i=1}^{\ell} d_i(\sigma) = N$, we see that $d_\ell(\sigma)$ is determined by the other cycle lengths, and is $\geq N/\ell$, so that $\log d_\ell(\sigma)$ is guaranteed to be very close to the end of the interval. It therefore makes sense to study the distribution of all but the largest part, and one can show that the points $\{\log d_i(\sigma)/\log N : \quad 1 \leq i \leq \ell-1\}$ are distributed on $(0,1)$ like $\ell-1$ random numbers, as we vary over such permutations σ.[12] In fact this holds provided $\ell \leq \frac{1}{2}\log\log N$. It similarly makes sense to study only the smallest $\ell-1$ prime factors of a typical integer with exactly ℓ prime factors for fixed (or small) values of ℓ. We can prove that the elements of $\{\log\log p_i(n)/\log\log n : \quad 1 \leq i \leq \ell-1\}$ are distributed on $(0,1)$ like $\ell-1$ random numbers, as we vary over the integers $n \leq x$ with exactly ℓ prime factors.

[10] In fact they estimated the Stirling numbers of the first kind, which can be defined as the number of $\sigma \in S_N$ with exactly ℓ cycles.

[11] As proved by HILDEBRAND and TENENBAUM (1988).

[12] More precisely a proportion $(\ell-1)!\epsilon^{\ell-1}$ of the permutations $\sigma \in S_N$ with exactly ℓ cycles have $\log d_i(\sigma)/\log N \in (\alpha_i, \alpha_i+\epsilon)$ for each $1 \leq i \leq \ell-1$, so long as $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{\ell-1} \leq \alpha_\ell = 1$ and the intervals $(\alpha_i, \alpha_i+\epsilon)$ are disjoint.

We would like to prove that the cycle lengths of permutations with exactly ℓ cycles, look random, in that the lengths are Poisson distributed, once $\ell \to \infty$ as $N \to \infty$. There are two different obstructions to this behaviour for various values of ℓ:

- Our definition of "Poisson distributed" involves a continuous distribution, and so we do not expect there to be any repeated cycle lengths. However if ℓ is around $\sqrt{N}$ or larger then most cycle lengths will be repeated for almost all such σ, so we have no hope of a Poisson-type distribution of cycle lengths. When ℓ is much smaller than $\sqrt{N}$ then there will be more than one cycle of each length up to $\ell/\log\nu$, so we should exclude these small cycle lengths from our considerations.
- There are no more than $\ell/\log\nu$ cycles of length $> (N/\ell)\log(N/\ell)$; that is a vanishing proportion of cycles are this large, so we should exclude these large cycle lengths from our considerations.

Taking these two observations into account, by only considering cycle lengths between ℓ and N/ℓ, one can indeed prove that the sets

$$\left\{ \frac{\log d_i(\sigma)}{\frac{1}{\ell}\log(N/\ell^2)} : \ d_i(\sigma) \in [\ell, N/\ell] \right\}$$

are Poisson distributed, for almost all $\sigma \in S_N$ with exactly ℓ cycles.[13]

When we look at the distribution of the prime divisors of integers then the analogy to the second of the two restrictions above applies, but not the first. Therefore we find that for almost all integers up to x with exactly ℓ distinct prime factors, the sets

$$\left\{ \frac{\log\log p}{\frac{1}{\ell}\log\log(n^{1/\ell})} : \ p|n, \ p \le n^{1/\ell} \right\}$$

are Poisson distributed, whenever ℓ and $\log x/\ell(\log\log x)^A \to \infty$ as $x \to \infty$ for every fixed A.

So in this question the ranges in which we can sensibly ask the question differ, but we do find, in both settings, that the parts are Poisson distributed throughout the feasible ranges.

2.3. The largest of the ℓ parts. In section 1.4 we saw that the distribution of the size of the largest cycle of a randomly chosen partition satisfies a complicated distribution function, the same as the distribution of the largest prime factor of a random integer. What if we restrict our attention to permutations with rather more cycles than is typical? We have more cycles, so more of them must be short indicating that there might be a bias towards having shorter cycles in general. However the longest cycle is not a typical cycle, and these permutations have more cycles than normal, so perhaps the longest cycle will be longer than usual? Which is it? How about for permutations with fewer cycles than is typical? And for integers with fewer prime factors than is typical? Or for integers with more prime factors than is typical?

In fact it is not difficult to show that for almost all permutations on N letters with exactly ℓ cycles, where ℓ is significantly smaller than is typical,[14] the longest cycle has

[13] Note that the average gap between the logarithm of cycle lengths in this interval is about $\frac{1}{\ell}\log(n/\ell^2)$.

[14] That is, $\ell/\log N \to 0$ as $N \to \infty$.

length close to N. And also that for almost all integers n up to x with exactly ℓ prime factors, where ℓ is significantly smaller than is typical, we have that $\log p$ is about $\log x$ for the largest prime factor p of n.

Now suppose that ℓ is significantly larger than is typical,[15] with $\ell \leq \sqrt{N}$, and let $\xi = (\ell - 1)/\log \nu$.[16] For almost all permutations on N letters with exactly ℓ cycles, the longest cycle has length about $\frac{\log \xi}{\xi} N$. (So the longest cycle of a permutation typically gets smaller the more cycles the permutation has).[17] Similarly if p is the largest prime factor of n then $\log p$ is about $\frac{\log \xi}{\xi} \log x$ for almost all integers n up to x with ℓ distinct prime factors.

Finally what if we fix ξ? Then the proportion of permutations on N letters with exactly ℓ cycles such that all cycles have length $\leq N/u$, is about $\rho_\xi(u)$; and this is also the proportion of integers $n \leq x$ with exactly ℓ distinct prime factors such that the largest prime factor of n is $\leq x^{1/u}$. Here $\rho_\xi(u) = 1$ for $0 \leq u \leq 1$, and $\rho_\xi(u) = 1 - \xi \int_1^u (u-t)^{\xi-1} \frac{\mathrm{d}t}{t}$ for $1 \leq u \leq 2$ with

$$\rho_\xi(u) = \frac{\xi}{u} \int_{u-1}^{u} \rho_\xi(t) \left(\frac{t}{u}\right)^{\xi-1} \mathrm{d}t \quad \text{for all } u \geq 2.$$

(One can easily verify that $\rho_1(u) = \rho(u)$.)

3. Whys, other relations, and uses.

Why are the anatomies of integers and permutations so similar? Could it be that they have the same DNA? Or, perhaps one is modelled on the other? There are two proposed explanations for why their anatomies are so similar, one from probability theory, the other from analytic combinatorics as we will discuss in section 3.1.[18] Moreover such frameworks suggest other organisms (that is, classes of mathematical objects) in which one finds similar anatomies, as we will find in section 3.2.

3.1. Any good explanations? Arratia, Barbour and Tavaré (1997) developed a probabilistic model, which yields a good approximation to the structure of randomly chosen permutations and randomly chosen integers, so that the properties of the model give accurate forensic predictions for their anatomies. This model considers the joint distribution of $(z_1, z_2, \dots)$ where each z_i is an independent random variable having a Poisson distribution with parameter $1/i$. This distribution (subject to the side condition $\sum_i iz_i = N$) is very close to the joint distribution of $(c_1(\sigma), c_2(\sigma), \dots)$ where we run through the permutations σ on N letters, and $c_i(\sigma)$ denotes the number of cycles of length i in σ. When we look at the whole population of permutations, as in section 1, this model mostly predicts things very well. However, it is not clear whether it works so well for the sub-populations considered in section 2, since the probability questions that arise when we

[15] That is, $\ell/\log N \to \infty$ as $N \to \infty$.

[16] Note that ξ is the correct generalization of r (from section 2.1) as ℓ varies.

[17] More precisely, every cycle has length $\leq \frac{1}{\xi} \log\left(\frac{\xi}{\lambda \log \xi}\right)$ for a proportion $e^{-\lambda}$ of the permutations on N letters with ℓ cycles.

[18] It could be that these two viewpoints are really the same, in disguise. It is often difficult to penetrate the different languages of mathematics and, in this case, one feels there are many elements in common without it being clear to me whether there are fundamental differences.

add in the condition $\sum_i z_i = \ell$ are somewhat more delicate. Moreover one cannot drop such conditions to simplify the calculations as the following example highlights: For m in the range $N/2 < m \leq N$, the probability that the largest cycle in σ has length m is precisely $1/m$; however the probability that $z_m = 1$ and $z_j = 0$ for $m < j \leq N$ is

$$\frac{1}{m} \prod_{j=m}^{N} e^{-\frac{1}{j}} \quad \text{which is about} \quad \frac{1}{N},$$

which is somewhat different.

In 2001, Panario and Richmond noted that many of the statistics of section 1 are true for a fairly general class of combinatorial objects for which the generating function takes the following form: The number of objects of size m is given by the coefficient of z^m in a generating function of the form $a(1 - z/\rho)^{-b} \exp(E(z))$ where $|E(z)| \leq |z - \rho|^{\epsilon}$ if z is sufficiently close to ρ. Their work appears to me to be more the development of an efficient calculating tool to prove that certain qualified organisms have rather similar anatomies than a reason for why they are so similar.

Both these methods can be applied to show several other organisms have remarkably similar anatomies to integers and permutations – we shall give some examples in the next subsection.

In 1994, Vershik provided perhaps the best explanation for this phenomenon: Fundamental mathematical structures should be organized in a natural way. There are a few outstanding possibilities for this "natural anatomy" (seven are listed in [15]), including the structure we see here. What is perhaps new in this article is the amazing amount of detail that these different anatomies share. The idea that wildly different objects should be organized along very similar lines has emerged recently in an area on the boundary between quantum chaos in mathematical physics and the theory of zeta functions in analytic number theory: Sets of eigenvalues of various naturally arising operators (for example, in quantum chaos), and zeros of L-functions also seem to always be organized in very similar ways, according to the distribution of the eigenvalues of matrices randomly selected from certain groups: In 1999, Katz and Sarnak showed that only a small set of possible groups seem to ever arise. I don't think anyone can say why, indeed it all seems unreasonably convenient, begging for a unifying explanation.

3.2. Other organisms with similar anatomies. At least one other important class of mathematical objects has a similar anatomy, namely the polynomials mod p. These factor into irreducible polynomials mod p, the indecomposable components. There are p^N monic polynomials mod p of degree N, of which roughly 1 in N are irreducible, and their anatomies seem to be similar to those of integers and of permutations, though much remains to be established.

And there are other class of objects which appear to be likely candidates to share similar anatomies:

- The connected components of the 2-regular graphs on N labeled vertices; that is, the vertices should all appear in a set of non-trivial disjoint cycles.

- The connected components of the directed graphs given by the edges $(i, f(i))$ of any map $f : \{1, 2, \ldots, N\} \to \{1, 2, \ldots, N\}$ (which typically have about $\frac{1}{2} \log N$ components).

- The equivalence classes of mappings $\{1, 2, \ldots, N\} \to \{1, 2, \ldots, N\}$, where π_1, π_2 are equivalent if there exist permutations σ, τ such that $\pi_2 = \sigma\pi_1\tau$.

- Additive arithmetic semigroups, and other algebraic objects generalizing the rational integers, or polynomials over finite fields.

3.3. Models for the primes. Arratia, Barbour and Tavaré gave a probabilistic model to represent the integers (so as to compare the integers to the model discussed in section 3.2) which can be described as a process to randomly select an integer $n \leq x$, so that it will be given fully factored: For each prime $p \leq x$ let e_p be a random variable for which

$$\text{Prob}(e_p = k) = \frac{1}{p^k}\left(1 - \frac{1}{p}\right) \quad \text{for each integer } k \geq 0.$$

Let $Y := \prod_{p \leq x} p^{e_p}$. For each integer $n \leq x$, define independent random variables u_n so that $\text{Prob}(u_n = 1) = n/x$, and $\text{Prob}(u_n = 0) = 1 - n/x$; for $n > x$ let $u_n = 0$. Then

$$\text{Prob}(Y = n \text{ and } u_n = 1 \mid Y \leq x \text{ and } u_Y = 1) = \frac{1}{x}.$$

So the process proceeds by selecting the random variables as described. The algorithm fails if $Y > x$ or if $u_n = 0$, which occur with probability $1 - c_x$. Otherwise the algorithm succeeds and we obtain an fully factored, randomly chosen integer $\leq x$. We expect this algorithm to succeed 1 in every $1/c_x \approx e^\gamma \log x$ times it is run.

This algorithm would run slowly on a computer since it requires the calculation of e_p for each prime $p \leq x$: Determining each e_p may be quick but it is the fact there are so many of them that causes this part of the algorithm to be so slow. However one can simplify matters using Vershik's 1997 observation that a random integer up to (integer) x can be constructed by what is, in essence, a Markov chain, picking at each step a random prime factor of our integer. Hence our algorithm to select a random factored integer $n \leq x$ runs as follows: The probability that we select $n = 1$ is $1/x$; if we do so then the algorithm terminates. Otherwise, the probability that prime p divides n is $\frac{1}{x-1}\left[\frac{x}{p}\right]$. So we select prime p with probability

$$\left(1 - \frac{1}{x}\right)\frac{1}{L}\left[\frac{x}{p}\right] \quad \text{where } L := \sum_{\substack{p \text{ prime} \\ p \leq x}} \left[\frac{x}{p}\right].$$

If we have selected prime p then we obtain $n = mp$, where m is a randomly selected integer $\leq [x/p]$. Now we repeat the process for m. Since the range in which we search gets at least halved each time we run this process, we will not have to run it more than $\log_2 x$ times. Thus the algorithm is fast provided we can select p rapidly, and this was achieved by Bach in his 1998 Ph.D. thesis [4], proving that random, fully factored integers up to x can be found in "polynomial time".[19] Bach's algorithm for selecting p quickly is clever but complicated.

[19] In Bach's original paper the primes selected were really pseudoprimes (so as to be chosen in polynomial time) but after the recently discovered Agrawal-Kayal-Saxena polynomial time primality test [1] this problem is easily avoided.

We now develop another approach to producing a random factored integer $\leq x$, using the fact that permutations and integers have such similar anatomies. Let $N = \log x$ and select a random permutation $\sigma \in S_N$.[20] Writing σ as a product of cycles of lengths $d_1(\sigma) \leq d_2(\sigma) \leq \cdots \leq d_k(\sigma)$, we then select random prime numbers $p_i \in (e^{d_i(\sigma)}, e^{d_i(\sigma)+1})$ and consider the product $p_1 p_2 \ldots p_k$.[21] With this algorithm, the probability that integer $n \leq x$ is produced is close to $1/x$ (up to a constant multiple), which is not quite what was required. What we want is that the probability that integer n is produced is exactly $1/x$. To fix this one can import Bach's ideas (as described in [4]) to "doctor the odds", and make our algorithm work as claimed.

Credits

0. Granville, A. and Granville, J. (2008), *Math Sciences Investigation (MSI): The Anatomy of Integers and Permutations* (to appear).
1. Agrawal, M., Kayal, N. and Saxena, N. (2004), *PRIMES is in* P, Ann. Math. **160**, 781-793.
2. Arratia, R., Barbour, A.D. and Tavaré, S. (1997), *Random combinatorial structures and prime factorizations*, Notices Amer. Math. Soc. **44**, 903-910.
3. Arratia, R., Barbour, A.D. and Tavaré, S., *Logarithmic Combinatorial Structures: a Probabilistic Approach* (at `http://www-hto.usc.edu/books/tavare/ABT/`) (to appear).
4. Bach, E. (1988), *How to generate factored random numbers*, SIAM J. Comput. **17**, 179-193.
5. Feller, W. (1968), *An Introduction to Probability Theory and Its Application (3rd ed)*, vol. 1, Wiley, New York.
6. Granville, A., *Prime divisors are Poisson distributed*, Int. J. Number theory **3** (2007), 1-18.
7. Granville, A., *Cycle lengths in a permutation are typically Poisson distributed*, Electr. J. Combinatorics **13 / R107** (2006), 23.
8. Hardy, G.H. and Wright, E.M., *Introduction to the theory of numbers*, Oxford, 1932.
9. Hildebrand, A. and Tenenbaum, G. (1988), *On the number of prime factors of an integer*, Duke Math. J **56**, 471501.
10. Katz, N.M. and Sarnak, P. (1999), *Zeroes of zeta functions and symmetry*, Bull. Amer. Math. Soc **36**, 1–26..
11. Knuth, D.E. and Trabb Prado, L. (1976), *Analysis of a simple factorization algorithm*, Theoret. Comput. Sci **3**, 321-348.
12. Moser, L. and Wyman, M. (1958), *Asymptotic development of the Stirling numbers of the first kind*, J. London Math. Soc **33**, 133-146.
13. Panario, D. and Richmond, B. (2001), *Smallest components in decomposable structures: exp-log class. Average-case analysis of algorithms*, Algorthmica **29**, 205–226.
14. Vershik, A.M. (1987), *The asymptotic distribution of factorizations of natural numbers into prime divisors*, Soviet Math. Dokl **34**, 57–61.
15. Vershik, A.M. (1995), *Asymptotic combinatorics and algebraic analysis*, Proc ICM Zurich (1994), 1384–1394.

4. Even more evidence.

There seem to be more than enough compelling analogies between the anatomies of integers and permutations in the article above, but we did find even more. We list another three below.

[20] It is easy to construct a random permutation by letting 1 be mapped to a random number $\sigma(1) \in \{1, \ldots, N\}$, then 2 be mapped to a random number $\sigma(2) \in \{1, \ldots, N\} \setminus \{\sigma(1)\}$, etcetera.

[21] By "random prime numbers" we mean we select each prime in the interval with roughly equal probability. To do this we select an integer at random in the interval and determine whether or not it is prime: if not, select another integer, and then another, until we find a prime.

4.1. Divisors.

Almost all integers n have $\log\log n$ prime factors, few of them repeated, and thus $2^{\{1+o(1)\}\log\log n} = (\log n)^{\log 2+o(1)}$ divisors. One might then guess that the average number of divisors of an integer is about this, but one's guess would be wrong. Indeed the average number of divisors of an integer $\leq x$ is $\log x + O(1)$. This is not hard to prove

$$\frac{1}{x}\sum_{n\leq x}\sum_{d|n} 1 = \frac{1}{x}\sum_{d\leq x}\sum_{\substack{n\leq x\\ d|n}} = \frac{1}{x}\sum_{d\leq x}\left[\frac{x}{d}\right] = \frac{1}{x}\sum_{d\leq x}\Big(\frac{x}{d}+O(1)\Big) = x(\log x+\gamma+O(1)),$$

and in fact DIRICHLET showed that the "$\gamma+O(1)$" could be replaced by $2\gamma-1+O(1/\sqrt{x})$ by counting over the divisors of n that are $\leq \sqrt{n}$.

What notion concerning permutations corresponds to divisors? Since cycles "correspond to primes", thus unions of "cycles" correspond to "divisors". In other words define $D(\sigma)$, the set of sub-divisors of $\sigma = C_1C_2\dots C_\ell$, to be $\{\cup_{i\in I}C_i : \ I\subset\{1,2,\dots,\ell\}\}$. We have seen that almost all permutations in S_N have $\sim \log N$ cycles, so have $2^{\{1+o(1)\}\log N} = N^{\{1+o(1)\}\log 2}$ sub-divisors, so what is the average going to be?

$$\frac{1}{N!}\sum_{\sigma\in S_N}|D(\sigma)| = \frac{1}{N!}\sum_{\sigma\in S_N}\sum_{C_1,\dots,C_k\in\sigma} 1 = \frac{1}{N!}\sum_{C_1,\dots,C_k}\sum_{\substack{\sigma\in S_N\\ C_1,\dots,C_k\in\sigma}} 1$$

$$= \sum_{C_1,\dots,C_k}\frac{(N-\sum_i d(C_i))!}{N!} = \sum_{m=1}^{N}\sum_{\substack{a_1,a_2,\dots\geq 0\\ \sum_i ia_i=m}}\prod_i\frac{1}{i^{a_i}a_i!} = \sum_{m=1}^{N} 1 = N;$$

exactly analogous again!

4.2. The Hardy-Ramanujan upper bound.

Hardy and Ramanujan showed that there exist constants $c_1, c_2 > 0$ such that the number of integers $\leq x$ with exactly ℓ prime factors is

$$\leq \frac{c_1x}{\log x}\frac{(\log\log x+c_2)^{\ell-1}}{(\ell-1)!}.$$

We will prove an analogous bound for permutations by induction: there are

$$\leq \frac{N!}{N}\frac{\mu_N^{\ell-1}}{(\ell-1)!} \quad \text{where} \quad \mu_N := \sum_{m=1}^{N}\frac{1}{m} \tag{1}$$

permutations in $S_{N,\ell}$. This is true for $\ell=1$ (with equality, as we noted in section 1). For $\ell=2$ we can write each permutation as the union of two cycles, one of length $m\leq N/2$ the other of length $N-m$. Therefore if N is odd,

$$\frac{|S_{N,2}|}{N!} = \frac{1}{N!}\sum_{\substack{\sigma\in S_N\\ \ell(\sigma)=2}} 1 = \sum_{1\leq m<N/2}\frac{1}{m(N-m)} = \frac{1}{N}\sum_{1\leq m<N/2}\Big(\frac{1}{m}+\frac{1}{N-m}\Big) = \frac{\mu_N}{N}$$

so (1) holds with equality. If N is even, we have to add in a term $1/2m^2$ where $m = N/2$, but which then contributes the missing $1/(N/2)$ term to the sum for μ_N, so (1) again holds with equality.

Now suppose (1) is true for $\ell = k \geq 2$ and consider the case $\ell = k+1$: Note that if σ has cycles of length $1 \leq d_1(\sigma) \leq \cdots \leq d_k(\sigma) \leq d_{k+1}(\sigma)$ then each $d_k(\sigma) \leq (1/2)(\sum_i d_i(\sigma) - d_1(\sigma)) \leq (N-1)/2 < N/2$. Therefore

$$k|S_{N,k+1}| \leq \sum_{\substack{\sigma \in S_N \\ \ell(\sigma)=k+1}} \sum_{\substack{C \in \sigma \\ d(C)<N/2}} = \sum_{C:\ d(C)<N/2} \sum_{\substack{\sigma \in S_N, C \in \sigma \\ \ell(\sigma)=k+1}} 1 = \sum_{C:\ d(C)<N/2} |S_{N-d(C),k}|,$$

and so, since there are $N!/(m(N-m)!)$ possible cycles of length m in S_N,

$$\begin{aligned} \frac{|S_{N,k+1}|}{N!} &\leq \frac{1}{k} \sum_{1 \leq m < N/2} \frac{1}{m} \frac{|S_{N-m,k}|}{(N-m)!} \leq \frac{1}{k} \sum_{1 \leq m < N/2} \frac{1}{m} \cdot \frac{1}{N-m} \frac{\mu_{N-m}^{k-1}}{(k-1)!} \\ &\leq \frac{\mu_N^{k-1}}{k!} \sum_{1 \leq m < N/2} \frac{1}{N} \left(\frac{1}{m} + \frac{1}{N-m} \right) \leq \frac{1}{N} \frac{\mu_N^k}{k!} \end{aligned}$$

by the induction hypothesis.

4.3. More functions on the parts.

Let f be a function defined on the length of a cycle, and then its value extended multiplicatively to permutations. In other words

$$f(\sigma) = f(C_1)f(C_2)\cdots f(C_\ell) = f(d_1(\sigma))f(d_2(\sigma))\cdots f(d_\ell(\sigma)).$$

We are interested in the mean value of f on S_N, that is $(1/N!)\sum_{\sigma \in S_N} f(\sigma)$. Note that the questions of section 1.4 are special cases of this. The number theory version of these problems can be efficiently handled by using sieve identities, which we imitate below: Define $\tau(N) = \tau_f(N) = (1/N!)\sum_{\sigma \in S_N} f(\sigma)$. Then, writing $\sigma = C \cup \phi$ when $C \in \sigma$,

$$\begin{aligned} N!\tau(N)N &= \sum_{\sigma \in S_N} f(\sigma)N = \sum_{\sigma \in S_N} f(\sigma) \sum_{C_i \in \sigma} d(C_i) \\ &= \sum_C d(C) \sum_{\substack{\sigma \in S_N \\ C \in \sigma}} f(\sigma) = \sum_C d(C) \sum_{C \cup \phi \in S_N} f(C)f(\phi) \\ &= \sum_C d(C)f(C) \sum_{\phi \in S_{N-d(C)}} f(\phi) = \sum_C d(C)f(d(C))(N-d(C))!\tau(N-d(C)) \\ &= \sum_{m=0}^{N} mf(m)(N-m)!\tau(N-m)\#\{\text{cycles } C \in S_N : d(C) = m\} \\ &= \sum_{m=0}^{N} mf(m)(N-m)!\tau(N-m)\frac{N(N-1)(N-2)\dots(N-m+1)}{m} \end{aligned}$$

and therefore we deduce that

$$\tau(N) = \frac{1}{N}\sum_{m=0}^{N} f(m)\tau(N-m).$$

Suppose that $f(n) = 1$ for all $n \leq Y$ so that $\tau(x) = 1$ for all $x \leq Y$; we re-normalize writing $\theta(t) = \tau(tY)$. Suppose f is well-approximated by a "smooth" (re-normalized) function χ, so that $f(m) \approx \chi(m/Y)$; then the above relation can be re-written as, for $N = uY$,

$$\theta(u) \approx \frac{1}{u}\int_0^u \chi(t)\theta(u-t)\mathrm{d}t \text{ for } u \geq 1. \tag{2}$$

In other words, given Y and f with $f(n) = 1$ we can construct χ (for example by extrapolating linearly between values $f(m)$ and $f(m+1)$) and then deduce a good approximation for τ using (2). In fact if $Y \to \infty$ as $N \to \infty$ then $\tau(N) \sim \theta(u)$.

Something very similar happens with integers. Function f is multiplicative if $f(mn) = f(m)f(n)$ whenever $\gcd(m,n) = 1$. We are interested in $(1/x)\sum_{n\leq x} f(n)$ which we denote by $F(u)$ when $x = y^u$ for some given y; and we suppose that $f(n) = 1$ whenever $n \leq y$. We define $\chi(t) = 1$ for $t \leq 1$; and then $\chi(t) := (1/y^t)\sum_{p\leq y^t} f(p)\log p$ where the sum is over primes p. Then $F(u) \sim \theta(u)$ where $\theta(u) = 1$ for $u \leq 1$ and otherwise $\theta(u)$ is determined by (2).

Another approach to understanding such equations comes simply from noting that there are $N!/\prod_i i^{a_i}a_i!$ permutations $\sigma \in S_N$ with a_i cycles of length i for each i (this holds if and only if $\sum_i ia_i = N$). Therefore

$$\sum_{N\geq 0}\left(\sum_{\sigma\in S_N} f(\sigma)\right)\frac{X^N}{N!} = \sum_{a_1,a_2,\dots\geq 0}\prod_{i\geq 1}\frac{f(i)^{a_i}X^{ia_i}}{i^{a_i}a_i!} = \exp\left(\sum_{i\geq 1}\frac{f(i)X^i}{i}\right).$$

Note also we can obtain the derivative of the logarithm of this equation from $\tau(N) = (1/N)\sum_{m=0}^{N} f(m)\tau(N-m)$ by multiplying both sides by NX^N, and summing over $N \geq 0$. One can do the analogous operations to (2): multiply both sides by ue^{su} and integrate over $u \geq 1$ to obtain $\mathcal{L}'(\theta,s) = \mathcal{L}(\chi,s)\mathcal{L}(\theta,s)$ where $\mathcal{L}(g,s) = \int_0^\infty g(u)e^{su}\mathrm{d}u$ is the Laplace transform of g; and thus, integrating,

$$\mathcal{L}(\theta,s)/\mathcal{L}(\theta,0) = \exp\left(\int_0^s \mathcal{L}(\chi,w)\mathrm{d}w\right),$$

the analogy to what appears above. Much discussion of such integral equations (in number theory) can be found in [17].

Another way to develop the above is to rewrite

$$\begin{aligned}\sum_{N\geq 0}\tau_f(N)X^N &= (1-X)^{-1}\exp\left(-\sum_{i\geq 1}\frac{(1-f(i))X^i}{i}\right)\\ &= \sum_{i\geq 0}X^i\cdot\sum_{j\geq 0}\frac{(-1)^j}{j!}\left(\sum_{i\geq Y+1}\frac{(1-f(i))X^i}{i}\right)^j\end{aligned}$$

if $f(i) = 1$ for all $i \leq Y$, so that

$$\tau_f(N) = \sum_{j\geq 0} \frac{(-1)^j}{j!} \sum_{\substack{i_1,i_2,\dots,i_j>Y \\ i_1+i_2+\cdots+i_j\leq N}} \frac{(1-f(i_1))\dots(1-f(i_j))}{i_1 i_2 \dots i_j},$$

which is in analogy to [17], Theorem 3.3.

Finally, for the ambitious reader; We estimated in section 4 the integers with no large prime factors/the permutations with no large cycles, and in section 5 the integers with no small prime factors/the permutations with no small cycles. Friedlander (1976) gave an asymptotic formula for the integers with neither large nor small prime factors – it would be interesting to prove the analogous result for permutations.

It would also be interesting to determine $\Psi[S_N, Y]$, the number of permutations in S_N all of whose cycle lengths are $\leq Y$, in a wide range. We have seen that this is $\sim N!\rho(u)$ if $Y = N/u$, but what is $u \to \infty$. It is known that $\Psi(x, y) \sim x\rho(u)$ for y in a very wide range (and indeed it holds for all $t > (\log x)^{2+o(1)}$ if and only if the Riemann Hypothesis is true). Another interesting case in where $\Psi(x, (\log x)^A) \sim x^{1-1/A}$ for any fixed $A > 1$; I believe that the analogous question is to estimate $\Psi[S_N, A \log N]$ but cannot even guess what that might equal.

[16] Friedlander, J.B. (1976), *Integers free of large and small primes*, Proc. London Math. Soc **3**, 565-576.

[17] Granville, A. and Soundararajan, K. (2001), *The Spectrum of Multiplicative Functions*, Ann. of Math **153**, 407–470.

Départment de Mathématiques et Statistique, Université de Montréal, CP 6128 succ Centre-Ville,
Montréal, Québec H3C 3J7, Canada.
andrew@dms.umontreal.ca
http://www.dms.umontreal.ca/∼andrew/