



MSI: Anatomy

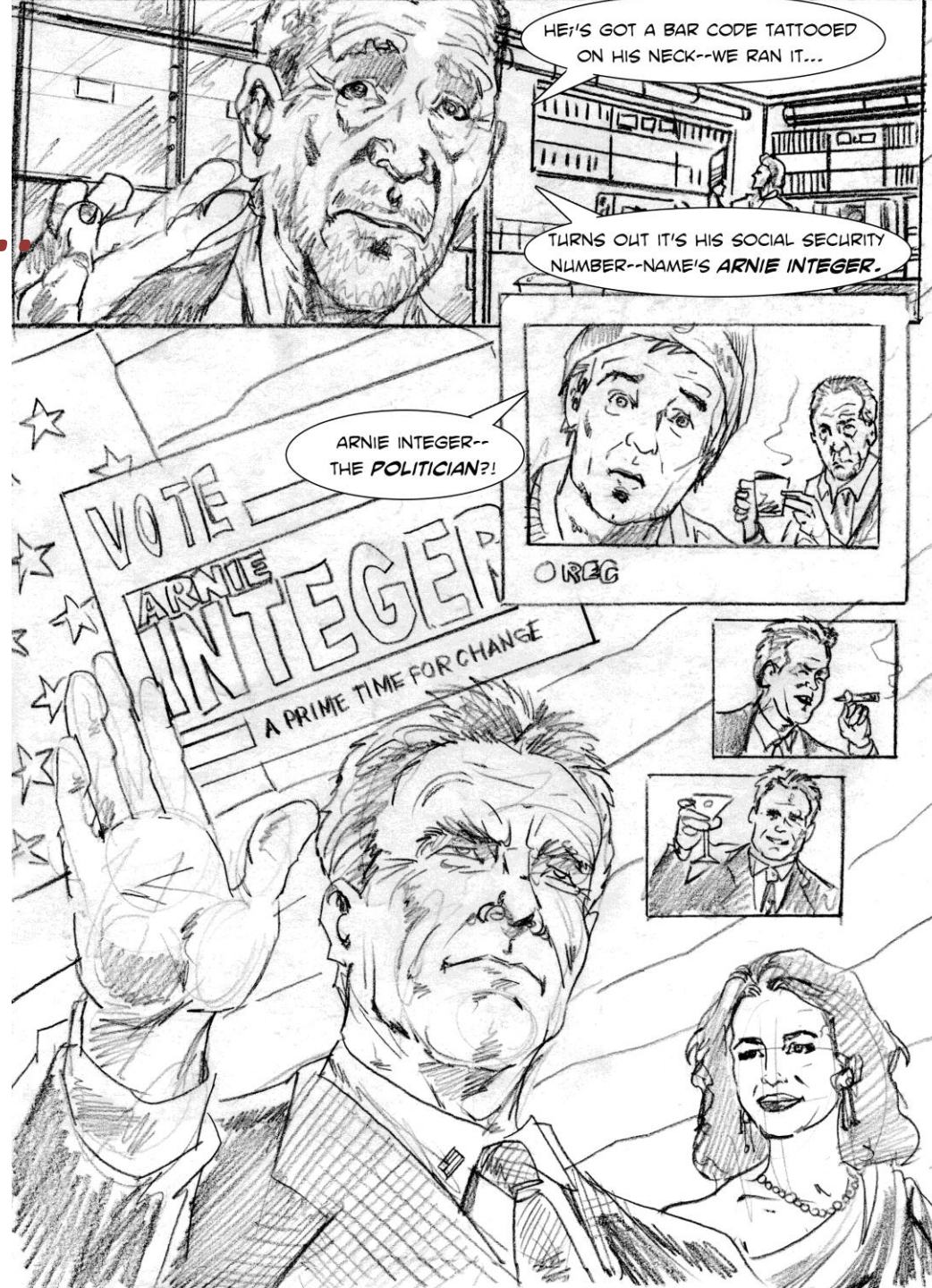
(of integers and
permutations)

Andrew Granville

(Université de Montréal)

*There have been
two homicides...*

An
integer:



*There have been
two homicides...*

And a
permutation

Daisy Permutation



anatomy [a-nat-o-my] noun

1. The scientific study of the shape and structure of an organism and the inter-relation of its various parts.
2. The art of separating the parts of an organism in order to ascertain their position, relations, structure, and function..

-On-Line Dictionary (2006)

*We need a mathematical
forensics expert*



Professeur **GAUSS**

And his two students / assistants

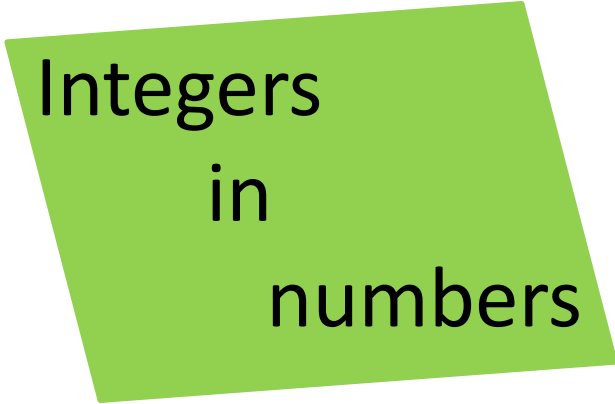


Emmy

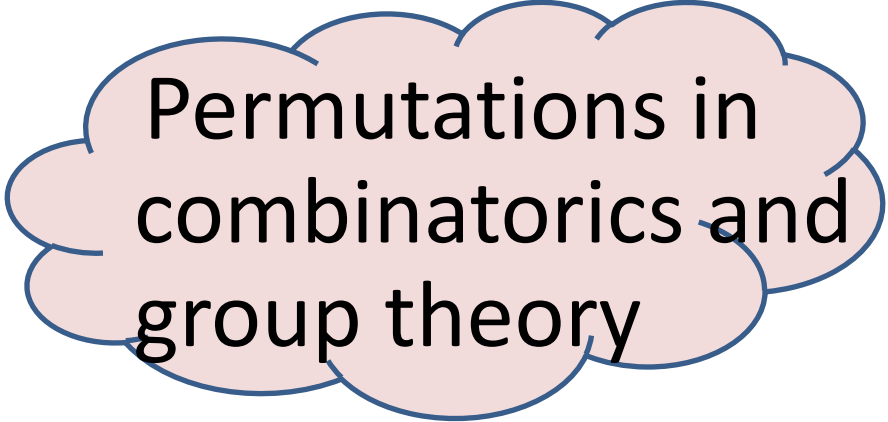


Langer

Different mathematical subjects
involve different basic objects; e.g.



Integers
in
numbers



Permutations in
combinatorics and
group theory

These objects come from very different worlds

– Can we compare them ?

A mathematical detective can compare and
contrast them, by studying their ``Anatomy''

Integers: The numbers $-3, -2, -1, 0, 1, 2, 3, \dots$

A prime number is an integer ≥ 2 , only divisible by 1 and itself.
All positive integers can be factored into a (unique) product of prime numbers.

The Fundamental Theorem of Arithmetic.

(Euclid's *Elements*, 4th century A.D.)

Example:

$$12 = 2 \times 2 \times 3 .$$

Each of 2 and 3 are primes.

No other way to factor 12 though

$$12 = 2 \times 3 \times 2 \text{ and } 12 = 3 \times 2 \times 2 .$$

Integers cannot be decomposed any further than into primes

The genetic code of Integers

The decomposition of an integer into primes cannot be broken down any further, so the primes are indeed the fundamental constituent parts of integers.

Every integer is composed of primes, and each integer is composed of a different set of primes (keeping track of how often each prime appears in the decomposition). Therefore you can just as accurately identify an integer through its set of prime factors as through the integer itself. It's like the DNA of the integer.

Primes are the fundamental constituent parts of integers, their genetic code, if you like. Any integer can be identified by the primes it contains, which ones and how many of each type.

Permutations: Re-organization of N objects

Playing card games at the casino:

You easily win if you know the order of the cards.

When the croupier shuffles one want to know how the cards are re-organized. (this is a permutation of the cards)



Useful fact 1: After seven riffle shuffles most of the $52!$ possible orders of the cards can occur, with roughly equal probability

Permutations: Re-organization of N objects



Playing card games at the casino:

You easily win if you know the order of the cards.

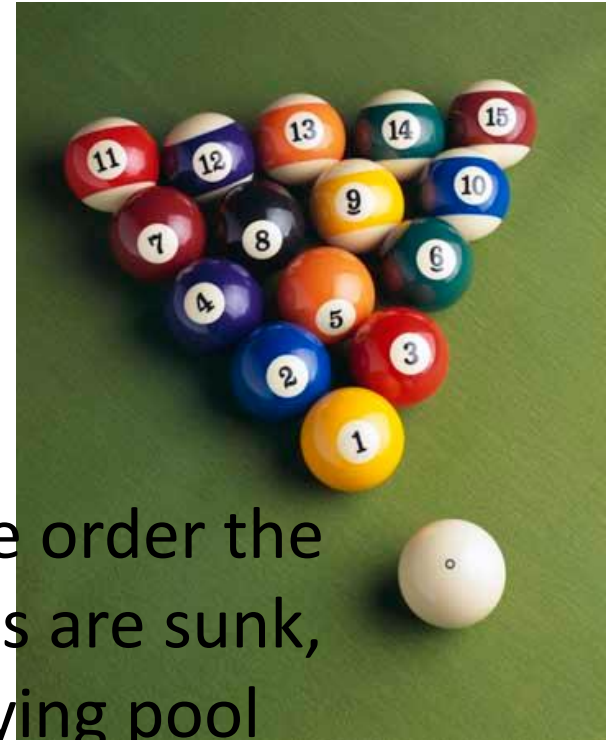
Useful fact 2: After eight perfect riffle shuffles the deck of cards returns to its starting position.

Permutations: Re-organization of N objects

Organizing objects comes up in many areas:



Where students sit in class



The order the balls are sunk, playing pool



The order of the competitors in a sports competition

Permutations: Re-organization of N objects

In the theory of re-organization, it is not the actual type of object that matters. We can label the objects $1, 2, 3, \dots, N$ in their starting order, and then look at the order of these numbers at the end.

This is a permutation σ :

The object in position 1 moves to position $\sigma(1)$

The object in position 2 moves to position $\sigma(2)$

.....

The object in position N moves to position $\sigma(N)$

Then the numbers $\sigma(1), \sigma(2), \dots, \sigma(N)$ is a re-arrangement of the numbers $1, 2, \dots, N$.

Persi Diaconis left home at 14 to travel with card magic legend [Dai Vernon](#), entertaining on cruise ships. Diaconis started creating his own card tricks based on mathematics. Discovered by Martin Gardner he started university at 24, getting a Ph.D. at 29, and is now Professor of Mathematical Statistics at Stanford.



Permutations: Re-organization of N objects

Example, $N=2$: Possible maps:

$1 \rightarrow 1$ and $2 \rightarrow 2$, the identity map;

or

$1 \rightarrow 2$ and $2 \rightarrow 1$,

which we can represent as

$1 \rightarrow 2 \rightarrow 1$ or $1 \leftrightarrow 2$.

All possible permutations

N=2: Possible maps

- $1 \leftrightarrow 1$ and $2 \leftrightarrow 2$, the identity map;

or

- $1 \rightarrow 2$ and $2 \rightarrow 1$

which we can represent as

$$1 \rightarrow 2 \rightarrow 1 \text{ or } 1 \leftrightarrow 2.$$

N=3: Six permutations:

- $1 \leftrightarrow 1, 2 \leftrightarrow 2, 3 \leftrightarrow 3$
- $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$
- $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$
- $1 \leftrightarrow 1, 2 \leftrightarrow 3$
- $2 \leftrightarrow 2, 1 \leftrightarrow 3$
- $3 \leftrightarrow 3, 1 \leftrightarrow 2$

Permutations break up into cycles

N=2: Possible maps

- $1 \leftrightarrow 1$ and $2 \leftrightarrow 2$, the identity map;
or
- $1 \rightarrow 2$ and $2 \rightarrow 1$

which we can represent as

$1 \rightarrow 2 \rightarrow 1$ or $1 \leftrightarrow 2$.

N=2: Two permutations

$(1) (2)$

or

$(1\ 2)$

N=3: Six permutations:

- $1 \leftrightarrow 1, 2 \leftrightarrow 2, 3 \leftrightarrow 3$
- $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$
- $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$
- $1 \leftrightarrow 1, 2 \leftrightarrow 3$
- $2 \leftrightarrow 2, 1 \leftrightarrow 3$
- $3 \leftrightarrow 3, 1 \leftrightarrow 2$

N=3: Six permutations:

$(1) (2) (3)$

$(1\ 2\ 3)$

$(1\ 3\ 2)$

$(1) (2\ 3)$

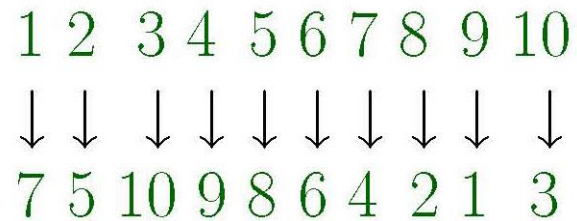
$(2) (1\ 3)$

$(3) (1\ 2)$

Permutations break up into cycles

All permutations break up into cycles in a unique way.

Example: The permutation



Permutations break up into cycles

All permutations break up into cycles in a unique way.

Example: The permutation

1	2	3	4	5	6	7	8	9	10
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
7	5	10	9	8	6	4	2	1	3

is more transparently written as $(1\ 7\ 4\ 9)\ (2\ 5\ 8)\ (3\ 10)\ (6)$

All permutations can be written into a product of cycles (each involving entirely different elements) in a unique way, apart from the order in which the cycles are written, and the element with which each cycle begins; e.g. the

above equals $(6)\ (2\ 5\ 8)\ (10\ 3)\ (7\ 4\ 9\ 1)$

or $(10\ 3)\ (9\ 1\ 7\ 4)\ (6)\ (8\ 2\ 5)$

The genetic code of Permutations

The decomposition of a permutation into cycles cannot be broken down any further, so the cycles are the fundamental constituent parts of permutations.

Every permutation is composed of them, and each permutation is composed of a different set of cycles. Therefore you can just as accurately identify a permutation through its set of cycles as through the permutation itself. It's like the DNA of the permutation.

Cycles are the fundamental constituent parts of permutations, their genetic code, if you like. Any permutation can be identified by the cycles that it contains.

Sounds familiar?

Comparing the genetic codes

Integers

The decomposition of an integer into primes cannot be broken down any further, so the primes are the fundamental constituent parts of integers.

Every integer is composed of them, and each integer is composed of a different set of primes. Therefore you can just as accurately identify an integer through its set of prime factors as through the integer itself. It's like the DNA of the integer.

Primes are the fundamental constituent parts of integers, their genetic code, if you like. Any integer can be identified by the primes that it contains.

Permutations

The decomposition of a permutation into cycles cannot be broken down any further, so the cycles are the fundamental constituent parts of permutations.

Every permutation is composed of them, and each permutation is composed of a different set of cycles. Therefore you can just as accurately identify a permutation through its set of cycles as through the permutation itself. It's like the DNA of the permutation.

Cycles are the fundamental constituent parts of permutations, their genetic code, if you like. Any permutation can be identified by the cycles that it contains.

Integers and Permutations: Chalk and cheese?

The fundamental
components



of Integers are primes

The fundamental components



of Permutations are cycles.

A vague qualitative analogy ----

Need a richer quantitative analogy.

A calibration to compare cycles and prime factors?

A calibration to compare cycles and prime factors?

médico-légal

1. Exercée pour aider la justice, en cas de crime.
2. Concernant l'utilisation de la science ou la technologie dans l'enquête et l'établissement des faits ou des éléments de preuve.

-Le Robert **micro** (2006)

Forensics – Science or Art ?

- When comparing the anatomies of two seemingly different organisms, the forensic scientist knows that one must calibrate their sizes else one might be misled into believing that they are different, whereas they might be twin organisms that have grown at different speeds in different environments. In order to do such a calibration, one needs to find some essential feature of the organisms, that allows one to better compare the two objects. So how does one identify what are the key constituents of each organism? Forensic scientists consider the selection and measurement of this key constituent to be as much an art as a science.
- In order to properly calibrate integers and permutations, we must therefore get a better idea of how they typically look. We have already identified their fundamental, indecomposable components, the question is how to compare them. We begin with a fundamental question:
- What proportion of integers, and of permutations, are fundamental?

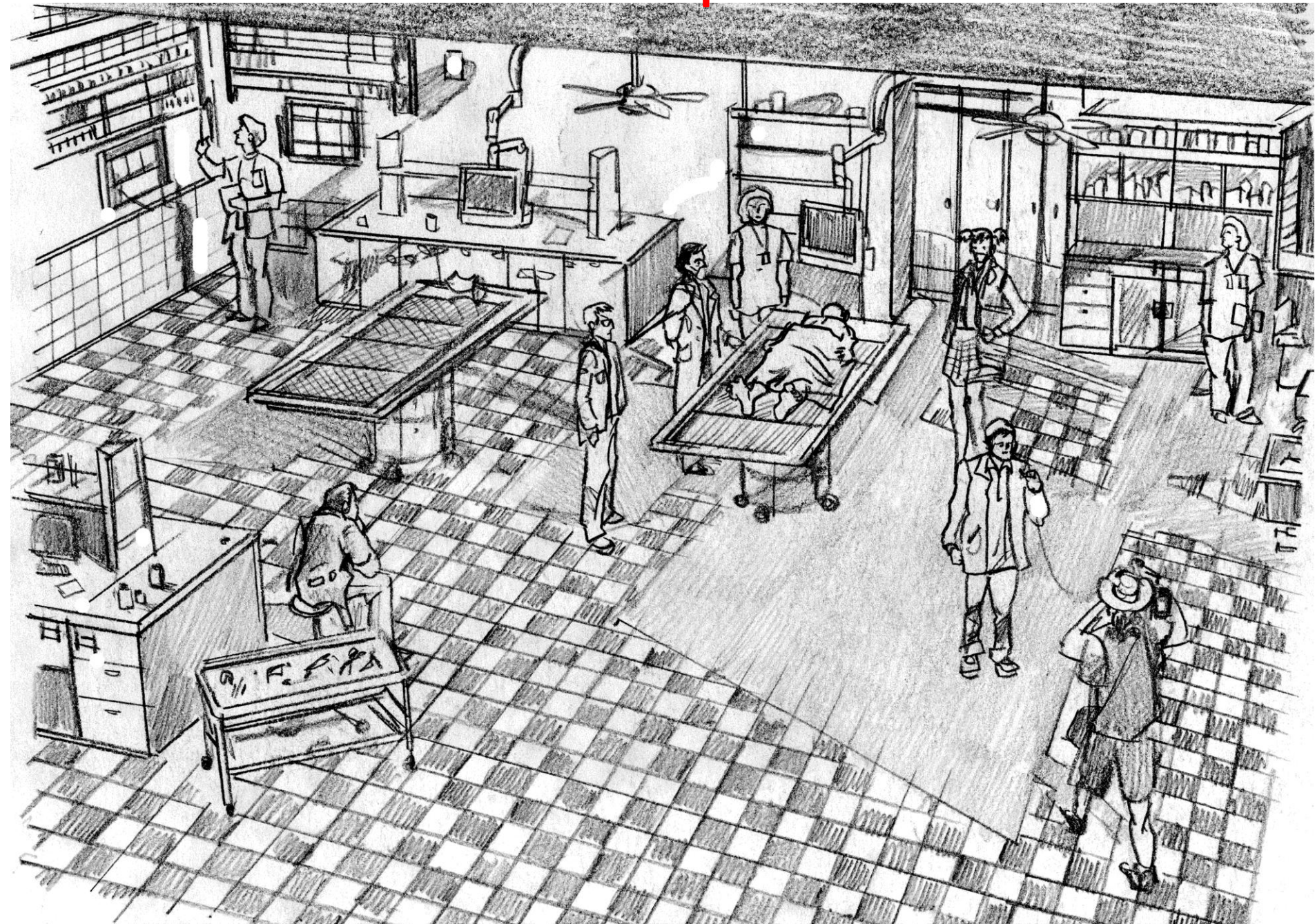
A possible calibration?

What proportion of integers, and of permutations, are *fundamental*?

That is:

- What proportion of integers are prime?
- What proportion of permutations are cycles?

The autopsies



What proportion are fundamental?

What proportion of permutations are fundamental?

(Have just one fundamental component? Is a cycle?)

How many permutations σ on N letters?

- N choices for $\sigma(1)$: $\sigma(1)=1$ or 2 or ... or N ;
- $N-1$ choices for $\sigma(2)$: $\sigma(2)=1$ or 2 or ... or N but not $\sigma(1)$;
- $N-2$ choices for $\sigma(3)$: $\sigma(3)=1$ or ... or N but not $\sigma(1)$ or $\sigma(2)$;
-
- 2 choices for $\sigma(N-1)$;
- 1 choices for $\sigma(N)$;

Total # of permutations

$$\text{Total \# of possible } \sigma = N \times (N-1) \times \dots \times 2 \times 1 = N!$$

What proportion of permutations are cycles?

Total # of permutations = $N!$

What is the total # cycles on N letters?

Idea: Trace the path of first element...

Cycle $\sigma = (1, \chi(1), \chi(2), \chi(3), \dots, \chi(N-1))$

Path does not cross to itself until the end:

That is $1, \chi(1), \chi(2), \dots, \chi(N-1)$ are all different:

- $N-1$ choices for $\chi(1)$: $\chi(1) = 1$ or 2 or ... or N but not 1 ;
- $N-2$ choices for $\chi(2)$: $\chi(2) = 1$ or ... or N but not 1 or $\chi(1)$;
-
- 2 choices for $\chi(N-2)$
- 1 choice for $\chi(N-1)$

Total # of cycles =
 $(N-1) \times (N-2) \times \dots \times 1 = (N-1)!$

What proportion of permutations are cycles?

permutations on N letters is $N!$

cycles on N letters is $(N-1)!$

$$\text{So proportion} = \frac{\text{\#cycles}}{\text{\#permutations}} = \frac{(N-1)!}{N!} = \frac{1}{N}$$

The proportion of permutations that are cycles is

$1/N$

The proportion of permutations that are indecomposable is $1/N$.

What proportion of integers are indecomposable?

What proportion of the integers up to x are prime?

This is a much deeper question for integers than for permutations....

GAUSS (at 16): The density of primes around x is about $1/\log x$

Took >100 years to prove.

$$\#\{\text{primes} \leq x\} \approx \int_2^x \frac{dt}{\log t}$$

x	$\pi(x) = \#\{\text{primes} \leq x\}$	Error: $\int_2^x \frac{dt}{\log t} - \pi(x)$
10^3	168	10
10^4	1229	17
10^5	9592	38
10^6	78498	130
10^7	664579	339
10^8	5761455	754
10^9	50847534	1701
10^{10}	455052511	3104
10^{11}	4118054813	11588
10^{12}	37607912018	38263
10^{13}	346065536839	108971
10^{14}	3204941750802	314890
10^{15}	29844570422669	1052619
10^{16}	279238341033925	3214632
10^{17}	2623557157654233	7956589
10^{18}	24739954287740860	21949555
10^{19}	234057667276344607	99877775
10^{20}	2220819602560918840	222744644
10^{21}	21127269486018731928	597394254
10^{22}	201467286689315906290	1932355208
10^{23}	1925320391606818006727	7236148412

Calibration?

One in every N permutations on N letters is a cycle

One in every $\log x$ integers up to x is prime.

Proposed Calibration

N

when we measure the
anatomy of a permutation

vs.

$\log x$

when we measure the
anatomy of an integer.

Let's check it out...

Does our calibration makes sense?

Proportion of permutations with exactly k cycles:

$$\approx \frac{1}{N} \frac{(\log N)^{k-1}}{(k-1)!}$$

Now replace N by $\log x$, to guess:

Proportion of integers with exactly k prime factors:

$$\approx \frac{1}{\log x} \frac{(\log \log x)^{k-1}}{(k-1)!}$$

(True: Hardy and Ramanujan)

Calibration?

How many indecomposable components is "typical"?

- A typical Permutation has about $\log N$ cycles
- A typical Integer has about $\log \log x$ prime factors

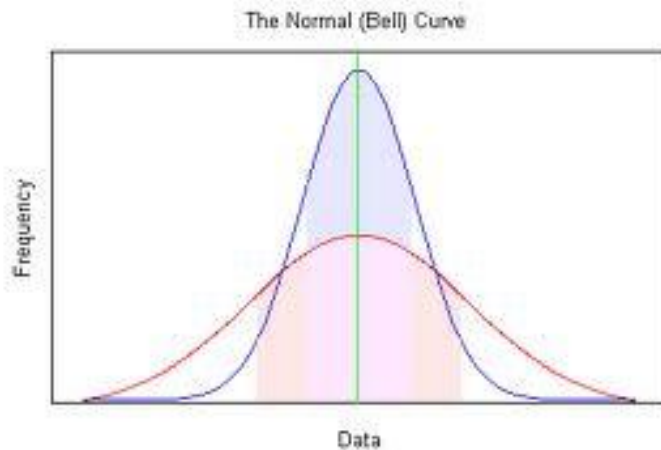
Not all integers have about $\log \log x$ prime factors: Primes have one, numbers like $2 \times 3 \times 5 \times 7 \times 11 \times \dots$ have many more.

Similarly not all permutations have about $\log N$ cycles; $(1\ 2\ \dots\ N)$ has one cycle and $(1)(2)\dots(N)$ has N cycles

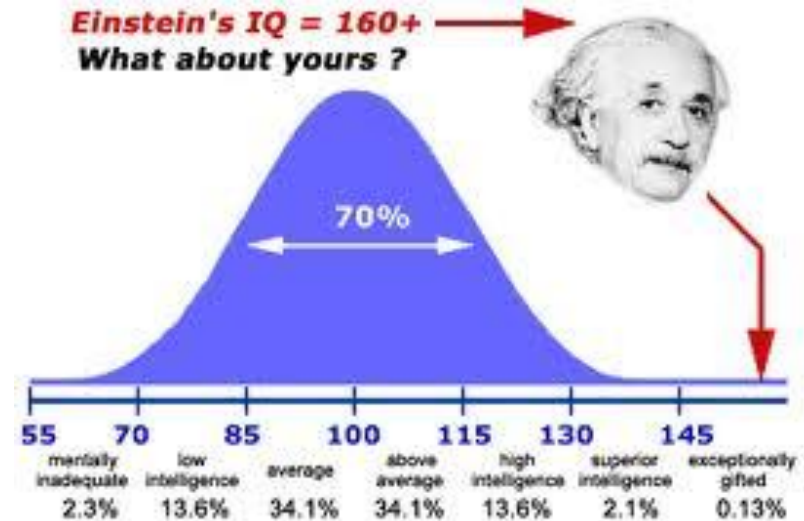
What about their distribution?

Distribution of the number of parts

Data that seems chaotic often organizes itself into certain recognizable patterns. The most common is where, when you graph the data, the plot is like a bell around the average.



Center of the bell is given by the mean
Width of the bell by the variance.



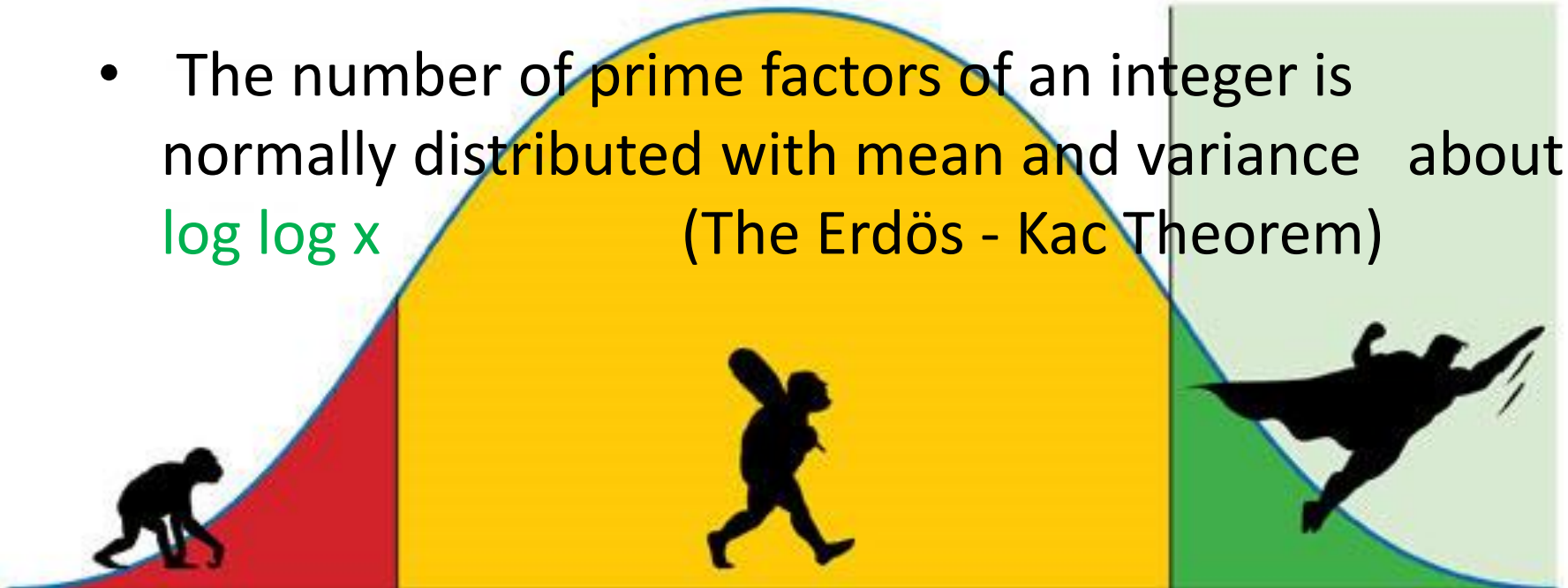
All the bells have the same basic shape, though the center may appear in different places, and some may be fatter than others

THE NORMAL DISTRIBUTION

A typical Permutation has about $\log N$ cycles
A typical Integer has about $\log \log x$ prime factors

What about their distribution?

- The number of cycles in a permutation is normally distributed with mean and variance about $\log N$
- The number of prime factors of an integer is normally distributed with mean and variance about $\log \log x$ (The Erdős - Kac Theorem)



Sizes of the indecomposable components?

There are $\log N$ cycles in a typical N letter permutation.

These $\log N$ integer lengths add up to N .

Can we predict the lengths of those cycles?

Occam's razor:

What is the simplest sequence of about $\log N$ numbers up to N ?



Occam's razor

What is the simplest
sequence of about $\log N$
elements up to N ?

$e^1, e^2, e^3, \dots, e^{\lceil \log N \rceil}$

But these are not integers;
and surely the cycle lengths
could not be that regular?

**IDEA: TAKE LOGS OF THE CYCLE LENGTHS
AND SEE HOW THESE ARE DISTRIBUTED?**

Occam's Razor

through the ages...



*Pluralitas non
est ponenda sine
necessitate.*

*(Plurality should not be
posited without necessity.)*

- William of Ockham

Everything should be
made as simple as
possible, but not
simpler.

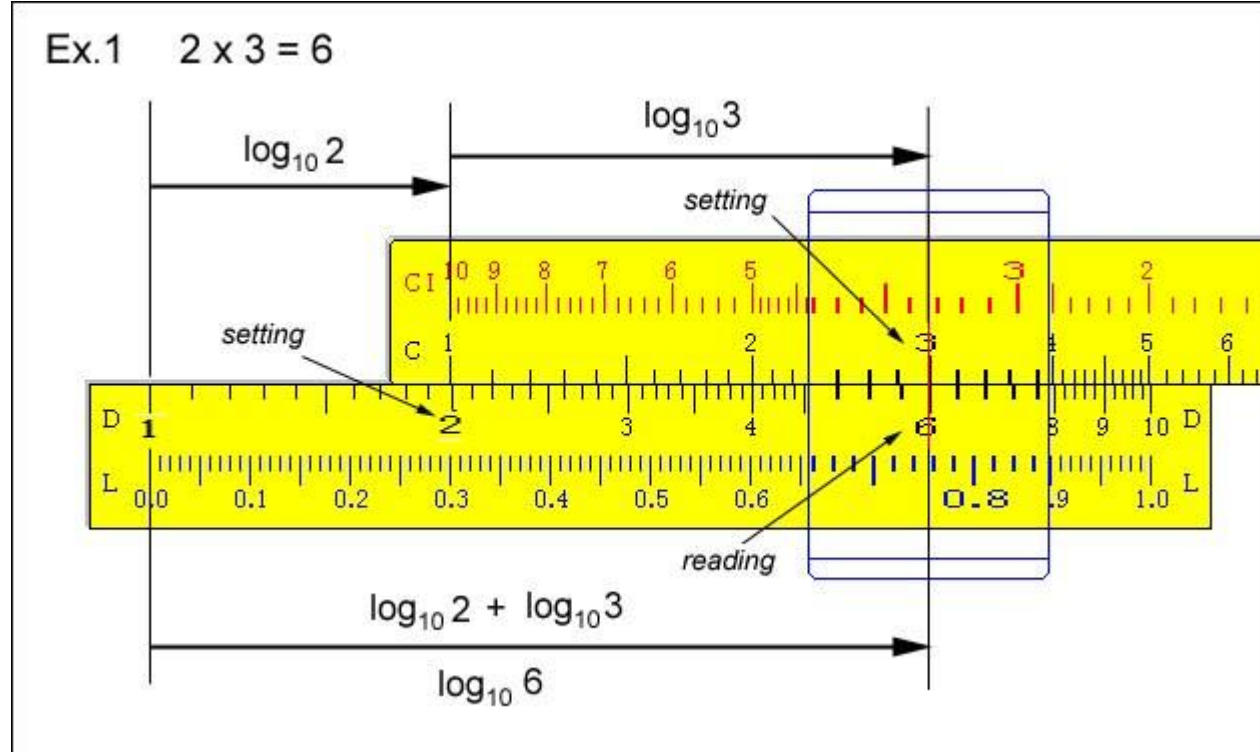
- Albert Einstein



Keep
It
Simple,
Stupid !

Sizes of the
indecomposable
components?

IDEA: TAKE
LOGS OF THE
CYCLE LENGTHS
AND SEE HOW
THESE ARE
DISTRIBUTED?



Now we have about $\log N$ numbers between 0 and $\log N$,
which add up to $\log N$. How are these distributed?

Randomly?

What is “randomly“?

How are random numbers distributed in an interval?

How are random numbers distributed in an interval?

3600 people open www.crm.umontreal.ca

in an hour.



Centre de recherche mathematiques X

Search

3600 hits in 3600 seconds

That's one hit per second.

Do we really expect one hit *every* second?

("One hit per second" is an average)

How are random numbers distributed in an interval?



3600 people open www.crm.umontreal.ca in an hour.

Do we really expect one hit every second?

Of course not! Experience shows that we should get a less evenly spaced distribution of hits. There should be:

Some seconds when there are lots of hits;
Other longer periods when there are no hits

How are random numbers distributed in an interval?

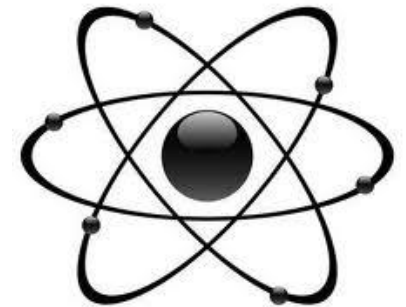


Spacings between cars on a freeway



The arrival of
customers in a queue.

The radioactive
decay of atoms



all are examples of a ...

Poisson Point Process

Poisson Point Process

If the average spacing between elements is 1 then we expect that the proportion of t second periods in which we get h hits is

$$e^{-t} \frac{t^h}{h!}$$

Expected number of secs with no hits: 1324

Number of secs with at least two hits: 951

Number of secs with at least five hits: 13

Five sec periods with no hits: 24

Poisson Point Process

So, how are the indecomposable components laid out?

The logarithms of the cycle lengths of a typical permutation form a

Poisson Point Process in $[0, \log N]$.

and

The logarithms of the logarithms of the prime factors of a typical integer form a

Poisson Point Process in $[0, \log \log x]$.

Could Integers and Permutations have the same anatomies?

When calibrated they have the same

- Proportion with **k** indecomposable components
- Typical number of indecomposable components
- Same (normal) distribution of indecomposable components
- Internal layout (Poisson Point Process)

Integers and Permutations -- the same

- Proportion with k indecomposable components
- Typical number of indecomposable components
- Same (normal) distribution of indecomposable components
- Internal layout (i.e. as a Poisson Point Process)



Twins?/ϕsniwT

"DNA" seems to form the same patterns at every feasible level...

Conclusive evidence that
Integers & Permutations
are twins?



"Twins"?

The cycle lengths and the prime factor sizes have to be distributed somehow - so perhaps it was obvious that it would be something random, like the normal and poisson distributions?



To get something interesting, perhaps we should look at unusual aspects of the anatomies of permutations and integers that are much less likely to be identical?

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?

No small components

The proportion of permutations on N letters that contain no cycle of length $< N/u$ is given by

$$\frac{u\omega(u)}{N}$$

The proportion of integers $< x$ with no prime factor p , with $\log p < (\log x)/u$ ($p < x^{1/u}$) is given by

$$\frac{u\omega(u)}{\log x}.$$

where $\omega(u)$, the Buchstab function is $1/u$ for $1 \leq u \leq 2$. For $u > 2$ we have

$$\omega(u) = \frac{1}{u} \left\{ 1 + \int_1^{u-1} \omega(t) dt \right\}$$

The value $\omega(u)$ depends on the history of $\omega(t)$ for $1 \leq t \leq u-1$.

Brain modeling

No large components

The proportion of permutations on N letters that contain only cycles of length $\leq N/u$ is given by

The proportion of integers $< x$ all of whose prime factors p , satisfy $\log p < (\log x)/u$ ($p < x^{1/u}$) is given by

$$\rho(u)$$

where $\rho(u)$, the Dickman function is 1 for $0 \leq u \leq 1$. For $u > 1$ we have

$$\rho(u) = \frac{1}{u} \int_{u-1}^u \rho(t) dt.$$

The value $\rho(u)$ depends on the history of $\rho(t)$ for $u-1 \leq t \leq u$.

Cryptography

Beyond mere co-incidence ?

Ridiculously complicated formulae for

- The proportion without small components
- The proportion without large components
- Exactly k components, with k near the mean
-

And my personal favourite :



My personal favourite

If there are more fundamental components, does the size of the largest component typically go up, or go down?

- ① More components / same space \Rightarrow Less room to be big?
- ② More components \Rightarrow More opportunities to be big?

① is correct: For almost all permutations with exactly k cycles, where $k/\log N$ is large, the longest cycle has length about

$$\frac{\log \tau}{\tau} N \quad \text{where} \quad \tau = \frac{k}{\log N}.$$

- For integers, same formula, replace N by $\log x$.

Other families with the same anatomy?

Polynomials mod p

A polynomial $f(x) \bmod p$ factors into irreducible polynomials; e.g.

$$x^2 + 1 \equiv (x + 2)(x + 3) \pmod{5}$$

Indecomposable components: The irreducible polynomials

There are p^d monic polynomials of degree d ;

of these $\approx p^d/d$ are irreducible,

Proportion: $1/d$.

Hence

Calibration: $N \iff \log x \iff d$

And it works!

Their anatomies are the same,

even though they appear differently on the outside

Integers/Permutations

Their
anatomies
seem to be
more-or-less
the same. All
of the
differences
are
superficial



jozefnovack33, Flickr

Also true of
polynomials
 $\text{mod } p$,
classes of
maps
between
sets,

This is true throughout mathematics:
Objects tend to organize themselves in certain special
patterns. It is the mathematician's job to identify and
recognize those patterns

Quickly producing random integers, factored
Quick algorithm known for factoring integers? No!
Quick algorithm (easy) for finding all cycles in a permutation.

To find a random factored integer around x ,
Find a random permutation for $N = \log x$
Determine the cycle lengths m
Find a random prime in $(e^{m-1}, e^m]$ for each m

Random factored integer: Product of these primes

MSI: ANATOMY (Graphic novel) -- Available Spring 2012

written by Jennifer and Andrew Granville Drawn by Robert J. Lewis

