

Multiplication Modulo n Along The Primorials With Its Differences And Variations Applied To The Study Of The Distributions Of Prime Number Gaps A.K.A. Introduction To The S Model

Russell Letkeman
r.letkeman@gmail.com
Dedicated to my son Panha

May 19, 2013

Abstract

The sequence of sets of Z_n on multiplication where n is a primorial gives us a surprisingly simple and elegant tool to investigate many properties of the prime numbers and their distributions through analysis of their gaps. A natural reason to study multiplication on these boundaries is a construction exists which evolves these sets from one primorial boundary to the next, via the sieve of Eratosthenes, giving us Just In Time prime sieving. To this we add a parallel study of gap sets of various lengths and their evolution all of which together informs what we call the S model.

We show by construction there exists for each prime number P a local finite probability distribution and it is surprisingly well behaved. That is we show the vacuum; ie the gaps, has deep structure.

We use this framework to prove conjectured distributional properties of the prime numbers by Legendre, Hardy and Littlewood and others. We also demonstrate a novel proof of the Green-Tao theorem. Furthermore we prove the Riemann hypothesis and show the results are perhaps surprising.

We go on to use the S model to predict novel structure within the prime gaps which leads to a new Chebyshev type bias we honorifically name the Chebyshev gap bias. We also probe deeper behavior of the distribution of prime numbers via ultra long scale oscillations about the scale of numbers known as Skewes numbers*.

0.0.0.1 Opening Remarks In 1901 Von Koch¹ proved that if the Riemann hypothesis were true, the absolute worst possible variation between the prime counting function $\pi(x)$ and $Li(x)$, the logintegral, could be was of order $\sqrt{x} \ln(x)$.

Without assuming RH, we'll show constructively the worst absolute error is $\sqrt{\frac{x}{\ln(x)}}$ or $Li(\sqrt{x})$. Furthermore, for large x , $\pi(x)$ spends as much time above $Li(x)$ as below.

1 Conventions And Notations And The Birth Of The S Model

We sequence the primes using the canonical order with $P_m < P_{m+1} < P_{m+2} \forall m \geq 0$ with $P_0=1$. That is 1 is the *zeroth* prime making 2 the *first* prime.

Primorials are to primes as factorials are to the positive naturals. Using the symbol $\#$ instead of $!$ we define the m^{th} primorial as $P_m\# = \prod_{i=0}^m P_i$ which gives us the familiar recursive form $P_{m+1}\# = P_{m+1}P_m\#$.

We also use 3 variants of the primorial, the **multiplicative moment**, the **additive moment** and the **minor**. The j^{th} primorial multiplicative moment is $\frac{P_m\#}{P_j}$, while the j^{th} primorial additive moment is $\frac{P_m\#}{P_j} \pm P_j \forall 0 < j < m$. This means the *zeroth* primorial has no additive moments while all others have their full compliment. This can be further generalized by using any subsequence of unique primes $< P_m$, for example $\frac{11\#}{7*3} \pm 7*3$ is the $7*3$ moment of $11\#$.

The r^{th} primorial minor is defined as $P_m^{-r}\# = \prod_{j=k}^m (P_j - r)$ such that k is the smallest prime where $P_k - r \geq 0$. The primorial minors also preserve the recursive relation $P_{m+1}^{-r}\# = P_{m+1}^{-r}P_m^{-r}\#$. In fact the basic primorials are simply the 0^{th} minor. As in factorials if $P_m=r$ we set $P_m^{-r}\#=1$ and $\forall m$ s.t. $P_m \geq r$ have our normal product while it is undefined $\forall m$ s.t. $P_m < r$.

The following is trivial to prove

$$P_{m+1}P_m^{-1}\# = (P_{m+1}-1+1)P_m^{-1}\# = (P_{m+1}^{-1}+1)P_m^{-1}\# = P_{m+1}^{-1}\# + P_m^{-1}\# \quad (1)$$

as is its generalization

$$P_{m+1}P_m^{-r}\# = P_{m+1}^{-r}\# + rP_m^{-r}\# \quad (2)$$

Since we will study the properties of Z_n where n is a primorial we will write $S^m = Z_{P_m\#}$ to save complex subscripting. We will use \mathbf{S}^m to represent S^m and all its residuals. We choose S because as we will see the processes we examine results from the common prime sieve. Membership in S^m and \mathbf{S}^m is denoted by a superscript P_j^m . It is very easy to show

$$\mathbf{S}^{m+1} \subset \mathbf{S}^m \forall m \geq 0 \quad (3)$$

We will also maintain a second set called the midden which contains all the primes up to our m^{th} and denoted as $\{P_m\}$. There can never be any confusion between membership in \mathbf{S}^m and the midden as it is straight forward to prove $\forall x > 1 x \in midden \Rightarrow x \notin \mathbf{S}^m$. The order of S^m is determined from the Euler totient function^{8,12}

$$|S^m| = \phi(P_m\#) = P_m^{-1}\# \quad (4)$$

The ring S^m is the direct product of the prime factors⁸

$$S^m = Z_2 X Z_3 X \dots X Z_{P_m} \quad (5)$$

and

$$\forall m > 0 \forall P_j^m \in S^m \rightarrow P_m\# - P_j^m \in S^m \quad (6)$$

That is restricting our multiplication to finite groups over supplies us with information, internal symmetries must supply half of the distributional information.

Beginning with a zero based index, instead of writing the usual

$$S^m = \{1, P_1^m, P_2^m, \dots, P_{P_m^{-1}\#-1}^m\}$$

we will write it in matrix form of n rows, the top row being our set of co-primes. The second row is the successive difference; written $g_j^m = P_{j+1}^m - P_j^m$ and occasionally represented separately as the gap class $G^m = \{g_j^m\}$. The third row is the difference between 2 successive gaps or difference of differences; written $c_j^m = P_{j+2}^m - P_j^m = g_{j+1}^m + g_j^m$ and represented separately as the coalescence class $C^m = \{c_j^m\}$. We have had to draw upon residuals to determine all the values. There is no limit for the number of successive differences and eventually we use them all but we will generally only refer to the first 3 as such

$$S^m = \begin{pmatrix} 1 & P_1^m & \dots & P_{P_m^{-1}\#-1}^m \\ g_0^m & g_1^m & \dots & g_{P_m^{-1}\#-1}^m \\ c_0^m & c_1^m & \dots & c_{P_m^{-1}\#-1}^m \end{pmatrix}$$

2 S Model Counting

2.1 The Gap Palindrome

S^m Gap Palindrome Theorem. $\forall m > 1$ or $P_m > 2$ S^m has a palindrome in its gaps of length $P_m^{-1}\# - 1$, the entire sequence always ends in a 2 and the middle of the palindrome is always a 4.

Proof. The following are trivially true, $P_0^m = 1 \forall m$, and if $P_j^m \in S^m$ than $P_m\# - P_j^m \in S^m$ therefore

$$P_m\# \pm 1 \in S^m \rightarrow \text{the last gap is a twin}$$

Since $P_m^{-1}\#\forall m > 1$ is even, the set minus the capping twin has an odd length so it has a well defined middle. It is also trivially true that

$$\frac{P_m\#}{2} \pm 2 \in S^m \rightarrow \text{the middle gap is a cousin}$$

and finally

$$g_j^m = P_{j+1}^m - P_j^m = (P_m\# - P_j^m) - (P_m\# - P_{j+1}^m) = g_{P_m^{-1}\#-j-1}^m \forall j < \frac{P_m^{-1}\#}{2}$$

□

2.2 The Construction Of The Construction Using A Construction

The word sieve is both a noun and a verb. Until now we have only considered the constructions in situ, as we found them. Now we show how to construct S^{m+1} from S^m , the construction of a construction using a construction. Instead of using the verbose construction of the construction using a construction we will also use the Composite Killing Machine or the CKM just because it sounds cool.

Now we'll introduce a JIT strategy of prime sieving. In our S^m notation we do the following

$$S^{m+1} = P_{m+1} \otimes S^m - P_{m+1}S^m \quad (7)$$

Where \otimes says to to use the first $P_{m+1} * |S^m|$ members of G^m and the second term is minus a scaler multiplication of the members of S^m .

That is with S^m , when we reframe our number span alignment to the next primorial we have **exactly** $|S^m|$ extra members composed against P_{m+1} . The recipe to remove them is using P_{m+1} multiples of S^m because $P_{m+1}^{-1}\# = P_{m+1}P_m^{-1}\# - P_m^{-1}\#$. This set of constructions is a manifestation of the primorial algebra. Or in the words of a friend of the author, "This thing writes itself".

For example $S^3 = Z_{30}$ with $P_3 = 5$ has $P_3^{-1}\# = 8$ members. To generate S^4 we lay 7 sets of (our next prime) gaps in a row. This set has $7 * 8 = 56$ members while our target has $P_4^{-1}\# = 48$ members. We know in advance we've polluted our set and each extra false member is simply one of 7 times our original co-prime members.

However using the above palindromic nature of each set's gaps gives us a remarkable 2^n speedup since using it reduces the above work load by half at each step. We need our sets to have at least 2 members to manifest both the capping twin and the pivotal cousin so we can use any $m > 1$. That is given we have computed some S^m using only the gaps proceed as follows

1. allocate an array of size $P_{m+1}^{-1}\#$ and set an accumulator to 1
2. systematically copy the gaps from S^m one at a time while adding them to the accumulator
3. in parallel compute the next composite via $P_{m+1}S^m$
4. if the accumulator equals our next composite skip the copy but add this gap to the next gap, we call this coalescence.
5. continue until you have completed $\frac{P_{m+1}^{-1}}{2} - 1$ steps
6. copy a mirror image of the new gap sequence to our sequence after adding the pivotal cousin gap (hence its' name)
7. append the capping twin

Of course there is no need to ever allocate the full sequence length since we know in advance it has this symmetry. If we desire we can simply read our progeny set forward and backward as needed.

2.3 The Induced P.M.F. Of S^m

We know our sets have exactly $P_m^{-1}\#$ members. We can immediately deduce

$$\sum_{j=0}^{P_m^{-1}\#-1} g_j^m = P_m\# \quad (8)$$

Which leads to the average gap size in S^m is

$$\hat{\mathbf{g}}^m = \frac{P_m\#}{P_m^{-1}\#} \quad (9)$$

We also have

$$\sum_{j=0}^{j < P_m^{-1}\#} c_j^m = 2P_m\# \rightarrow \hat{\mathbf{c}}^m = 2\hat{\mathbf{g}}^m \quad (10)$$

That is on average successive pairs sum to twice the average gap. This immediately leads to the general form

$$\sum_{j=0}^{j < P_m^{-1}\#} S_j^{m,n} = nP_m\# \rightarrow \langle S^{m,n} \rangle = n\hat{\mathbf{g}}^m \quad (11)$$

where $S_j^{m,n}$ are chosen from the n^{th} row, where for convenience index the first row from 0 counting the number of successive differences. Because the basis of S^m is modular we also have if $n > 1$

$$\mathbf{S}_{j+kP_m^{-1}\#}^{m,n} = S_j^{m,n} \forall k \quad (12)$$

That is though the co-prime row is most definitely not periodic every row of successive differences is.

The astute reader may have already sensed that our random variable sounds remarkably like $\ln(P_m)$. We now consider Merten's theorem³ which is very elegant in our notation.

Merten's Theorem And $\hat{\mathbf{g}}^m$.

$$\lim_{m \rightarrow \infty} \frac{\ln(P_m)}{\hat{\mathbf{g}}^m} = e^{-\gamma}$$

with γ being the Euler–Mascheroni constant[?] 0.5772156649... That is for large m we have an asymptotic approximation

$$\hat{\mathbf{g}}^m \simeq e^\gamma \ln(P_m) \quad (13)$$

If we are to reconcile our work with the PNT⁹ we need to scale our set counts with e^γ . That is given a subsequence of gap from S^m which sums to S , we have

$$\text{count} = \frac{e^\gamma S}{\hat{\mathbf{g}}^m} \simeq \frac{e^\gamma S}{e^\gamma \ln(P_m)} = \frac{S}{\ln(P_m)}$$

That is we can (and do) causally switch between the different models as long as we account for the appropriate conditions. For example \mathbf{G}^m is periodic with period $P_m^{-1}\#$ and here we need an exact rational form, but generally this is not the case, especially for large numbers.

2.3.1 S^m IS NOT the PNT

Each S^m is quasi linear and periodic in the gaps. By theorems of Chebyshev and others they must cover all the primes co-prime to $P_m\#$ but they are also highly polluted with composites.

For example by Chebyshev $P_m\#\simeq O(e^{P_m})$ so $P_m^{-1}\#\simeq O\left(\frac{e^{P_m-\gamma}}{\ln(P_m)}\right)$ while the PNT predicts $\pi(P_m\#)\simeq\frac{e^{P_m}}{P_m}$.

That is if you draw a member at random you will be prime about once in $\frac{e^{-\gamma}P_m}{\ln(P_m)}=e^{-\gamma}m$. It's much better than the completely random selection of only a once in P_m chance of drawing a prime among the integers in our range.

Also $\hat{\mathbf{g}}^m$ is a constant across all of S^m , it thinks the average gap at P^n is $\hat{\mathbf{g}}^m$. The PNT informs us the average gap at $P^n=n\ln(P)$.

However we have an extremely important property, all members less than P_{m+1}^2 are prime by construction. That is there is a small natural bias in the first few members being purely prime with this region growing as the square becoming arbitrarily large. But the overall set size grows exponentially so its relative importance becomes arbitrarily small.

That is given some P_{m-1} , the spacing between members in the region $[P_m, P_m^2)$ is randomly linear with a smooth logarithmic decay.

2.4 Survivorship Of Gap Sequences Across The CKMs

Without performing the multiplication we don't know where $P_{m+1}P_j^m$ is but we know where ever it is its current gap is g_j^m . And when it is removed as a composite we know that the new gap will be exactly c_j^m . Equation 10 also tells us that on average the new gaps have doubled in size because the sum across the whole of the set of possible changes is double the gap sum of the progenitor set. That is $P_{m+1}^{-2}P_m^{-1}\#$ of our original $P_{m+1}P_m^{-1}\#$ members survived coalescence while $2P_m^{-1}\#$ members were coalesced in pairs to form new gaps.

The easiest way to picture this is consider the fact that every gap is a pair of co-prime members, each pair has 2 ways to die and each way to die happens with certainty. The same being true for any gap tuple length of r will have $r+1$ ways to die, hence the name CKM.

Given S^m , consider *any* consecutive sequence of gaps of length $r=1$ up to $r=P_{m+1}^{-2}$. During the construction of S^{m+1} we make exactly P_{m+1} exact copies of your random selection. However the final coalescence stage will *kill* all but $P_{m+1}-(r-1)$ of them.

That is, as we showed above for single gaps, P_{m+1}^{-2} copies survive intact. This generalizes to P_{m+1}^{-3} pairs, P_{m+1}^{-4} triples and so on up to $P_{m+1}^{-P_{m+1}^{-1}}$ tuples of length P_{m+1}^{-2} . This

implies the primorial minors form tuple duplication machinery, acting as conveyors for future gap cluster production, becoming active as soon as P_m becomes large enough.

2.5 A Few Examples

We stop with theory and examine the first few m , but let's begin with a question. You are given a zero based array and told it has been sieved to some m^{th} prime and it contains some unknown number of residuals, can you determine m ?

Perhaps some starting points would be after the zeroth member 1, the next number is the next prime and the first composite is the next prime squared.

2.5.1 The Curious Case of S^0

Chapter $m = 0$, the genesis of the primes.

In the beginning; before the first big bang of the composite killing machine, the primes were almost without form.

We consider the first and always ignored $Z_1 = S^0 = \{1\}$ which generates $\mathbf{S}^0 = \{1 + n \mid \forall n \geq 0\}$. The first few rows of S^0 are

$$\left\{ \begin{array}{l} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \right\} \left\{ \begin{array}{l} 2 \ 3 \ 4 \ \dots \\ 2 \ 2 \ 2 \ \dots \\ 3 \ 3 \ 3 \ \dots \\ 4 \ 4 \ 4 \ \dots \\ 5 \ 5 \ 5 \ \dots \end{array} \right\}$$

where we have added some residuals to make the form more apparent. From this perspective everything looks prime and we have no multiplication, no fast way to get to a next number in a sequence; all that remains is addition by 1. But in fact we have more, since the next number must be prime, not simply co-prime. In fact all consecutive members are prime until this next prime squared or $\{1, 2, 3\}$.

That 1 is the first (*zeroth*) prime and equal to its own square, i.e. $1 = 1^2$ is the basis of its total degeneracy in multiplication among the naturals. However it is clear that the positive naturals are *born sieved* by 1. Also it's why we must set $P_0^{-1}\# = 1$ since $\phi(1) = (1 - 1)\# = 1$ as $|\mathbf{S}^0| = 1$.

The fact it is so degenerate is also why its most primitive minor doesn't exist. The positive side exists since $P_m\# + 1$ is always co-prime to $P_m\#$, however $1 - 1 = 0 \notin S^0$ so in this case the negative side makes no sense.

Since the next (*first*) prime is 2, the largest subset of gaps that can be successfully replicated is zero, that is all gaps of size 1 are consumed in S^1 so we can never see another gap of unit length after 4.

However since our next prime is 2 $P_{m+1}^{-2}\# = 1$ will exist/exists by definition, and since the next generation will first use a pair of 1's to create the first twin gap a single virtual span of 2 already exists.

2.5.2 S^1 , The Death Of A Subring

The first few rows of S^1 are

$$\left\{ \begin{array}{c} 1 \\ 2 \\ 4 \\ 6 \\ 8 \end{array} \right\} \left\{ \begin{array}{cccc} 3 & 5 & 7 & \dots \\ 2 & 2 & 2 & \dots \\ 4 & 4 & 4 & \dots \\ 6 & 6 & 6 & \dots \\ 8 & 8 & 8 & \dots \end{array} \right\}$$

Following our construction to generate S^1 from S^0 we have consumed all gaps of unit length to produce a single twin gap. Define T_2^m to be the count of twin gaps in any S^m , $T_2^0 = 1$, at least as a virtual span which can/will coalesce into twins. We already know for each target gap in S^m that P_{m+1}^{-2} copies **will** survive in the next set. We don't know what will happen in any future coalescence event without the requisite *hard work*. Since the next prime is 3 we will see our single twin survive along with all its children. However 3 is small enough that only a single twin can survive even though 3 copies will be made. That is after this set is expanded we will never see pairs of twin gaps again.

At this point we can now count all twins because the $P_m^{-2}\#$ copy machine is running and we trivially have a single gap type to track, that is

$$T_2^m = P_m^{-2}\#\forall m > 0 \quad (14)$$

Referring to equation 5 we see the ring of any S^m is the product of rings of our primorial but Z_1 is absent for very good reasons, yet $Z_2 = S^1$ would not be constructible without it. That is $Z_1 = S^0$ has a unique property of acting like a bootstrap, letting us build our next generation in the series before gracefully disappearing and we will never see its impact again. We don't know of any other examples of bootstrap structures in nature. They may however, be interesting to suppose.

2.5.3 S^2 , The Birth Of Deep Gap Structure

The first few rows of S^2 are

$$\left\{ \begin{array}{cc} 1 & 5 \\ 4 & 2 \\ 6 & 6 \\ 8 & 10 \\ 12 & 12 \end{array} \right\} \left\{ \begin{array}{cccc} 7 & 11 & 13 & 17 & \dots \\ 4 & 2 & 4 & 2 & \dots \\ 6 & 6 & 6 & 6 & \dots \\ 8 & 10 & 8 & 10 & \dots \\ 12 & 12 & 12 & 12 & \dots \end{array} \right\}$$

giving us our pivotal cousin and capping twin though the palindrome is the trivial sequence $\{4\}$. As predicted, a single twin survived and the virtual span of the single event of pairs of twins was consumed to produce our pivotal cousin. It seems it is no coincidence $(2 - 2)\# = (3 - 2)\# = 1$. The single possibility for pairs of twins to enter the gap formation machinery ended at their birth. We now have 2 distinct gap types; twins and cousins each count synchronized to $P_m^{-2}\#$, that is

$$T_4^m = T_2^m = P_m^{-2}\#\forall m > 1 \quad (15)$$

Since the next prime is 5 we know all spans of length 3 survive with certainty. However since the current prime is 3 the mechanics of pair reproduction is available and the only pairs currently in existence are $\{\{4, 2\}, \{2, 4\}\}$, that is

$$T_{\{4,2\}}^m = T_{\{2,4\}}^m = P_m^{-3} \# \forall m > 1 \quad (16)$$

Also just as cousin gaps can no longer percolate from coalescence, the same is true for our new pairs. These are what we call primal atoms, they form lowest density regions, the sum of gaps is the smallest number for all space and time as it were. Once these atoms are created in a particular S^m they will faithfully reproduce and nowhere after this can any region of matching size versus span achieve a lower sum for the same tuple count. Not only will twins exist forever but they are as small as are allowed, the same is true for the atomic pairs $\{\{4, 2\}\{2, 4\}\}$, all future iterations of the S model will inherit them.

2.5.3.1 Construction, What's In A Name The S model is a construction in the sense it is a coherent set of rules that generate a set of structured numbers. It does not however construct the primes, it uses the known primes to find new primes.

We bootstrap with '1' to generate the sequence of sequences the leads to the primes, and we find the next prime *by looking* at what is next. The best description we have is it is the first nontrivial member of the set. Out of the gate it's trivial, there is no choice, the same holds at step 1. But as we've seen, by step 2 we have choice. In fact every S^m generates $P_m^{-1} \#$ equations

$$P_j^m + nP_m \# \forall 0 \leq j \leq P_m^{-1} \# - 1$$

which generate all the primes $> P_m$.

The S model is in a sense sculpting by cutting away the parts that don't belong. Perhaps an interesting self portrait of the S model shows an important component which is a milling machine whose cutting head grows (explodes) exponentially at each turn.

2.5.4 S^3 And The Explosion Of Complexity

We apologize in advance for the short treatment we present for S^3 as it truly deserves a substantial investment in time while the rewards it returns are beyond description. The first few rows of S^3 are

$$\left(\begin{array}{cccccccc} 1 & 7 & 11 & 13 & 17 & 19 & 23 & 29 \\ 6 & 4 & 2 & 4 & 2 & 4 & 6 & 2 \\ 10 & 6 & 6 & 6 & 6 & 10 & 8 & 8 \\ 12 & 10 & 8 & 10 & 12 & 12 & 14 & 12 \\ 16 & 12 & 12 & 16 & 14 & 18 & 18 & 14 \\ 18 & 16 & 18 & 18 & 20 & 22 & 20 & 18 \\ 22 & 22 & 20 & 24 & 24 & 26 & 24 & 20 \\ 28 & 24 & 26 & 28 & 26 & 28 & 26 & 24 \\ 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 \end{array} \right)$$

Notice how G^3 forms our beautiful palindrome about the pivotal cousin with a capping twin.

It is now clear to see how gaps and spans of 6 or sexy size evolve, for brevity we will simply state the results. Let $\mathcal{S}_s^{m,n}$ represent the set of the sequences of gaps of length n such that their sum is s . Let \mathcal{S}_s^m be the sum of all possible sequence lengths. For example the sum of all ways to span a gulf of 6 with any length of consecutive gaps in \mathcal{S}^m is $\mathcal{S}_6^m = 2P_m^{-2}\#$. While individual gaps of 6 are counted as $T_6^m = 2(P_m^{-2}\# - P_m^{-3}\#)$.

This leads to a remarkable conclusion, there must exist an underlying algebra based on the primorial minors whose linear combinations generate finite equations describing the counts of patterns of gap subsequences. Or perhaps not so remarkable as this has been long foretold¹³.

2.6 Primorial Minor Linear Algebra And Gap Evolution

We have seen that subsequences of gaps of length up to P_{m+1}^{-2} in \mathcal{S}^m will be faithfully replicated with certainty with $P_{m+1}^{-1} - P_{m+1}^{-2} = 1$, a single copy surviving. We have also seen that the accessible replication machinery, the primorial minors don't so much as copy specific contents but rather subsequences of specific length are copied without regard to content. That $T_2^m = T_4^m = P_m^{-2}\#$ have simple forms is birth order luck.

Let $\Omega(m, r)$ be an rxr matrix starting from some m with the members $\omega_{i,j} = P_{m+j}^{-(1+i)}\#\forall 1 \leq i, j \leq r$. That is each column is a particular primorial minor at some index, and successive columns are deeper minors, while rows represent the index to the primes. It's name comes from Primorial Minor Matrix shortened to Oriial Matrix and is pronounced 'om', an old Sanskrit term chanted at the start of meditating the perfection of the universe.

$$\Omega(m, r) = \begin{pmatrix} P_m^{-2}\# & P_m^{-3}\# & \dots & P_m^{-(r+1)}\# \\ P_{m+1}^{-2}\# & P_{m+1}^{-3}\# & \dots & P_{m+1}^{-(r+1)}\# \\ \vdots & \vdots & \ddots & \vdots \\ P_{m+r-1}^{-2}\# & P_{m+r-1}^{-3}\# & \dots & P_{m+r-1}^{-(r+1)}\# \end{pmatrix} \quad (17)$$

Also, let $T_s^{m,r}$ be a sequence s of interest such that we form a vector of length $r + 1$ where each member is the count of the sequence in $\mathcal{S}^{m+k}\forall 0 \leq k \leq r$. For each such sequence there exists by construction a sequence of rational numbers C_s such that

$$T_s^m = \Omega(m, r) C_s \forall m > m_0 \quad (18)$$

Or given the type counts we can invert this to get our constants

$$C_s = \Omega^{-1}(m, r) T_s^m \forall m > m_0 \quad (19)$$

finally

$$T_s^m = \sum_{r=1}^R c_{sr} P_m^{-(r+1)}\#\forall m > m_0 \quad (20)$$

The first gap types we have been able to count are

$$T_2^m = T_4^m = P_m^{-2} \# \forall m > 0 \quad (21)$$

$$T_6^m = 2P_m^{-2} \# - 2P_m^{-3} \# \forall m > 1 \quad (22)$$

The terms are getting too large so we'll just write out the coefficients

$$T_8^m = \{1, -2, 1\} \forall m > 2 \quad (23)$$

$$T_{10}^m = \left\{ \frac{4}{3}, -3, 2 \right\} \forall m > 2 \quad (24)$$

$$T_{12}^m = \{2, -7, 10, -2\} \forall m > 2 \quad (25)$$

$$T_{14}^m = \left\{ \frac{6}{5}, -5, \frac{28}{3}, -3 \right\} \forall m > 2 \quad (26)$$

$$T_{16}^m = \{1, -5, 12, -6, 1\} \forall m > 2 \quad (27)$$

$$T_{18}^m = \left\{ 2, -\frac{23}{2}, \frac{100}{3}, -22, 6 \right\} \forall m > 2 \quad (28)$$

$$T_{20}^m = \left\{ \frac{4}{3}, -\frac{39}{4}, \frac{116}{3}, -40, 24, -2 \right\} \forall m > 3 \quad (29)$$

$$T_{22}^m = \left\{ \frac{10}{9}, -\frac{63}{8}, \frac{632}{21}, -\frac{175}{6}, \frac{72}{5} \right\} \forall m > 2 \quad (30)$$

$$T_{24}^m = \left\{ 2, -17, \frac{1738}{21}, -\frac{344}{3}, 108, -21 \right\} \forall m > 3 \quad (31)$$

$$T_{26}^m = \left\{ \frac{12}{11}, -\frac{209}{20}, \frac{11090}{189}, -\frac{1536}{16}, \frac{4224}{35}, -\frac{119}{3}, \frac{28}{5} \right\} \forall m > 3 \quad (32)$$

$$T_{28}^m = \left\{ \frac{6}{5}, -\frac{185}{16}, \frac{456}{7}, -\frac{325}{3}, \frac{662}{5}, -42, \frac{16}{3} \right\} \forall m > 3 \quad (33)$$

Notice that gaps of size 20 appear out of order, that is when $P_m=13$ we have gaps of 22 while gaps of 20 are conspicuously absent, appearing instead when $P_m=17$.

We've computed a great deal of tuples of 2 and more gap sets as well, for example

$$T_{\{2,4,6,2,6,4\}}^m = \frac{5}{2} P_m^{-7} \# \forall m > 3$$

2.6.1 How Big Is $P_m^{-2} \#$?

Since $P_m^{-2} \#$ is central to prime gap evolution it helps up to understand how big it is relative to other features of S^m . We will show

$$P_{m-1}^{-1} \# < P_m^{-2} \# < P_m^{-1} \# \forall m > 1 \quad (34)$$

Second Minor Growth Lemma.

Proof. The right hand side is trivial so proceeding by induction looking at the left hand side we can check the first few values.

m	P_m	$P_{m-1}^{-1}\#$	$P_m^{-2}\#$
1	2	1	1
2	3	1	2
3	5	2	3
4	7	8	15
5	11	48	135

Assume $P_{m-1}^{-1}\# < P_m^{-2}\#$, the next iteration we multiply the lhs by P_m^{-1} and the rhs by P_{m+1}^{-2} . But $P_{m+1} \geq P_m + 2$ so $P_{m+1}^{-2} \geq P_m > P_m^{-1}$, therefore $P_m^{-1}\# < P_{m+1}^{-2}\# \forall m > 1$. \square

This has profound consequences on our distributions. Because of this there will always be more twin gaps in S^{m+1} than there were all gaps combined in S^m . The same is true for all gap types whose growth rates are $\geq P_m^{-2}\#$ which are $\{2, 4, 6, 8, 10, 12\}$ for large enough m . An immediate conclusion we can draw from this is $P_m^{-2}\#$ grows at a rate very close to $P_m^{-1}\#$.

Dividing 34 by $P_m^{-1}\#$ gives us

$$\frac{1}{P_m^{-1}} < \frac{P_m^{-2}\#}{P_m^{-1}\#} < 1 \forall m > 1 \quad (35)$$

2.7 The Frequency Distribution

We can in theory count all the members of S^m and build a frequency distribution, having at least learned to count up to gap sizes of 32.

That is, for each S^m there is a maximum gap $2max_m$ such that

$$P_m\# = \sum_{j=1}^{max_m} 2jt_j^m$$

where t_j^m is the count of gaps of type $2j$, that is $t_1^m = T_2^m$. Together they form the type class T^m .

Define the Frequency class

$$F^m = \frac{T^m}{P_m^{-1}\#} = \left\{ f_j^m = \frac{t_j^m}{P_m^{-1}\#} \right\}$$

to be the frequency of gaps of size $2j$.

That is

$$\hat{\mathbf{g}}^m = \sum_{j=1}^{max_m} 2jf_j^m \simeq e^\gamma \ln(P_m)$$

This means $\hat{\mathbf{g}}^m$ is twice the first moment of frequency distribution F^m

$$\sum_{j=1}^{max_m} f_j^m = 1 \quad \forall m \quad (36)$$

with

$$\bar{\mathbf{f}}^m = \frac{1}{max_m}$$

That is whatever the local maximum; across the local range, and all things being equal, the average count of each gap type would be about 1 in max_m .

2.7.1 Packing Rules

Because of the CKM, there is a densest any pair of gaps can be because it is certain the CKM will cause some of those pairs to coalesce to form max_{m+1} . Similarly the densest triples will be copied forward some of which will coalesce to pairs, be copied forward yet again and collapse to max_{m+2} .

Now all the new gaps up to max_{m+k} will participate in newer gap formations in ways S^m can not know, the same holds true for all future predictions. Nature will inevitably find more efficient ways to pack primes, it must if it is to fit within its own constraints. But these place firm constraints themselves, all actual future values are bounded from below by the guess values.

2.8 Hardy-Littlewood Prime Distribution Model

The Hardy-Littlewood twin prime constant in our notation is

$$\lim_{m \rightarrow \infty} \frac{T_2^m}{P_m^{-1} \# \hat{\mathbf{g}}^m} = T_2 = 2C_2 \simeq \frac{4}{3} \quad (37)$$

except they computed it over the odd primes while we do so over all primes, the result is our constant is twice as large. This is more natural as we need no extra constants in our description. Earlier we pointed out the fact that $T_2 = P_m^{-1} \#$ is a coincidence of birth order, we need a notation to distinguish our primorial minor constants at infinity. We'll define them as

$$E_r = \lim_{m \rightarrow \infty} \frac{P_m^{-r} \# \hat{\mathbf{g}}^m}{P_m^{-1} \#}$$

read as the limiting value of the Euler product of a primorial minor of depth r .

With Merten's theorem they have taken the limit of the twin frequency against the logarithm noticing it converges they divide back by the logarithm to estimate the local frequency distribution as follows

$$f_1^m \hat{\mathbf{g}}^m \simeq f_1^m e^{\gamma \ln(P_m)} = E_2$$

Dividing by $\ln(P_m)$ gives us

$$\frac{f_1^m e^{\gamma \ln(P_m)}}{\ln(P_m)} = f_1^m e^{\gamma} = \frac{E_2}{\ln(P_m)}$$

We have a e^{γ} scale factor between our estimator and actual predicted frequency. That is it seems

$$f_1^m \simeq \frac{E_2}{e^{\gamma \ln(P_m)}}$$

while this is not observed experimentally.

Experimentally we observe

$$f_1^m e^\gamma \simeq \frac{E_2}{\ln(P_m)}$$

That is the densities of the various gaps begin as their pure primorial minor form and asymptotically grow by an e^γ scale factor. While within S^m we must use the pure form.

To generalize this further

$$\frac{P_m^{-r} \#}{P_m^{-1} \#} e^\gamma \simeq \frac{E_r}{\ln(P_m)} \quad (38)$$

Of course using limiting values for deeper minors will over estimate counts because the deeper the minor the slower it converges to its constant. So being able to compute the actual Euler products for specific indexes will always yield a closer result than the limiting estimate.

2.8.1 Evolution Of The Deeper Minors

Higher order or deeper minors can be studied via the ratio of their Euler product with the Euler product of $P_m^{-2} \#$ as follows

$$R_r^m = \frac{P_m^{-r} \# P_m^{-1} \#}{P_m^{-1} \# P_m^{-2} \#} = \frac{P_m^{-r} \#}{P_m^{-2} \#} \quad (39)$$

The ratio being defined where the products are defined.

Both original products tend to constants as m grows large so their ratio also tends to a constant. We have found no general form for them, experimentally however $e^{-(2r+1)}$ seems a good fit.

$$R_r^m \simeq \bigcirc \left(e^{-(2r+1)} \right) \quad (40)$$

That is though any length of gap sequences of length r can be found to repeat infinitely often, they will occur at a rate of roughly 1 over twice the exponential.

For example consider a sequences of length 100 which could be found around $x = 103^2 = 10609$ and will repeat on the distance scale of $103 \#$ however your odds of finding another are about 1 in $2 * 10^{87}$ at each trial.

There is a powerful motivation to explore this deeper as we discussed earlier primal atoms are sequences of gaps whose equations have a single term. That is *atoms* = $CP_m^{-r} \#$ and these act as building blocks of lowest density regions which dominate most of S^m .

Finally, experimentally at least we found

$$\frac{P_m^{-r} \#}{P_m^{-1} \#} \simeq \frac{E_2}{\ln(P_m)} \bigcirc \left(e^{-(2r+1)} \right) \quad (41)$$

2.9 The Relative Size Of Gap Covariances

Every gap has a left and right neighbor. For example the sum of twins is $P_m^{-2}\#$, we can rewrite this as the sum of twins and their neighbors, which summing to some local maximum we get

$$P_m^{-2}\# = \sum_{j=1} (2|2j) = \sum_{j=1} (2j|2) \quad (42)$$

That is we're adding the number of twins given all the possible neighbors on some side of each twin. We already know how to count some of the pairs since $T_{\{4,2\}}^m = T_{\{2,4\}}^m = P_m^{-3}\#$ so

$$P_m^{-2}\# = 2P_m^{-3}\# + \dots \quad (43)$$

or

$$\frac{P_m^{-2}\#}{P_m^{-2}\#} = 1 = 2\frac{P_m^{-3}\#}{P_m^{-2}\#} + \dots \quad (44)$$

which for large m becomes

$$1 = 0.18 + \dots \quad (45)$$

That is over the entire set of the primes if you find a twin gap it is paired with a cousin about 18% of the time, everything else through infinity must share the remaining distributional space. Not only is the induced p.m.f. fat headed and a long tailed but so are all of its covariances.

2.10 The Generalized Hardy-Littlewood Prime Gap Conjecture

HL went on to argue that the sum of the set of all prime gap spans of order n ; all subsequences of gaps that sum to n ; written χ_n^m will be found with frequencies

$$\chi_n^m = P_m^{-2}\# \prod_{q|n} \frac{q-1}{q-2} \quad (46)$$

We refer to

$$\prod_{q|n} \frac{q-1}{q-2}$$

as the *HL boost* factor. For example, gap subset spans of size 2^n have the lowest boost of 1 while spans of size $2^n 3^m$ have a boost of 2, that is they will be twice as common.

The reason it works in the S model and why it's actually true, besides their original clever argument, is because it fits. Inverting the boost divides into every set, basically the $q-1$ factors cancel as long as the largest prime factor is $\leq P_m$. If the largest prime factor is greater it simply can't exist yet.

To go from a gap span of size n to a gap of size n we need to find all the subspan counts and subtract them away from $P_m^{-2}\#$.

For example, the entirety of spans of 6 is $\chi_6^m = 2P_m^{-2}\#$ while it has 3 *decay paths*, $\{4, 2\}$, $\{2, 4\}$ and $\{6\}$. We arrive at T_6^m by starting with the boost times the 2^{nd} minor

and subtract all pairs (or larger tuples as needed) and our gap count is what remains. Since each pair occurs at $P_m^{-3}\#$ we get

$$T_6^m = \chi_6^m - 2P_m^{-3}\# = 2P_m^{-2}\# - 2P_m^{-3}\#$$

As n increases it also evolves many more decay path possibilities which grow (apparently) exponentially in count leaving little count space for individual gaps.

That is, except for the HL boost, gaps of size n occur with an (apparently) exponentially decaying rate. This correlates well with the top heavy fat tail observation though it needs (substantially) further investigation.

2.10.1 A Note On Compressibility

It's trivial now to see the distribution of gaps is locally compressible within any reasonable iteration. For example because S^0 and S^1 have single members you only need to store which primordial you use and the maximum value, and we can trivially reconstruct the entire set.

Even when the sets get relatively large compression is useful, it certainly helped us when we studied sets with billions of members.

However, because the CKM is constantly creating new rules, the value of any lookup table decreases in value logarithmically.

And this is even more dramatic on the primes because most of the patterns seen in a particular S^m eventually evaporate due to the CKM.

That is, if you use a compression lookup table and it gives you a compression quality $q_m = \frac{P_m^{-1}\#}{C(S^m)}$ at S^m , with C being the compressed size; we want a number > 1 , than

$$\lim_{k \rightarrow \infty} q_{m+k} = 1 \tag{47}$$

Compression is only possible if new data contains either constant or growing relative rates of members of your lookup table. This is simply not the case in either S^m or the naturals.

2.11 Gap Behaviors At Different Ranges

The group structure manifests a handedness due to Z_6 and this will be inherited by all future sets. All primes must now either be of the form $1 + 6n$ (left) or $5 + 6n$ (right). As well, all prime gaps must now evolve from various sequences of gaps which began life as sequences of various lengths of $\{4, 2\}$ or their symmetric cousin $\{2, 4\}$ while never allowing the possibility of ever having had $\{2, 2\}$ or $\{4, 4\}$ in any subsequence (except S^1).

For example gaps of 8 began as $\{2, 4, 2\}$ and gaps of 10 as $\{4, 2, 4\}$. It is instantly apparent that gaps of 8 could never pair with themselves, nor could gaps of 10 while they could pair together. Also gaps of 8 $\{2, 4, 2\}$ could never pair with twins and gaps of 10 $\{4, 2, 4\}$ can never pair with cousins.

Every gap from now on (post S^2) is of the form $g_j=6n$ or $g_j=2+6n$ or $g_j=4+6n$ for some n . That is

- If n is even, if its anchor prime is
 - left handed than to the right is a right handed gap or it is
 - right handed and to the right is a left handed gap
- Or if n is odd, if its anchor prime is
 - left handed and has a left handed neighbor to the right or
 - it is right handed and has a right handed neighbor to the right

We also have the following rule, only different hands can touch. This protects us from ever seeing $\{2, 2\}$ and $\{4, 4\}$. Similarly we see frequent pairing of 6's but close inspection reveals the must be different handed.

We refer to these as S(hort) R(ange) R(ule)(s).

2.11.1 Immediate Consequences Of The SRRs

The obvious first rule is that a gap of size $6n$ can pair with anything while gaps of $2+6n$ or $4+6n$ must mismatch, unless it's with a gap of size $6n$. That is if these were the only rules for tuples of arbitrary length, choosing a random sequence of length n has about a 1 in $\frac{2}{3}^n$ chance of existing.

Examples of higher banned pairs are $\{2, 8\}$, $\{8, 8\}$, $\{4, 10\}$, $\{10, 10\}$ and their mirrors and so on. Also any pair of numbers which does not violate the SRRs will also occur, the same is true for triples. But it breaks here.

2.11.2 Longer Range Rules

There is an infinite set of these because for each length scale there is an absolute lower bound on minimum densities as a result of the primordial copy machinery. That is, though the average gap is growing and so is the average size of tuples of length say n we still have tuples of length n which originated long before and so retain their lower density; in fact they dominate the space.

As well, experimentally the gaps seem as defined by what is banned as by what is allowed. For example, though gaps of 6 are the most popular, as are pairs and triples, 4 consecutive gaps of 6 is banned as its feedstock was killed in S^3 . Consider its gap set $\{6, 4, 2, 4, 2, 4, 6, 2\}$, you can not construct $\{6, 6, 6, 6\}$.

For the first few larger prime gap sequences we could study experimentally, about one half of 'expected' sequences; i.e. not violating the SRRs, were missing. That is there maybe infinitely many gap combinations of counts of 4 or more which pass the SRRs that are ultimately banned by construction.

2.12 Induced Topology

Recall

$$\mathbf{S}^{m+1} \subset \mathbf{S}^m \quad \forall m \geq 0 \quad (48)$$

That is given G^m , the set of gaps in S^m , we say the set of its ordered subsets is open for all $G^{m+k} \forall k > 0$ and closed $\forall -m \geq k \leq 0$.

This is in the sense that earlier gap subsets can cover most but not all of each model because we're always creating more rules. While at the same time it has more information than needed to reconstruct earlier sets. For example the first step in constructing S^{m+1} is building P_{m+1} copies of S^m gaps which is followed by consuming $P_m^{-1} \#$ sets of pairs to create $P_m^{-1} \#$ new gaps that can not be covered by older rules.

That is S^m can accurately describe

$$\frac{P_{m+1}^{-2} P_m^{-1} \#}{P_{m+1}^{-1} \#} = \frac{P_{m+1}^{-2}}{P_{m+1}^{-1}}$$

parts of S^{m+1} . That is the degree of self-similarity comes arbitrarily close to 100% as m becomes arbitrarily large.

But it is also recursive such that given we understand S^m , its knowledge of S^{m+k} is

$$\frac{P_{m+k}^{-2} \# P_m^{-1} \#}{P_{m+k}^{-1} \# P_m^{-2} \#} \simeq \frac{\ln(P_m)}{\ln(P_{m+k})}$$

Realistically we can never have very much knowledge because the sets grow too quickly but knowing that first little bit does at least begin to describe a portion of about

$$\frac{1}{\ln(P_{m+k})}$$

2.13 How Random Is S^m

Given that we can exactly predict at least the count of twin gaps we can ask what would happen at some m if instead of the composites being kicked out rationally it went haywire and the members kicked out were chosen randomly? At some m we have $P_m^{-2} \#$ twins from a collection of $P_m^{-1} \#$ gaps, we make P_{m+1} copies and randomly choose $P_m^{-1} \#$ pairs of gaps to coalesce into new gaps. Twins can only be eaten and never born during coalescence so the count in the next generation is expected to be

$$P_{m+1} P_m^{-2} \# - \frac{P_m^{-2} \# P_m^{-1} \#}{P_{m+1}^{-1} \#} = P_m^{-2} \# \left(P_{m+1} - \frac{1}{P_{m+1}} \right)$$

or

$$P_{m+1}^{-2} \# \left(\frac{P_{m+1}}{P_{m+1}^{-2}} - \frac{1}{P_{m+1}^{-2} P_{m+1}} \right)$$

That is the S model approaches a uniform random behavior as m becomes arbitrarily large.

3 The S Model P.M.F. Structure And Geometry

We now ask what is the maximum allowed gap max_m in S^m . We have argued the gap distribution is highly random with its unique distribution for each m , now we assert this as an axiom. That is we assert the distribution of the primes is everywhere random with a local mean beginning as $\hat{\mathbf{g}}^m$ and asymptotically limiting to $\ln(p)$.

That is the distribution of gaps is a random bag and the particular order we measure in the primes is a single sample drawn from the bag. In other words the sequence of intervening gaps is a random walk of gaps of *unit length* selected from the derived frequency table F^m from S^m where instead of selecting direction we choose a particular realized instance of the *unit step*.

In yet other words given exact knowledge of the first m primes, it doesn't give us enough information to exactly predict the next prime. At best the guess is the last (current) prime plus the logarithm of that number ranging from some minimum error (it can't be worse than 2 and 2 is always a possibility) to a maximum. Of course knowledge of the SRRs can cut by a third the possibilities, but it remains fundamentally uncertain.

3.0.1 Everything Becomes Beautiful

We have to assume the first gap can always be the largest gap, and we need some outside force to constrain local growth. That is, we need S model plus something else. Given this we note Bertrand's lemma² which states

Bertrand's Lemma². $\forall n > 1 \exists P \in \text{primes such that } n < P < 2n$

If max_m is the largest gap and the gaps are chosen at random the only possibility allowed which satisfies all S^m is if $max_m \leq 2 P_m^{-1} < 2P_m$. That is since

$$P_m < P_{m+1} < 2P_m$$

or

$$P_m - 1 < P_{m+1} - 1 = g_0^m \leq 2P_m - 2 < 2P_m - 1$$

or finally since the smallest gap is a twin we get

$$P_m + 2 \leq P_{m+1} \leq 2P_m - 2 = 2P_m^{-1} + 1 = max_m + 1$$

While this protects the first non-trivial member's gap from violating the rule it does so at the price of capping the entire set.

The flavor of random is not uniform because we clearly have stable covariance and memory via the SRRs. This forces the largest span between any 3 co-primes (sum of 2 gaps) to be $2P_{m+1}^{-1}$, since the CKM will cause such pairs to coalesce in the next iteration of the model. The same is true for the next few tuple lengths as well such as gap triples and quads. S^m must know at least something about the future of other primorial modular sets because they evolve from it. Nature's solution is to make the gaps random but bounded.

The following table gives the first few local maximums.

m	P_m	g_{max_m}	$2P_m^{-1}$
0	1	1	undef
1	2	2	2
2	3	4	4
3	5	6	8
4	7	10	12

We see gap max is undefined in S^0 because $2(1 - 1) = 0 \notin S^0$.

Our p.m.f. now has a local maximum size though we see not all the tables are necessarily fully filled.

3.1 Returning To Our Gap P.M.F.

Recall $T_{2j}^m = \{t_j \mid \forall 0 > j \leq P_m^{-1}\}$ is the collection of counts of even gap types of size $2j$ in S^m . That is

$$\sum_{j=1}^{j \leq P_m^{-1}} 2jt_j^m = P_m\# \quad (49)$$

Since S^m has a count of $P_m^{-1}\#$ let $F^m = \{f_j^m = \frac{t_j^m}{P_m^{-1}\#}\}$ be the frequency density of $2j$ in S^m . That is

$$\sum_{j=1}^{j \leq P_m^{-1}} f_j^m = 1 \quad (50)$$

While its first additive moment is half the average gap

$$\sum_{j=1}^{j \leq P_m^{-1}} 2j f_j^m = 2 \sum_{j=1}^{j \leq P_m^{-1}} j f_j^m = 2 \langle F^{m,1} \rangle = \hat{\mathbf{g}}^m = \frac{P_m\#}{P_m^{-1}\#} \quad (51)$$

3.1.1 Structural Summary

The logarithm is not simply the local average gap at infinity. Life begins that way at the birth of the primes, as demonstrated by the JIT sieving method of the S model. Beginning from S^0 with $\hat{\mathbf{g}}^0=1=successor^0=successor$, our Euler product is constructively perfect. Successor must be considered as a random operator, meaning given some large prime P, we can never really know what the next prime is. Within the S model $\hat{\mathbf{g}}^m$ acts as successor while among the primes it begins as our product and evolves quickly to $\ln(P)$.

But we're not blind, on average; given many primes to test, the next prime will be about $\ln(P)$ away. Though it can range from 2 up to $2(\sqrt{p}-1)$ more often than not it will be $<\ln(P)$ because of the top heavy nature of each local distribution. We can also suggest an efficient search algorithm, check the gaps with the greatest local frequency first, in order $\{6, 12, 10, 2, 4, 8, \dots\}$, but don't forget the SRRs.

3.2 A Probabilistic Look At The Primes

Just as $\pi(x)$ is the member counting function for primes on the naturals below some value x we can define a similar function in the S model

$$\pi^m(P_j^m) = j$$

that is our member counting function returns the index of the nearest member not greater than some value. By doing this we rid ourselves of the pesky issue of counting the fully degenerate and emotionally discordant zeroth prime, 1 as we had cleverly set its index to zero. To compare this co-prime member counting function with the natural version we need to add the m members lost to the midden and remember our first composite is P_{m+1}^2 . That is

$$\pi(x) = m + \pi^m(x) \forall x < P_{m+1}^2$$

and in particular

$$\pi(P_{m+1}^2) = m + \pi^m(P_{m+1}^2) - 1$$

because P_{m+1}^2 is counted in the local function. So

$$\pi(P_{m+1}^2) - \pi(P_m^2) = \pi^m(P_{m+1}^2) - \pi^m(P_m^2) - 1$$

Define

$$\Delta(P_m) = \pi^m(P_{m+1}^2) - \pi^m(P_m^2)$$

as the expected count in which we naively assume the span between our anchors is uniformly filled with the effective unit operator. Notice $\ln(P_m)$ is the mean of the local distribution in this region and the average local gap is twice the local mean. Next we have 2 paths; the first is no knowledge of the next gap and the second is knowing the proceeding gap.

$$\Delta(P_m) = \frac{(P_m + \ln(P_m))^2 - P_m^2}{2\ln(P_m)}$$

this becomes

$$\Delta(P_m) = \frac{2P_m \ln(P_m) - \ln^2(P_m)}{2\ln(P_m)}$$

and finally

$$\Delta(P_m) = P_m + \frac{\ln(P_m)}{2} \tag{52}$$

That is we can expect P_m with a half gap contribution when sampled over many m . However the last term is a small contribution which vanishes since $P_m \gg \ln(P_m)$ for large m .

We also examine

$$\Delta(P_m) = \frac{(P_m + g_m)^2 - P_m^2}{2\ln(P_m)}$$

which becomes

$$\Delta(P_m) = P_m \frac{g_m}{\ln(P_m)} + \frac{\ln(P_m)}{2}$$

Once again we can make sense of this probabilistically as a large enough sample of m will average out our surviving product term or in other words

$$\lim_{m \rightarrow \infty} \frac{1}{n} \sum \frac{g_m}{\ln(P_m)} = 1 \quad (53)$$

So both models are the same in the end as they must be.

Also, as a sanity check we have the following by construction

$$\lim_{m \rightarrow \infty} \inf \left(\frac{g_j}{\ln(P_j)} \right) = \lim_{m \rightarrow \infty} \frac{2}{\ln(P_j)} = 0 \quad (54)$$

and

$$\lim_{m \rightarrow \infty} \sup \left(\frac{g_j}{\ln(P_j)} \right) = \lim_{m \rightarrow \infty} \frac{2(P_j - 1)}{\ln(P_j)} = \infty \quad (55)$$

3.3 Chebyshev's Function⁵

We can find the $m^{\text{th}} + 1$ prime if we have knowledge of the first m gaps as follows

$$P_{m+1} = 1 + \sum_{j=0}^m g_j \quad (56)$$

Replace the actual gap with its random variable

$$P_{m+1} \simeq 1 + \sum_{j=0}^m \ln(P_j) \quad (57)$$

and we get Chebyshev's function. We can extend this for prime squares as follows. We notice

$$P_{m+1}^2 = 1 + \sum_{j=0}^m (P_{j+1}^2 - P_j^2) \quad (58)$$

We can reformulate our definition as follows

$$P_{m+1}^2 = 1 + \sum_{j=0}^m (P_{j+1} + P_j) (P_{j+1} - P_j)$$

but

$$\sum_{j=0}^m (P_{j+1} + P_j) (P_{j+1} - P_j) = \sum_{j=0}^m (P_j + g_j + P_j) g_j = \sum_{j=0}^m (2P_j g_j + g_j^2)$$

Replacing our gap with the random variable, simplifying and taking square roots gives us

$$P_{m+1} \simeq \sqrt{\sum_{j=0}^m (2P_j \ln(P_j) + \ln^2(P_j))} \simeq \sqrt{\sum_{j=0}^m 2P_j \ln(P_j)} \quad (59)$$

The 2 different forms tell us 2 different stories, the first that each gap has a unique meaning, different from the previous and different from the next. The second shows us on the scale of the primes squared they appear locally the same.

3.4 Backbones And Hard-Stops

3.4.1 The Backbones

We can compare the actual primal counts against something we call the primal backbone function, our line of expected outcomes. Let $x=1 + j\hat{\mathbf{g}}^m$ be our estimator for the j^{th} member, define

$$L^m(x) = \left\lfloor \frac{x-1}{\hat{\mathbf{g}}^m} \right\rfloor$$

to be the expected count for some x . Notice once again we ignore 1 by subtracting it to maintain alignment with our member counting function. That is we have by construction

$$L^m(x) = \langle \pi^m(x) \rangle \quad (60)$$

That is, the backbone is the expectation of the member counting function. It has the interesting property that

$$L^m(x + \hat{\mathbf{g}}^m) = L^m(x) + 1$$

That is $\hat{\mathbf{g}}^m$ is also the local successor along the backbone.

It is straight forward to show

$$L^m(P_{m+1}^2) - L^m(P_m^2) = 2P_m + \hat{\mathbf{g}}^m \quad (61)$$

This has diverged from 52 because $\hat{\mathbf{g}}^m$ is the gap everywhere across S^m while the PNT enforces a factor of $\frac{1}{n}$ when $P_x = P^n$.

We need to generalize this so we'll analytically continue the backbone constrained with the PNT to first order.

Since the expectation operator is a function of m via P_m , we have an infinite number of backbones, for each unique m . We can ask 2 questions of the backbone, that it satisfy both forms of the Chebyshev functions. That is we want

$$L(x + \ln(x)) - L(x) = 1$$

while enforcing the PNT at x^2 ; the 2^{nd} form of Chebyshev's function, gives us

$$L((x + \ln(x))^2) - L(x^2) = x$$

Both generate the same function when Taylor expanded to first order, it is easy to show

$$\frac{dL(x)}{dx} = \frac{1}{\ln(x)} \quad (62)$$

for both, and

$$L(x) = Li(x) \quad (63)$$

That is, the analytic continuation of the backbone of the S model is $Li(x)$, the logarithmic integral function⁷. This means $Li(x)$ is indeed the best possible representation of the primes. Just as the expectation of the members of S^m is $L^m(x)$ we have

$$Li(x) = \langle \pi(x) \rangle \quad (64)$$

By construction the Logintegral is the expected value of the count of primes less than a number.

3.4.2 The Hard-Stop

We prove the following

Legendre's Prime Density Conjecture⁶. $\forall n > 1 \exists$ a prime P s.t. $n^2 < P < (n+1)^2$

Proof. The span between our bounds is $2n + 1$ while the largest gap in the population of gaps in this neighborhood is $\leq 2(n - 1) = 2n - 2$ so we can always fit at least one gap or 2 primes in this space. \square

Since Legendre's prime density is true we have a natural bound on the uncertainty, we expect roughly P_m members between successive primes squared and there must always be at least 2

$$\pi^m(P_{m+1}^2) - \pi^m(P_m^2) - P_m > 0 \quad (65)$$

or

$$-P_m > \pi^m(P_{m+1}^2) - L^m(P_m^2) \quad (66)$$

If the order of the error were any different from 1, say $1 + \epsilon$ than for large enough P_m the expected error would eventually lead to an infinite number of violations. This only represents the tightest density case which given that S^m as a p.m.f. is symmetric leads to

$$-P_m > \pi^m(P_{m+1}^2) - L^m(P_{m+1}^2) < P_m \quad (67)$$

A reasonable continuation of 67 is

$$|\Psi(x)| = |\pi((x + \ln(x))^2) - Li((x + \ln(x))^2)| \leq x \quad (68)$$

or

$$|\Psi(x)| = |\pi(x^2) - Li(x^2)| < x \quad (69)$$

This also rules out Von Koch's estimate as being too large.

We'll now show even this is too large.

3.4.3 Another Approach

Returning to examine tuples of a certain length and their variability, because the tuple rows of S^m are periodic $S_j^{m,n} = S_{j+nP_m^{-1}}^{m,n} \forall n$ we can draw a ray joining each of the periodic points and it's slope will be $\ln(P_m)$. This means we have a family of P_m^{-1} parallel rays and we want to find the intercepts furthest from the backbone.

We note all tuples of length 2 are size bounded from above because the densest a pair can be in S^m is $2P_{m+1}^{-2}$ since it is certain such pairs will coalesce in S^{m+1} . The same is true for all tuples up to length P_m , actually $P_{m+1} - 2 \geq P_m + 2 - 2$ because of the CKM. Thus we replace all the gaps with multiples of the random variable which in this case the upper bounding values become

$$max(n) = P_m + n \ln(P_m) \quad (70)$$

Identifying the lower bounding envelop is a bit more complex. For example the smallest pairs in each set $\{\{4, 2\}, \{2, 4\}\}$ are independent of the current index, this is true for all tuples up to length P_m . The first few smallest tuples sum to $\{2, 6, 8, 12, 16\}$. These early gap tuples grow as $\ln(P_l)$ in the iteration $m = l$ they were created. For example twins when $m = 1$, pairs summing to 6 when $m = 2$, tuples of 3,4 and 5 when $m = 3$ and so on. By the time we reach the region of tuples bigger than those that can sneak past the CKM they must catch up to the stable growth rate of the backbone because that is its feedstock, this occurs around P_{m+1}^2 . That is, the rays below the backbone slowly spread away as the tuples grow until they reach the average expected size driven by the local mean. On the scale of S^m this is actually rapid as the size of the set is roughly exponential while this catchup game is played on order P^2 . So by symmetry

$$\min(n) = -P_m + n \ln(P_m) \quad (71)$$

That is we have 3 parallel lines, a central backbone and 2 equidistant neighbors which represent the expected bounding limits of of tuple count to span. It also gives us our earlier result.

However when we ask questions of the primes they are always of the form given a span what is the count, this while we can only currently answer given a count to predict the span. We can translate from count to span by checking our limits from normal projections of the expected line instead of normal to the axes. That is we simply divide by the slopes and the ranges now become

$$\max(n) = m + n \quad (72)$$

$$\min(n) = -m + n \quad (73)$$

so

$$|\Psi(P_m)| \leq m \forall m \quad (74)$$

or when continued

$$|\Psi(x)| < \frac{x}{\ln(x)} \quad (75)$$

However a little insight into the behavior of $Li(x)$ lets us rewrite equation 75 as

$$|\Psi(x)| < Li(x) \quad (76)$$

3.4.4 The Riemann Prime Counting Function¹

$$\pi(x) = Li(x) - \frac{Li(\sqrt{x})}{2} - \sum_{\rho}^P Li\left(\frac{x^{\rho}}{2}\right) \quad (77)$$

The first term is as expected the primal backbone next we offset by the lower middle of the middle of the range, we add a sum of factors involving the zeros of the zeta plus some small terms we'll ignore. We do a variable transform to x^2

$$\pi(x^2) - Li(x^2) = \Psi(x - \ln(x)) = -\frac{Li(x)}{2} - \sum_{\rho}^P Li\left(\frac{x^{2\rho}}{2}\right) \quad (78)$$

referring to 76 we get

$$-Li(x) < \Psi(x - \ln(x)) = -\frac{Li(x)}{2} - \sum_{\rho}^P Li\left(\frac{x^{2\rho}}{2}\right) < Li(x) \quad (79)$$

and finally

$$-1 < \frac{\Psi(x - \ln(x))}{Li(x)} < 1 \quad (80)$$

We refer to the above as the normalized error function.

3.5 Life Along The Backbone

At the birth of the primes we have $P_0 = 1$, $P_1 = 2$ and $\pi^0(P_{m+1}^2) - \pi^0(P_m^2)$ for $m = 0$ becomes $\pi(4) = 2$ which is exactly on the backbone since S^0 is degenerate. The same holds for $m = 1$. This means the prime distribution begins on the central backbone. Naively given S^m generates a random walk periodic on $P_m^{-1}\#$ and that we have completely symmetric local absolute bounds we expect by symmetry “frequent” crossings.

Littlewood showed that $\pi(x)$ crosses the primal backbone infinitely often¹¹ and our work not only fully substantiates this but given all the symmetries expect nothing less. That is we’re modeling a random walk about a central expected moment, little more than a coin toss experiment. What is completely impossible to explain is the frequency of crossings, occurring at numbers referred to as Skewes’ numbers¹¹. The smallest known crossing other than 1 is near $e^{727.95133}$ as opposed to say 10 or even 100.

3.6 Experimental Observations And Sub-Spines Revealed

We used Mathematica to generate the table $\{m, P_m, P_m^2, \pi(P_m^2), Li(P_m^2)\} \forall 0 < m \leq 200000$. Given the above hypothesis we found

$$E\left(\frac{\Psi(x)}{Li(x)}\right) \simeq -\frac{1}{2}$$

in the region of study. That is the distribution of the primes are actually random valued between the backbone and the negative hard-stop with an expectation along the negative sub-spine at least until they spontaneously cross the backbone at the next Skewes number. We conjecture the following, if the last Skewes number’s index is even

$$E\left(\frac{\Psi(x)}{Li(x)}\right) \simeq -\frac{1}{2} \quad (81)$$

for all large enough samples until the next crossing. While if the index is odd we would have

$$E\left(\frac{\Psi(x)}{Li(x)}\right) \simeq \frac{1}{2} \quad (82)$$

for all large enough samples in the region until the next crossing. While sampling over the entire space for large even valued Skewes indexes we would have

$$E \left(\frac{\Psi(x)}{Li(x)} \right) \simeq 0 \quad (83)$$

The above possibility means the oscillatory component of $\Psi(x)$ has 2 hidden strange attractors,

$$\frac{\Psi(x)}{Li(x)} = \pm \frac{1}{2} + \text{“noise”}$$

and it switches between the attractors on perhaps the longest length scale studied anywhere to date.

And this also means we can also begin to give meaning to the $\frac{Li(\sqrt{x})}{2}$ term in Riemann’s prime counting function, it’s a sub-spine.

4 Generalized Modular Multiplication And Member Gaps

Let n be a random number whose largest divisor is P_m . That is

$$n=r * \frac{P_m\#}{\{P_j\}}$$

where $\{P_j\}$ is the set of missing members of $P_m\#$ and r is the product of other factors not participating in the primorial. By construction the largest gap is still $2(P_m-1)$. That is a generalized multiplication is still bounded at worst by the largest prime factor.

In fact multiplication modulo a semiprime P_aP_b with $P_a < P_b$ looks identical to Z_1 until P_a is suddenly observed missing leaving a gap of 2. The number of numbers $\leq P_aP_b$ is P_aP_b while Z_n has $(P_a-1)(P_b-1)$. That is it is different from the integers by a portion of

$$\frac{P_a+P_b-1}{P_aP_b} \simeq \frac{1}{P_a}$$

and this disappears for large primes. It is also highly compressible since it mostly remains long strings of gaps of 1.

It’s only with each possible prime participating that all possible coalescences can occur. Any other factor simply adds positional requirements (copy number) without coalescence, that is one need only need knowledge of the primal core gap set and how many copies to make in order to restore the entire gap set. For example Z_n with $n=2^l P_m\#$ is just 2^l copies of $P_m\#$ ’s gap set making such sets highly compressible. As hinted at by Euler’s phi function for a general composite, it is only the first occurrence of a prime which contributes to the size of the core pattern of member gaps, the rest tells us how many copies to make.

Another point is of all multiplication systems we can choose, the natural numbers and the primes form the most complex one.

5 The Chebyshev Bias

Given the deep symmetries which exist in the primes we want to account for the Chebyshev bias which states primes of the form $4k + 3$ are more frequent than primes of the form $4k + 1$.

The remarkable fact is that S^m is completely unbiased to this form. Even though it is nowhere used in this calculus it is easy to show and we leave it to the reader this is the case. And yet we have the remarkable calculation of Landau and Ramanujan⁴¹⁰

$$\frac{\pi}{4} \prod_{p=1 \bmod 4} \left(1 - \frac{1}{p^2}\right)^{\frac{1}{2}} = \frac{1}{\sqrt{2}} \prod_{p=3 \bmod 4} \left(1 - \frac{1}{p^2}\right)^{-\frac{1}{2}}$$

which clearly indicates a bias so what gives?

The S model shows us that equations of Euler products; if properly combined, describe counts of the structure of interest when viewed in the context of the limiting count of P_m in the region just above P_m^2 .

The disappearance of the single P_m^2 caused a disappearance on the other side of the pivotal cousin since $P_m \# - P_m^2 \notin S^m$ so S^m remains unbiased. That is while all S^m remain unbiased the midden set; the ashes of the S model, retains a small bias.

Apparently P_m^2 is so close to being prime it got counted in some sense in nature's strive to maintain neutrality and *snuck by*. Importantly it's of the same scale of other Euler products like E_2 which itself is of the order $\frac{1}{P_m}$.

This also leaves a remarkable bias in the prime gap distribution. Consider in S^0 all gaps were of size one and suddenly *almost randomly* the CKM exerts its damage and as quickly there is a new gap where 4 used to be and it's twice as big as its parent. From S^1 all gaps were of size 2 until suddenly the CKM kills 9 and there is a gap of 4, twice it's parent. Where 25 stood is $6 = 2\hat{g}^3$.

This is methodical in the beginning but becomes probabilistic as m grows so for an arbitrary P_m , by the time it's scheduled for the incinerator the gaps are on average the same on either side around p^2 ; a restatement of the Chebyshev function's convergence. In the ashes we see this as follows, if p is a prime, i.e. $\in midden$ find the m s.t. P_m is the largest prime $< p^2$, on average

$$\frac{g_{m-1}}{\ln(P_m)} \simeq \frac{g_{m+1}}{\ln(P_m)} \simeq 2 \frac{g_m}{\ln(P_m)} \quad (84)$$

That is if you look at where p^2 used to be, on average the gap will be twice the expected value. We propose naming the gap bias the Chebyshev gap bias in his honor as it is a direct consequence of the bias and the S model.

Of course since Chebyshev many more biases have been found. The general rule is if the counting correct, the result is not theoretical, it's concrete.

6 Acknowledgments

We thank the authors and contributors of Wikipedia and MathWorld. This work could never have been completed without the vast store of human knowledge so beautifully categorized by all the wikipedians and contributors to MathWorld. We are all the more better for them.

We thank Dany at Lonely Beach Resort who gave us refuge from another of life's many storms. This work is not the result of a singular pursuit but the result of the solitude a great escape can offer. We thank Rudy and Gil at Paradise Bungalows, the best escape artists on the planet.

Also no small part was played by the author's ASUS netbook. It was asked to compute far more than it was ever expected to perform.

“Are there any more questions?” - The Buddha's penultimate words.

“All composite things are impermanent” - The Buddha's last words.

References

- [1] Len Goodman and Eric W. Weisstein. Riemann hypothesis, April 2012. URL <http://mathworld.wolfram.com/RiemannHypothesis.html>.
- [2] Jonathan Sondow and Eric W. Weisstein. Bertand's postulate, April 2012. URL <http://mathworld.wolfram.com/BertrandsPostulate.html>.
- [3] Jonathan Sondow and Eric W. Weisstein. Mertens theorem., April 2012. URL <http://mathworld.wolfram.com/MertensTheorem.html>.
- [4] Eric W. Weisstein. Chebyshev bias., April 2012. URL <http://mathworld.wolfram.com/ChebyshevBias.html>.
- [5] Eric W. Weisstein. Chebyshev functions, April 2012. URL <http://mathworld.wolfram.com/ChebyshevFunctions.html>.
- [6] Eric W. Weisstein. Legendre's conjecture., April 2012. URL <http://mathworld.wolfram.com/LegendresConjecture.html>.
- [7] Eric W. Weisstein. Logarithmic integral., April 2012. URL <http://mathworld.wolfram.com/LogarithmicIntegral.html>.
- [8] Eric W. Weisstein. Modulo multiplication group., April 2012. URL <http://mathworld.wolfram.com/ModuloMultiplicationGroup.html>.
- [9] Eric W. Weisstein. Prime number theorem, April 2012. URL <http://mathworld.wolfram.com/PrimeNumberTheorem.html>.

- [10] Eric W. Weisstein. Landau-ramanujan constant., April 2012. URL <http://mathworld.wolfram.com/Landau-RamanujanConstant.html>.
- [11] Eric W. Weisstein. Skewes number, April 2012. URL <http://mathworld.wolfram.com/SkewesNumber.html>.
- [12] Eric W. Weisstein. Totient function., April 2012. URL <http://mathworld.wolfram.com/TotientFunction.html>.
- [13] Eric W. Weisstein. Twin primes, April 2012. URL <http://mathworld.wolfram.com/TwinPrimes.html>.

7 Notes

*In this work, except for what we introduce, we use no modern (post 1930) mathematics. The author has no access to pay per view scholarly journals so all references are wiki based. This seems reasonable as all of the research required for this investigation has long since become a part of history and is well diffused among the community. In fact S model plus Von Koch's¹ 1901 paper is all that is required to prove RH. We reference later papers such as those by Hardy, Ramanujan et al to probe deeper structures.