Divisor Functions and the Distribution of the Primes

Take a look at this recursive function.

$$f_k(n,1)=0 \qquad\qquad m=\{1,-1,-1,0,-1,1,-1,\ldots\}$$

$$f_k(n,j)=\frac{1}{k}-f_{k+1}(\lfloor\frac{n}{j}\rfloor,\lfloor\frac{n}{j}\rfloor)+f_k(n,j-1) \qquad\qquad p(n)=\sum_{k=1}^{\lfloor\log_2 n\rfloor}\frac{m_k}{k}f_1(\lfloor n^{\frac{1}{k}}\rfloor,\lfloor n^{\frac{1}{k}}\rfloor)$$

Evaluate $p(n)$, and you'll find it bumps up by 1 every time n is prime. It stays constant otherwise. It's a function that counts primes. Here's some Mathematica code to watch $p(n)$ work.

```
f[ n_, j_, k_ ] := 1 / k - f[ Floor[ n / j ], Floor[ n / j ], k + 1 ] + f[ n, j - 1, k ]; f[ n_, 1, k_ ] := 0
p[ n_ ] := Sum[ m[[ k ]] / k f[ Floor[ n ^ ( 1 / k )], Floor[ n ^ ( 1 / k )], 1 ], { k, Log[ 2, n ]}]; m := {1,-1,-1,0,-1,1,-1,0,0,1,-1,0,-1,1,1,0}
{ $RecursionLimit = 10000 }; Table[ { n, p[ n ], p[ n ] - p[ n - 1 ]}, { n, 2, 100 }] // TableForm
```

So, why exactly does $p(n)$ work? How is that equation counting primes?

$f_k(n,j)$, doing the heavy lifting, doesn't seem *that* complicated. It's a simple recursive function, only a bit more complex than the Fibonacci sequence definition $F(n)=F(n-1)+F(n-2)$ with $F(1)=1$ and $F(0)=0$.

Perhaps looking at some mainstays of prime counting will help.

Some prime counting functions manually conduct primeness tests on ranges of numbers, summing the results. Is $f_k(n,j)$ doing that? As an example, with $\pi(n)$ the notation for $\#\ primes \le n$, the identity

$$\pi(n)=-1+\sum_{j=1}^{n}\lfloor\cos^2(\pi\frac{(j-1)!+1}{j})\rfloor$$

```
Table[{n, -1 + Sum[ Floor[Cos[Pi ((j-1)!+1)/j]^2], {j,1,n}]},{n,1,100}]//TableForm
```

uses Wilson's Theorem to identify primes. It doesn't rely on properties of the distribution of primes at all; it just identifies individual primes and counts them. There are a lot of functions like this, but let's disregard them all. $f_k(n,j)$ clearly isn't testing any n or j for primeness directly, so let's move on.

Another really important general approach to prime counting works by sieving ranges of numbers by smaller primes. The Sieve of Eratosthenes is probably the most familiar such approach. Adrien-Marie Legendre gave a formula using a similar idea, but in a less algorithmic, more mathematical form.

$$\varphi(n,0)=n \qquad\qquad (p_j=\text{the } j^{th} \text{ prime})$$

$$\varphi(n,j)=-\varphi(\lfloor\frac{n}{p_j}\rfloor,j-1)+\varphi(n,j-1) \qquad\qquad \pi(n)=\pi(\lfloor n^{\frac{1}{2}}\rfloor)+\varphi(n,\pi(\lfloor n^{\frac{1}{2}}\rfloor))-1$$

```
Phi[ x_ ,a_ ] := -Phi[ Floor[ x / Prime[ a ]], a - 1 ] + Phi[ x, a - 1 ]; Phi[ x_, 0 ] := x
Table[ { n,  Phi[ n, PrimePi[ n ^ ( 1 / 2 )]] + PrimePi[ n ^ ( 1 / 2 )] - 1 }, { n, 1, 100 }] // TableForm
```

That seems more promising! In particular, $\varphi(n, j)$ and $f_k(n, j)$ are visually quite similar. And $\varphi(n, j)$ relies on combinatorial properties of the primes to count them; maybe $f_k(n, j)$ is doing something similar? Still, $\varphi(n, j)$ is sieving with primes. There's not a prime in sight for $f_k(n, j)$. So this general approach must not be that of $f_k(n, j)$.

Let's turn to another family of prime counting approaches. Now we're entering really challenging territory. In 1859, Bernhard Riemann gave a prime counting formula using logarithmic integrals and the complex zeroes of his soon-to-be very famous zeta function.

$$m=\{1,-1,-1,0,-1,1,-1,...\}$$

$$\pi(n)=\sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{m_k}{k}\left(\frac{\Lambda(\lfloor n^{\frac{1}{k}} \rfloor)}{2\log\lfloor n^{\frac{1}{k}}\rfloor}+li(\lfloor n^{\frac{1}{k}}\rfloor)-\sum_{\rho} li(\lfloor n^{\frac{1}{k}}\rfloor^{\rho})-0.693147...+\int_{\lfloor n^{\frac{1}{k}}\rfloor}^{\infty}\frac{dt}{t(t^2-1)\log t}\right)$$

```
m := {1,-1,-1,0,-1,1,-1,0,0,1,-1,0,-1,1,1,0}
pi[n_] := Sum[ m[[ k ]] / k ( MangoldtLambda[ a = Floor[ n ^ ( 1 / k )]] / Log[ a ] / 2 + LogIntegral[ a ] - N[ 2
Re[ Sum[ ExpIntegralEi[ ZetaZero[ k ] Log[ a ]], { k, 1, 200 }]]] + NIntegrate[ 1 / (( y ^ 3 - y ) Log[ y ]), { y, a, Infinity}] - Log[ 2 ]), { k, 1,
Log[ 2, n ]}]
Table[ { n, pi[ n ]}, { n, 2, 100 }] // TableForm
```

Why this works is beyond our scope, but notice the $\sum_{\rho}$. Each ρ is one of an infinite number of complex numbers of the form $.5+x\,i$ (but only .5 *if* the Riemann Hypothesis is true). Still, $f_k(n, j)$ doesn't use any complex analysis – or any calculus at all – and there are no zeta zeroes ρ to be found. So this isn't how $f_k(n, j)$ works either. Both $f_k(n, j)$ and this formula begin with the same term, $\sum_{k=1}^{\lfloor \log_2 n \rfloor}\frac{m_k}{k}$, though. That warrants an eyebrow raise at the very least.

So why, then, does $f_k(n, j)$ work?
And, for that matter, why does the following count primes, too?

$$g_{0,a}(n)=1 \qquad\qquad m=\{1,-1,-1,0,-1,1,-1,...\}$$

$$g_{k,a}(n)=\sum_{j=0}^{k}\binom{k}{j} g_{k-j,a+1}\left(\frac{n}{a^j}\right) \qquad\qquad \pi(n)=\sum_{k=1}^{\lfloor \log_2 n\rfloor}\sum_{j=1}^{\lfloor \log_2 \lfloor n^{\frac{1}{k}}\rfloor\rfloor}\frac{(-1)^{j-1} m_k}{j\cdot k} g_{j,2}(\lfloor n^{\frac{1}{k}}\rfloor)$$

$$g_{k,a}(n)=0 \text{ if } n<a^k$$

```
g[ n_, k_, a_] := If[ n < a ^ k, 0, Sum[ Binomial[ k, j ] g[ n / a ^ j, k - j, a + 1 ], { j, 0, k }]]; g[ n_, 0, a_ ] := 1
m := {1,-1,-1,0,-1,1,-1,0,0,1,-1,0,-1,1,1,0}
pi[ n_ ] := Sum[ m[[ k ]] / ( j k ) ( -1 ) ^ ( j + 1 ) g[ Floor[ n ^ ( 1 / k )], j, 2 ], { k, Log[ 2, n ]}, { j, Log[ 2, Floor[ n ^ ( 1 / k )]]}];
{ $RecursionLimit = 10000 }; Table[ { n, pi[ n ], pi[ n ] - pi[ n - 1 ]}, { n, 2, 100 }] // TableForm
```

It's even worse than $f_k(n, j)$. It doesn't manually examine numbers to identify primes one by one. It doesn't rely on zeta zeroes or complex analysis. It doesn't sieve by primes, and it doesn't even have a visual parallel to Legendre's formula this time, either. And what's that binomial doing there?

These are leading questions, of course. Both $f_k(n, j)$ and $g_{k,a}(n)$ rely on developments of a certain identity first discovered by Yuri Linnik in the early 1960's. That identity connects specific number-of-divisor functions to a key “is this number prime?” function, but the developments, which are what we'll be excited by here, let us say and do quite a lot more.

Linnik's identity, with development, sits halfway between Legendre's function and Riemann's approach. It's purely combinatorial in nature, so superficially, it produces identities that look like those associated with Legendre's function. But its core idea is deeply related to the connection between $\zeta(s)$ and $\log \zeta(s)$ from Riemann's work, which can easily be used to prove Linnik's identity, in fact.

This paper has two audiences in mind.

One is general readers who like and are inclined towards math, who find the distribution of primes intriguing, but who find confronting the sheer difficulty of analytic number theory daunting. If you're such a reader, most of the results here, largely combinatorial in nature, might be easier to grasp. In particular, if power series from calculus were manageable, you might be surprised to find a host of closely related techniques work with prime counting, too. This document is empirically-oriented and somewhat casual. It's more a survey of interesting mathematical territory, with high-level justifications and explanations along the way, than a rigorous collection of lemmas and proofs. To that end, Mathematica code accompanies nearly formula so you can experiment with the ideas for yourself to try to see how they work.

References

Contents

2

Linnik's Identity

Let's start with Linnik's identity for a single number. It's a simpler first step, even if it won't be our focus. This section introduces divisor-style functions, showing several place where they mirror rules and properties of power series. It also introduces the prime identifying function $[\nabla \zeta(0)]_n$, a strict count of divisor function $[\nabla(\zeta-1)^k]_n$, and then immediately connects the two.

That will be Linnik's identity.

2-1: Our "is this prime?" function

We need to define our prime-identifying function. We'll call it $[\nabla \log \zeta(0)]_n$ (it lacks standard notation). It's defined like so:

$$\boxed{[\nabla \log \zeta(0)]_n = \frac{1}{p} \, if \; n = \text{prime}^{\text{positive integer } p}, 0 \, \text{otherwise}}$$

$[\nabla \log \zeta(0)]_n$ isn't a pure "is this number prime?" function.

For $[\nabla \log \zeta(0)]_{3,} \nabla[\nabla \log \zeta(0)]_{17,} [\nabla \log \zeta(0)]_{31}$, or other primes, its value is 1. For $\kappa(6), \kappa(15), \kappa(42)$, and other composites divisible by at least two primes, its value is 0. That's enough to be very nearly an "is this number prime?" function.

However, given primes to powers, its value is a fraction, not 0. $\kappa(6) = \kappa(15) = \kappa(42) = \frac{1}{2}$. $\kappa(8) = \kappa(27) = \kappa(125) = \frac{1}{3}$. $\kappa(2^{12}) = \frac{1}{12}$. We'll have a method later to deal with this peculiarity. Additionally, $\kappa(1)$ is 0.

So, why this function? Well, much as e just is the natural base for exponentiation, and 2π is the natural unit for trigonometric functions, the plumbing of math just strongly, strongly favors $\kappa(n)$ as the natural "is this number prime?" function.

Mathematica doesn't provide $\kappa(n)$, so for reference, we evaluate it like this:

$$[\nabla \log \zeta(0)]_n = \frac{\Lambda(n)}{\log n}$$

```
PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
Table[{ n, PrimeK[ n ]}, { n, 1, 100}] // TableForm
```

2-2: Counting Divisors

Linnik's identity connects primes to counts of divisor, so we also need divisor counting functions. We'll use two, one that doesn't treat 1 as a divisor, and then one that does.

2-2a: The Strict Count of Divisors Function, $[\nabla(\zeta(0)-1)^k]_n$

First, let's define our strict count of divisors function as

$$\boxed{[\nabla(\zeta(0)-1)^k]_n = \sum_{a_1 \cdot a_2 \cdot \ldots \cdot a_k = n;\, a_i > 1} 1}$$

$[\nabla(\zeta(0)-1)^k]_n$ is the count of solutions to $a_1 \cdot a_2 \cdot \ldots \cdot a_k = n$ where a_i is an integer greater than 1 (excluding 1 makes it strict). So, for example, since 12 can be expressed as $2\cdot 6, 3\cdot 4, 4\cdot 3, \text{and}\, 6\cdot 2$, $d_2{}'(12)=4$. As another example, 30 can be expressed as $2\cdot 3\cdot 5$ with 6 possible combinations (as 3!, in fact), so $d_3{}'(30)=6$.

$[\nabla(\zeta(0)-1)^k]_n$ is easier to calculate using this identity:

$$\boxed{\begin{array}{l} [\nabla(\zeta(0)-1)^k]_n = \sum_{j|n} [\nabla(\zeta(0)-1)^{k-1}]_j \cdot [\nabla(\zeta(0)-1)]_{n\cdot j^{-1}} \\ [\nabla(\zeta(0)-1)]_n = 1 \mathit{\ if\ } n>1, 0 \mathit{\ otherwise} \\ [\nabla(\zeta(0)-1)^0]_n = 1 \mathit{\ if\ } n=1, 0 \mathit{\ otherwise} \end{array}}$$

```
d2[n_,k_] := Sum[ d2[ j, k - 1 ] d2[ n / j, 1 ], { j, Divisors[ n ]}]; d2[n_, 1] := 1; d2[1,1]:=0; d2[ n_, 0] := 0; d2[ 1, 0]:=1
Grid[ Table[ d2[ n, k ], { n, 1, 100 }, { k, 1, 7 }]]
```

Note that this first identity can be generalized further to $[\nabla(\zeta(0)-1)^{a+b}]_n = \sum_{j|n} [\nabla(\zeta(0)-1)^a]_j \cdot [\nabla(\zeta(0)-1)^{*b}]_{n\cdot j^{-1}}$, where $a, b \geq 0$. Note too the similarity here to $x^{a+b} = x^a \cdot x^b$.

2-2b: The Count of Divisors Function, $[\nabla\zeta(0)^k]_n$

That's one count of divisor function we need. The other is the standard number theory count of divisor function, $d_k(n)$ (sometimes referred to as $\tau_k(n)$). It treats 1 as a factor when counting solutions.

$$[\nabla\zeta(0)^k]_n = \sum_{a_1 \cdot a_2 \cdot \ldots \cdot a_k = n} 1 \qquad [\nabla\zeta(0)^k]_n = \sum_{j|n} [\nabla\zeta(0)^{k-1}]_j \qquad [\nabla\zeta(0)]_n = 1 \qquad [\nabla\zeta(0)^0]_n = 1 \mathit{\ if\ } n=1, 0 \mathit{\ otherwise}$$

```
d[n_,k_] := Sum[ d[ j, k - 1 ] d[ n / j, 1 ], { j, Divisors[ n ]}]; d[ n_, 1 ] := 1; d[ n_, 0 ] := 0; d[ 1, 0 ] := 1
Grid[ Table[ d[ n, k ], { n, 1, 100 }, { k, 1, 10 }]]
```

2-2c: Generalizing $[\nabla\zeta(0)^k]_n$

For our purposes, we'll prefer two generalized identities for $[\nabla\zeta(0)^k]_n$ for complex k (rename k to z). The first uses the prime factorization $n=p_1^{a_1}\cdot p_2^{a_2}\cdot\ldots$:

$$\boxed{[\nabla\zeta(0)^z]_n=\prod_{p^a|n}\frac{z(z+1)\ldots(z+a-1)}{a!}}$$

```
d[ n_, k_ ] := Sum[ d[ j, k - 1 ] d[ n / j, 1 ], { j, Divisors[ n ]}]; d[ n_, 1 ] := 1; d[ n_, 0 ] := 0; d[ 1, 0 ] := 1
dAlt[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[ n ]}]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
Grid[ Table[{ d[ n, k ], dAlt[ n, k ]}, { n, 1, 100 }, { k, 1, 10 }]]
```

This identity makes clear that $[\nabla\zeta(0)^k]_n$ is multiplicative: if a and b share no common factors, $[\nabla\zeta(0)^z]_{a\cdot b}=[\nabla\zeta(0)^z]_a\cdot[\nabla\zeta(0)^z]_b$. $[\nabla(\zeta(0)-1)^k]_n$ doesn't share this property.

With this first generalized identity, we can also mention $[\nabla\zeta(0)^{z+w}]_n=\sum_{j|n}[\nabla\zeta(0)^z]_j\cdot[\nabla\zeta(0)^w]_{n\cdot j^{-1}}$, where z and w can be any complex numbers, mirroring $x^{z+w}=x^z\cdot x^w$ even more closely.

2-2d: Computing a generalized $[\nabla\zeta(0)^z]_n$ with $[\nabla(\zeta(0)-1)^k]_n$

The second generalized identity expresses $[\nabla\zeta(0)^z]_n$ as a sum of $[\nabla(\zeta(0)-1)^k]_n$ terms.

$$\boxed{[\nabla\zeta(0)^z]_n=\sum_{a=0}\frac{(z)(z-1)\ldots(z-a+1)}{a!}[\nabla(\zeta(0)-1)^a]_n}$$

```
d[n_,z_]:=Product[Pochhammer[z,a=p[[2]]]/a!,{p,FI[n]}];FI[n_]:=FactorInteger[n];FI[1]:={}
d2[n_,k_]:=Sum[d2[j,k-1] d2[n/j,1],{j,Divisors[n]}];d2[n_,1]:=1;d2[1,1]:=0;d2[n_,0]:=0;d2[1,0]:=1
dAlt[n_,z_]:=Sum[FactorialPower[z,a]/a! d2[n,a],{a,0, Log[2,n]}]
Grid[Table[{d[n,2.2+(-2+.3k)I],dAlt[n,2.2+(-2+.3k)I]},{n,7,100,5},{k,1,8}]]
```

Note the similarity to the Taylor series at 0 for $x^z=\sum_{a=0}\frac{(z)(z-1)\ldots(z-a+1)}{a!}(x-1)^a$. As we'll see, $d_k(n)$ and $d_k{}'(n)$ sometimes relate to each other the way that x^k and $(x-1)^k$ do.

2-3: Connecting $[\nabla\zeta(0)^z]_n$ and $\kappa(n)$

Here's a way to derive $[\nabla\log\zeta(0)]_n$. We'll use the standard count of divisor function, $[\nabla\zeta(0)^z]_n$.

2-3a: Evaluating a limit of $[\nabla\zeta(0)^z]_n$ *at z=0*

Start with our generalized function for $[\nabla\zeta(0)^z]_n$ from 2-2c, $[\nabla\zeta(0)^z]_n=\prod_{p^\alpha|n}\frac{z(z+1)\ldots(z+\alpha-1)}{\alpha!}$.

Now, let's explore this limit: $\lim_{z\to 0} \frac{[\nabla \zeta(0)^z]_n}{z}$.
If n is 1, because an empty product is 1,

$$\lim_{z\to 0} \frac{[\nabla \zeta(0)^z]_n}{z} = \lim_{z\to 0} \frac{1}{z}\prod = \infty$$

If n is p^α, a single prime base to a power, then

$$\lim_{z\to 0} \frac{[\nabla \zeta(0)^z]_n}{z} = \lim_{z\to 0} \frac{1}{z} \frac{z(z+1)\dots(z+\alpha-1)}{\alpha!} = \lim_{z\to 0} \frac{(z+1)\dots(z+\alpha-1)}{\alpha!} = \frac{1}{\alpha}$$

And if n is $p^\alpha \cdot q^\beta \cdot \dots$, with more than 1 prime base,

$$\lim_{z\to 0} \frac{[\nabla \zeta(0)^z]_n}{z} = \lim_{z\to 0} \frac{1}{z} \frac{z(z+1)\dots(z+\alpha-1)}{\alpha!} \cdot \frac{z(z+1)\dots(z+\beta-1)}{\beta!} \cdot \dots =$$
$$\lim_{z\to 0} \frac{(z+1)\dots(z+\alpha-1)}{\alpha!} \cdot \frac{z(z+1)\dots(z+\beta-1)}{\beta!} \cdot \dots = 0$$

2-3b: The Result of That Limit

The results of this limit should ring a bell. In fact, with a touch of term rearrangement,

$$\boxed{[\nabla \log \zeta(0)]_n = \lim_{z\to 0} \frac{[\nabla \zeta(0)^z]_n - [\nabla \zeta(0)^0]_n}{z}}$$

```
PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[ n ]}]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
dlimit[ n_, z_ ] := Round[ N[ 1 / z (d[ n, z ]-d[n,0])], .0000001 ]
Table[{ n, dlimit[ n, 10^(-120) ] ,PrimeK[ n ]}, { n, 2, 100 }] // TableForm
```

Compare this to a slight restatement of the standard log identity, $\log x = \lim_{z\to 0} \frac{x^z - x^0}{z}$. In fact, in several important ways, the relationship between our prime counting function and count of divisors function, $[\nabla \log \zeta(s)]_n$ to $[\nabla \zeta(s)^z]_n$, mirrors that of $\log x$ to x^z.

2-4: Linnik's Identity

Now we can derive Linnik's identity. Start with the generalized expression for $d_k(n)$ from 2-2d.

$$[\nabla \zeta(s)^z]_n = \sum_{k=0} \frac{(z)(z-1)\dots(z-k+1)}{a!} [\nabla (\zeta(s)^k - 1)]_n$$

Now, take the limit of both sides divided by z as z goes to 0.

$$\lim_{z\to 0}\frac{[\nabla\zeta(0)^z]_n}{z}=\lim_{z\to 0}\frac{1}{z}\sum_{a=0}\frac{(z)(z-1)\ldots(z-k+1)}{k!}[\nabla(\zeta(0)-1)^k]_n$$

$$\lim_{z\to 0}\frac{[\nabla\zeta(0)^z]_n}{z}=\lim_{z\to 0}\frac{1}{z}\frac{1}{0!}[\nabla(\zeta(0)-1)^0]_n+\frac{1}{z}\frac{z}{1!}[\nabla\zeta(0)-1]_n+\frac{1}{z}\frac{z(z-1)}{2!}[\nabla(\zeta(0)-1)^2]_n+\frac{1}{z}\frac{z(z-1)(z-2)}{3!}[\nabla(\zeta$$

$$\lim_{z\to 0}\frac{[\nabla\zeta(0)^z]_n}{z}=(\lim_{z\to 0}\frac{1}{z}\frac{1}{0!}[\nabla(\zeta(0)-1)^0]_n)+\frac{1}{1!}[\nabla\zeta(0)-1]_n+\frac{(0-1)}{2!}[\nabla(\zeta(0)-1)^2]_n+\frac{(0-1)(0-2)}{3!}[\nabla(\zeta(0)-1)$$

$$\lim_{z\to 0}\frac{[\nabla\zeta(0)^z]_n-[\nabla\zeta(0)^0]_n}{z}=[\nabla\zeta(0)-1]_n-\frac{1}{2}[\nabla(\zeta(0)-1)^2]_n+\frac{1}{3}[\nabla(\zeta(0)-1)^3]_n-\frac{1}{4}[\nabla(\zeta(0)-1)^4]_n+\ldots$$

This derivation leaves us with Linnik's identity:

$$\boxed{[\nabla\log\zeta(0)]_n=[\nabla\zeta(0)-1]_n-\frac{1}{2}[\nabla(\zeta(0)-1)^2]_n+\frac{1}{3}[\nabla(\zeta(0)-1)^3]_n-\frac{1}{4}[\nabla(\zeta(0)-1)^4]_n+\frac{1}{5}\ldots}$$

```
PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
d2[ n_, k_ ] := Sum[ d2[ j, k - 1 ] d2[ n / j, 1 ], { j, Divisors[ n ]}]; d2[ n_, 1 ] := If[ n < 2, 0, 1 ]
Linnik[ n_ ] := Sum[ (-1)^( k + 1 ) / k d2[ n, k ], { k, 1, Log[ 2, n ]}]
Table[{ n, PrimeK[ n ], Linnik[ n ]}, { n, 2, 100 }] // TableForm
```

Only up to $[\nabla(\zeta(0)-1)^{\lfloor\frac{\log n}{\log 2}\rfloor}]_n$ needs to be computed because $[\nabla(\zeta(0)-1)^k]_n=0$ when $n<2^k$.

If we took the previously mentioned series $x^z=\sum_{a=0}\frac{(z)(z-1)\ldots(z-a+1)}{a!}(x-1)^a$, divided both sides by z, and evaluated its limit as z approached 0, the exact same computation would lead to the well known Taylor series for the logarithm, $\log x=\lim_{z\to 0}\frac{x^z-x^0}{z}=\frac{(x-1)}{1}-\frac{(x-1)^2}{2}+\frac{(x-1)^3}{3}+\ldots$

The take away here is this: if we have $[\nabla(\zeta(0)-1)^k]_n$ for n, we can trivially compute the prime identifying $[\nabla\log\zeta(0)]_n$ too. At first glance, this hardly seems useful, as $[\nabla(\zeta(0)-1)^k]_n$ is a pain to compute as well. In sections 4 and 5, though, when we explore Linnik's identity applied in bulk, we'll see why this idea is so powerful.

Intriguing, I hope. But none of this has anything to say about the distribution of prime numbers. We need to think bigger.

2-5: References, Notes, Further Questions

- Our function $\kappa(n)$ is introduced as (2.9) on p.28 of P. Bateman and H. Diamond's 2004 edition of "Analytic Number Theory – An Introductory Course" in a section covering Linnik's Identity style identities from a more generalized algebraic perspective.

- The generalized equation for $d_k(n)$ from 2-2c can be found as identity (14.122) from p.421 of the 2003 edition of A. Ivic's "The Riemann Zeta-Function: Theory and Practice".

- The generalized equation for $d_k(n)$ *from 2-2d can be found by taking formula (13.34) from p.343 of the 2004 edition of H. Iwaniec and E. Kowalski's "Analytic Number Theory",* $d_k'(n)=\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}d_j(n)$, *inverting it to* $d_k(n)=\sum_{j=0}^{k}\binom{k}{j}d_j(n)$ *then generalizing the binomial with* $\binom{k}{j}=\frac{(k)(k-1)\ldots(k-j+1)}{j!}$ *, essentially copying the mechanics of Newton's generalized binomial theorem.*

- Linnik's identity, as given in 2-4 here, can be found as identity (13.32) from p. 343 of the 2004 edition of H. Iwaniec and E. Kowalski's "Analytic Number Theory". (17.1) from p.346 of the 2010 edition of J. Friedlander and H. Iwaniec's "Opera de Cribro" provides similar coverage.

3

Summing Linnik's Identity

In the last section, we encountered Linnik's identity. It expressed the prime identifying function $\kappa(n)$ in terms of the strict count of divisors function $d_k'(n)$ in a form that looked uncannily like the Taylor series for log x. Those identities didn't say anything about the distribution of prime numbers, though.

Now, we'll change that by looking at the relationship between $\kappa(n)$ and $d_k'(n)$ for ranges of numbers.

3-1: The Riemann Prime Counting Function

Let's introduce the Riemann Prime Counting Function, which we'll write as $[\log\zeta(0)]_n$ (it lacks standard notation, also referred to as $J(n)$, $\Pi(n)$, $\pi^*(n)$, and $\pi_1(n)$). It can be expressed as

$$\boxed{[\log\zeta(0)]_n = \sum_{j=1}^{n} [\nabla\log\zeta(0)]_j}$$

```
PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
RiePrimeCnt[ n_ ] := Sum[ PrimeK[ j ], { j, 2, n }]
Table[{ n, N[ RiePrimeCnt[ n ]]}, { n, 1, 100 }] // TableForm
```

It first appeared, as far as I know, in Riemann's epoch-making 1859 paper unveiling his Zeta function. In fact, the paper's end goal, and the development of the Zeta function and its zeroes, were all in the service of providing an explicit formula for $[\log\zeta(0)]_n$ (which Riemann called f(x)). The Zeta function, its zeroes, and the associated techniques have gone on to be much, much more broadly important, of course.

3-1a: $[\log\zeta(0)]_n$ in terms of the number of primes, $\pi(n)$

Fortunately, the connection between $[\log\zeta(0)]_n$ and the number of primes less than n, $\pi(n)$, is easily computed. $[\log\zeta(0)]_n$ in terms of $\pi(n)$ is

$$\boxed{\begin{array}{l}\pi(n) = \text{number of primes} \le n \\ [\log\zeta(0)]_n = \sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{1}{k}\pi(n^{\frac{1}{k}})\end{array}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimeK[ j ], { j, 1, n}]; PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
RiePrimeCntAlt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 /k )] / k, { k, 1, Log[ 2, n ]}]
Table[{ n, RiePrimeCntAlt[ n ], RiePrimeCnt[ n ]}, { n, 1, 100}] // TableForm
```

The sum stops at $\log_2 n$ because $\pi(1)=0$.

3-1b: The number of primes, $\pi(n)$ *, in terms of* $[\log\zeta(0)]_n$

Mathematica doesn't implement $[\log\zeta(0)]_n$, but $\pi(n)$ is PrimePi[n], so this identity is our reference way to compute $[\log\zeta(0)]_n$. More importantly, this identity is easily inverted through Moebius inversion, so $\pi(n)$ in terms of $[\log\zeta(0)]_n$ is

$$\boxed{\pi(n) = \sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{\mu(k)}{k}[\log\zeta(0)]_{n^{\frac{1}{k}}}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimeK[ j ], { j, 1, n }]; PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
PrimePiAlt[ n_ ] := Sum[ MoebiusMu[ k ] RiePrimeCnt[ n ^ ( 1 / k )] / k, { k, 1, Log[ 2, n ]}]
Table[{ n, PrimePi[ n ], PrimePiAlt[ n ]}, { n, 1, 100 }] // TableForm
```

Moebius inversion is a standard combinatorial technique, so you can read about it elsewhere at your leisure.

(As a quick sketch, though, note that $\mu(n)$ is equal to $[\nabla\zeta(0)^{-1}]_n$ - in Mathematica, try

```
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[ n ] }]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
Table[{ n, d[ n, -1 ], MoebiusMu[ n ]}, { n, 1, 100 }] // TableForm
```

and our two sums are $[\log\zeta(0)]_n = \sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{1}{k}[\nabla\zeta(0)]_k \cdot \pi(n^{\frac{1}{k}})$ and $\pi(n) = \sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{1}{k}[\nabla\zeta(0)^{-1}]_k \cdot [\log\zeta(0)]_{n^{\frac{1}{k}}}$.)

And so with that Moebius inversion, we finally accommodate $\kappa(n)$'s counting of prime powers, not primes.

Here's the really important point. If we can compute $[\log\zeta(0)]_n$ quickly or approximate it well, we can trivially do likewise for $\pi(n)$, the count of primes, through a simple Moebius inversion.

Turning to $[\log\zeta(0)]_n$ (and thus $\pi(n)$) let's us talk about the distribution of the primes.

3-2: The Divisor Summatory Functions

Let's introduce two more functions, the sum-of-number-of-divisors $[\zeta(0)^k]_n$ and the strict sum-of-number-of-divisors, $[(\zeta(s)-1)^k]_n$. They can be expressed as

$$[(\zeta(0)-1)^k]_n=\sum_{j=1}^{n}[\nabla(\zeta(0)-1)^k]_n \qquad [\zeta(0)^k]_n=\sum_{j=1}^{n}[\nabla\zeta(0)^k]_n$$

3-2a: $D_k(n)$ and $D_k'(n)$ recursively

For computation and approximation, the following identities are often much, much more useful.

$$D_k'(n)=\sum_{j=2}^{n}D_{k-1}'(\lfloor\frac{n}{j}\rfloor) \qquad [\zeta(0)^k]_n=\sum_{j=1}^{n}[\zeta(0)^{k-1}]_{n\,j^{-1}}$$
$$D_0'(n)=1 \qquad [\zeta(0)^0]_n=1$$

```
d2[ n_, k_ ] := Sum[ d2[ j, k - 1 ] d2[ n / j, 1 ], { j, Divisors[ n ]}]; d2[ n_, 1] := 1; d2[ 1, 1 ] := 0; d2[ n_, 0 ] := 0; d2[ 1, 0 ] := 1
D2[ n_, k_ ] := Sum[ d2[ j, k ], { j, 2, n }]
D2Alt[ n_, k_ ] := Sum[ D2Alt[ Floor[ n / j ], k - 1 ], { j, 2, n }]; D2Alt[ n_, 0 ] := 1
Grid[ Table[{ D2[ n, k ], D2Alt[ n, k ]}, { n, 7, 100, 5 }, { k, 1, 8 }]]
```

```
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[n] }]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
DD[ n_, k_ ] := Sum[ d[ j, k ], { j, 1, n}]
DDAlt[ n_, k_ ] := Sum[ DDAlt[ Floor[ n / j ], k - 1 ], { j, 1, n }]; DDAlt[ n_, 0 ] := 1
Grid[ Table[{ DD[ n, k], DDAlt[ n, k]}, { n, 7, 100, 5}, { k, 1, 8 }]]
```

These are both special cases of the much more general

$$D_k'(n)=\sum_{j=2}^{n}d_a'(j)\,D_{k-a}'(\lfloor\frac{n}{j}\rfloor) \qquad D_k(n)=\sum_{j=1}^{n}d_a(j)\,D_{k-a}(\lfloor\frac{n}{j}\rfloor)$$

It's worth noting, once again, the parallel here to standard exponent behavior, $x^k=x^a\cdot x^{k-a}$.

Getting away from recursive notation for a moment, in terms of arithmetic, this means

$$D_1'(n)=\sum_{j=2}^{n}1 \qquad D_2'(n)=\sum_{j=2}^{n}\sum_{k=2}^{\lfloor\frac{n}{j}\rfloor}1 \qquad D_3'(n)=\sum_{j=2}^{n}\sum_{k=2}^{\lfloor\frac{n}{j}\rfloor}\sum_{m=2}^{\lfloor\frac{n}{jk}\rfloor}1$$
$$D_1(n)=\sum_{j=1}^{n}1 \qquad D_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor\frac{n}{j}\rfloor}1 \qquad D_3(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor\frac{n}{j}\rfloor}\sum_{m=1}^{\lfloor\frac{n}{jk}\rfloor}1$$

and so on.

3-2b: $D_k(n)$ and $D_k'(n)$ expressed in terms of each other

The two functions can be interchanged like so:

$$D_k'(n)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}D_{k-j}(n) \qquad D_k(n)=\sum_{j=0}^{k}\binom{k}{j}D_{k-j}'(n)$$

```
D2[ n_, k_ ] := Sum[ D2[ Floor[ n / j ], k - 1 ], { j, 2, n }]; D2[ n_, 0 ] := 1
```

```
DD[ n_, k_ ] := Sum[ DD[ Floor[ n / j ], k - 1 ], { j, 1, n }]; DD[ n_, 0 ] := 1
D2Alt[ n_, k_ ] := Sum[ ( -1 ) ^ j Binomial[ k, j ] DD[ n, k-j ], { j, 0, k }]
DDAlt[ n_, k_ ] := Sum[ Binomial[ k, j ] D2[ n, k-j ], { j, 0, k }]
Grid[ Table[{ D2[ n, k ], D2Alt[ n, k ]}, { n, 7, 100, 5 }, { k, 1, 8 }]]
Grid[ Table[{ DD[ n, k ], DDAlt[ n, k ]}, { n, 7, 100, 5 }, { k, 1, 8 }]]
```

Much as with $d_k(n)$ and $d_k'(n)$, this can be generalized to

$$D_z(n)=\sum_{a=0}^{\log_2 n}\frac{(z)(z-1)\ldots(z-a+1)}{a!}D_a'(n)$$

```
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[n] }]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
DD[ n_, k_ ] := Sum[ d[ j, k ], { j, 1, n}]
D2[ n_, k_ ] := Sum[ D2[ Floor[ n / j ], k - 1 ], { j, 2, n }]; D2[ n_, 0 ] := 1
DDAlt[n_,z_]:=Sum[FactorialPower[z,a]/a! D2[n,a],{a,0, Log[2,n]}]
Grid[Table[{DD[n,2.2+(-2+.3k)I],DDAlt[n,2.2+(-2+.3k)I]},{n,7,100,5},{k,1,8}]]
```

It's worth taking a moment to comment on this identity. The sum's limit is $\log_2 n$ because $D_k'(n)=0$ if $n<2^k$. Now, computing $D_k(n)$ for *n > 10^15* or so and for larger k is a difficult task. $D_k(n)$ is an unruly, jittery, unpredictable function with a lot of high frequency information. So it's surprising that if you've computed $\log_2 n$ terms of $D_k'(n)$, you can exactly compute $D_z(n)$ for any complex value of *z* nearly as easily as you would a polynomial with $\log_2 n$ terms. And because $D_k'(n)=0$ if $n<2^k$, that's with cleanly guaranteed convergence. More surprising still, $D_k'(n)$ isn't special - there are other functions, each with their own charms, similar to $D_k'(n)$ that, if computed, will quickly yield $D_z(n)$ for any complex *z*. It's another consequence of the parallel to power series mechanics here. We'll discuss that in chapter 6.

So, now we've got a summed prime counting function, $\Pi(n)$. And we've got a summed divisor function, $D_k'(n)$. Let's connect them.

3-3: Linnik's Identity Summed

With our new notation, we can write Linnik's identity summed.

$$\boxed{\Pi(n)=D_1'(n)-\frac{1}{2}D_2'(n)+\frac{1}{3}D_3'(n)-\frac{1}{4}D_4'(n)+\frac{1}{5}\ldots}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j ) ] / j, { j, 1, Log[ 2, n ]}]
D2[ n_, k_ ] := Sum[ D2[ Floor[ n / j ], k - 1 ], { j, 2, n }]; D2[ n_, 0 ] := 1
LinnikSum[ n_ ] := Sum[ ( -1 ) ^ ( k + 1 ) / k  D2[ n, k ], { k, 1, Log[ 2, n ]}]
Table[{ n, RiePrimeCnt[ n ], LinnikSum[ n ]}, { n, 1, 100 }] // TableForm
```

Swapping to summation notation here might seem trivial, barely worthy of note, but in fact it's very significant, really the central point in this entire paper.

3-3a: Why Linnik's Identity Summed Matters

Specifically, it's very significant because reasoning about this function $D_k'(n)$ opens up many, many more doors than reasoning about the sum $\sum_{j=2}^{n} d_k'(j)$. $D_k'(n)$, as an idea, is "the count of solutions to $a_1 \cdot a_2 \cdot \ldots a_k \leq n$ with each $a >= 2$." That idea lends itself to many different approaches. In fact, the heart of this paper, in sections 4 and 5, catalogs the results of exploring some of those approaches.

3-3b: Another Parallel to log x

As an aside, let's pause for a moment and take a look at the familiar power series for the logarithm, $\log x = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \frac{1}{5.}...$. Expand it, and we have

$$\log x = (x^1 - x^0) - \frac{1}{2}(x - 2x^1 + x^0) + \frac{1}{3}(x^3 - 3x^2 + 3x^1 - x^0)$$
$$\frac{-1}{4}(x^4 - 4x^3 + 6x^2 - 4x^1 + x^0) + \frac{1}{5.}..$$

Now, let's look at Linnik's identity summed in terms of $D_k(n)$, not $D_k'(n)$. Section 3-2 mentioned that $D_k'(n) = \sum_{j=0}^{k} (-1)^{(k-j)} \binom{k}{j} D_j(n)$. Apply that to Linnik's identity summed, and we have

$$\Pi(n) = (D_1(n) - D_0(n)) - \frac{1}{2}(D_2(n) - 2D_1(n) + D_0(n)) + \frac{1}{3}(D_3(n) - 3D_2(n) + 3D_1(n) - D_0(n)) -$$
$$\frac{1}{4}(D_4(n) - 4D_3(n) + 6D_2(n) - 4D_1(n) + D_0(n)) + \frac{1}{5}...$$

Once again, that's a strong parallel to the logarithm power series.

3-4: D and $\Pi(n)$ mirroring x and log x

Here's another parallel to the relationship between x and log x. Section 2-3 noted that $\kappa(n) = \lim_{z \to 0} \frac{d_z(n) - d_0(n)}{z}$. Sum this from 1 to n, use our $\Pi(n)$ and $D_k(n)$ notation, and subtract 1 to prevent $d_0(1)/0$ from going to infinity, and we also have

$$\boxed{\Pi(n) = \lim_{z \to 0} \frac{D_z(n) - 1}{z}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[ n ]}]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ]:={}
DD[ n_, z_ ] := Sum[ d[ j, z ], { j, 1, n }]
Dlimit[ n_, z_ ] := ( DD[ n, z ] - 1 ) / z
```

```
Table[{ n, N[ RiePrimeCnt[ n ]], N[ Dlimit[ n, 10 ^ ( -120 ) ]]}, { n, 1, 100 }] // TableForm
```

This also bears a tidy resemblance to the familiar $\log x = \lim_{z \to 0} \frac{x^z - 1}{z}$.

3-5: References, Notes, Further Questions

- Riemann's prime counting function $\Pi(n)$ *, with a slight modification, is covered as the function J, starting in section 1.11 on pp.22-33 of the 2000 edition of H. M. Edwards' "Riemann's Zeta Function".*

- Mobius Inversion is an extremely common combinatorial technique. It's covered on p.236 of the 2000 edition (5th) of G.H. Hardy and E. M. Wright's "An Introduction to the Theory of Numbers".

- The specific mechanics of using Moebius inversion to recover $\pi(n)$ *from* $\Pi(n)$ *is detailed as a footnote in p.34 of the 2000 edition of H. M. Edwards' "Riemann's Zeta Function".*

4

Counting the Distribution of Primes with Linnik's Identity Summed

By this point, I hope I've convinced you that $\Pi(n)$, $D_k(n)$, and $D_k'(n)$ are all deeply connected.

Now we're finally ready to use those connections to talk about the distribution of the primes. In this section, we'll show three ways to compute Riemann's Prime Counting Function $\Pi(n)$, based on different approaches to computing $D_k'(n)$. We'll end with a mention of wheels, a way to speed up prime counting, too.

4-1: The Simplest Way to Compute

Here's the first approach to evaluating $\Pi(n)$, giving slow-to-compute but extremely concise expressions for $\Pi(n)$.

In section 3-3, we had this identity:

$$\Pi(n) = D_1'(n) - \frac{1}{2} D_2'(n) + \frac{1}{3} D_3'(n) - \frac{1}{4} D_4'(n) + \frac{1}{5} \ldots$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
D2[ n_, k_ ] := Sum[ D2[ Floor[ n / j ], k - 1 ], { j, 2, n }]; DD[ n_, 0 ] := 1
LinnikSum[ n_ ] := Sum[ ( -1 ) ^ ( k +1 ) / k D2[ n, k ], { k, 1, Log[ 2, n ]}]
Table[{ n, RiePrimeCnt[ n ], LinnikSum[ n ]}, { n, 1, 100 }] // TableForm
```

4-1a: As Nested Sums

Given that $D_k'(n) = \sum_{j=2}^{n} D_{k-1}'(\lfloor \frac{n}{j} \rfloor)$ and $D_0'(n) = 1$, we can rewrite this more concretely to make absolutely explicit the heart of the computation.

$$\Pi(n) = \sum_{j=2}^{n} 1 - \frac{1}{2} \sum_{j=2}^{n} \sum_{k=2}^{\lfloor \frac{n}{j} \rfloor} 1 + \frac{1}{3} \sum_{j=2}^{n} \sum_{k=2}^{\lfloor \frac{n}{j} \rfloor} \sum_{l=2}^{\lfloor \frac{n}{j \cdot k} \rfloor} 1 - \frac{1}{4} \sum_{j=2}^{n} \sum_{k=2}^{\lfloor \frac{n}{j} \rfloor} \sum_{l=2}^{\lfloor \frac{n}{j \cdot k} \rfloor} \sum_{m=2}^{\lfloor \frac{n}{j \cdot k \cdot l} \rfloor} 1 + \frac{1}{5} \ldots$$

4-1b: A Recursive Function

There's recursive structure in there. Let's pull it out.

$$\boxed{\begin{gathered}\Pi_k(n)=\sum_{j=2}^{n}\frac{1}{k}-\Pi_{k+1}(\lfloor\frac{n}{j}\rfloor)\\ \Pi(n)=\Pi_1(n)\end{gathered}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
PI[ n_, k_] := Sum[ 1 / k - PI[ Floor[ n / j ], k + 1 ], { j, 2, n}]
Table[{ n, RiePrimeCnt[n], PI[ n, 1 ]}, { n, 1, 100 }] // TableForm
```

4-1c: Another Recursive Function

One final step: we can rewrite the sum in 4-1b as a fully recursive definition.

$$\boxed{\begin{gathered}\Pi_k(n,j)=\frac{1}{k}-\Pi_{k+1}(\lfloor\frac{n}{j}\rfloor,\lfloor\frac{n}{j}\rfloor)+\Pi_k(n,j-1)\\ \Pi_k(n,1)=0\\ \Pi(n)=\Pi_1(n,n)\end{gathered}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
PI[ n_, j_, k_ ] := 1 / k - PI[ Floor[ n / j ], Floor[ n / j ], k + 1 ] + PI[ n, j - 1, k ]; PI[ n_, 1, k_ ] := 0
Table[{ n, RiePrimeCnt[ n ], PI[ n, n, 1 ]}, { n, 1, 100 }] // TableForm
```

And so finally we recover $f_k(n,j)$ from the introduction, which is $\Pi_k(n,j)$ here. None of these identities are fast to compute, but they all have the advantage of relative clarity.

4-2: The Hyperbola Method

That's a good start, but $D_k'(n)$ and $D_k(n)$ have other combinatorial properties for much faster computation. Let's turn to those. An example will show one such property in action.

4-2a: A Worked Example: Counting Permutations for $D_3'(n)$

Look at $D_3'(n)$, the count of solutions to $a\cdot b\cdot c\leq n; a,b,c\geq 2$. In previous sections, we counted these solutions one by one with $D_3'(n)=\sum_{a=2}^{n}\sum_{b=2}^{\lfloor\frac{n}{a}\rfloor}\sum_{c=2}^{\lfloor\frac{n}{ab}\rfloor}1$.

But look closer at the actual computation here. If *n* >= 24, we end up manually counting the set of numbers $2\cdot3\cdot4$ 6 times as $\{2\cdot3\cdot4, 2\cdot4\cdot3, 3\cdot2\cdot4, 3\cdot4\cdot2, 4\cdot2\cdot3, 4\cdot3\cdot2\}$, This is a basic permutation problem, $\binom{3}{1}\binom{2}{1}\binom{1}{1}=6$. If *n* >= 12, we count the set $2\cdot2\cdot3$ 3 times, as $\binom{3}{1}\binom{1}{1}=3$, another a basic permutation problem.

We previously computed $D_3'(n)$ by simply counting the answers to $a \cdot b \cdot c \le n; a, b, c \ge 2$ one by one. Now let's count these permutations instead. We need a way to avoid counting a permutation more than once, so here's a new constraint for counting: $a \le b \le c$ Considering both the constraints $a \le b \le c$ and $a \cdot b \cdot c \le n$, clearly $a \le \lfloor n^{\frac{1}{3}} \rfloor$ and $b \le \lfloor (\frac{n}{a})^{\frac{1}{2}} \rfloor$. Also, since the number of permutations is affected by whether $a = b$ or $a < b$, and by whether $b = c$ or $b < c$, let's split our counts into four categories.

$$D_3'(n) = \binom{3}{1}\binom{2}{1}\binom{1}{1} \#\{a \cdot b \cdot c \le n; a, b, c \ge 2; a < b < c\} + \binom{3}{1}\binom{2}{2} \#\{a \cdot b \cdot c \le n; a, b, c \ge 2; a < b = c\} + \binom{3}{2}\binom{1}{1} \#\{a \cdot b \cdot c \le n; a, b, c \ge 2; a = b < c\} + \binom{3}{3} \#\{a \cdot b \cdot c \le n; a, b, c \ge 2; a = b = c\}$$

$$D_3'(n) = \binom{3}{1}\binom{2}{1}\binom{1}{1} \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \sum_{b=a+1}^{\lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor} \sum_{c=b+1}^{\lfloor \frac{n}{a\,b} \rfloor} 1 + \binom{3}{1}\binom{2}{2} \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \sum_{b=a+1}^{\lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor} \sum_{c=b}^{b} 1 + \binom{3}{2}\binom{1}{1} \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \sum_{b=a}^{a} \sum_{c=b+1}^{\lfloor \frac{n}{a\,b} \rfloor} 1 + \binom{3}{3} \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \sum_{b=a}^{a} \sum_{c=b}^{b} 1$$

```
D2k3[ n_ ] := Sum[ 1, { a, 2, n }, { b, 2, n / a }, { c, 2, n / (a b) }]
D2k3Alt[ n_ ] := 6 Sum[ 1, { a, 2, n ^ ( 1 / 3 )}, { b, a + 1, Floor[ n / a ] ^ ( 1 / 2 )}, { c, b + 1, Floor[ n / (a b)]}]+
3 Sum[ 1, { a, 2, n ^ ( 1 / 3 )}, { b, a + 1, Floor[ n / a ] ^ ( 1 / 2 )}, { c, b, b }]+
3 Sum[ 1, { a, 2, n ^ ( 1 / 3 )}, { b, a, a }, { c, b + 1, n / (a b )}]+
Sum[ 1, { a, 2, n ^ ( 1 / 3 )}, { b, a, a }, { c, b, b}]
Table[{ n, D2k3[ n ], D2k3Alt[ n ]}, { n, 10, 1000, 10 }] // TableForm
```

That right-hand identity simplifies to

$$D_3'(n) = \left(6 \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \sum_{b=a+1}^{\lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor} \lfloor \frac{n}{a\,b} \rfloor - b\right) + \left(3 \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor - a\right) + \left(3 \sum_{a=2}^{\lfloor n^{\frac{1}{3}} \rfloor} \left(\lfloor \frac{n}{a^2} \rfloor - a\right)\right) + \left(\lfloor n^{\frac{1}{3}} \rfloor - 1\right)$$

This process can produce identities for any $D_k'(n)$, resulting in 2^{k-1} terms (here, *k=3* led to $2^{3-1} = 4$ terms to count). If counting started with $a = 1$ rather than $a = 2$, $D_k(n)$ could be computed this way too. For larger *k*, though, this gets complicated.

4-2b: The Formula for Counting Permutations to Compute $D_{k,a}(n)$

Fortunately, a really concise recursive formula exists expressing this idea for any *k*. One point about notation for this recursive formula, though, which generalizes $D_k'(n)$ and $D_k(n)$: it uses another parameter, *a*, as the smallest allowed factor, giving us $D_{k,a}(n) = \#\{n_1 \cdot n_2 \cdot \ldots n_k \le n; n_i \ge a\}$. In this notation, $D_k'(n) = D_{k,2}(n)$ and $D_k(n) = D_{k,1}(n)$.

That formula is this:

$$D_{k,a}(n)=\sum_{j=0}^{k}\binom{k}{j}D_{k-j,a+1}\left(\frac{n}{a^j}\right) \text{ with } D_{0,a}(n)=1 \text{ and } D_{k,a}(n)=0 \text{ when } n<a^k$$

```
D2[ n_, k_ ] := Sum[ D2[ Floor[ n / j ], k - 1 ], { j, 2, n }]; D2[ n_, 0 ] := 1
DD[ n_, k_ ] := Sum[ DD[ Floor[ n / j ], k - 1 ], { j, 1, n }]; DD[ n_, 0 ] := 1
Dhyp[ n_, k_, a_ ] := If[ n < a ^ k, 0, Sum[ Binomial[ k, j ] Dhyp[ n / a ^ j, k - j, a + 1 ], { j, 0, k }]]; Dhyp[ n_, 0, a_ ] := 1
Grid[ Table[{ D2[ n, k ], Dhyp[ n, k, 2 ]}, { n, 7, 100, 5 }, { k, 1, 8 }]]
Grid[ Table[{ DD[ n, k ], Dhyp[ n, k, 1 ]}, { n, 7, 100, 5 }, { k, 1, 8 }]]
```

This is $g_{k,a}(n)$ from the introduction. $D_k(n)=\sum_{j=0}^{k}\binom{k}{j}D_{k-j}'(n)$ from 3-2b is also a specific case of this identity.

4-2c: Improving the Performance of $D_{k,a}(n)$

Two improvements speed this identity up for computational purposes.

First, work through the arithmetic for $D_{1,a}(n)$, and clearly $D_{1,a}(n)=\lfloor n\rfloor-a+1$. Let's take note of that.

Second, if we take advantage of the condition $D_{k,a}(n)=0$ when $n<a^k$, we can replace most recursion here with iteration. We split $D_{k,a}(n)$ into $D_{k,a}(n)=D_{k,a+1}(n)+\sum_{j=1}^{k}\binom{k}{j}D_{k-j,a+1}\left(\frac{n}{a^j}\right)$, then replace $D_{k,a+1}(n)$ with $D_{k,a+2}(n)+\sum_{j=1}^{k}\binom{k}{j}D_{k-j,a+2}\left(\frac{n}{(a+1)^j}\right)$. Repeat this until $a^k>n$, at which point $D_{k,a}(n)=0$, and we will find we have $D_{k,a}(n)=\sum_{j=1}^{k}\binom{k}{j}\sum_{m=a}^{\lfloor n^{\frac{1}{k}}\rfloor}D_{k-j,m+1}\left(\frac{n}{m^j}\right)$.

4-2d: Using $D_{k,a}(n)$ to Compute $\Pi(n)$

So let's use this new $D_{k,a}(n)$ formula to compute the Riemann Prime Counting Function via Linnik's identity summed.

$$\boxed{\begin{gathered} D_{k,a}(n)=\sum_{j=1}^{k}\binom{k}{j}\sum_{m=a}^{\lfloor n^{\frac{1}{k}}\rfloor}D_{k-j,m+1}\left(\frac{n}{m^j}\right) \\ D_{1,a}(n)=\lfloor n\rfloor-a+1 \\ D_{0,a}(n)=1 \\ \Pi(n)=D_{1,2}(n)-\frac{1}{2}D_{2,2}(n)+\frac{1}{3}D_{3,2}(n)-\frac{1}{4}D_{4,2}(n)+\frac{1}{5}\ldots \end{gathered}}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
Dhyp[ n_, k_, a_ ] := Sum[ Binomial[ k, j ] Dhyp[ n / ( m ^ ( k - j )), j, m + 1 ], { m, a, n ^ ( 1 / k )}, { j, 0, k - 1 }]
Dhyp[ n_, 1, a_ ] := Floor[ n ] - a + 1; Dhyp[ n_, 0, a_ ] := 1
LinnikSumHyp[ n_ ] := Sum[ ( -1 ) ^ ( k + 1 ) / k Dhyp[ n, k, 2 ], { k, 1, Log[ 2, n ]}]
Table[{ n, RiePrimeCnt[ n ], LinnikSumHyp[ n ]}, { n, 1, 100}] // TableForm
```

By empirical evidence, this approach seems to compute $\Pi(n)$ (and thus $\pi(n)$, the number of primes) a bit faster than $O(n)$ and uses $O(\log n)$ memory.

4-2c: Connection to the Dirichlet Hyperbola Method

This idea, counting permutations instead of individual solutions, is exactly the approach used by the well-known Dirichlet hyperbola method. By that method,

$$D_2(n) = -\lfloor n^{\frac{1}{2}} \rfloor^2 + 2 \sum_{a=1}^{\lfloor n^{\frac{1}{2}} \rfloor} \lfloor \frac{n}{a} \rfloor$$

$$D_3(n) = \lfloor n^{\frac{1}{3}} \rfloor^3 + 3 \sum_{a=1}^{\lfloor n^{\frac{1}{3}} \rfloor} \lfloor \frac{n}{a^2} \rfloor - \lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor^2 + 2 \sum_{b=a+1}^{\lfloor \lfloor \frac{n}{a} \rfloor^{\frac{1}{2}} \rfloor} \lfloor \frac{n}{a\, b} \rfloor$$

```
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[2]]] / a!, { p, FI[ n ]}]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
DD[ n_, k_ ] := Sum[ d[ j, k ], { j, 1, n }]
Dk2Alt[ n_ ] := -Floor[ n ^ ( 1 / 2 )] ^ 2 + 2 Sum[ Floor[ n / m], { m, 1, Floor[ n ^ ( 1 / 2 )]}]
Dk3Alt[ n_ ] := Floor[ n ^ ( 1 / 3 )] ^ 3 + 3 Sum[ Floor[ n / ( j ^ 2 ) ] - Floor[ Floor[ n / j ] ^ ( 1 / 2 )] ^ 2 + 2 Sum[ Floor[ n / k / j ],{ k, j + 1,
Floor[ Floor[ n / j] ^ ( 1 / 2 )]}], { j, 1, Floor[ n ^ ( 1 / 3 )]}]
Table[{ DD[ n, 2 ], Dk2Alt[ n ]}, { n, 1, 100 }] // TableForm
Table[{ DD[ n, 3 ], Dk3Alt[ n ]}, { n, 1, 100 }] // TableForm
```

and more general identities for $D_k(n)$ are possible.

The connection between $D_{k,a}(n)$ and the Dirichlet hyperbola method might seem obscure, but if you begin with $D_{2,1}(n) = \sum_{j=1}^{2} \binom{2}{j} \sum_{m=1}^{\lfloor n^{\frac{1}{2}} \rfloor} D_{2-j,m+1}(\frac{n}{m^j})$ and manually swap in the values for $D_{1,a}(n) = \lfloor n \rfloor - a + 1$ and $D_{0,a}(n) = 1$, you arrive (after some extra simplifying) at $D_{2,1}(n) = -\lfloor n^{\frac{1}{2}} \rfloor^2 + 2 \sum_{j=1}^{\lfloor n^{\frac{1}{2}} \rfloor} \lfloor \frac{n}{j} \rfloor$. Likewise for $D_{3,1}(n) = \sum_{j=1}^{3} \binom{3}{j} \sum_{m=1}^{\lfloor n^{\frac{1}{3}} \rfloor} D_{3-j,m+1}(\frac{n}{m^j})$, though it's more work.

4-3: A Faster Combinatorial Method

4-3a: Another Combinatorial Identity for $D_{k,a}(n)$

Let's keep using our generalized divisor sum from 4-2, $D_{k,a}(n) = \#\{ n_1 \cdot n_2 \cdot \ldots n_k \le n ; n_i \ge a \}$, with $D_k(n) = D_{k,1}$ and $D_k'(n) = D_{k,2}$, and a similar divisor function, $d_{k,a}(n) = \#\{ n_1 \cdot n_2 \cdot \ldots n_k = n ; n_i \ge a \}$.

For *k >= 2*, $D_{k,a}(n)$ has another useful combinatorial identity. It needs one more parameter, *t,* with $a < t < n$.

$$D_{k,a,t}(n)= \sum_{j=t+1}^{n} d_{1,a}(j)\, D_{k-1,a}\left(\frac{n}{j}\right) + \sum_{j=a}^{t} d_{k-1,a}(j)\, D_{1,a}\left(\frac{n}{j}\right) + \sum_{j=a}^{t} \sum_{s=\lfloor\frac{t}{j}\rfloor+1}^{\lfloor\frac{n}{j}\rfloor} \sum_{m=1}^{k-2} d_{1,a}(s)\, d_{m,a}(j)\, D_{k-m-1,a}\left(\frac{n}{js}\right)$$

```
Dhyp[ n_, k_, a_ ] := Sum[ Binomial[ k, j ] Dhyp[ n / ( m ^ ( k - j )), j, m + 1 ], { m, a, n ^ ( 1 / k )}, { j, 0, k - 1 }]
Dhyp[ n_, 1, a_ ] := If[ n < a, 0, Floor[ n ] - a + 1 ]; Dhyp[ n_, 0, a_ ] := 1
dhyp[ n_, k_, a_ ] := dhyp[ n, k, a ] = Dhyp[ n, k, a ] - Dhyp[ n - 1, k, a ]
D2Alt[ n_, k_, a_, t_ ] := Sum[ dhyp[ j, 1, a ] Dhyp[ Floor[ n / j ], k - 1, a ], { j, t + 1, n }] + Sum[ dhyp[ j, k - 1, a ] Dhyp[ Floor[ n / j ], 1, a ],{ j,
a, t }] + Sum[ dhyp[ s, 1, a ] dhyp[ j, m, a ] Dhyp[ Floor[ n / (j s) ], k - m - 1, a ], { j, a, t }, { s, Floor[ t / j ] + 1, Floor[ n / j ]}, { m, 1, k - 2 }]
Table[ Grid[ Table[{ Dhyp[ n, k, a ], D2Alt[ n, k, a, Floor[ n ^ ( 1 / 3 )]], D2Alt[ n, k, a, Floor[ n ^ ( 1 / 2 )]]}, { n, 7, 300, 21}, { k, 2, 6 }]], { a, 1,
4 }]// TableForm
```

This identity's utility takes some explanation.

We want to compute $D_{k,2}(n)$ as quickly as possible so that, through Linnik's identity summed, we will compute $\Pi(n)$ as quickly as possible, and all with a reasonable memory bound, too.

For computational purposes, $d_{1,a}(n)$ and $D_{1,a}(n)$ are special because they can be computed instantly for any value of *n*. $d_{1,a}(n)=1\, if\ n \geq a,\, 0\, otherwise$ and $D_{1,a}(n)=\lfloor n\rfloor - a+1$.

Look closely at $D_{k,a,t}(n)$, and you'll notice that $D_{k,a}(\frac{n}{t})$ is the largest argument for $D_{k,a}$, and $d_{k,a}(t)$ is the largest argument for $d_{k,a}$, except in the cases of $d_{1,a}(n)$ and $D_{1,a}(n)$, which can be computed instantly for any *n*. An example should make this more clear. If we pick $t=n^{\frac{1}{2}}$, and we had a precomputed table with values of $D_{k,a}(n)$ up to $n^{\frac{1}{2}}$ and $d_{k,a}(n)$ up to $n^{\frac{1}{2}}$, computing $D_{k,a,t}(n)$ would reduce to performing the basic arithmetic of those sums. We would be able to look up any reference to $D_{k,a}(n)$ or $d_{k,a}(n)$, except for $d_{1,a}(n)$ and $D_{1,a}(n)$, which we would compute instantly.

Let's get more specific about the case we're concerned with, where *a=2*.

Our arithmetic can be reduced with $\sum_{j=2}^{n} f(\lfloor\frac{n}{j}\rfloor)=\sum_{j=2}^{n^{\frac{1}{2}}} f(\lfloor\frac{n}{j}\rfloor)+\sum_{j=1}^{\lfloor\frac{n}{n^{\frac{1}{2}}}\rfloor-1} (\lfloor\frac{n}{j}\rfloor-\lfloor\frac{n}{j+1}\rfloor) f(j)$, a consequence of division in the floor function.

$$\Pi(n)=D_1'(n)+\sum_{j=\lfloor n^{\frac{1}{3}}\rfloor+1}^{\lfloor n^{\frac{1}{2}}\rfloor}\sum_{k=2}^{\lfloor\log_2 n\rfloor}\frac{-1^{k+1}}{k}D_{k-1}'(\lfloor\frac{n}{j}\rfloor)+\sum_{j=1}^{\lfloor n^{\frac{1}{2}}\rfloor}(D_1'(\lfloor\frac{n}{j}\rfloor)-D_1'(\lfloor\frac{n}{j+1}\rfloor))\sum_{k=2}^{\lfloor\log_2 n\rfloor}\frac{-1^{k+1}}{k}D_{k-1}'(j)$$
$$+\sum_{j=2}^{\lfloor n^{\frac{1}{3}}\rfloor}\sum_{k=2}^{\lfloor\log_2 n\rfloor}\frac{-1^{k+1}}{k}d_{k-1}'(j)D_1'(\lfloor\frac{n}{j}\rfloor)+\sum_{j=2}^{\lfloor n^{\frac{1}{3}}\rfloor}\sum_{s=\lfloor\frac{\lfloor n^{\frac{1}{3}}\rfloor}{j}\rfloor+1}^{\lfloor\frac{n}{j}\rfloor}\sum_{k=2}^{\lfloor\log_2 n\rfloor}\frac{-1^{k+1}}{k}\sum_{m=1}^{k-2}d_m'(j)D_{k-m-1}'(\lfloor\frac{n}{js}\rfloor)$$
$$+\sum_{j=2}^{\lfloor n^{\frac{1}{3}}\rfloor}\sum_{s=1}^{\lfloor\lfloor\frac{n}{j}\rfloor^{\frac{1}{2}}\rfloor-1}(D_1'(\lfloor\frac{n}{js}\rfloor)-D_1'(\lfloor\frac{n}{j(s+1)}\rfloor))\cdot\sum_{k=2}^{\lfloor\log_2 n\rfloor}\frac{-1^{k+1}}{k}\sum_{m=1}^{k-2}d_m'(j)D_{k-m-1}'(s)$$

```
D2Cache[ n_, k_, s_ ] := Sum[ Binomial[ k, j ] D2Cache[ n / ( m ^ ( k - j )), j, m + 1 ], { m, s, n ^ ( 1 / k )}, { j, 0, k - 1 }]
D2Cache[ n_, 1, s_ ] := Floor[ n ] - s + 1; D2Cache[ n_, 0, s_ ] := 1
d2cache[ n_, k_ ] := D2Cache[ n, k, 2 ] - D2Cache[ n - 1, k, 2 ]

D2Fast[ n_, k_ ] :=
Sum[ D2Cache[ n / j, k - 1, 2 ], { j, Floor[ n ^ ( 1 / 3 )] + 1, n ^ ( 1 / 2 )}] + Sum[( Floor[ n / j ] - ( Floor[ n / ( j + 1 )])) D2Cache[ j, k - 1, 2 ],{ j,
1, n / Floor[ n ^ ( 1 / 2 )] - 1}] + Sum[ d2cache[ j , k - 1 ] ( Floor[ n / j ] - 1 ), { j, 2, n ^ ( 1 / 3 )}] + Sum[ d2cache[ j, m ] D2Cache[ n / ( j  s ),
k - m - 1, 2 ], { j, 2, n ^ ( 1 / 3 )}, { s, Floor[ Floor[ n ^ ( 1 / 3 )] / j ] + 1, Floor[ n / j ] ^ ( 1 / 2 )}, { m, 1, k - 2 }] + Sum[ (Sum[ 1, { m, Floor[ n / (
j ( s + 1 ))] + 1, n / ( j s )}]) ( Sum[ d2cache[ j, m ]  D2Cache[ s, k - m - 1, 2 ], { m, 1, k - 2 }]), { j, 2, n ^ ( 1 / 3 )}, { s, 1, Floor[ n / j ] /
Floor[ Floor[ n / j ] ^ ( 1 / 2 )] - 1 }]
D2Fast[ n_, 1 ] := Floor[ n ] - 1

LinnikSumFast[ n_ ] := Sum[ ( -1 ) ^ ( k + 1 ) / k D2Fast[ n, k ], { k, 1, Log[ 2, n ]}]
RiePrimeCnt[ n_ ]  := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
Table[{ n, LinnikSumFast[ n ], RiePrimeCnt[ n ]}, { n, 1, 100 }] // TableForm
```

4-4: Wheels

$$\varphi(x,a)=\varphi(x,a-1)-\varphi(\frac{x}{prime_a},a-1)$$
$$\varphi(x,0)=\lfloor x\rfloor$$

$$D_k'(n)=\sum_{j=2;\,2,3,\ldots p_a\nmid j}^{n}D_{k-1}'(\lfloor\frac{n}{j}\rfloor)$$
$$D_1'(n)=\varphi(n \bmod m,a)+\varphi(m,a)\lfloor n/m\rfloor \text{ where } m=(2\cdot 3\cdot\ldots p_a)$$
$$D_0'(n)=1$$

```
(* SETTING UP THE WHEEL *)

(* Wheel Initialization *)
WheelEntries:=WheelEntries= 5
WheelSize:=WheelSize=Product[Prime[j],{j,1,WheelEntries}]
ExcludedPrimes[n_]:=Sum[1/k,{j,1,WheelEntries},{k,1,Log[Prime[j],n]}]

(* Use[n] is 1 if n isn't rejected by the wheel, and 0 otherwise *)
CoprimeCache:=CoprimeCache=Table[CoprimeQ[WheelSize,n],{n,1,WheelSize}]
```

```
Use[n_]:=Use[n]=If[ CoprimeCache[[Mod[n-1,WheelSize]+1]]==True,1,0]

(* Coprimes[n] is the count of numbers <= n coprime to our wheel *)
LegendrePhi[x_,a_]:=LegendrePhi[x,a-1]-LegendrePhi[x/Prime[a],a-1];LegendrePhi[x_,0]:=Floor[x]
LegPhiCache:=LegPhiCache=Table[LegendrePhi[n,WheelEntries],{n,1,WheelSize}]
FullWheel:=FullWheel= LegendrePhi[WheelSize,WheelEntries]
Coprimes[n_]:=Coprimes[n]=LegPhiCache[[Mod[n-1,WheelSize]+1]]+Floor[(n-1)/WheelSize]FullWheel

(* VARIATIONS OF LINNIK'S IDENTITY SUMMED *)

(* Linnik's Identity Summed with the Default Method *)
DD[n_,k_]:=Sum[If[Use[j]==0,0,DD[Floor[n/j],k-1]],{j,2,n}]
DD[n_,1]:=DD[n,1]=Coprimes[n]-1
DD[n_,0]:=DD[n,0]=1
LinnikSum[n_]:=Sum[(-1)^(k+1)/k DD[n,k],{k,1,Log[2,n]}]

(* Linnik's Identity Summed Recursively *)
PI[n_,k_]:=PI[n,k]=Sum[If[Use[j]==0,0,1/k-PI[Floor[n/j],k+1]],{j,2,n}]

(* Linnik's Identity Summed with the Hyperbola Method *)
Dhyp[n_,k_,s_]:=Sum[If[Use[m]==0,0,Binomial[k,j]Dhyp[Floor[n/(m^(k-j))],j,m+1]],{m,s,n^(1/k)},{j,0,k-1}]
Dhyp[n_,1,s_]:=Dhyp[n,1,s]=Coprimes[n]-Coprimes[s-1]
Dhyp[n_,0,s_] := Dhyp[n,0,s]=1
LinnikSumHyp[n_]:=Sum[(-1)^(k+1)/k Dhyp[n,k,2],{k,1,Log[2,n]}]

(* The Test Bed *)
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
Table[{n,RiePrimeCnt[n],LinnikSum[n]+ExcludedPrimes[n],PI[n,1]+ExcludedPrimes[n],LinnikSumHyp[n]+ExcludedPrimes[n]},
{n,1,100,1}]//TableForm
```

WHEELS ARE CURRENTLY BUSTED.

4-5: References, Notes, Further Questions

- Comparison to Legendre identities

- Hyperbola Method Write Up?

- Mertens function: http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.em/1047565447 Marc Deléglise and Joöl Rivat "Computing the Summation of the Mobius Function"

- Wheels?
- Other prime counting approaches.

5

Estimating the Distribution of Primes with Linnik's Identity Summed

Since antiquity we've known the primes are limitless. It's taken longer to resolve questions about their frequency among natural numbers.

In the decades around 1800, Legendre speculated that $\pi(x) \approx \frac{x}{\log x + C}$, and Gauss, in turn, that $\pi(x) \approx \int_0^x \frac{dt}{\log t}$, which is the logarithmic integral $li(x)$. These good estimates were empirical, based on tables of prime counts, meaning they provided no particular insight into the difference between the approximations and $\pi(x)$.

Riemann, in his 1859 paper, massively advanced knowledge about the entire topic. He showed that $li(x)$ approximated a slightly different prime counting function, $\Pi(n)$. And he gave an incredible exact formula for the difference between $li(x)$ and $\Pi(n)$. That equation relies directly on the so called non-trivial zeroes of the Riemann Zeta function and leads immediately to the importance of the Riemann Hypothesis. The deep nature of those non-trivial zeroes directly impacts the difference between $li(x)$ and $\Pi(n)$.

If that last paragraph sounds like gibberish to you, hang on! Riemann's work, and all the analytic number theory that followed it, is amazing and beautiful and a supreme achievement. But it's also hugely complex and demanding, requiring deep knowledge of complex analysis to follow.

This section provides a different route to reasoning about the difference between $li(x)$ and $\Pi(n)$, one that relies only on combinatorial arguments and some real-valued calculus.

Our plan of attack will be this: we'll approximate $D_k'(n)$ and then, using Linnik's identity summed and those approximations, also approximate $\Pi(n)$, Riemann's prime counting function.

5-1: Approximating $D_k'(n)$

Here's one way to approximate $D_k'(n)$. Take this earlier definition $D_k'(n) = \sum_{j=2}^{n} D_{k-1}'(\lfloor \frac{n}{j} \rfloor),\ D_0'(n) = 1$, and write out the sums explicitly for different values of k

$$D_1'(n) = \sum_{j=2}^{n} 1 \quad D_2'(n) = \sum_{j=2}^{n} \sum_{k=2}^{\lfloor \frac{n}{j} \rfloor} 1 \quad D_3'(n) \sum_{j=2}^{n} \sum_{k=2}^{\lfloor \frac{n}{j} \rfloor} \sum_{l=2}^{\lfloor \frac{n}{j \cdot k} \rfloor} 1$$

and so on.

5-1a: The Smooth Volume Bounding $D_k'(n)$

Now, let's replace those discrete sums with continuous integrals. Remove the floor function on the upper bounds and change the lower bounds from 2 to 1, too.

$$D_1'(n) \approx \int_1^n dx \qquad D_2'(n) \approx \int_1^n \int_1^{\frac{n}{x}} dy\, dx \qquad D_3'(n) \approx \int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\, dy\, dx$$

```
Dif1[ n_ ] := Integrate[ 1, { x, 1, n }] - Sum[ 1, { j, 2, n }]
Dif2[ n_ ]  := Integrate[ 1, { x, 1, n }, { y, 1, n / x }] - Sum[ 1, { j, 2, n }, { k, 2, n / j }]
Dif3[ n_ ]  := Integrate[ 1, { x, 1, n }, { y, 1, n / x }, { z, 1, n / (x y)}] - Sum[ 1, { j, 2, n }, { k, 2, n / j }, { l, 2, n / (j k)}]
Dif4[ n_ ]  := Integrate[ 1, { x, 1, n }, { y, 1, n / x }, { z, 1, n / (x y)}, { w, 1, n / (x y z)}]
    - Sum[ 1, { j, 2, n }, { k, 2, n / j }, { l, 2, n / (j k)}, { m, 2, n / ( j k l)}]
Table[{ n, N[ Dif1[ n ]], N[ Dif2[ n ]], N[ Dif3[ n ]], N[Dif4[ n ]]}, { n, 2, 10 }] // TableForm
```

This approximation is not at all tight. On the other hand, every lattice square / cube / hypercube / etc in $D_k'(n)$ is entirely bounded by these curves, and no lattice squares / cubes / hypercubes / etc not in $D_k'(n)$ are entirely bounded by these curves.

5-1b: Evaluating the Integrals for the Approximation of $D_k'(n)$

Here's a method for evaluating these integrals. $D_1'(n) \approx \int_1^n dx = n-1$, obviously. The integral in that identity is the inner integral of $D_2'(n) \approx \int_1^n \int_1^{\frac{n}{x}} dy\, dx$, with n/x replacing x, so $D_2'(n) \approx \int_1^n \int_1^{\frac{n}{x}} dy\, dx = \int_1^n \frac{n}{x} - 1\, dx$. Then integrate for $D_2'(n) \approx \int_1^n \int_1^{\frac{n}{x}} dy\, dx = n \log n - n + 1$. But the double integral of that identity is the inner double integral of $D_3'(n) \approx \int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\, dy\, dx$, with n/x replacing n, so $D_3'(n) \approx \int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\, dy\, dx = \int_1^n \frac{n}{x} \log \frac{n}{x} - \frac{n}{x} + 1\, dx$. Integrate term by term to get $D_3'(n) \approx \int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\, dy\, dx = \frac{1}{2} n (\log n)^2 - n \log n + n - 1$. This approach is easy to repeat. The general pattern here is

$$\boxed{D_j'(n) \approx (-1)^j \left(1 - n \sum_{k=0}^{j-1} \frac{(-\log n)^k}{k!}\right)}$$

```
D2Approx[ j_ ] := Expand[ ( -1 ) ^ j( 1 - n Sum[(-Log[ n ]) ^ k / ( k! ), { k, 0, j - 1 }])]
{ D2Approx[ 1 ],  Integrate[ 1, { x, 1, n }]}
{ D2Approx[ 2 ], Expand[ Integrate[ 1, { x, 1, n }, { y, 1, n / x }]]}
{ D2Approx[ 3 ], Expand[ Integrate[ 1, { x, 1, n }, { y, 1, n / x }, { z, 1, n / (x y)}]]}
{ D2Approx[ 4 ], Expand[ Integrate[ 1, { x, 1, n }, { y, 1, n / x }, { z, 1, n / (x y)}, { w, 1, n / (x y z)}]]}
```

5-1c: Generalizing the Approximation of $D_k'(n)$

This approximation can be made to handle non-integer values for *j*. The upper incomplete gamma function $\Gamma(j,n)$ can be written as $\Gamma(j,n)=\Gamma(j)e^{-n}\sum_{k=0}^{j-1}\frac{n^k}{k!}$ and thus $\Gamma(j,-\log n)=\Gamma(j)n\sum_{k=0}^{j-1}\frac{(-\log n)^k}{k!}$ So, we can extend our approximation to

$$\boxed{D_z'(n)\approx(-1)^z(1-\frac{\Gamma(z,-\log n)}{\Gamma(z)})}$$

```
D2Approx[ n_, j_ ] := ( -1 ) ^ j ( 1 - n Sum[(-Log[ n ]) ^ k / ( k! ), { k, 0, j - 1 }])
D2ApproxAlt[ n_, j_ ] := ( -1 ) ^ j ( 1 - ( Gamma[ j, -Log[ n ]]) / Gamma[ j ])
Grid[ Table[ {N[ D2Approx[ n, k ]], N[ Round[ D2ApproxAlt[ n, k ], 10 ^ -120 ]]}, { n, 7, 1000, 41 }, { k, 0, 8 }]]
```

5-2: Approximating $\Pi(n)$

Now let's see what our approximations for $D_k'(n)$ are good for.
Section 3-4 gave this identity for the Riemann Prime Counting Function.

$$\Pi(n)=\lim_{z\to 0}\frac{(D_z(n)-1)}{z}$$

Can we approximate $\Pi(n)$ by approximating $D_0(n)$ somehow? We know $D_k(n)=\sum_{j=0}^{k}\binom{k}{j}D_j'(n)$ from 3-2, so $D_0(n)=D_0'(n)$. So we might see interesting results if we approximate $\Pi(n)=\lim_{z\to 0}\frac{(D_z'(n)-1)}{z}$.

Let's use $D_z'(n)\approx(-1)^z(1-\frac{\Gamma(z,-\log n)}{\Gamma(z)})$ from 5-1 for $D_0'(n)$ in our limit, giving us

$$\Pi(n)\approx\lim_{z\to 0}\frac{1}{z}((-1)^z-\frac{(-1)^z\Gamma(z,-\log n)}{\Gamma(z)}-1)=\lim_{z\to 0}-\frac{\Gamma(0,-\log n)}{z\cdot\Gamma(z)}=\lim_{z\to 0}-\frac{\Gamma(0,-\log n)}{\Gamma(z+1)}=$$

$$\boxed{\Pi(n)\approx-\Gamma(0,-\log n)}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
```

```
LinnikApprox[ n_ ] := -Gamma[ 0, -Log[ n ]]
Table[{ n, a = N[ RiePrimeCnt[ 10 ^ n ]], b = N[ LinnikApprox[ 10 ^ n ]], a - b }, { n, 1, 12 }] // TableForm
```

which looks like a pretty good approximation. In fact, given that the exponential integral is connected to the incomplete gamma function by $\Gamma(0,x)=-Ei(-x)-\pi i$ when x < 0, and the logarithmic integral is connected to the exponential integral via $li(z)=Ei(\log z)$, that leads immediately to

$$\boxed{\Pi(n)\approx li(n)+\pi i}$$

```
Table[{a=10^n,N[-Gamma[0,-Log[a]]],N[ExpIntegralEi[Log[a]]+Pi I],N[LogIntegral[a]+Pi I]},{n,1,12}]//TableForm
```

5-3: Approximating $\Pi(n)$ Using Linnik's Identity Summed

Let's try another way to approximate $\Pi(n)$. Start with Linnik's Identity Summed:

$$\Pi(n)=D_1'(n)-\frac{1}{2}D_2'(n)+\frac{1}{3}D_3'(n)-\frac{1}{4}D_4'(n)+\frac{1}{5}...$$

Now insert our approximation from 5-1, $D_j'(n)\approx(-1)^j(1-n\sum_{k=0}^{j-1}\frac{(-\log n)^k}{k!})$, giving

$$\boxed{\Pi(n)\approx\sum_{j=1}^{\infty}\frac{(-1)^{(j+1)}}{j}D_j'(n)=\sum_{j=1}^{\infty}\frac{(-1)^{(j+1)}}{j}(-1)^j(1-n\sum_{k=0}^{j-1}\frac{(-\log n)^k}{k!})}$$

Simplify powers of -1:

$$\Pi(n)\approx-\sum_{j=1}^{\infty}\frac{1}{j}(1-n\sum_{k=0}^{j-1}\frac{(-\log n)^k}{k!})$$

Now, here are two known identities for the lower incomplete gamma function, $\gamma(n)$. $\gamma(j,x)=(j-1)!(1-e^{-x}\sum_{k=0}^{j-1}\frac{x^k}{k!})$ and $\gamma(j,x)=\int_0^x t^{j-1}e^{-t}dt$. If we say $x=-\log n$, we have

$$\Pi(n)\approx-\sum_{j=1}^{\infty}\frac{1}{j}\cdot\frac{1}{(j-1)!}\gamma(j,-\log n)$$

and then

$$\Pi(n)\approx-\sum_{j=1}^{\infty}\frac{1}{j!}\int_0^{-\log n}t^{j-1}e^{-t}dt$$

Swap the sum and integral for

$$\Pi(n)\approx-\int_0^{-\log n}\sum_{j=1}^{\infty}(\frac{t^{j-1}}{j!})e^{-t}dt$$

which is

$$\Pi(n) \approx -\int_0^{-\log n} \left(\frac{e^t - 1}{t}\right) e^{-t}\, dt$$

Multiply to get

$$\Pi(n) \approx -\int_0^{-\log n} \frac{1 - e^{-t}}{t}\, dt$$

With a bit of sweat, simplify out minus signs

$$\Pi(n) \approx \int_0^{\log n} \frac{e^t - 1}{t}\, dt$$

Now, the exponential integral can be expressed as $Ei(x) = \int_0^x \frac{e^t - 1}{t}\, dt + \log x + .577215\ldots$, and the logarithmic integral is connected to the exponential integral via $li(x) = Ei(\log x)$, leaving us with

$$\boxed{\Pi(n) \approx li(n) - \log\log n - .577215\ldots}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
LinnikApprox[ n_ ] := Sum[ ( -1 ) ^ ( j + 1 ) / j ( -1 ) ^ j ( 1 - n Sum[ (-Log[ n ]) ^ k / k!, { k, 0, j – 1 }]), { j, 1, Infinity }]
Table[ { 10 ^ n, a = N[ RiePrimeCnt[ 10 ^ n ]], b = N[ LinnikApprox[ 10 ^ n ]], N[ LogIntegral[ a ] - Log[ Log[ a ]] - EulerGamma] }, { n, 1,
12 }] // TableForm
```

5-4: More on Approximating $\Pi(n)$ Using Linnik's Identity Summed

5-4a: Generalizing

The introduction mentioned that Riemann provided an explicit formula for his prime counting function, which is

$$\Pi(n) = \frac{1}{2}\kappa(n) + li(n) - \sum_{\rho} li(n^{\rho}) - 0.693147\ldots + \int_x^{\infty} \frac{dt}{t(t^2-1)\log t}$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j )] / j, { j, 1, Log[ 2, n ]}]
RieExplicitForumla[ x_, t_ ] := MangoldtLambda[x]/2/Log[x] +LogIntegral[ x ] - N[ 2Re[ Sum[ ExpIntegralEi[ ZetaZero[ k ] Log[ x ]], { k, 1,
t }]]] + NIntegrate[ 1 / (( y ^ 3 - y ) Log[ y ]), { y, x, Infinity }] - Log[ 2 ]
Table[ { n, N[ RiePrimeCnt[ n ]], RieExplicitForumla[ n, 200 ]}, { n, 2, 100 }] // TableForm
```

where ρ are the zeroes of his zeta function and $li(n)$ is the logarithmic integral. As mentioned, this formula not only expresses the count of primes, it makes the difference between $\Pi(n)$ and $li(n)$ explicit, with properties of those zeta zeros giving estimates for how big the difference is.

5-4b: Generalizing

Our approach here is not so powerful in that regard. Nevertheless, using our formula for $\Pi(n)$ from 4-1a and our formula for $li(n)$ from 5-3, we see we have the picturesque

$$\Pi(n)= \sum_{j=2}^{n} 1 - \frac{1}{2}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor \frac{n}{j}\rfloor} 1 + \frac{1}{3}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor \frac{n}{j}\rfloor}\sum_{l=2}^{\lfloor \frac{n}{j\cdot k}\rfloor} 1 - \frac{1}{4}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor \frac{n}{j}\rfloor}\sum_{l=2}^{\lfloor \frac{n}{j\cdot k}\rfloor}\sum_{m=2}^{\lfloor \frac{n}{j\cdot k\cdot l}\rfloor} 1+...$$

$$li(n)= \log\log n+.577215...+\int_1^n dx-\frac{1}{2}\int_1^n\int_1^{\frac{n}{x}} dy\,dx+\frac{1}{3}\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\frac{1}{4}\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx+...$$

From this perspective, $\Pi(n)$ and $li(n)$ bear roughly the same relationship to each other, in an elaborate sense, that $\lfloor n\rfloor$ and n do.

$$li(n)-\Pi(n)$$

$$= \sum_{\rho} li(n^{\rho})+\frac{1}{2}\kappa(n)+0.693147...-\int_x^{\infty}\frac{dt}{t(t^2-1)\log t}$$

$$= \log\log n+.577215...$$
$$-\frac{1}{2}\left(\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor \frac{n}{x}\rfloor} 1\right)$$
$$+\frac{1}{3}\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor \frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor \frac{n}{xy}\rfloor} 1\right)$$
$$-\frac{1}{4}\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor \frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor \frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor \frac{n}{xyz}\rfloor} 1\right)$$
$$+\frac{1}{5}...$$

```
RiePrimeCnt[n_]:=Sum[PrimePi[n^(1/j)]/j,{j,1,Log[2,n]}]
RiemannError[x_,t_]:=-MangoldtLambda[x]/2/Log[x]+N[2Re[Sum[ExpIntegralEi[ZetaZero[k] Log[x]],{k,1,t}]]]-NIntegrate[1/((y^3-y)
Log[y]),{y,x,Infinity}]+Log[2]
D2Approx[n_,j_]:=(-1)^j (1-n Sum[(-Log[n])^k/(k!),{k,0,j-1}])
Dhyp[n_,k_,a_]:=Sum[Binomial[k,j] Dhyp[n/(m^(k-j)),j,m+1],{m,a,n^(1/k)},{j,0,k-1}]
Dhyp[n_,1,a_]:=If[n<a,0,Floor[n]-a+1];Dhyp[n_,0,a_]:=1
dhyp[n_,k_,a_]:=dhyp[n,k,a]=Dhyp[n,k,a]-Dhyp[n-1,k,a]
LinnikError[n_]:=Log[Log[n]]+EulerGamma+Sum[(-1)^(k+1)/k (D2Approx[n,k]-Dhyp[n,k,2]),{k,2,100}]
Table[{n,N[LogIntegral[n]-RiePrimeCnt[n]],N[RiemannError[n,200]], N[LinnikError[n]]},{n,2,100}]//TableForm
```

5-4c: Generalizing

$$
\begin{aligned}
li(n) - \Pi(n) = \\
&\log\log n + .577215\ldots \\
&-\frac{1}{2}\left(n\log n - n + 1 - \sum_{x=1}^{n} d_2(x) + 2\sum_{x=1}^{n} d_1(x) - \sum_{x=1}^{n} d_0(x)\right) \\
&+\frac{1}{3}\left(\frac{n}{2}(\log n)^2 - n\log n + n - 1 - \sum_{x=1}^{n} d_3(x) + 3\sum_{x=1}^{n} d_2(x) - 3\sum_{x=1}^{n} d_1(x) + \sum_{x=1}^{n} d_0(x)\right) \\
&-\frac{1}{4}\left(\frac{n}{6}(\log n)^3 - \frac{n}{2}(\log n)^2 + n\log n - n + 1 - \sum_{x=1}^{n} d_4(x) + 4\sum_{x=1}^{n} d_3(x) - 6\sum_{x=1}^{n} d_2(x) + 4\sum_{x=1}^{n} d_1(x) - \sum_{x=1}^{n} d_0(x)\right) \\
&+\frac{1}{5}\ldots
\end{aligned}
$$

5-5: References, Notes, Further Questions

- Reference for history of PNT

- On Riemann's formula

- Cauchy Principle Value of D_k' through residues leading experimentally to li. Connection to the Dirichlet Divisor Problem.

- Voronoi summation?

- Gamma and incomplete gamma

6

Identities like Linnik's Identity

L

6-1: Linnik's Identity Scaled

$a \log x = \log x^a$

$$\boxed{a\,\Pi(n) = \lim_{z \to 0} \frac{D_{a \cdot z}(n) - 1}{z}}$$

$$\boxed{a\,\Pi(n) = (D_a(n) - D_0(n)) - \frac{1}{2}(D_{2a}(n) - 2D_a(n) + D_0(n)) + \frac{1}{3}(D_{3a}(n) - 3D_{2a}(n) + 3D_a(n) - D_0(n)) - \frac{1}{4}\ldots}$$

```
RiemannPrimeCounting[n_] := Sum[ PrimePi[ n^(1/j)]/j,{j,1,N[Log[n]/Log[2]]}]
d[n_,z_]:=Product[1/(p[[2]]!) Pochhammer[z,p[[2]]],{p,FI[n]}]; FI[n_]:=If[n==1,{},FactorInteger[n]]
DD[n_,k_]:=Sum[d[j,k],{j,1,n}]
LinniksIdentityExpanded[n_, a_ ] := Sum[ (-1)^(k+1)/k Sum[(-1)^(k-j) Binomial[k,j] DD[n,a j],{j,0,k}],{k,1,N[Log[n]/Log[2]]}]
Table[{ n, N[RiemannPrimeCounting[n]], N[LinniksIdentityExpanded[n,1]]},{n,1,100}] // TableForm
```

6-2: Inverting Linnik's Identity and D

$$\kappa_k(n) = \sum_{a_1 \cdot a_2 \cdot \ldots \cdot a_k = n} \kappa(a_1) \cdot \kappa(a_2) \cdot \ldots \cdot \kappa(a_k)$$

$$\kappa_0(n) = 1 \; if \; n = 1, 0 \; otherwise$$

$$d_a(n) = \frac{a^0}{0!}\kappa_0(n) + \frac{a^1}{1!}\kappa_1(n) + \frac{a^2}{2!}\kappa_2(n) + \frac{a^3}{3!}\kappa_3(n) + \frac{a^4}{4!}\kappa_4(n) + \frac{a^5}{5!}\ldots$$

```
PrimeKappa[n_,0]:=If[n==1,1,0]
PrimeKappa[n_,1]:=If[n==1,0,FullSimplify[MangoldtLambda[n]/Log[n]]]
PrimeKappa[n_,k_]:=Sum[PrimeKappa[j,k-1] PrimeKappa[n/j,1],{j,Divisors[n]}]
dv2[n_,k_]:=Sum[k^j/(j!) PrimeKappa[n,j],{j,0,N[Log[n]/Log[2]]}]

FactInteger[n_]:=If[n==1,{},FactorInteger[n]]
d[n_,z_]:=Product[1/(p[[2]]!) Pochhammer[z,p[[2]]],{p,FactInteger[n]}]
Grid[Table[{d[n,k],dv2[n,1,k]},{n,7,100,5},{k,1,8}]]
Grid[Table[{d[n,2.2+(-2+.3k)I],dv2[n,1,2.2+(-2+.3k)I]},{n,7,100,5},{k,1,8}]]
```

$$n-1=\sum_{j=2}^{n}\kappa(j)+\frac{1}{2}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor\frac{n}{j}\rfloor}\kappa(j)\kappa(k)+\frac{1}{6}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor\frac{n}{j}\rfloor}\sum_{l=2}^{\lfloor\frac{n}{j\cdot k}\rfloor}\kappa(j)\kappa(k)\kappa(l)$$
$$+\frac{1}{24}\sum_{j=2}^{n}\sum_{k=2}^{\lfloor\frac{n}{j}\rfloor}\sum_{l=2}^{\lfloor\frac{n}{j\cdot k}\rfloor}\sum_{m=2}^{\lfloor\frac{n}{j\cdot k\cdot l}\rfloor}\kappa(j)\kappa(k)\kappa(l)\kappa(m)+\frac{1}{120}\ldots$$

$$F_k(n)=\sum_{j=2}^{n}\kappa(j)\left(\frac{1}{k!}+F_{k+1}\left(\lfloor\frac{n}{j}\rfloor\right)\right)$$
$$n-1=F_1(n)$$

```
Kappa[n_] := N[MangoldtLambda[n]/Log[n]]
F[n_, k_ ] := Sum[ Kappa[j]( 1/(k!) + F[Floor[n/j],k+1]),{j,2,n}]
F[n_] := F[n,1]

Kappa[n_]:=FullSimplify[MangoldtLambda[n]/Log[n]]
dv2[n_,k_,a_]:=a/k( Kappa[n]+Sum[If[ j==1 || j==n,0,Kappa[j] ( dv2[n/j,k+1,a])],{j,Divisors[n]}])

FactInteger[n_]:=If[n==1,{},FactorInteger[n]]
d[n_,z_]:=Product[1/(p[[2]]!) Pochhammer[z,p[[2]]],{p,FactInteger[n]}]
Grid[Table[{d[n,k],dv2[n,1,k]},{n,7,100,5},{k,1,8}]]
Grid[Table[{d[n,2.2+(-2+.3k)I],dv2[n,1,2.2+(-2+.3k)I]},{n,7,100,5},{k,1,8}]]
```

```
Kappa[n_]:=FullSimplify[MangoldtLambda[n]/Log[n]]
NewD[n_, k_,a_ ] := Sum[ Kappa[j]a /k(1+NewD[ Floor[n/j],k+1,a]),{j,2,n}]
TestNewD[n_,a_] := 1+NewD[n,1,a]

FactInteger[n_]:=If[n1,{},FactorInteger[n]]
d[n_,z_]:=Product[1/(p[[2]]!) Pochhammer[z,p[[2]]],{p,FactInteger[n]}]
DD[n_,k_]:=Sum[d[j,k],{j,1,n}]
Grid[Table[{DD[n,k],TestNewD[n,k]},{n,7,100,5},{k,1,8}]]
Grid[Table[{DD[n,2.2+(-2+.3k)I],TestNewD[n,2.2+(-2+.3k)I]},{n,7,100,5},{k,1,8}]]
```

6-3: D from Other Bases

$$((x-1)+1)^a=x^a$$
$$e^{a\log x}=x^a$$

$$\left(\frac{1}{\left(\frac{1}{x-1}\right)+1}\right)^{a}=x^{a}$$

6-4: Linnik's Identities for Other Arithmetic Functions

```
PrimeK[n_] := If[n==1,0,FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
d[ f_, n_, k_ ] := Sum[ d[ f, j, k - 1 ] d[ f, n / j, 1 ], { j, Divisors[ n ]}]; d[ f_, n_, 1 ] := f[ n ]; d[ f_, n_, 0 ] := 0; d[ f_, 1, 0 ] := 1
d2[ f_, n_, k_ ] := Sum[ ( -1 ) ^ ( k - j ) Binomial[ k, j ] d[ f, n, j ], { j, 0, k }]
Linnik[ f_, n_ ] := Sum[ ( -1 ) ^ ( k + 1 ) / k d2[ f, n, k ], { k, 1, Log[ 2, n ]}]

MuSquared[ n_ ] := MoebiusMu[ n ] ^ 2
Divisor1[ n_ ] := DivisorSigma[ 1, n ]
PrimeExp[ n_, z_ ] := Product[ z ^ p[[2]] / (p[[2]]!), { p, FI[ n ]}]; FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
PrimeExp1[ n_ ] := PrimeExp[ n, 1 ]

Table[{ n,
  Linnik[ MoebiusMu, n ],        -PrimeK[ n ]," ",
  Linnik[ LiouvilleLambda, n ], PrimeK[ n ] LiouvilleLambda[ n ]," ",
  Linnik[ MuSquared, n ],        -PrimeK[ n ] LiouvilleLambda[ n ]," ",
  Linnik[ Divisor1, n ],         PrimeK[ n ]( n + 1 )," ",
  Linnik[ EulerPhi, n ],         PrimeK[ n ]( n - 1 )," ",
  Linnik[ PrimeExp1, n ],        If[ PrimeQ[ n ], 1, 0 ]," "
  }, { n, 2, 100 }] // TableForm
```

6-5: References, Notes, Further Questions

- *General power series reference*

- *Bateman reference?*

7

Linnik's Identity Summed and Alternating Series

Another

Here's what we're doing in this section.

First, we'll define an alternating sign count of divisors function, $E_k(n)$.

Because we want to say interesting things about $D_k(n)$, specifically the Riemann prime counting function $\lim_{z \to 0} \frac{D_z(n)-1}{z} = \Pi(n)$, we'll then express $D_k(n)$ in terms of $E_k(n)$.

Then, to extend the domain of $E_k(n)$ to real values of *k,* we'll introduce a strict alternating count of divisors function, $E_k'(n)$ and express $E_k(n)$ in terms of it.

Finally, we'll generalize the concept of alternating signs, introducing the function $E_{k,b}(n)$, for which our original alternating series $E_k(n)$ is the special case of *b=2* and our count of divisors function $D_k(n)$ is the limit as $b \to \infty$, as well as a corresponding $E_{k,b}'(n)$. Because $D_k(n)$ can be expressed in terms of $E_{k,b}(n)$, and a generalized $E_{k,b}(n)$ can be expressed in terms of $E_{k,b}'(n)$, we'll see that the limit as $b \to 1$ for $E_{k,b}'(n)$ let's us say important things about $D_k(n)$.

7-1: Defining E

Let's look at an alternating series divisor-style sum, given by

$$E_1(n) = \sum_{j=1}^{n} (-1)^{j+1} , \; E_2(n) = \sum_{j=1}^{n} \sum_{k=1}^{\lfloor \frac{n}{j} \rfloor} (-1)^{j+k} , \; E_3(n) = \sum_{j=1}^{n} \sum_{k=1}^{\lfloor \frac{n}{j} \rfloor} \sum_{m=1}^{\lfloor \frac{n}{jk} \rfloor} (-1)^{j+k+m+1}$$

and more generally as

$$E_k(n) = \sum_{j=1}^{n} (-1)^{j+1} E_{k-1}\left(\frac{n}{j}\right) \text{ with } E_0(n) = 1$$

7-2: E in terms of D

$E_k(n)$ resembles the count of divisor function $D_k(n)$ from before, but with alternating positive/negative signs rather than the static value 1. How are these functions related?

Let's work through an example. From the above definition

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}(-1)^{j+k}$$

The goal here will be to express $E_2(n)$ in terms of $D_2(n)$, which requires replacing the -1 with sums involving 1.

First, extract the effect of the first sum's contribution to the power of the -1 like so:

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}(-1)^{k+1}-2\sum_{j=2,2|j}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}(-1)^{k+1}$$

Next, do the same thing to both of the second sums.

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}1-2\sum_{j=1}^{n}\sum_{k=2,2|k}^{\lfloor \frac{n}{j}\rfloor}1-2\sum_{j=2,2|j}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}1+4\sum_{j=2,2|j}^{n}\sum_{k=2,2|k}^{\lfloor \frac{n}{j}\rfloor}1$$

The middle pairs of sums are equivalent, so

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}1-4\sum_{j=1}^{n}\sum_{k=2,2|k}^{\lfloor \frac{n}{j}\rfloor}1+4\sum_{j=2,2|j}^{n}\sum_{k=2,2|k}^{\lfloor \frac{n}{j}\rfloor}1$$

The sums over even integers simplify to

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}1-4\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{2j}\rfloor}1+4\sum_{j=1}^{\lfloor \frac{n}{2}\rfloor}\sum_{k=1}^{\lfloor \frac{n}{4j}\rfloor}1$$

Now, remember, $D_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor \frac{n}{j}\rfloor}1$ .Thus, the right hand side can be rewritten in terms of $D_k(n)$ as

$$E_2(n)=D_2(n)-4\,D_2\left(\frac{n}{2}\right)+4\,D_2\left(\frac{n}{4}\right)$$

By extending this process, this relationship can be generalized to

$$E_k(n)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}2^j\,D_k\left(\frac{n}{2^j}\right)$$

If you're familiar with the Riemann Zeta function and the Dirichlet Eta function, this is roughly analogous to the relationship

$$\eta(s)=(1-2^{1-s})\zeta(s)$$

7-3: D in terms of E

The previous section noted that $E_2(n)=D_2(n)-4\,D_2(\frac{n}{2})+4\,D_2(\frac{n}{4})$. Let's invert this.

$$D_2(n)=E_2(n)+4\,D_2(\frac{n}{2})-4\,D_2(\frac{n}{4})$$

Applying this identity to itself, to the $D_2(\frac{n}{2})$ and $D_2(\frac{n}{4})$ terms, gives

$$D_2(n)=E_2(n)+4\,E_2(\frac{n}{2})+16\,D_2(\frac{n}{4})-16\,D_2(\frac{n}{8})-4\,E_2(\frac{n}{4})-16\,D_2(\frac{n}{4})+16\,D_2(\frac{n}{8})$$

Collect like terms, repeat this, and eventually $D_2(n)=0$ when $n<1$, leaving

$$D_2(n)=\sum_{j=0}^{\log_2 n} 2^j\,E_2(\frac{n}{2^j})$$

Extend this process, and this relationship generalizes to

$$D_k(n)=\sum_{j=0}^{\log_2 n} \binom{k+j-1}{k-1} 2^j\,E_k(\frac{n}{2^j})$$

7-4: Extending E with E'

The previous section expressed $D_k(n)$ in terms of $E_k(n)$. $E_k(n)$'s recursive definition in 7-1 only works for positive integers *k*, however. Handling Riemann's prime counting function, $\Pi(n)=\lim_{k\to 0}\frac{D_k(n)-1}{k}$, requires an expression for $E_k(n)$ where *k* can be real values.

Fortunately, the techniques back from 3-2b work here. Given a strict alternating count of divisors function

$$E_k{}'(n)=\sum_{j=2}^{n}(-1)^{j+1}E_{k-1}{}'(\frac{n}{j}) \text{ with } E_0{}'(n)=1$$

(note that *j* starts at 2, not 1, and $E_k{}'(n)=0$ when $n<2^k$), then a generalized $E_z(n)$ can be expressed as

$$E_z(n)=\sum_{a=0}^{\log_2 n}\frac{(z)(z-1)...(z-a+1)}{a!}E_a'(n)$$

Section 7-3 showed that $D_k(n)=\sum_{j=0}^{\log_2 n}\binom{k+j-1}{k-1}2^j E_k(\frac{n}{2^j})$. To let *k* take real (or complex) values, the binomial can be rewritten using the identity $\binom{k}{j}=\frac{k!}{(k-j)!\,j!}=\frac{(k)(k-1)...(k-j+1)}{j!}$, valid for complex *k*. Thus,

$$D_z(n)=\sum_{j=0}^{\log_2 n}\sum_{a=0}^{\log_2 n}\binom{z+j-1}{z-1}\frac{(z)(z-1)...(z-a+1)}{a!}2^j E_z'(\frac{n}{2^j})$$

7-5: Extending E to a More General Alternating Series

In the previous set of equations, you might wonder at the factor of 2 showing up. It does seem arbitrary, and a clue that we might be able to generalize our alternating series sums to interesting ends.

Let's define following function

$$t(n,b)=(n\,mod\,b)-((n-1)\,mod\,b)$$

where mod runs from 0 to $b-1$.

As a function, $t(n,b)$ yields 1 when *n* isn't divisible by *b*, and $b-1$ when *n* is divisible by *b*.
A moment's thought should make clear that $t(n,2)$ is, in fact, $(-1)^{n+1}$, which suggests $t(n,b)$ might let us generalize $E_k(n)$. Add a second parameter to give us $E_{k,b}(n)$, where *b* was 2 for our previous definition, and we have

$$E_{k,b}(n)=\sum_{j=1}^{n}t(j,b)E_{k-1,b}(\frac{n}{j}) \text{ with } E_{0,b}(n)=1$$

In previous sections, we expressed $E_k(n)$ and $D_k(n)$ in terms of each other. We can do likewise with this new generalization, giving us

$$E_{k,b}(n)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}b^j D_k(\frac{n}{b^j})$$

and

$$D_k(n)=\sum_{j=0}^{\log_2 n}\binom{k+j-1}{k-1}b^j E_{k,b}(\frac{n}{b^j})$$

As a quick aside, $t(n,b)$ is interesting in its own right. It provides one method for finding intermediary sums between alternating sums and non-alternating sums.

For example, it is well known that $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$ converges to log 2, while $\sum_{n=1}^{\infty}\frac{1}{n}$, the harmonic series, doesn't converge at all. Using $t(n,b)$ provides the intermediary values between these two series, $\sum_{n=1}^{\infty}\frac{t(n,b)}{n}=\log b$.

This immediately suggests the limit $\lim_{n\to\infty} H_{kn}-H_n=\log k$, where H_n are the harmonic numbers.

7-6: Extending E to Reals > 1 for *b*

In the previous section, we extended $E_k(n)$ from the special case where *b=2* to a more general identity for *b* a positive integer greater than 1. Is there some way to extend to even more values of *b*?

Back in section 7-2, when connecting $E_2(n)$ to $D_2(n)$, we noticed in passing that

$$E_2(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor\frac{n}{j}\rfloor}1-2\sum_{j=1}^{\lfloor\frac{n}{2}\rfloor}\sum_{k=1}^{\lfloor\frac{n}{2j}\rfloor}1$$

We might wonder what would happen if we replaced 2 here with some real valued variable *b* greater than 1, as

$$E_{2,b}(n)=\sum_{j=1}^{n}\sum_{k=1}^{\lfloor\frac{n}{j}\rfloor}1-b\sum_{j=1}^{\lfloor\frac{n}{b}\rfloor}\sum_{k=1}^{\lfloor\frac{n}{bj}\rfloor}1$$

Though we won't show it here, our identities from section 7-4, expressing $E_{k,b}(n)$ and $D_k(n)$ in terms of each other, hold for this definition of $E_{2,b}(n)$. In fact, we can generalize this equation to give us

$$E_{k,b}(n)=\sum_{j=1}^{n}E_{k-1,b}\left(\frac{n}{j}\right)-b\sum_{j=1}^{\lfloor\frac{n}{b}\rfloor}E_{k-1,b}\left(\frac{n}{bj}\right)\ \text{ with }\ E_{0,b}(n)=1$$

8

Extending Linnik's Identity
to Sum Primes

L

8-1: Sums of Primes

```
SumPrimes[n_,a_] := Sum[ Prime[j]^a,{j,PrimePi[n]}]
```

References

- *Mathoverflow link?*

9

More Divisor / Prime Power Connections

Some well-known results about the Moebius $\mu(n)$ function (which is $d_{-1}(n)$) and the von Mangoldt $\Lambda(n)$ function have analogues with $\kappa(n)$. But first, as an identity cheat sheet,

$$d_z(n)=\prod_{p^a|n}\frac{z(z+1)\dots(z+a-1)}{a!}=\sum_{a=0}^{\log_2 n}\frac{(z)(z-1)\dots(z-a+1)}{a!}d_a{}'(n)$$

$$d_k{}'(n)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}d_{k-j}(n)$$

$$\mu(n)=d_{-1}(n)=\sum_{k=0}^{\log_2 n}(-1)^k d_k{}'(n)$$

$$\kappa(n)=\frac{\Lambda(n)}{\log n}=\lim_{z\to 0}\frac{d_z(n)-d_0(n)}{z}=\sum_{k=0}^{\log_2 n}\frac{(-1)^{k+1}}{k}d_k{}'(n)$$

$$\Lambda(n)=\kappa(n)\log n=\lim_{z\to 0}\frac{(d_z(n)-d_0(n))(n^z-1)}{z\cdot z}$$

$$D_z(n)=\sum_{j=1}^{n}d_z(j)\qquad M(n)=\sum_{j=1}^{n}\mu(j)\qquad \Pi(n)=\sum_{j=2}^{n}\kappa(j)\qquad \psi(n)=\sum_{j=2}^{n}\Lambda(j)$$

It's well known that $\sum_{j|n}\mu(j)=0$ unless n is 1, where it's one. Also well known is $\sum_{j|n}\mu(j)\log\frac{n}{j}=\Lambda(n)$. Keeping in mind $\mu(n)=d_{-1}(n)$, here we have the analogous

$$\lim_{x\to 0}\frac{1}{x}\left(\left(\sum_{j|n}d_{-1+x}(j)\right)-d_0(n)\right)=\kappa(n)$$

```
PrimeK[ n_ ] := If[ n == 1, 0, FullSimplify[ MangoldtLambda[ n ] / Log[ n ]]]
d[ n_, z_ ] := Product[ Pochhammer[ z, a = p[[ 2 ]]] / a!, { p, FI[ n ]}];FI[ n_ ] := FactorInteger[ n ]; FI[ 1 ] := {}
dlimit[ n_, x_ ] := Round[ N[( 1 / x (Sum[ d[ j, -1 + x ], { j, Divisors[ n ]}] - d[ n, 0 ] ))],.0000001]
Table[{ n, dlimit[ n, 10^( -120 ) ], N[ PrimeK[ n ]]}, { n, 1, 100 }]//TableForm
```

It's also a standard result that $\sum_{j=1}^{n}M(\frac{n}{j})=1.$, and Mertens himself gave $\sum_{j=1}^{n}M(\frac{n}{j})\log j=\psi(n)$. Here we have the analogous

$$\lim_{x\to 0}\frac{1}{x}((\sum_{j=1}^{n} D_{-1+x}(\frac{n}{j}))-1)=\Pi(n)$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j ) ] / j, { j, 1, Log[ 2, n ]}]
d[n_,z_]:=Product[Pochhammer[z,a=p[[2]]]/a!,{p,FI[n]}];FI[n_]:=FactorInteger[n];FI[1]:={}
DD[ n_, k_ ] := Sum[ d[ j, k ], { j, 1, n}]
Dlimit[ n_, x_ ] := Round[ 1 / x N[Sum[ DD[ n/j, -1+x ],{j, 1,n }]  - 1 ], .0000001 ]
Table[{ n, Dlimit[ n, 10^(-120) ] ,N[RiePrimeCnt[ n ]]}, { n,1, 100 }] // TableForm
```

with the corresponding sum $\sum_{j=2}^{n}\mu(j)\log(\lfloor\frac{n}{j}\rfloor!)=\psi(n)$ matching

$$\lim_{x\to 0}\frac{1}{x}\sum_{j=2}^{n}\mu(j)D_{1+x}(\frac{n}{j})=\Pi(n)$$

```
RiePrimeCnt[ n_ ] := Sum[ PrimePi[ n ^ ( 1 / j ) ] / j, { j, 1, Log[ 2, n ]}]
d[n_,z_]:=Product[Pochhammer[z,a=p[[2]]]/a!,{p,FI[n]}];FI[n_]:=FactorInteger[n];FI[1]:={}
DD[ n_, k_ ] := Sum[ d[ j, k ], { j, 1, n}]
Dlimit[ n_, x_ ] := Round[ 1 / x N[ -1 + Sum[MoebiusMu[j] DD[ n/j, 1+x ],{j, 1,n }]], .0000001 ]
Table[{ n, Dlimit[ n, 10^(-120) ] , N[RiePrimeCnt[ n ]]}, { n, 1, 100 }] // TableForm
```

References

- *The identity* $\sum_{j|n}\mu(j)=0$ *is identity (1.18) from p. 12 of the 2004 edition of H. Iwaniec and E. Kowalski's "Analytic Number Theory".* $\sum_{j|n}\mu(j)\log\frac{n}{j}=\Lambda(n)$ *is identity (1.40) from p. 15 of that same work.*

10

Conclusion

L