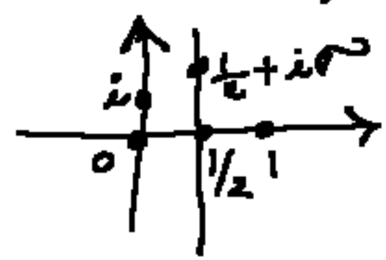
The Möbius Function and the Riemann Hypothesis

L Kauffman, UIC

Recall background on Zeta function.

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}$$

The Riemann Hypothesis (RH) says that the (non-trivial) zeroes of ζ are all on the line $\{\frac{1}{2} + i\sigma\}$ in the complex plane.



$$\prod_p \left(1-\frac{1}{p^s}\right)^{-1} = \prod_p \left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\frac{1}{p^{3s}}+\cdots\right)$$

$$= \sum_n \frac{1}{n^s}$$ by unique factorization of integers.

$$= \zeta(s).$$

$$\frac{1}{\zeta(s)} = \prod_p \left(1-\frac{1}{p^s}\right) = \sum_n \frac{\mu(n)}{n^s}$$

$$\begin{cases} \mu(p) = -1 \; ; \; \mu(1) = 1 \\ \mu(p_1 p_2 \cdots p_n) = (-1)^n \text{ when } \{p_1, \ldots, p_n\} \text{ all distinct.} \\ \mu(n) = 0 \text{ if } p^2 \mid n \text{ for some prime } p. \end{cases}$$

$$\mu : \{1, 2, 3, \ldots\} \longrightarrow \{-1, 0, 1\}$$

<u>Möbius Function</u>

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \sum_n \frac{\mu(n)}{n^s}$$

We'll see on next slides that the <u>cumulative Möbius function</u>

$$M(x) = \sum_{n<x} \mu(n)$$

is relevant to the Riemann Hypothesis.

<u>Theorem</u>: RH $\Longleftrightarrow$ $\forall \varepsilon > 0$, $M(x)\, x^{-(\frac{1}{2})-\varepsilon} \longrightarrow 0$ as $x \longrightarrow \infty$

and (remarkably) this is how a fair coin ($H=+1$, $T=-1$) behaves. But $\mu(n)$ is <u>not</u> a fair coin. So $\boxed{??}$

12.1 THE RIEMANN HYPOTHESIS AND THE GROWTH OF $M(x)$

Let dM be the Stieltjes measure such that the formula

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \qquad (\operatorname{Re} s > 1) \tag{1}$$

[(1) of Section 5.6] takes the form

$$\frac{1}{\zeta(s)} = \int_0^\infty x^{-s}\, dM(x) \qquad (\operatorname{Re} s > 1).$$

Then $M(x) = \int_0^x dM$ is a step function which is zero at $x = 0$, which is constant except at positive integers, and which has a jump of $\mu(n)$ at n. As usual, the value of M at a jump is by definition $\frac{1}{2}[M(n - \varepsilon) + M(n + \varepsilon)] = \sum_{j=1}^{n-1} \mu(j) + \frac{1}{2}\mu(n)$. Integration by parts gives for $\operatorname{Re} s > 1$

$$\begin{aligned}\frac{1}{\zeta(s)} &= \int_0^\infty d[x^{-s}M(x)] - \int_0^\infty M(x)\, d(x^{-s}) \\ &= \lim_{X\to\infty}\left[X^{-s}M(X) + s\int_0^X M(x)x^{-s-1}\, dx\right] \\ &= s\int_0^\infty M(x)x^{-s-1}\, dx\end{aligned}$$

because the obvious inequality $|M(x)| \leq x$ implies that $x^{-s}M(x) \to 0$ as $x \to \infty$ and that $\int_0^\infty M(x)x^{-s-1}\, dx$ converges, both provided $\operatorname{Re} s > 1$. Now if $M(x)$ grows less rapidly than x^a for some $a > 0$, then this integral for $1/\zeta(s)$ converges for all s in the halfplane $\{\operatorname{Re}(a - s) < 0\} = \{\operatorname{Re} s > a\}$, and therefore, by analytic continuation, the function $1/\zeta(s)$ is analytic in this halfplane. Since $1/\zeta(s)$ has poles on the line $\operatorname{Re} s = \frac{1}{2}$, this shows that $M(x)$ *does not* *grow less rapidly than x^a for any $a < \frac{1}{2}$*. Moreover, it shows that *in order to prove the Riemann hypothesis, it would suffice to prove that $M(x)$ grows less rapidly than $x^{(1/2)+\varepsilon}$ for all $\varepsilon > 0$*. Littlewood in his 1912 note [L12] on the three circles theorem proved that this sufficient condition for the Riemann hypothesis is also necessary; that is, he proved the following theorem.

Theorem The Riemann hypothesis is equivalent to the statement that for every $\varepsilon > 0$ the function $M(x)x^{-(1/2)-\varepsilon}$ approaches zero as $x \longrightarrow \infty$.

Proof It was shown above that the second statement implies the Riemann hypothesis. Assume now that the Riemann hypothesis is true. Then Backlund's proof in Section 9.4 shows [using the Riemann hypothesis to conclude that $F(s) = \zeta(s)$] that for every $\varepsilon > 0$, $\delta > 0$, and $\sigma_0 > 1$ there is a T_0 such that $|\log \zeta(\sigma + it)| < \delta \log t$ whenever $t \geq T_0$ and $\frac{1}{2} + \varepsilon \leq \sigma \leq \sigma_0$. Since $|\log \zeta(s)|$ is bounded on the halfplane $\{\text{Re } s \geq \sigma_0\}$, this implies that on the quarterplane $\{s = \sigma + it \colon \sigma = \frac{1}{2} + \varepsilon, t \geq T_0\}$ there is a constant K such that $|1/\zeta(s)| \leq Kt^\delta$. This is the essential step of the proof. Littlewood omits the remainder of the proof, stating merely that it follows from known theorems. One way of completing the proof is as follows.

The estimates (2) and (3) of Section 3.3 show that the error in the approximation

$$M(x) = \sum_{n<x} \mu(n) \sim \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left[\sum_{n=1}^{\infty} \mu(n) \left(\frac{x}{n} \right)^s \right] \frac{ds}{s}$$
$$= \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \frac{x^s}{\zeta(s)} \frac{ds}{s},$$

12.3 DENJOY'S PROBABILISTIC INTERPRETATION OF THE RIEMANN HYPOTHESIS

One of the things which makes the Riemann hypothesis so difficult is the fact that there is no plausibility argument, no hint of a reason, however unrigorous, why it should be true. This fact gives some importance to Denjoy's probabilistic interpretation of the Riemann hypothesis which, though it is quite absurd when considered carefully, gives a fleeting glimmer of plausibility to the Riemann hypothesis.

Suppose an unbiased coin is flipped a large number of times, say N times. By the de Moivre–Laplace limit theorem the probability that the number of heads deviates by less than $KN^{1/2}$ from the expected number of $\frac{1}{2}N$ is nearly equal to $\int_{-(2K^2/\pi)^{1/2}}^{(2K^2/\pi)^{1/2}} \exp(-\pi x^2)\, dx$ in the sense that the limit of these probabilities as $N \longrightarrow \infty$ is equal to this integral. Thus if the total number of heads is subtracted from the total number of tails, the probability that the resulting number is less than $2KN^{1/2}$ in absolute value is nearly equal to $2\int_0^{(2K^2/\pi)^{1/2}} \exp(-\pi x^2)\, dx$, and therefore the probability that it is less than $N^{(1/2)+\varepsilon}$ for some fixed $\varepsilon > 0$ is nearly $2\int_0^{N^\varepsilon(2\pi)^{1/2}} \exp(-\pi x^2)\, dx$. The fact that this approaches 1 as $N \longrightarrow \infty$ can be regarded as saying that *with probability one the number of heads minus the number of tails grows less rapidly than* $N^{(1/2)+\varepsilon}$.

Consider now a very large square-free integer n, that is, a very large integer n with $\mu(n) \neq 0$. Then $\mu(n) = \pm 1$. It is perhaps plausible to say that $\mu(n)$ is plus or minus one "with equal probability" because n will normally have a large number of factors (the density of primes $1/\log x$ approaches zero) and there seems to be no reason why either an even or an odd number of factors would be more likely. Moreover, by the same principle it is perhaps plausible to say that successive evaluations of $\mu(n) = \pm 1$ are "independent" since knowing the value of $\mu(n)$ for one n would not seem to give any† information about its values for other values of n. But then the evaluation of $M(x)$ would be like flipping a coin once for each square-free integer less than x and subtracting the number of heads from the number of tails. It was shown above that for any given $\varepsilon > 0$ the outcome of this experiment for a large‡ number of flips is, with probability nearly one, less than the number of flips raised to the power $\frac{1}{2} + \varepsilon$ and *a fortiori* less than $x^{(1/2)+\varepsilon}$. Thus these probabilistic assumptions about the values of $\mu(n)$ lead to the conclusion, ludicrous as it seems, that $M(x) = O(x^{(1/2)+\varepsilon})$ with probability one and hence that the Riemann hypothesis is true with probability one!

(These pages are from "The Riemann Zeta Function" by Harold Edwards.

Lets Continue with $\mu(n)$.

Theorem. $\displaystyle\sum_{d|n} \mu(d) = \phi.$

e.g. $n = 2\cdot3\cdot5$

$$\{d \mid d|n\} = \{\underset{+1}{1}, \underset{-1}{2}, \underset{-1}{3}, \underset{-1}{5}, \underset{+1}{2\cdot3}, \underset{+1}{2\cdot5}, \underset{+1}{3\cdot5}, \underset{-1}{2\cdot3\cdot5}\}$$

$$\underbrace{+1-1-1-1+1+1+1-1}=\phi$$

Proof. Consider

$$(1+2)(1+3)(1+5) = 1+2+3+5 + 2\cdot3+2\cdot5+3\cdot5 + 2\cdot3\cdot5$$

and $(1-1)(1-1)(1-1)=\phi$.

This proves the result for all d that have distinct prime factorizations. Rest is easy. //

Note $\sum_{d|n} \mu(n) \Rightarrow \mu(n) = -\sum_{\substack{d|n \\ d \neq n}} \mu(d)$.

This certainly shows that $\mu(n)$ is not random, but we can do better.

Prime Numbers

$6 = \vdots\vdots\vdots = 2 \times 3$, not prime.

A number ($\neq 1$) is said to be prime if it has no factors other than itself and one.

$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$

Euclid (~500 BC): There are infinitely many primes.

Proof. Suppose $\{P_1, P_2, \ldots, P_n\}$ is any finite list of primes. Form the number $N = (P_1 \times P_2 \times \cdots \times P_n) + 1$.
No P_i divides N. Hence either N is a (new) prime, or N has (new) prime factors.

Is 137 prime?

If $a \times b = 137$, then

either $a < \sqrt{137}$

or $b < \sqrt{137}$.

otherwise $ab > \sqrt{137}\sqrt{137} = 137$.

So we need only search for prime factors up to $\sqrt{137}$.

$10^2 = 100$, $11^2 = 121$, $13^2 = 169$.

So we can look at the primes $\{2, 3, 5, 7, 11\}$

Since none of them divide 137 we conclude that

<u>137 is prime</u>.

The Seive of Eratosthenes

To find the primes.

- strike out all multiples of 2 then 3, then 5, ...
- Knowing all primes up to N, your strikes will reveal all primes up to N^2. "!"

e.g. You know $\{2, 3, 5\}$, $5^2 = 25$.
Striking multiples of 2, 3, 5 will REVEAL all primes up to 25.

2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11

~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19

~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ ~~25~~

2 3 $\cancel{4}$ 5 $\cancel{6}$ 7 $\cancel{8}$ $\cancel{9}$ $\cancel{10}$ 11

$\cancel{12}$ 13 $\cancel{14}$ $\cancel{15}$ $\cancel{16}$ 17 $\cancel{18}$ 19

$\cancel{20}$ $\cancel{21}$ $\cancel{22}$ 23 $\cancel{24}$ $\cancel{25}$

Thus $\{2,3,5\} \Rightarrow$ The remaing primes < 25 are $\{7, 11, 13, 17, 19, 23\}$.

Let $\pi(n) = \#$ primes $\leq n$.

$$\pi(5^2) = \pi(5) + 6 = 9$$

<u>Legendre</u>: $\pi(n^2) - \pi(n) = \left(\sum_{\substack{d: \\ d \text{ gen by } \{p \leq n\}}} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor \right) - 1$

$2 \quad 3 \quad \cancel{4} \quad 5 \quad \cancel{6} \quad 7 \quad \cancel{8} \quad \cancel{9} \quad \cancel{10} \quad 11$

$\cancel{12} \quad 13 \quad \cancel{14} \quad \cancel{15} \quad \cancel{16} \quad 17 \quad \cancel{18} \quad 19$

$\cancel{20} \quad \cancel{21} \quad \cancel{22} \quad 23 \quad \cancel{24} \quad \cancel{25}$

The formula of Legendre for $\pi(n^2) - \pi(n)$ is based on the following idea: First list ~~all~~ 25 numbers except 1. $(25-1)$. Then we strike all mults of $2, 3, 5$.

$\frac{25}{2} = 12\frac{1}{2}$. $\left\lfloor \frac{25}{2} \right\rfloor = 12$ greatest integer in $25/2$.

$12 = \#$ of strikes by 2. Then $\left\lfloor \frac{25}{3} \right\rfloor = 8$ number of 3-strikes. $\left\lfloor \frac{25}{5} \right\rfloor = 5 = \#$ 5 strikes.

But then some strikes are multiple (e.g. $10 = 2\times 5$). We add these back: $\mu(2) = \mu(3) = \mu(5) = -1$

$\mu(2\times 3) = \mu(2\times 5) = \mu(3\times 5) = +1$

$2\times 3\times 5$ is too big! So ...

$2 \quad 3 \quad \cancel{4} \quad 5 \quad \cancel{6} \quad 7 \quad \cancel{8} \quad \cancel{9} \quad \cancel{10} \quad 11$

$\cancel{12} \quad 13 \quad \cancel{14} \quad \cancel{15} \quad \cancel{16} \quad 17 \quad \cancel{18} \quad 19$

$\cancel{20} \quad \cancel{21} \quad \cancel{22} \quad 23 \quad \cancel{24} \quad \cancel{25}$

$25 - 1$

$2 : \mu(2)\left\lfloor \frac{25}{2} \right\rfloor = -12$

$3 : \mu(3)\left\lfloor \frac{25}{3} \right\rfloor = -8$

$5 : \mu(5)\left\lfloor \frac{25}{5} \right\rfloor = -5$

$2\cdot 3 : \mu(2\cdot 3)\left\lfloor 25/6 \right\rfloor = 4$

$2\cdot 5 : \mu(2\cdot 5)\left\lfloor 25/10 \right\rfloor = 2$

$3\cdot 5 : \mu(3\cdot 5)\left\lfloor \frac{25}{15} \right\rfloor = 1$

$12 - 8 - 5 + 4 + 2 + 1$

$= 6 = \pi(25) - \pi(5)$

$$\boxed{\pi(N^2) - \pi(N) = \sum_{\substack{d \text{ sqfr} \\ \text{up to } \sqrt{N}}} \mu(d)\left\lfloor \frac{N}{d} \right\rfloor}$$

2 3 $\cancel{4}$ 5 $\cancel{6}$ 7 $\cancel{8}$ $\cancel{9}$ $\cancel{10}$ 11

$\cancel{12}$ 13 $\cancel{14}$ $\cancel{15}$ $\cancel{16}$ 17 $\cancel{18}$ 19

$\cancel{20}$ $\cancel{21}$ $\cancel{22}$ 23 $\cancel{24}$ $\cancel{25}$

$\mu(n)$ is the <u>Moebius Function</u>

$\mu(p) = -1$, p prime

$\mu(p_1 p_2 \cdots p_n) = (-1)^n$ if $p_1, \ldots, p_n$ are <u>distinct</u> primes.

$\mu(n) = 0$ if $p^2 | n$ for a prime p.

$$\zeta(s) = \sum_n \frac{1}{n^s}$$

Riemann Zeta Function.

$$\frac{1}{\zeta(s)} = \sum_n \frac{\mu(n)}{n^s}$$

GSB observes that

$$\text{Legendre} \Rightarrow \sum_{k=1}^{n} \mu(k) \left\lfloor \frac{n}{k} \right\rfloor = 0$$

$$\left(\text{adjust } \left\lfloor \frac{n}{1} \right\rfloor \longrightarrow n-1 \right)$$

and that this means that

$$\mu(n) = -\sum_{k=1}^{n-1} \mu(k) \left\lfloor \frac{n}{k} \right\rfloor$$

a <u>negative feedback</u> equation for the Moebius Function.

$$\mu(n) = -\sum_{k=1}^{n-1} \mu(k) \left\lfloor \frac{n}{k} \right\rfloor \quad \text{(GSB)}$$

Convention
$\left(\left\lfloor \frac{n}{1} \right\rfloor = n-1 \right)$

$$\text{e.g. } \overset{\overset{+1}{\shortparallel}}{\overset{\mu(2\cdot 3)}{\shortparallel}}\mu(6) = -\sum_{k=1}^{5} \mu(k) \left\lfloor \frac{6}{k} \right\rfloor \quad (\mu(4) = 0)$$

$$= -\mu(1)\left\lfloor \frac{6}{1} \right\rfloor - \mu(2)\left\lfloor \frac{6}{2} \right\rfloor - \mu(3)\left\lfloor \frac{6}{3} \right\rfloor - \mu(5)\left\lfloor \frac{6}{5} \right\rfloor$$

$$= -(6-1) + 3 + 2 + 1 = +1 \checkmark$$

Note that the GSB/Legendre formula computes values of $\mu(n)$ without any factorizations.

The Mäbius Seive

Here is a way to see how GSB formula comes from the Seive.

$$\boxed{\sum_{k=1}^{n} \mu(n) \left\lfloor \frac{n}{k} \right\rfloor = \phi}$$

$\left\lfloor \frac{n}{1} \right\rfloor = n-1$ $\Leftarrow$ All Cols Sum to Zero by $\sum_{d|n} \mu(n) = \phi$

	1	2	3	4	5	6	7	8	9	10	
1	$\boxed{0}$	1	1	1	1	1	1	1	1	1	$10-1 = 9$
2	0	−1	0	−1	0	−1	0	−1	0	−1	$-\left\lfloor \frac{10}{2} \right\rfloor = -5$
3	0	0	−1	0	0	−1	0	0	−1	0	$-\left\lfloor \frac{10}{3} \right\rfloor = -3$
4	0	0	0	$\boxed{0}$	0	0	0	$\boxed{0}$	0	0	$0 \cdot \left\lfloor \frac{10}{4} \right\rfloor = 0$
5	0	0	0	0	−1	0	0	0	0	−1	$-\left\lfloor \frac{10}{5} \right\rfloor = -2$
6	0	0	0	0	0	1	0	0	0	0	$1 \cdot \left\lfloor \frac{10}{6} \right\rfloor = 1$
7	0	0	0	0	0	0	−1	0	0	0	$-1 \cdot \left\lfloor \frac{10}{7} \right\rfloor = -1$
8	0	0	0	0	0	0	0	$\boxed{0}$	0	0	$0 \cdot \left\lfloor \frac{10}{8} \right\rfloor = 0$
9	0	0	0	0	0	0	0	0	$\boxed{0}$	0	$0 \cdot \left\lfloor \frac{10}{9} \right\rfloor = 0$
10	0	0	0	0	0	0	0	0	0	1	$1 \cdot \left\lfloor \frac{10}{10} \right\rfloor = 1$

$n=6$

	1	2	3	4	5	6	
1	$\boxed{0}$	1	1	1	1	1	$6-1=5 \quad = 5$
2	0	-1	0	-1	0	-1	$(-1)\left\lfloor \frac{6}{2} \right\rfloor = -3$
3	0	0	-1	0	0	-1	$(-1)\left\lfloor \frac{6}{3} \right\rfloor = -2$
4	0	0	0	$\boxed{0}$	0	0	$0\left\lfloor \frac{6}{4} \right\rfloor = 0$
5	0	0	0	0	-1	0	$(-1)\left\lfloor 6/5 \right\rfloor = -1$
6	0	0	0	0	0	1	$(1)\left\lfloor 6/6 \right\rfloor = 1$

Sum $=\phi$

The Möbius Seive gives an elegant proof that $\sum_{k=1}^{n} \mu(k)\left\lfloor \frac{n}{k} \right\rfloor = \phi$

- all column sums $=\phi$ (each col sum $=\sum_{d|k}\mu(d)=\phi$).
- $\therefore$ the sum of all column sums $=\phi$.
- the sum of all col sums = sum of all row sums.
- k^{th} row sum $=\mu(k)\left\lfloor \frac{n}{k} \right\rfloor$.
- $\therefore \ \phi = \sum_{k=1}^{n} \mu(k)\left\lfloor \frac{n}{k} \right\rfloor$.

$$\mu(n) = -\sum_{k<n} \mu(k) \left\lfloor \frac{n}{k} \right\rfloor \qquad \underline{\text{GSB formula}}$$

If you think of $\mu(n)$ as a "coin toss", then it cannot have arbitrarily long runs of $(+1)$ or (-1) due to the <u>negative feedback</u> in the GSB formula.

This means that the swings of $M(x) = \sum_{k<x} \mu(k)$ will be <u>more restrained</u> than the swings of a cumulative coin toss.

The Moebius Function

is like a "**cointoss**" where the coin remembers the past and <u>cannot</u> behave randomly.

A run of +1's will eventually be stopped. A run of −1's will eventually be stopped.

GSB's work on Riemann Hypothesis is based on this property of the Mobius Function.

The last slide illustrates how the cumulative Mobius function $M(n) = \sum_{k \leq n} \mu(k)$

varies much less wildly than a random coin.

$\boxed{\mu(n) \text{ is a Magic Coin}}$

and this is the key to the Riemann Hypothesis.

See papers & papers to come by: J. Flagg, LK, D. Sahoo

Recall that $\boxed{C(k) = \pm 1, 0 \text{ random}}$ $M(x) = \sum_{k<x} \mu(k)$

$$\frac{M(x)}{x^{\frac{1}{2}+\varepsilon}} \longrightarrow \phi \Rightarrow RH$$

$RCoin(x) = \sum_{k<x} C(k)$

and $\boxed{\dfrac{RCoin(x)}{x^{\frac{1}{2}+\varepsilon}} \longrightarrow \phi \text{ is a mathematical fact.}}$

Thus our understanding that $\mu(n)$ is a restrained (by neg feedback) coin. Suggests that one should be able to "see" how

$\dfrac{M(x)}{x^{\frac{1}{2}+\varepsilon}} \to \phi$ relative to $\dfrac{RCoin(x)}{x^{\frac{1}{2}+\varepsilon}} \to \phi$.

Examine the Graphics

```
DiscretePlot[{Marton[n], Merton[n]}, {n, 1, 10 000}, PlotRange → {-300, 300}]
```