The Iterated Carmichael Lambda Function

Nick Harland
University of British Columbia

PIMS/SFU/UBC Number Theory Seminar
Simon Fraser University
October 20, 2011

Outline

Definition

Definition of Carmichael Lambda Function

$\lambda(n)$ is the smallest natural number m such that

$$a^m \equiv 1 \pmod{n}$$

for all $(a, n) = 1$.

Clearly $\lambda(n) \mid \phi(n)$ and they are equal when n has a primitive root, i.e. when $n = 2, 4, p^k, 2p^k$ for an odd prime p.

Definition

Definition of Carmichael Lambda Function

$\lambda(n)$ is the smallest natural number m such that

$$a^m \equiv 1 \pmod{n}$$

for all $(a, n) = 1$.

Clearly $\lambda(n) \mid \phi(n)$ and they are equal when n has a primitive root, i.e. when $n = 2, 4, p^k, 2p^k$ for an odd prime p.

Calculating $\lambda(n)$

On odd prime powers, $\lambda(p^k) = \phi(p^k) = (p-1)p^{k-1}$.

On odder prime powers

$$\lambda(2) = 1, \lambda(4) = 2 \text{ and } \lambda(2^k) = \frac{1}{2}\phi(2^k) = 2^{k-2}$$

for $k \geq 3$.

Question

What if n is not a prime power?

By the Chinese Remainder Theorem we can get that

$$\lambda(\mathrm{lcm}\{a, b\}) = \mathrm{lcm}\{\lambda(a), \lambda(b)\}.$$

Calculating $\lambda(n)$

On odd prime powers, $\lambda(p^k) = \phi(p^k) = (p-1)p^{k-1}$.
On odder prime powers

$$\lambda(2) = 1, \lambda(4) = 2 \text{ and } \lambda(2^k) = \frac{1}{2}\phi(2^k) = 2^{k-2}$$

for $k \geq 3$.

Question

What if n is not a prime power?

By the Chinese Remainder Theorem we can get that

$$\lambda(\text{lcm}\{a,b\}) = \text{lcm}\{\lambda(a), \lambda(b)\}.$$

Calculating $\lambda(n)$

On odd prime powers, $\lambda(p^k) = \phi(p^k) = (p-1)p^{k-1}$.
On odder prime powers

$$\lambda(2) = 1, \lambda(4) = 2 \text{ and } \lambda(2^k) = \frac{1}{2}\phi(2^k) = 2^{k-2}$$

for $k \geq 3$.

Question

What if n is not a prime power?

By the Chinese Remainder Theorem we can get that

$$\lambda(\operatorname{lcm}\{a, b\}) = \operatorname{lcm}\{\lambda(a), \lambda(b)\}.$$

Calculating $\lambda(n)$

On odd prime powers, $\lambda(p^k) = \phi(p^k) = (p-1)p^{k-1}$.
On odder prime powers

$$\lambda(2) = 1, \lambda(4) = 2 \text{ and } \lambda(2^k) = \frac{1}{2}\phi(2^k) = 2^{k-2}$$

for $k \geq 3$.

Question

What if n is not a prime power?

By the Chinese Remainder Theorem we can get that

$$\lambda(\text{lcm}\{a, b\}) = \text{lcm}\{\lambda(a), \lambda(b)\}.$$

Analytic Properties of $\lambda(n)$

$\lambda(n)$ has a trivial upper bound of $\frac{2}{n}\sum_{i=1}^{n-1} i$ which is reached whenever n is prime. For a lower bound,

Theorem (Erdős, Pomerance, Schmutz (1991))

For any increasing sequence (n_i), for sufficiently large i

$$\lambda(n_i) > (\log n_i)^{c_0 \log\log\log n_i}$$

for any constant $0 < c_0 < 1/\log 2$.

They also showed that this can be acheived with some different effective constant in place of c_0.

Analytic Properties of $\lambda(n)$

$\lambda(n)$ has a trivial upper bound of $\frac{2}{n}\sum_{i=1}^{n-1} i$ which is reached whenever n is prime. For a lower bound,

Theorem (Erdős, Pomerance, Schmutz (1991))

For any increasing sequence (n_i), *for sufficiently large* i

$$\lambda(n_i) > (\log n_i)^{c_0 \log\log\log n_i}$$

for any constant $0 < c_0 < 1/\log 2$.

They also showed that this can be acheived with some different effective constant in place of c_0.

That's Typical

Question

What is the "typical" value of $\lambda(n)$?

Theorem (Erdős, Pomerance, Schmutz (1991))

There exists a set S of asymptotic density 1, where for all $n \in S$

$$\lambda(n) = n/(\log n)^{\log\log\log n + A + o(1)}$$

where $A = 0.2269688...$

That's Typical

Question

What is the "typical" value of $\lambda(n)$?

Theorem (Erdős, Pomerance, Schmutz (1991))

There exists a set S of asymptotic density 1, where for all $n \in S$

$$\lambda(n) = n/(\log n)^{\log\log\log n + A + o(1)}$$

where $A = 0.2269688...$

$2 > 1$

Question

What about $\lambda_2(n) = \lambda(\lambda(n))$?

Theorem (Martin, Pomerance (2005))

As $n \to \infty$ through a set of asymptotic density 1

$$\lambda_2(n) = n \exp\big(-(1+o(1))(\log\log n)^2 \log\log\log n \big).$$

Question

What happens for more iterations?!?!?!

$2 > 1$

Question

What about $\lambda_2(n) = \lambda(\lambda(n))$?

Theorem (Martin, Pomerance (2005))

As $n \to \infty$ through a set of asymptotic density 1

$$\lambda_2(n) = n \exp\Big(-(1+o(1))(\log\log n)^2 \log\log\log n\Big).$$

Question

What happens for more iterations?!?!?!

$2 > 1$

Question

What about $\lambda_2(n) = \lambda(\lambda(n))$?

Theorem (Martin, Pomerance (2005))

As $n \to \infty$ through a set of asymptotic density 1

$$\lambda_2(n) = n \exp\Big(-(1 + o(1))(\log\log n)^2 \log\log\log n \Big).$$

Question

What happens for more iterations?!?!?!

Why do 2 when you can do them all?

In the same paper, Martin and Pomerance gave the following conjecture

Conjecture (Martin, Pomerance (2005))

For any fixed $k \geq 1$,

$$\lambda_k(n) = n \exp\left(- \left(\frac{1}{(k-1)!} + o(1) \right) (\log\log n)^k \log\log\log n \right)$$

for almost all n.

Why do 2 when you can do them all?

In the same paper, Martin and Pomerance gave the following conjecture

Conjecture (Martin, Pomerance (2005))

For any fixed $k \geq 1$,

$$\lambda_k(n) = n \exp\left(- \left(\frac{1}{(k-1)!} + o(1) \right) (\log \log n)^k \log \log \log n \right)$$

for almost all n.

Why do 2 when you can do them all?

In the same paper, Martin and Pomerance gave the following conjecture

Conjecture (Martin, Pomerance (2005))

For any fixed $k \geq 1$,

$$\lambda_k(n) = n \exp\left(- \left(\frac{1}{(k-1)!} + o(1) \right) (\log\log n)^k \log\log\log n \right)$$

for almost all n.

However, now it's

Why do 2 when you can do them all?

In the same paper, Martin and Pomerance gave the following conjecture

Theorem (H. (2011))

For any fixed $k \geq 1$,

$$\lambda_k(n) = n \exp\left(- \left(\frac{1}{(k-1)!} + o(1) \right) (\log\log n)^k \log\log\log n \right)$$

for almost all n.

$\lambda(n)$ and $\phi(n)$ are friends

We are looking for the normal order of $\log(n/\lambda_k(n))$. However the relationship between n and $\lambda_k(n)$ is hard to see. It would be easier to look at the relationship between $\lambda_k(n)$ and $\phi_k(n)$. We do this by

$$\log\left(\frac{n}{\lambda_k(n)}\right) = \log\left(\frac{n}{\phi(n)}\right) + \log\left(\frac{\phi(n)}{\phi_2(n)}\right) + \dots + \log\left(\frac{\phi_{k-1}(n)}{\phi_k(n)}\right) + \log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right).$$

The other terms are $O(\log\log\log n)$ and get sucked into our error.

$\lambda(n)$ and $\phi(n)$ are friends

We are looking for the normal order of $\log(n/\lambda_k(n))$. However the relationship between n and $\lambda_k(n)$ is hard to see. It would be easier to look at the relationship between $\lambda_k(n)$ and $\phi_k(n)$. We do this by

$$\log\left(\frac{n}{\lambda_k(n)}\right) = \log\left(\frac{n}{\phi(n)}\right) + \log\left(\frac{\phi(n)}{\phi_2(n)}\right) + \dots + \log\left(\frac{\phi_{k-1}(n)}{\phi_k(n)}\right) + \log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right).$$

The other terms are $O(\log\log\log n)$ and get sucked into our error.

$\lambda(n)$ and $\phi(n)$ are friends

We are looking for the normal order of $\log(n/\lambda_k(n))$. However the relationship between n and $\lambda_k(n)$ is hard to see. It would be easier to look at the relationship between $\lambda_k(n)$ and $\phi_k(n)$. We do this by

$$\log\left(\frac{n}{\lambda_k(n)}\right) = \log\left(\frac{n}{\phi(n)}\right) + \log\left(\frac{\phi(n)}{\phi_2(n)}\right) + \dots + \log\left(\frac{\phi_{k-1}(n)}{\phi_k(n)}\right) + \log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right).$$

The other terms are $O(\log\log\log n)$ and get sucked into our error.

Why have one log when you can have many sums?

Let q be a prime and $a = v_q(n)$ be the largest power of q such that $q^a \mid n$. Let $y = \log\log x$. Then

$$\log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right) = \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))=1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q + \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))\geq 2}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q + \sum_{q\leq y^k} \nu_q(\phi_k(n)) \log q - \sum_{q\leq y^k} \nu_q(\lambda_k(n)) \log q.$$

Why have one log when you can have many sums?

Let q be a prime and $a = v_q(n)$ be the largest power of q such that $q^a \mid n$. Let $y = \log\log x$. Then

$$\log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right) = \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))=1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n)))\log q + \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))\geq 2}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n)))\log q + \sum_{q\leq y^k} \nu_q(\phi_k(n))\log q - \sum_{q\leq y^k} \nu_q(\lambda_k(n))\log q.$$

Which sum matters?

Of the 4 summations, only one matters enough to give us our main term. That summation is

$$\sum_{q \le y^k} \nu_q(\phi_k(n)) \log q$$

Regardless, in light of the appearance of v_q, it's very important to see how primes divide $\phi_k(n)$ and $\lambda_k(n)$.

Which sum matters?

Of the 4 summations, only one matters enough to give us our main term. That summation is

$$\sum_{q \le y^k} \nu_q(\phi_k(n)) \log q$$

Regardless, in light of the appearance of v_q, it's very important to see how primes divide $\phi_k(n)$ and $\lambda_k(n)$.

Examples

Example 1

Fix a prime q. How many n can have q dividing $\phi_3(n)$?

The short answer is many ways. One obvious case would have

$$q^4 \mid n$$

another would be the supersquarefree case

$$q \mid r-1, r \mid s-1, s \mid p-1 \text{ where } p \mid n$$

since if $p \mid n$, then $s \mid \phi(n)$ so $r \mid \phi_2(n)$ leading to $q \mid \phi_3(n)$.

How many n can have q dividing $\phi_3(n)$ in these cases?

Examples

Example 1

Fix a prime q. How many n can have q dividing $\phi_3(n)$?

The short answer is many ways. One obvious case would have

$$q^4 \mid n$$

another would be the supersquarefree case

$$q \mid r-1, r \mid s-1, s \mid p-1 \text{ where } p \mid n$$

since if $p \mid n$, then $s \mid \phi(n)$ so $r \mid \phi_2(n)$ leading to $q \mid \phi_3(n)$.

How many n can have q dividing $\phi_3(n)$ in these cases?

Examples

Example 1

Fix a prime q. How many n can have q dividing $\phi_3(n)$?

The short answer is many ways. One obvious case would have

$$q^4 \mid n$$

another would be the supersquarefree case

$$q \mid r-1, r \mid s-1, s \mid p-1 \text{ where } p \mid n$$

since if $p \mid n$, then $s \mid \phi(n)$ so $r \mid \phi_2(n)$ leading to $q \mid \phi_3(n)$.

How many n can have q dividing $\phi_3(n)$ in these cases?

Brun–Titchmarsh is our friend

Recall the Brun–Titchmarsh inequality

$$\pi(x;m,a) \leq \frac{2x}{\phi(m)\log(x/m)}$$

where $\pi(x;m,a)$ is the number of primes up to x congruent to a modulo m.

Using partial summation we can obtain

$$\sum_{\substack{p\leq x\\ p\in\mathcal{P}_m}} \frac{1}{p} \leq \frac{c\log\log x}{m}$$

where $\mathcal{P}_m$ is the set of primes congruent to 1 modulo m, provided $m/\phi(m)$ is bounded.

Brun–Titchmarsh is our friend

Recall the Brun–Titchmarsh inequality

$$\pi(x; m, a) \leq \frac{2x}{\phi(m)\log(x/m)}$$

where $\pi(x; m, a)$ is the number of primes up to x congruent to a modulo m.

Using partial summation we can obtain

$$\sum_{\substack{p \leq x \\ p \in \mathcal{P}_m}} \frac{1}{p} \leq \frac{c \log\log x}{m}$$

where $\mathcal{P}_m$ is the set of primes congruent to 1 modulo m, provided $m/\phi(m)$ is bounded.

Counting Cases

While the number of $n \le x$ such that $q^4 \mid n$ is clearly at most x/q^4, the number such n in our second way is

$$\sum_{n \le x} \sum_{r \in \mathcal{P}_q} \sum_{s \in \mathcal{P}_r} \sum_{\substack{p \in \mathcal{P}_s \\ p \mid n}} 1 = \sum_{r \in \mathcal{P}_q} \sum_{s \in \mathcal{P}_r} \sum_{p \in \mathcal{P}_s} \sum_{\substack{n \le x \\ p \mid n}} 1$$

$$\ll \sum_{r \in \mathcal{P}_q} \sum_{s \in \mathcal{P}_r} \sum_{p \in \mathcal{P}_s} \frac{x}{p}$$

$$\ll \sum_{r \in \mathcal{P}_q} \sum_{s \in \mathcal{P}_r} \frac{x \log\log x}{s}$$

Counting Cases

While the number of $n \le x$ such that $q^4 \mid n$ is clearly at most x/q^4, the number such n in our second way is

$$\sum_{n\le x}\sum_{r\in\mathcal{P}_q}\sum_{s\in\mathcal{P}_r}\sum_{\substack{p\in\mathcal{P}_s\\ p\mid n}} 1 = \sum_{r\in\mathcal{P}_q}\sum_{s\in\mathcal{P}_r}\sum_{p\in\mathcal{P}_s}\sum_{\substack{n\le x\\ p\mid n}} 1$$

$$\ll \sum_{r\in\mathcal{P}_q}\sum_{s\in\mathcal{P}_r}\sum_{p\in\mathcal{P}_s} \frac{x}{p}$$

$$\ll \sum_{r\in\mathcal{P}_q}\sum_{s\in\mathcal{P}_r} \frac{x\log\log x}{s}$$

Counting Cases

$$\ll \sum_{r \in \mathcal{P}_q} \frac{x(\log\log x)^2}{r}$$

$$\ll \frac{x(\log\log x)^3}{q}$$

In general, we can show

Lemma

Suppose $q > y^k$. For any such way for $q^a \mid \phi_k(n)$, the number of $n \leq x$ is that case is

$$O\left(\frac{xy^{ak}}{q^a}\right),$$

where $y = \log\log x$.

Counting Cases

$$\ll \sum_{r \in \mathcal{P}_q} \frac{x(\log\log x)^2}{r}$$
$$\ll \frac{x(\log\log x)^3}{q}$$

In general, we can show

Lemma

Suppose $q > y^k$. *For any such way for* $q^a \mid \phi_k(n)$, *the number of* $n \leq x$ *is that case is*

$$O\left(\frac{xy^{ak}}{q^a}\right),$$

where $y = \log\log x$.

Goodbye Sums

Let $\psi(x)$ be a function growing to infinity where $\psi(x) = o(\log\log\log x) = o(\log y)$. Recall

$$\log\left(\frac{\phi_k(n)}{\lambda_k(n)}\right) = \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))=1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n)))\log q + \sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))\geq 2}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n)))\log q + \sum_{q\leq y^k} \nu_q(\phi_k(n))\log q - \sum_{q\leq y^k} \nu_q(\lambda_k(n))\log q.$$

Goodbye Sums

The lemma can show

$$\sum_{n\leq x}\sum_{\substack{q>y^k\\ \nu_q(\phi_k(n))=1}}(\nu_q(\phi_k(n))-\nu_q(\lambda_k(n)))\log q\ll xy^k$$

which yields

$$\sum_{\substack{q>y^k\\ \nu_q(\phi_k(n))=1}}(\nu_q(\phi_k(n))-\nu_q(\lambda_k(n)))\log q\ll y^k\psi(x)$$

for all $n\leq x$ outside a set of size $O(x/\psi(x))$.

Goodbye Sums

The lemma can show

$$\sum_{n \le x} \sum_{\substack{q > y^k \\ \nu_q(\phi_k(n)) = 1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q \ll xy^k$$

which yields

$$\sum_{\substack{q > y^k \\ \nu_q(\phi_k(n)) = 1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q \ll y^k \psi(x)$$

for all $n \le x$ outside a set of size $O(x/\psi(x))$.

Goodbye Sums

With enough care, we can do the same thing to get

$$\sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))>1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q \ll y^k \psi(x)$$

for all $n \leq x$ outside a set of size $O(x/\psi(x))$.

and by analyzing the easier cases for $q^a \mid \lambda_k(n)$ we can also get

$$\sum_{q \leq y^k} \nu_q(\lambda_k(n)) \log q \ll y^k \psi(x)$$

for all $n \leq x$ outside a set of size $O(x/\psi(x))$.

Goodbye Sums

With enough care, we can do the same thing to get

$$\sum_{\substack{q>y^k \\ \nu_q(\phi_k(n))>1}} (\nu_q(\phi_k(n)) - \nu_q(\lambda_k(n))) \log q \ll y^k \psi(x)$$

for all $n \leq x$ outside a set of size $O(x/\psi(x))$.

and by analyzing the easier cases for $q^a \mid \lambda_k(n)$ we can also get

$$\sum_{q \leq y^k} \nu_q(\lambda_k(n)) \log q \ll y^k \psi(x)$$

for all $n \leq x$ outside a set of size $O(x/\psi(x))$.

Last Sum Standing

We are left with the main sum which is

$$\sum_{q \le y^k} \nu_q(\phi_k(n)) \log q.$$

Note that

$$v_q(\phi(m)) = \max\{0, v_q(m) - 1\} + \sum_{p|m} v_q(p-1)$$

yielding

$$\sum_{p|m} v_q(p-1) \le v_q(\phi(m)) \le v_q(m) + \sum_{p|m} v_q(p-1)$$

Last Sum Standing

We are left with the main sum which is

$$\sum_{q \le y^k} \nu_q(\phi_k(n)) \log q.$$

Note that

$$v_q(\phi(m)) = \max\{0, v_q(m) - 1\} + \sum_{p|m} v_q(p-1)$$

yielding

$$\sum_{p|m} v_q(p-1) \le v_q(\phi(m)) \le v_q(m) + \sum_{p|m} v_q(p-1)$$

Last Sum Standing

We are left with the main sum which is

$$\sum_{q \le y^k} \nu_q(\phi_k(n)) \log q.$$

Note that

$$v_q(\phi(m)) = \max\{0, v_q(m) - 1\} + \sum_{p|m} v_q(p-1)$$

yielding

$$\sum_{p|m} v_q(p-1) \le v_q(\phi(m)) \le v_q(m) + \sum_{p|m} v_q(p-1)$$

Last Sum Standing

We would expect the sum to contribute more here. Repeatedly using this yields

$$\sum_{p|\phi_{k-1}(n)} v_q(p-1) \leq v_q(\phi_k(n)) \leq \sum_{p|\phi_{k-1}(n)} v_q(p-1) + \sum_{p|\phi_{k-2}(n)} v_q(p-1) + \cdots + \sum_{p|\phi(n)} v_q(p-1) + v_q(n).$$

Last Sum Standing

We can do better. The supersquarefree case is the case which yields the most n. so we split the sum into the (ssf) case and the non (ssf) case, we get

$$\sum_{ssf} v_q(p-1) \le v_q(\phi_k(n)) \le \sum_{ssf} v_q(p-1) + \sum_{nssf} v_q(p-1) + \sum_{p|\phi_{k-2}(n)} v_q(p-1) + \cdots + \sum_{p|\phi(n)} v_q(p-1) + v_q(n).$$

Subtracting the sum on the left, multiplying by $\log q$ and summing over $q \le y^k$ gives us

Last Sum Standing

We can do better. The supersquarefree case is the case which yields the most n. so we split the sum into the (ssf) case and the non (ssf) case, we get

$$\sum_{ssf} v_q(p-1) \leq v_q(\phi_k(n)) \leq \sum_{ssf} v_q(p-1) + \sum_{nssf} v_q(p-1) + \sum_{p|\phi_{k-2}(n)} v_q(p-1) + \cdots + \sum_{p|\phi(n)} v_q(p-1) + v_q(n).$$

Subtracting the sum on the left, multiplying by $\log q$ and summing over $q \leq y^k$ gives us

Last Sum Standing

We can do better. The supersquarefree case is the case which yields the most n. so we split the sum into the (ssf) case and the non (ssf) case, we get

$$\sum_{ssf} v_q(p-1) \leq v_q(\phi_k(n)) \leq \sum_{ssf} v_q(p-1) + \sum_{nssf} v_q(p-1) + \sum_{p|\phi_{k-2}(n)} v_q(p-1) + \cdots + \sum_{p|\phi(n)} v_q(p-1) + v_q(n).$$

Subtracting the sum on the left, multiplying by $\log q$ and summing over $q \leq y^k$ gives us

Last Sum Standing

$$0 \leq \sum_{q \leq y^k} \nu_q(\phi_k(n)) \log q - h_k(n) \leq F_k(x)$$

where $h_k(n)$ is the additive function

$$h_k(n) = \sum_{p_1 | n} \sum_{p_2 | p_1 - 1} \cdots \sum_{p_k | p_{k-1} - 1} \sum_{q \leq y^k} \nu_q(p_k - 1) \log q$$

and $F_k(x)$ is the combination of a bunch of sums which can all be shown to be small via similar techniques as before.

Turán–Kubilius

Since $h_k(n)$ is additive, we can use the Turán–Kubilius inequality, which says

Turán-Kubilius Inequality

If $f(n)$ is an complex additive function, then there exists an absolute constant C such that

$$\sum_{n \leq x} |f(n) - M_1(x)|^2 \leq CxM_2(x)$$

where $M_1(x) = \sum_{p \leq x} \frac{|f(p)|}{p}$ and $M_2(x) = \sum_{p \leq x} \frac{|f(p)|^2}{p}$.

Turán–Kubilius

Since $h_k(n)$ is additive, we can use the Turán–Kubilius inequality, which says

Turán-Kubilius Inequality

If $f(n)$ is an complex additive function, then there exists an absolute constant C such that

$$\sum_{n\leq x}|f(n) - M_1(x)|^2 \leq CxM_2(x)$$

where $M_1(x) = \sum_{p\leq x}\frac{|f(p)|}{p}$ and $M_2(x) = \sum_{p\leq x}\frac{|f(p)|^2}{p}$.

Turán–Kubilius

Since we want $h_k(n)$ to have normal order $\frac{1}{(k-1)!}y^k \log y$, we will show it's true for $M_1(x)$. If we can show

$$M_1(x) = \frac{1}{(k-1)!}y^k \log y + O(y^k)$$

$$M_2(x) \ll y^{2k-1} \log^2 y$$

then if N is the number of $n \le x$ such that $|h_k(n) - M_1(x)| > y^k$, then

$$Ny^{2k} \ll xy^{2k-1} \log^2 y \Rightarrow N \ll \frac{x \log^2 y}{y} = o(x).$$

Turán–Kubilius

Since we want $h_k(n)$ to have normal order $\frac{1}{(k-1)!}y^k \log y$, we will show it's true for $M_1(x)$. If we can show

$$M_1(x) = \frac{1}{(k-1)!}y^k \log y + O(y^k)$$

$$M_2(x) \ll y^{2k-1} \log^2 y$$

then if N is the number of $n \leq x$ such that $|h_k(n) - M_1(x)| > y^k$, then

$$Ny^{2k} \ll xy^{2k-1} \log^2 y \Rightarrow N \ll \frac{x \log^2 y}{y} = o(x).$$

Turán–Kubilius

Since we want $h_k(n)$ to have normal order $\frac{1}{(k-1)!}y^k \log y$, we will show it's true for $M_1(x)$. If we can show

$$M_1(x) = \frac{1}{(k-1)!}y^k \log y + O(y^k)$$

$$M_2(x) \ll y^{2k-1} \log^2 y$$

then if N is the number of $n \leq x$ such that $|h_k(n) - M_1(x)| > y^k$, then

$$Ny^{2k} \ll xy^{2k-1} \log^2 y \Rightarrow N \ll \frac{x \log^2 y}{y} = o(x).$$

Evaluating $M_1(x)$ (Now it gets ugly)

$$M_1(x) = \sum_{p \le x} \frac{h_k(p)}{p}$$

$$= \sum_{p \le e^e} \frac{h_k(p)}{p} + \sum_{e^e < p \le x} \frac{h_k(p)}{p}$$

$$= O(1) + \sum_{e^e < p \le x} h_k(p) \left(\frac{1}{x} + \int_p^x \frac{dt}{t^2} \right)$$

$$= O(1) + \frac{1}{x} \sum_{e^e < p \le x} h_k(p) + \int_{e^e}^x \frac{dt}{t^2} \sum_{e^e < p \le t} h_k(p).$$

Evaluating $M_1(x)$ (Now it gets ugly)

$$\begin{aligned} M_1(x) &= \sum_{p \le x} \frac{h_k(p)}{p} \\ &= \sum_{p \le e^e} \frac{h_k(p)}{p} + \sum_{e^e < p \le x} \frac{h_k(p)}{p} \\ &= O(1) + \sum_{e^e < p \le x} h_k(p) \left(\frac{1}{x} + \int_p^x \frac{dt}{t^2} \right) \\ &= O(1) + \frac{1}{x} \sum_{e^e < p \le x} h_k(p) + \int_{e^e}^x \frac{dt}{t^2} \sum_{e^e < p \le t} h_k(p). \end{aligned}$$

Evaluating $M_1(x)$ (Now it gets ugly)

$$\begin{aligned}
M_1(x) &= \sum_{p\le x} \frac{h_k(p)}{p} \\
&= \sum_{p\le e^e} \frac{h_k(p)}{p} + \sum_{e^e<p\le x} \frac{h_k(p)}{p} \\
&= O(1) + \sum_{e^e<p\le x} h_k(p)\left(\frac{1}{x} + \int_p^x \frac{dt}{t^2}\right) \\
&= O(1) + \frac{1}{x}\sum_{e^e<p\le x} h_k(p) + \int_{e^e}^x \frac{dt}{t^2} \sum_{e^e<p\le t} h_k(p).
\end{aligned}$$

Evaluating $M_1(x)$ (Now it gets ugly)

$$\begin{aligned} M_1(x) &= \sum_{p \le x} \frac{h_k(p)}{p} \\ &= \sum_{p \le e^e} \frac{h_k(p)}{p} + \sum_{e^e < p \le x} \frac{h_k(p)}{p} \\ &= O(1) + \sum_{e^e < p \le x} h_k(p) \left(\frac{1}{x} + \int_p^x \frac{dt}{t^2} \right) \\ &= O(1) + \frac{1}{x} \sum_{e^e < p \le x} h_k(p) + \int_{e^e}^x \frac{dt}{t^2} \sum_{e^e < p \le t} h_k(p). \end{aligned}$$

Evaluating $\sum_{p\le x} h_k(p)$ (Oh so ugly)

Now using our definition of $h_k(n)$, we can get that

$$\begin{aligned}
\sum_{p\le t} h_k(p) &= \sum_{p\le t}\sum_{p_1|p}\sum_{p_2|p_1-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\nu_q(p_k-1)\log q \\
&= \sum_{p\le t}\sum_{p_2|p-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ a\in\mathbb{N}}}\log q \\
&= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\sum_{\substack{p\le t\\ p\in\mathcal{P}_{p_2}}}1 \\
&= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\pi(t;p_2,1).
\end{aligned}$$

Evaluating $\sum_{p\leq x} h_k(p)$ (Oh so ugly)

Now using our definition of $h_k(n)$, we can get that

$$\sum_{p\leq t} h_k(p) = \sum_{p\leq t}\sum_{p_1|p}\sum_{p_2|p_1-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\leq y^k}\nu_q(p_k-1)\log q$$

$$= \sum_{p\leq t}\sum_{p_2|p-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\leq y^k}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ a\in\mathbb{N}}}\log q$$

$$= \sum_{q\leq y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\sum_{\substack{p\leq t\\ p\in\mathcal{P}_{p_2}}} 1$$

$$= \sum_{q\leq y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\pi(t;p_2,1).$$

Evaluating $\sum_{p\le x} h_k(p)$ (Oh so ugly)

Now using our definition of $h_k(n)$, we can get that

$$\sum_{p\le t} h_k(p) = \sum_{p\le t}\sum_{p_1|p}\sum_{p_2|p_1-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\nu_q(p_k-1)\log q$$
$$= \sum_{p\le t}\sum_{p_2|p-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ a\in\mathbb{N}}}\log q$$
$$= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\sum_{\substack{p\le t\\ p\in\mathcal{P}_{p_2}}} 1$$
$$= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\pi(t;p_2,1).$$

Evaluating $\sum_{p\le x} h_k(p)$ (Oh so ugly)

Now using our definition of $h_k(n)$, we can get that

$$\begin{aligned}\sum_{p\le t} h_k(p) &= \sum_{p\le t}\sum_{p_1|p}\sum_{p_2|p_1-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\nu_q(p_k-1)\log q\\ &= \sum_{p\le t}\sum_{p_2|p-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ a\in\mathbb{N}}}\log q\\ &= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\sum_{\substack{p\le t\\ p\in\mathcal{P}_{p_2}}}1\\ &= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\pi(t;p_2,1).\end{aligned}$$

Evaluating $\sum_{p \le x} h_k(p)$ (Oh so ugly)

Now using our definition of $h_k(n)$, we can get that

$$\begin{aligned}
\sum_{p\le t} h_k(p) &= \sum_{p\le t}\sum_{p_1|p}\sum_{p_2|p_1-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\nu_q(p_k-1)\log q \\
&= \sum_{p\le t}\sum_{p_2|p-1}\cdots\sum_{p_k|p_{k-1}-1}\sum_{q\le y^k}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ a\in\mathbb{N}}}\log q \\
&= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\sum_{\substack{p\le t\\ p\in\mathcal{P}_{p_2}}}1 \\
&= \sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{p_k\in\mathcal{P}_{q^a}}\sum_{p_{k-1}\in\mathcal{P}_{p_k}}\cdots\sum_{p_2\in\mathcal{P}_{p_3}}\pi(t;p_2,1).
\end{aligned}$$

Evaluating $\sum_{p\leq x} h_k(p)$ (Still ugly)

By a whole lot of estimation involving the Brun Sieve and evaluating nasty summations involving arithmetic progressions of the Euler phi function, we can strip off large values of the primes leaving us with

Lemma

For all $x > e^{e^e}$ and $t > e^e$,

$$\sum_{p\leq t} h_k(p) = \sum_{q\leq y^k} \log q \sum_{a\in\mathbb{N}} \sum_{\substack{p_k\in\mathcal{P}_{q^a} \\ p_k\leq t^{1/3^{k-1}}}} \sum_{\substack{p_{k-1}\in\mathcal{P}_{p_k} \\ p_{k-1}\leq t^{1/3^{k-2}}}} \cdots \sum_{\substack{p_2\in\mathcal{P}_{p_3} \\ p_2\leq t^{1/3}}} \pi(t;p_2,1)$$
$$+ O\left(t^{1-1/3^k}\log t(\log\log t)^{k-2}y^k + \frac{t(\log\log t)^{k-2}\log y}{\log t}\right).$$

Evaluating $\sum_{p\le x} h_k(p)$ (Still ugly)

By a whole lot of estimation involving the Brun Sieve and evaluating nasty summations involving arithmetic progressions of the Euler phi function, we can strip off large values of the primes leaving us with

Lemma

For all $x > e^{e^e}$ and $t > e^e$,

$$\sum_{p\le t} h_k(p) = \sum_{q\le y^k} \log q \sum_{a\in\mathbb{N}} \sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ p_k\le t^{1/3^{k-1}}}} \sum_{\substack{p_{k-1}\in\mathcal{P}_{p_k}\\ p_{k-1}\le t^{1/3^{k-2}}}} \cdots \sum_{\substack{p_2\in\mathcal{P}_{p_3}\\ p_2\le t^{1/3}}} \pi(t; p_2, 1)$$

$$+ O\left(t^{1-1/3^k}\log t(\log\log t)^{k-2}y^k + \frac{t(\log\log t)^{k-2}\log y}{\log t}\right).$$

Evaluating $\sum_{p \le x} h_k(p)$ (This doesn't look so bad)

Question

Why did we strip off the large values?

We want to use the Bombieri–Vinogradov Theorem to freely change $\pi(t; p_2, 1)$ to $\frac{\mathrm{li}(t)}{p_2 - 1}$, then change that to $\frac{\mathrm{li}(t)}{p_2}$, where the errors will remain small.

Evaluating $\sum_{p \le x} h_k(p)$ (This doesn't look so bad)

Question

Why did we strip off the large values?

We want to use the Bombieri–Vinogradov Theorem to freely change $\pi(t; p_2, 1)$ to $\frac{\mathrm{li}(t)}{p_2-1}$, then change that to $\frac{\mathrm{li}(t)}{p_2}$, where the errors will remain small.

Evaluating $\sum_{p\leq x} h_k(p)$ (Back to ugly)

We define a similar function $g_{k,l}(u)$ to be

$$\sum_{q\leq y^k} \log q \sum_{a\in\mathbb{N}} \sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ p_k\leq u^{1/3^{l-1}}}} \sum_{\substack{p_{k-1}\in\mathcal{P}_{p_k}\\ p_{k-1}\leq u^{1/3^{l-2}}}} \cdots \sum_{\substack{p_{k-l+2}\in\mathcal{P}_{p_{k-l+3}}\\ p_{k-l+2}\leq u^{1/3}}} \pi(u; p_{k-l+2}, 1).$$

noting that when $u = t$ and $l = k$ we get the sum remaining in the lemma. We'll evaluate this by starting with $g_{k,2}(u)$ and then recursively working our way to $g_{k,k}(u)$.

Evaluating $\sum_{p\leq x} h_k(p)$ (Back to ugly)

We define a similar function $g_{k,l}(u)$ to be

$$\sum_{q\leq y^k} \log q \sum_{a\in\mathbb{N}} \sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ p_k\leq u^{1/3^{l-1}}}} \sum_{\substack{p_{k-1}\in\mathcal{P}_{p_k}\\ p_{k-1}\leq u^{1/3^{l-2}}}} \cdots \sum_{\substack{p_{k-l+2}\in\mathcal{P}_{p_{k-l+3}}\\ p_{k-l+2}\leq u^{1/3}}} \pi(u; p_{k-l+2}, 1).$$

noting that when $u = t$ and $l = k$ we get the sum remaining in the lemma. We'll evaluate this by starting with $g_{k,2}(u)$ and then recursively working our way to $g_{k,k}(u)$.

Evaluating $\sum_{p\le x} h_k(p)$ (Back to ugly)

We define a similar function $g_{k,l}(u)$ to be

$$\sum_{q\le y^k}\log q\sum_{a\in\mathbb{N}}\sum_{\substack{p_k\in\mathcal{P}_{q^a}\\ p_k\le u^{1/3^{l-1}}}}\sum_{\substack{p_{k-1}\in\mathcal{P}_{p_k}\\ p_{k-1}\le u^{1/3^{l-2}}}}\cdots\sum_{\substack{p_{k-l+2}\in\mathcal{P}_{p_{k-l+3}}\\ p_{k-l+2}\le u^{1/3}}}\pi(u;p_{k-l+2},1).$$

noting that when $u=t$ and $l=k$ we get the sum remaining in the lemma. We'll evaluate this by starting with $g_{k,2}(u)$ and then recursively working our way to $g_{k,k}(u)$.

Evaluating $g_{k,2}$

$$g_{k,2}(u) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \pi(u; p_k, 1)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \frac{1}{p_k} + O(Error)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \left(\frac{\log\log u^{1/3}}{\phi(q^a)} \right) + O(ERror)$$

Evaluating $g_{k,2}$

$$g_{k,2}(u) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \pi(u; p_k, 1)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \frac{1}{p_k} + O(Error)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \left(\frac{\log \log u^{1/3}}{\phi(q^a)} \right) + O(ERror)$$

Evaluating $g_{k,2}$

$$g_{k,2}(u) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \pi(u; p_k, 1)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le u^{1/3}}} \frac{1}{p_k} + O(Error)$$

$$= \mathrm{li}(u) \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \left(\frac{\log \log u^{1/3}}{\phi(q^a)} \right) + O(ERror)$$

Evaluating $g_{k,2}$

$$= \mathrm{li}(u)(\log\log u)\sum_{q\le y^k}\frac{\log q}{q} + O(ERROr)$$

$$= \frac{ku\log\log u\log y}{\log u} + O(ERROR)$$

Evaluating $g_{k,2}$

$$= \mathrm{li}(u)(\log\log u) \sum_{q \le y^k} \frac{\log q}{q} + O(ERROr)$$

$$= \frac{ku \log\log u \log y}{\log u} + O(ERROR)$$

Evaluating $g_{k,l}$

As for the recursion, we get

$$g_{k,l}(v) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \pi(v; p_{k-l+2}, 1).$$

$$= \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \frac{\mathrm{li}(v)}{p_{k-l+2}} + O(error)$$

Evaluating $g_{k,l}$

As for the recursion, we get

$$g_{k,l}(v) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \pi(v; p_{k-l+2}, 1).$$

$$= \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \frac{\mathrm{li}(v)}{p_{k-l+2}} + O(error)$$

Evaluating $g_{k,l}$

As for the recursion, we get

$$g_{k,l}(v) = \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \pi(v; p_{k-l+2}, 1).$$

$$= \sum_{q \le y^k} \log q \sum_{a \in \mathbb{N}} \sum_{\substack{p_k \in \mathcal{P}_{q^a} \\ p_k \le v^{1/3^{l-1}}}} \cdots \sum_{\substack{p_{k-l+2} \in \mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2} \le v^{1/3}}} \frac{\mathrm{li}(v)}{p_{k-l+2}} + O(error)$$

Evaluating $g_{k,l}$

Then using partial summation, we can get that

$$\sum_{\substack{p_{k-l+2}\in\mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2}\leq v^{1/3}}} \frac{1}{p_{k-l+2}} = \frac{\pi(v^{1/3};p_{k-l+3},1)}{v^{1/3}} + \int_2^{v^{1/3}} \frac{\pi(u;p_{k-l+3},1)}{u^2}du$$

The first term can be estimated trivially and added to the error, yielding the recursion

Evaluating $g_{k,l}$

Then using partial summation, we can get that

$$\sum_{\substack{p_{k-l+2}\in\mathcal{P}_{p_{k-l+3}} \\ p_{k-l+2}\leq v^{1/3}}} \frac{1}{p_{k-l+2}} = \frac{\pi(v^{1/3}; p_{k-l+3}, 1)}{v^{1/3}} + \int_2^{v^{1/3}} \frac{\pi(u; p_{k-l+3}, 1)}{u^2} du$$

The first term can be estimated trivially and added to the error, yielding the recursion

Evaluating $g_{k,l}$

Lemma

Let $3 \leq l \leq k$, *then*

$$g_{k,l}(v) = \mathrm{li}(v) \int_2^{v^{1/3}} \frac{1}{u^2} g_{k,l-1}(u)du + O\left(\frac{v(\log\log v)^{l-2}\log y}{\log v}\right).$$

Evaluating $g_{k,l}$

Hence

$$g_{k,2}(v) = \frac{kv \log\log v \log y}{\log v} + O(error)$$

$$g_{k,3}(v) = \mathrm{li}(v) \int_2^{v^{1/3}} \frac{k \log\log u \log y du}{u \log u} + O(error)$$
$$= \frac{kv(\log\log v)^2 \log y}{2 \log v} + O(error)$$

and so on to get

Evaluating $g_{k,l}$

Hence

$$g_{k,2}(v) = \frac{kv \log\log v \log y}{\log v} + O(error)$$

$$\begin{aligned} g_{k,3}(v) &= \mathrm{li}(v) \int_2^{v^{1/3}} \frac{k \log\log u \log y du}{u \log u} + O(error) \\ &= \frac{kv(\log\log v)^2 \log y}{2 \log v} + O(error) \end{aligned}$$

and so on to get

Evaluating $g_{k,l}$

$$g_{k,l}(v) = \frac{kv(\log\log v)^{l-1}\log y}{(l-1)!\log v} + O(error)$$

Using $v = t$ and $l = k$ which implies (when we include the errors

Lemma

Let $2 \leq k$, then

$$\sum_{p\leq t} h_k(p) = \frac{kt(\log\log t)^{k-1}\log y}{(k-1)!\log t} + O\bigg(\frac{t(\log\log t)^{k-1}}{\log t} + \frac{t(\log\log t)^{k-2}\log^2 y}{\log t} + t^{1-1/3^k}\log t(\log\log t)^{k-2}y^k\bigg).$$

Evaluating $g_{k,l}$

$$g_{k,l}(v) = \frac{kv(\log\log v)^{l-1}\log y}{(l-1)!\log v} + O(error)$$

Using $v = t$ and $l = k$ which implies (when we include the errors

Lemma

Let $2 \leq k$, *then*

$$\sum_{p\leq t} h_k(p) = \frac{kt(\log\log t)^{k-1}\log y}{(k-1)!\log t} + O\Bigg(\frac{t(\log\log t)^{k-1}}{\log t} + \frac{t(\log\log t)^{k-2}\log^2 y}{\log t} + t^{1-1/3^k}\log t(\log\log t)^{k-2}y^k\Bigg).$$

Evaluating $M_1(x)$ (Yes kids, we're almost here)

Hence if we go back to $M_1(x)$

$$M_1(x) = O(1) + \frac{1}{x}\sum_{e^e < p \le x} h_k(p) + \int_{e^e}^{x} \frac{dt}{t^2} \sum_{e^e < p \le t} h_k(p).$$

$$= \frac{1}{x} O\left(\frac{xy^{k-1}\log y}{\log x}\right) + \int_{e^e}^{x} \frac{k(\log\log t)^{k-1}\log y dt}{(k-1)! t \log t} + O(y^k)$$

$$= \frac{y^k \log y}{(k-1)!} + O(y^k)$$

HOORAY!

Evaluating $M_1(x)$ (Yes kids, we're almost here)

Hence if we go back to $M_1(x)$

$$M_1(x) = O(1) + \frac{1}{x}\sum_{e^e<p\le x} h_k(p) + \int_{e^e}^{x} \frac{dt}{t^2} \sum_{e^e<p\le t} h_k(p).$$

$$= \frac{1}{x}O\left(\frac{xy^{k-1}\log y}{\log x}\right) + \int_{e^e}^{x} \frac{k(\log\log t)^{k-1}\log y dt}{(k-1)!t\log t} + O(y^k)$$

$$= \frac{y^k \log y}{(k-1)!} + O(y^k)$$

HOORAY!

Evaluating $M_1(x)$ (Yes kids, we're almost here)

Hence if we go back to $M_1(x)$

$$M_1(x) = O(1) + \frac{1}{x}\sum_{e^e<p\le x} h_k(p) + \int_{e^e}^{x} \frac{dt}{t^2}\sum_{e^e<p\le t} h_k(p).$$

$$= \frac{1}{x}O\left(\frac{xy^{k-1}\log y}{\log x}\right) + \int_{e^e}^{x} \frac{k(\log\log t)^{k-1}\log y dt}{(k-1)!t\log t} + O(y^k)$$

$$= \frac{y^k\log y}{(k-1)!} + O(y^k)$$

HOORAY!

Evaluating $M_1(x)$ (Yes kids, we're almost here)

Hence if we go back to $M_1(x)$

$$\begin{aligned} M_1(x) &= O(1) + \frac{1}{x}\sum_{e^e < p \le x} h_k(p) + \int_{e^e}^{x} \frac{dt}{t^2} \sum_{e^e < p \le t} h_k(p). \\ &= \frac{1}{x} O\left(\frac{xy^{k-1}\log y}{\log x}\right) + \int_{e^e}^{x} \frac{k(\log\log t)^{k-1}\log y dt}{(k-1)! t \log t} + O(y^k) \\ &= \frac{y^k \log y}{(k-1)!} + O(y^k) \end{aligned}$$

HOORAY!

Evaluating $M_2(x)$ (Just kidding)

With similar crazy sieve type lemma and evaluation of crazy sums involving the Euler Phi function we can take care of $M_2(x)$.

Further Questions

Question

We related n to $\lambda_k(n)$. What about compositions of λ and ϕ?

The short answer is that if $f(n)$ is k compositions of λ and ϕ beginning with λ, then the relation to n is the same as $\lambda_k(n)$. Any deviation would be part of the error term. If it starts with ϕ, then get rid of all the $\phi' s$ it starts with and ask the question again. i.e.

$$\log\left(\frac{n}{\phi(\phi(\lambda(\phi(n))))}\right) \sim \log\left(\frac{n}{\lambda_2(n)}\right) \sim (\log\log x)^2 \log\log\log x.$$

Further Questions

Question

We related n to $\lambda_k(n)$. What about compositions of λ and ϕ?

The short answer is that if $f(n)$ is k compositions of λ and ϕ beginning with λ, then the relation to n is the same as $\lambda_k(n)$. Any deviation would be part of the error term. If it starts with ϕ, then get rid of all the $\phi's$ it starts with and ask the question again. i.e.

$$\log\left(\frac{n}{\phi(\phi(\lambda(\phi(n))))}\right) \sim \log\left(\frac{n}{\lambda_2(n)}\right) \sim (\log\log x)^2 \log\log\log x.$$

The end

Thanks for your attention. These slides will be available at my website at www.nickharland.com