Math 571 Chapter 4 The Selberg Sieve

Robert C. Vaughan

January 27, 2023

The Sieve of Eratosthenes c200BC

		a	b		c		d		
	1	2	3	4a	5	6a	7	8a	9b
10a	11	12a	13	14a	15b	16a	17	18a	19
20a	21b	22a	23	24a	25c	26a	27b	28a	29a
30a	31	32a	33b	34a	35c	36a	37	38a	39b
40a	41	42a	43	44a	45b	46a	47	48a	49d
50a	51b	52a	53	54a	55c	56a	57b	58a	59
60a	61	62a	63b	64a	65c	66a	67	68a	69b
70a	71	72a	73	74a	75b	76a	77d	78a	79
80a	81b	82a	83	84a	85c	86a	87b	88a	89
90a	91d	92a	93b	94a	95c	96a	97	98a	99b

The Sieve of Eratosthenes c200BC

		a	b		c		d		
	1	2	3	4a	5	6a	7	8a	9
10a	11	12a	13	14a	15b	16a	17	18a	19
20a	21b	22a	23	24a	25c	26a	27b	28a	29a
30a	31	32a	33b	34a	35c	36a	37	38a	39b
40a	41	42a	43	44a	45b	46a	47	48a	49d
50a	51b	52a	53	54a	55c	56a	57b	58a	59
60a	61	62a	63b	64a	65c	66a	67	68a	69b
70a	71	72a	73	74a	75b	76a	77d	78a	79
80a	81b	82a	83	84a	85c	86a	87b	88a	89
90a	91d	92a	93b	94a	95c	96a	97	98a	99b

After 1, 2, 3, 5, 7, the numbers remaining are

$$11, 13, 17, 19, 23, 31, 37, 41, 43, 47, 53,$$
$$59, 61, 67, 71, 73, 79, 83, 89, 97.$$

and are precisely the primes in the range $(7, 100]$.

The Sieve of Eratosthenes c200BC

		a	b		c		d		
	1	2	3	4a	5	6a	7	8a	9
10a	11	12a	13	14a	15b	16a	17	18a	19
20a	21b	22a	23	24a	25c	26a	27b	28a	29a
30a	31	32a	33b	34a	35c	36a	37	38a	39b
40a	41	42a	43	44a	45b	46a	47	48a	49d
50a	51b	52a	53	54a	55c	56a	57b	58a	59
60a	61	62a	63b	64a	65c	66a	67	68a	69b
70a	71	72a	73	74a	75b	76a	77d	78a	79
80a	81b	82a	83	84a	85c	86a	87b	88a	89
90a	91d	92a	93b	94a	95c	96a	97	98a	99b

After 1, 2, 3, 5, 7, the numbers remaining are

$$11, 13, 17, 19, 23, 31, 37, 41, 43, 47, 53,$$
$$59, 61, 67, 71, 73, 79, 83, 89, 97.$$

and are precisely the primes in the range $(7, 100]$. Also we find that

$$\pi(100) = 25.$$

- Can we use this to prove the prime number theorem?

- Can we use this to prove the prime number theorem?
- We can formalise this as

$$\pi(x) - \pi(\sqrt{x}) + 1 = \text{card}\{n \le x : \nexists p | n, p \le \sqrt{x}\}.$$

- Can we use this to prove the prime number theorem?
- We can formalise this as

$$\pi(x) - \pi(\sqrt{x}) + 1 = \mathrm{card}\{n \le x : \nexists p | n, p \le \sqrt{x}\}.$$

- Another way to write this is

$$\pi(x) - \pi(\sqrt{x}) + 1 = \mathrm{card}\{n \le x : (n, P) = 1\}$$

where

$$P = P(x) = \prod_{p \le \sqrt{x}} p.$$

- Can we use this to prove the prime number theorem?
- We can formalise this as

$$\pi(x) - \pi(\sqrt{x}) + 1 = \mathrm{card}\{n \le x : \nexists p | n, p \le \sqrt{x}\}.$$

- Another way to write this is

$$\pi(x) - \pi(\sqrt{x}) + 1 = \mathrm{card}\{n \le x : (n, P) = 1\}$$

where

$$P = P(x) = \prod_{p \le \sqrt{x}} p.$$

- Sylvester noticed that this can be realised as a form of the inclusion-exclusion principle. For example, given two statements $Q(n)$ and $R(n)$ about integers n in some finite set $\mathcal{N}$, the number of n for which both $Q(n)$ and $R(n)$ are false is equal to the cardinality of $\mathcal{N}$ minus the number of n for which $Q(n)$ is true, minus the number for which $R(n)$ is true plus the number for which both $Q(n)$ and $R(n)$ are true.

- Here is a way of setting this in our situation, and this easily generalises to general statements about sets.

- Here is a way of setting this in our situation, and this easily generalises to general statements about sets.
- Let

$$\eta_d(n) = \begin{cases} 1 & \text{when } d|n, \\ 0 & \text{when } d \nmid n. \end{cases}$$

and consider

$$\prod_{p\le\sqrt{x}} (1-\eta_p(n)) =$$

$$1 - \sum_{p\le\sqrt{x}} \eta_p(n) + \sum_{p_1<p_2\le\sqrt{x}} \eta_{p_1p_2}(n)$$

$$+ \cdots + (-1)^k \sum_{p_1<P_2<\ldots<p_k\le\sqrt{x}} \eta_{p_1\ldots p_k}(n) + \cdots .$$

- Here is a way of setting this in our situation, and this easily generalises to general statements about sets.
- Let

$$\eta_d(n) = \begin{cases} 1 & \text{when } d|n, \\ 0 & \text{when } d\nmid n. \end{cases}$$

and consider

$$\prod_{p\le\sqrt{x}} (1-\eta_p(n)) = $$

$$1 - \sum_{p\le\sqrt{x}} \eta_p(n) + \sum_{p_1<p_2\le\sqrt{x}} \eta_{p_1p_2}(n)$$

$$+\cdots+(-1)^k \sum_{p_1<P_2<\ldots<p_k\le\sqrt{x}} \eta_{p_1\ldots p_k}(n) + \cdots .$$

- We also have

$$\mathrm{card}\{n\le x : \nexists p|n, p\le\sqrt{x}\} = \sum_{n\le x} \prod_{p\le\sqrt{x}} (1-\eta_p(n))$$

- This starts to look combinatorially very complicated.

- This starts to look combinatorially very complicated.
- Here is an alternative way of expressing the same thing. Recall that

$$\sum_{d|m}\mu(d)=\begin{cases}1 \text{ when } m=1,\\ 0 \text{ when } m>1.\end{cases}$$

- This starts to look combinatorially very complicated.
- Here is an alternative way of expressing the same thing. Recall that

$$\sum_{d|m} \mu(d) = \begin{cases} 1 \text{ when } m = 1, \\ 0 \text{ when } m > 1. \end{cases}$$

- Then

$$\sum_{d|(n,P)} \mu(d) = \begin{cases} 1 \text{ when } (n,P) = 1, \\ 0 \text{ when } (n,P) > 1. \end{cases}$$

- This starts to look combinatorially very complicated.
- Here is an alternative way of expressing the same thing. Recall that

$$\sum_{d|m}\mu(d)=\begin{cases}1 \text{ when } m=1,\\ 0 \text{ when } m>1.\end{cases}$$

- Then

$$\sum_{d|(n,P)}\mu(d)=\begin{cases}1 \text{ when } (n,P)=1,\\ 0 \text{ when } (n,P)>1.\end{cases}$$

- Thus

$$\begin{aligned}\operatorname{card}\{n\le x:\nexists p|n, p\le\sqrt{x}\}&=\sum_{n\le x}\sum_{d|(n,P)}\mu(d)\\ &=\sum_{d|P}\mu(d)\sum_{n\le x}\eta_d(n)\\ &=\sum_{d|P}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor.\end{aligned}$$

- Thus we have

$$\pi(x)-\pi(\sqrt{x})+1=\sum_{d|P}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor.$$

- Thus we have

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor .$$

- Suppose we approximate $\lfloor x/d \rfloor$ by x/d. The error introduced in each term is less than 1. We obtain

- Thus we have
$$\pi(x)-\pi(\sqrt{x})+1=\sum_{d|P}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor.$$
- Suppose we approximate $\lfloor x/d\rfloor$ by x/d. The error introduced in each term is less than 1. We obtain
- $$\pi(x)-\pi(\sqrt{x})+1\approx x\sum_{d|P}\frac{\mu(d)}{d}=x\prod_{p\le\sqrt{x}}\left(1-\frac{1}{p}\right)$$

- Thus we have

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor .$$

- Suppose we approximate $\lfloor x/d \rfloor$ by x/d. The error introduced in each term is less than 1. We obtain
-

$$\pi(x) - \pi(\sqrt{x}) + 1 \approx x \sum_{d|P} \frac{\mu(d)}{d} = x \prod_{p \le \sqrt{x}} \left(1 - \frac{1}{p}\right)$$

- and by Merten's theorem this is

$$\sim \frac{2xe^{-B}}{\log x}.$$

- We would have
$$\pi(x) - \pi(\sqrt{x}) + 1 \approx \frac{2xe^{-B}}{\log x}.$$

- We would have
$$\pi(x) - \pi(\sqrt{x}) + 1 \approx \frac{2xe^{-B}}{\log x}.$$
- Oh, wait a minute, $2e^{-B} \neq 1$, so this would contradict the prime number theorem.

- We would have

$$\pi(x) - \pi(\sqrt{x}) + 1 \approx \frac{2xe^{-B}}{\log x}.$$

- Oh, wait a minute, $2e^{-B} \neq 1$, so this would contradict the prime number theorem.
- The problem is that the number of terms in

$$\sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor,$$

that is, the number of choices for d, is huge,

$$2^{\pi(\sqrt{x})}$$

and we cannot afford an error as large as 1 in each term.

- A French mathematician, Merlin, found a clever way of truncating the terms to give upper and lower bounds.

- A French mathematician, Merlin, found a clever way of truncating the terms to give upper and lower bounds.
- Basically

$$\sum_{\substack{d|P\\ \omega(d)\le 2k-1}} \mu(d)\left\lfloor\frac{x}{d}\right\rfloor \le \sum_{d|P}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor \le \sum_{\substack{d|P\\ \omega(d)\le 2k}} \mu(d)\left\lfloor\frac{x}{d}\right\rfloor .$$

- However he was killed in WWI before he could develop the idea and it was taken up and developed by Brun.

- A French mathematician, Merlin, found a clever way of truncating the terms to give upper and lower bounds.
- Basically

$$\sum_{\substack{d|P\\ \omega(d)\le 2k-1}} \mu(d)\left\lfloor\frac{x}{d}\right\rfloor \le \sum_{d|P}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor \le \sum_{\substack{d|P\\ \omega(d)\le 2k}} \mu(d)\left\lfloor\frac{x}{d}\right\rfloor.$$

- However he was killed in WWI before he could develop the idea and it was taken up and developed by Brun.
- The method became combinatorially very complicated and very few people understood it. Perhaps really only Paul Erdős.

- So far I have just described things in relation to the prime number theorem, but sieve theory is very adaptable.

- So far I have just described things in relation to the prime number theorem, but sieve theory is very adaptable.
- Consider the twin prime problem.

$$3,5;5,7;11,13;17,19;29,31;\ldots;101,103;107,109;\ldots$$

It looks as though there are infinitely many primes p for which $p+2$ is prime, but no proof is known.

- So far I have just described things in relation to the prime number theorem, but sieve theory is very adaptable.
- Consider the twin prime problem.

  $$3,5;5,7;11,13;17,19;29,31;\ldots;101,103;107,109;\ldots$$

  It looks as though there are infinitely many primes p for which $p+2$ is prime, but no proof is known.
- One way of counting them is to consider

  $$\begin{aligned}&\operatorname{card}\{\sqrt{x}<p\le x : p+2 \text{ prime}\}\\ &\qquad\qquad = \operatorname{card}\{n\le x : \big(n(n+2), P(\sqrt{x})\big)=1\}.\end{aligned}$$

- So far I have just described things in relation to the prime number theorem, but sieve theory is very adaptable.
- Consider the twin prime problem.

  $$3, 5; 5, 7; 11, 13; 17, 19; 29, 31; \ldots; 101, 103; 107, 109; \ldots$$

  It looks as though there are infinitely many primes p for which $p + 2$ is prime, but no proof is known.
- One way of counting them is to consider

  $$\begin{aligned}&\operatorname{card}\{\sqrt{x} < p \le x : p + 2 \text{ prime}\}\\ &\qquad\qquad = \operatorname{card}\{n \le x : \big(n(n+2), P(\sqrt{x})\big) = 1\}.\end{aligned}$$

- Another famous question is the Goldbach binary problem, to show that every even number $N > 2$ is the sum of two primes.

- So far I have just described things in relation to the prime number theorem, but sieve theory is very adaptable.
- Consider the twin prime problem.

$$3, 5; 5, 7; 11, 13; 17, 19; 29, 31; \dots ; 101, 103; 107, 109; \dots$$

It looks as though there are infinitely many primes p for which $p + 2$ is prime, but no proof is known.
- One way of counting them is to consider

$$\begin{aligned}&\operatorname{card}\{\sqrt{x} < p \le x : p + 2 \text{ prime}\}\\ &\qquad = \operatorname{card}\{n \le x : \big(n(n + 2), P(\sqrt{x})\big) = 1\}.\end{aligned}$$

- Another famous question is the Goldbach binary problem, to show that every even number $N > 2$ is the sum of two primes.
- This can be set up by considering

$$\begin{aligned}&\operatorname{card}\{\sqrt{N} < p_1, p_2 : p_1 + p_2 = N\}\\ &\qquad = \operatorname{card}\{1 < n < N - 1 : \big(n(N - n), P(\sqrt{N})\big) = 1\}.\end{aligned}$$

- Yet another open problem concerns the frequency with which n^2+1 is prime, and this could be set up by looking at

$$\begin{aligned}\operatorname{card}\{\sqrt{x} < n \le x : n^2+1 \text{ prime}\}\\ = \operatorname{card}\{n \le x : \bigl(n^2+1, P(x)\bigr) = 1\}.\end{aligned}$$

- Yet another open problem concerns the frequency with which n^2+1 is prime, and this could be set up by looking at

$$\begin{aligned}&\operatorname{card}\{\sqrt{x}<n\le x : n^2+1 \text{ prime}\}\\ &\qquad\qquad = \operatorname{card}\{n\le x : \big(n^2+1, P(x)\big)=1\}.\end{aligned}$$

- In view of the great generality of the concepts I want to set up some notation.

- Let

$$a : \mathbb{Z} \to \mathbb{R}^+,$$

$$A = \sum_n a(n) < \infty,$$

$$A_d = \sum_n a(dn),$$

and suppose that

$$A_d = f(d)X + R_d,$$

where

$$f \in \mathcal{M},$$

the set $\mathcal{M}$ of multiplicative functions.

- Let
$$a:\mathbb{Z}\to\mathbb{R}^+,$$
$$A=\sum_n a(n)<\infty,$$
$$A_d=\sum_n a(dn),$$
and suppose that
$$A_d=f(d)X+R_d,$$
where
$$f\in\mathcal{M},$$
the set $\mathcal{M}$ of multiplicative functions.
- It is also convenient to assume that $0\le f(p)<1$ for each prime p.

- Let
$$a : \mathbb{Z} \to \mathbb{R}^+,$$
$$A = \sum_n a(n) < \infty,$$
$$A_d = \sum_n a(dn),$$
and suppose that
$$A_d = f(d)X + R_d,$$
where
$$f \in \mathcal{M},$$
the set $\mathcal{M}$ of multiplicative functions.
- It is also convenient to assume that $0 \le f(p) < 1$ for each prime p.
- We define
$$S(A, P) = \sum_{\substack{n \\ (n,P)=1}} a(n).$$

- $A_d = f(d)X + R_d.$

- $A_d = f(d)X + R_d$.
- In principle we suppose that X is “large” and R_d is “small” compared with $f(d)X$ when d is relatively small.

- $A_d = f(d)X + R_d$.
- In principle we suppose that X is "large" and R_d is "small" compared with $f(d)X$ when d is relatively small.
- **Example 1** Let $a(n) = 1$ when $Y < n \le Y + X$ and $a(n) = 0$ otherwise. Then

$$A_d = \frac{X}{d} + R_d, \quad |R_d| \le 1.$$

This corresponds to counting primes in an interval.

- $A_d = f(d)X + R_d$.
- In principle we suppose that X is "large" and R_d is "small" compared with $f(d)X$ when d is relatively small.
- **Example 1** Let $a(n) = 1$ when $Y < n \le Y + X$ and $a(n) = 0$ otherwise. Then

  $$A_d = \frac{X}{d} + R_d, \quad |R_d| \le 1.$$

  This corresponds to counting primes in an interval.
- **Example 2** If $P = \prod_{p \le \sqrt{X}} p$, and a is as in Example 1, then

  $$\pi(X + Y) - \pi(Y) \le \pi(\sqrt{X}) + S(A, P).$$

  This is a formalisation of the sieve of Erathosthenes–Legendre.

- $A_d = f(d)X + R_d$.
- In principle we suppose that X is "large" and R_d is "small" compared with $f(d)X$ when d is relatively small.
- **Example 1** Let $a(n) = 1$ when $Y < n \le Y + X$ and $a(n) = 0$ otherwise. Then

  $$A_d = \frac{X}{d} + R_d, \quad |R_d| \le 1.$$

  This corresponds to counting primes in an interval.
- **Example 2** If $P = \prod_{p \le \sqrt{X}} p$, and a is as in Example 1, then

  $$\pi(X + Y) - \pi(Y) \le \pi(\sqrt{X}) + S(A, P).$$

  This is a formalisation of the sieve of Erathosthenes–Legendre.
- Any method which deduces estimates for $S(A, P)$ is called a sieve.

- Other examples.

- Other examples.
- **Example 3 (Primes in arithmetic progression)** Suppose that $(q,r)=1$ and $a(n)=1$ when $y<n\le x+y$ and $n\equiv r \pmod q$. Since we already have $(n,q)=1$ we can suppose
$$P=P_q=\prod_{\substack{p\le z\\ p\nmid q}} p.$$

- Other examples.
- **Example 3 (Primes in arithmetic progression)** Suppose that $(q,r)=1$ and $a(n)=1$ when $y<n\le x+y$ and $n\equiv r \pmod q$. Since we already have $(n,q)=1$ we can suppose

$$P=P_q=\prod_{\substack{p\le z\\ p\nmid q}} p.$$

- Then for $d|P$ we have by the Chinese remainder theorem

$$A_d=\frac{x/q}{d}+O(1)$$

and we can take $X=x/q$ and $f(d)=1/d$.

- **Example 4 (Twin primes)** Let $a(n) = 1$ when $n = m(m+2)$ for some $m \le X$ and $a(n) = 0$ otherwise and P as before. Then

$$\pi_2(X) := \sum_{\substack{p \le X \\ p+2 \text{ prime}}} 1 \le \pi(\sqrt{X}) + S(A,P).$$

- **Example 4 (Twin primes)** Let $a(n) = 1$ when $n = m(m+2)$ for some $m \le X$ and $a(n) = 0$ otherwise and P as before. Then

$$\pi_2(X) := \sum_{\substack{p \le X \\ p+2 \text{ prime}}} 1 \le \pi(\sqrt{X}) + S(A, P).$$

- It is easily verified that $A_d = f(d)X + R_d$ holds with $f(d) = \rho(d)/d$, $\rho \in \mathcal{M}$, $\rho(2) = 1$, $\rho(p) = 2$ $(p > 2)$, and with $|R_d| \le \rho(d)$.

- **Example 5 (Goldbach binary problem)** Let X be an even positive integer and let

$$a(n) = \mathrm{card}\{m : n = m(X - m), m < X\}$$

and P as before.

- **Example 5 (Goldbach binary problem)** Let X be an even positive integer and let

$$a(n) = \mathrm{card}\{m : n = m(X-m), m < X\}$$

and P as before.

- Then

$$\mathrm{card}\{p < X : X - p \text{ prime}\} \le 2\pi(\sqrt{X}) + S(\mathcal{A}, P).$$

Again it is easily verified that $A_d = f(d)X + R_d$ holds with $f(d) = \rho(d)/d$, $\rho \in \mathcal{M}$, $\rho(p) = 1$ when $p|X$, $\rho(p) = 2$ when $p \nmid X$, and with $|R_d| \le \rho(d)$ once more.

- **Example 5 (Goldbach binary problem)** Let X be an even positive integer and let

  $$a(n) = \mathrm{card}\{m : n = m(X-m), m < X\}$$

  and P as before.
- Then

  $$\mathrm{card}\{p < X : X - p \text{ prime}\} \le 2\pi(\sqrt{X}) + S(A,P).$$

  Again it is easily verified that $A_d = f(d)X + R_d$ holds with $f(d) = \rho(d)/d$, $\rho \in \mathcal{M}$, $\rho(p) = 1$ when $p|X$, $\rho(p) = 2$ when $p \nmid X$, and with $|R_d| \le \rho(d)$ once more.
- **Example 6** Let $a(n) = 1$ when $n = m^2 + 1$ for some $m \le X$ and $P = \prod_{p \le X}$. Then

  $$\mathrm{card}\{m \le X : m^2 + 1 \text{ prime}\} \le \pi(\sqrt{X}) + S(A,P).$$

  Also $A_d = f(d)X + R_d$ holds with $f(d) = \rho(d)/d$ with $\rho \in \mathcal{M}$ and $\rho(2) = 1$, $\rho(p) = 2$ when $p \equiv 1 \pmod 4$ and $\rho(p) = 0$ otherwise, and $|R_d| \le \rho(d)$.

- A more sophisticated version of Example 4 is

- A more sophisticated version of Example 4 is
- **Example 7 (twin primes revisited)** Let $a(n) = 1$ when $n-2$ is a prime $p \le Y$ and 0 otherwise and let $P = \prod_{p\le\sqrt{Y}} p$. Then

$$\sum_{\substack{p\le Y\\ p+2 \text{ prime}}} 1 \le \pi(\sqrt{Y}) + S(A,P).$$

- A more sophisticated version of Example 4 is
- **Example 7 (twin primes revisited)** Let $a(n) = 1$ when $n-2$ is a prime $p \le Y$ and 0 otherwise and let $P = \prod_{p\le\sqrt{Y}} p$. Then

$$\sum_{\substack{p\le Y\\ p+2 \text{ prime}}} 1 \le \pi(\sqrt{Y}) + S(A,P).$$

- Now $A_d = \pi(Y; d, -2)$ and we have

$$A_d = f(d)X + R_d$$

where $f(d) = 0$ when d is even and $f(d) = \frac{1}{\phi(d)}$ when d is odd, and where now

$$X = \mathrm{li}(Y) = \int_2^Y \frac{dt}{\log t}$$

and where R_d is relatively small ($\ll Y^{\frac{1}{2}+\varepsilon}$ on GRH).

- It is perhaps worth remarking that although in all the cases considered so far one gets a multiplicative function f this is not always the case.

- It is perhaps worth remarking that although in all the cases considered so far one gets a multiplicative function f this is not always the case.
- Romanov proved [1934] that a positive proportion of positive integers can be written as the sum of a prime and a power of 2.

$$\liminf_{x\to\infty} x^{-1}\,\mathrm{card}\{n \le x : n = p + 2^k\} > 0$$

- It is perhaps worth remarking that although in all the cases considered so far one gets a multiplicative function f this is not always the case.
- Romanov proved [1934] that a positive proportion of positive integers can be written as the sum of a prime and a power of 2.

$$\liminf_{x\to\infty} x^{-1}\operatorname{card}\{n\le x : n = p + 2^k\} > 0$$

- The underlying problem with this is that the $\operatorname{ord}_2(q)$, the order of 2 mod q is not a multiplicative function of q. For example if 2 is a primitive root modulo p_1 and p_2 (both odd), then $\operatorname{ord}_2(p_j) = p_j - 1$, but $\operatorname{ord}_2(p_1p_2) = lcm(p_1-1, p_2-1) \le (p_1-1)(p_2-1)/2$.

- Modern sieve theory attempts to overcome the problem of having too many choices for $d|P$ by seeking functions $\lambda_d^{\pm}$ such that
$$\sum_{d|m} \lambda_d^- \le \sum_{d|m} \mu(d) \le \sum_{d|m} \lambda_d^+$$
but the support for the $\lambda_d^{\pm}$ is restricted.

- Modern sieve theory attempts to overcome the problem of having too many choices for $d|P$ by seeking functions $\lambda_d^\pm$ such that
$$\sum_{d|m} \lambda_d^- \le \sum_{d|m} \mu(d) \le \sum_{d|m} \lambda_d^+$$
but the support for the $\lambda_d^\pm$ is restricted.
- We will not be concerned with lower bound sieves, where the theory is more delicate.

- Modern sieve theory attempts to overcome the problem of having too many choices for $d|P$ by seeking functions $\lambda_d^{\pm}$ such that
$$\sum_{d|m}\lambda_d^- \le \sum_{d|m}\mu(d) \le \sum_{d|m}\lambda_d^+$$
but the support for the $\lambda_d^{\pm}$ is restricted.
- We will not be concerned with lower bound sieves, where the theory is more delicate.
- Selberg introduced a very simple and elegant upper bound sieve which is very effective in many situations, and also has the merit of great flexibility. It has also lead to some recent sensational developments.

- Let
$$\lambda_1 = 1$$
and suppose that the $\lambda_q \in \mathbb{R}$ are otherwise at our disposal. Then
$$\sum_{d|m} \mu(d) \le \left(\sum_{d|m} \lambda_d \right)^2 .$$

- Let
$$\lambda_1 = 1$$
and suppose that the $\lambda_q \in \mathbb{R}$ are otherwise at our disposal. Then
$$\sum_{d|m} \mu(d) \le \left(\sum_{d|m} \lambda_d \right)^2 .$$
- In order to retain some structure we suppose that the support $\mathcal{D}$ of the λ_d is a divisor closed set of squarefree numbers.

- Let
$$\lambda_1 = 1$$
and suppose that the $\lambda_q \in \mathbb{R}$ are otherwise at our disposal. Then
$$\sum_{d|m} \mu(d) \le \left(\sum_{d|m} \lambda_d \right)^2 .$$
- In order to retain some structure we suppose that the support $\mathcal{D}$ of the λ_d is a divisor closed set of squarefree numbers.
- Thus for each $d \in \mathcal{D}$, $\mu(d) \neq 0$ and if $q|d$, then $q \in \mathcal{D}$.

- **Example 8** We often suppose in applications that $\mathcal{D} = \{d|P : d \leq D\}$ where $P = \prod_{p \leq Z} p$ for some Z.

- **Example 8** We often suppose in applications that $\mathcal{D} = \{d|P : d \le D\}$ where $P = \prod_{p \le Z} p$ for some Z.
- We recall that

$$S(A,P) = \sum_{\substack{n \\ (n,P)=1}} a(n) \text{ and } \lambda_1 = 1.$$

- **Example 8** We often suppose in applications that $\mathcal{D} = \{d|P : d \le D\}$ where $P = \prod_{p \le Z} p$ for some Z.
- We recall that

$$S(A,P) = \sum_{\substack{n \\ (n,P)=1}} a(n) \text{ and } \lambda_1 = 1.$$

- Thus

$$\begin{aligned} S(A,P) &\le \sum_n a(n) \left(\sum_{d|n} \lambda_d \right)^2 \\ &= \sum_d \sum_e \lambda_d \lambda_e \sum_m a(m[d,e]) \\ &= X \sum_d \sum_e \lambda_d \lambda_e f([d,e]) + R \end{aligned}$$

where

$$R = \sum_d \sum_e \lambda_d \lambda_e R_{[d,e]}.$$

- Thus
$$S(A,P) \le X \sum_d \sum_e \lambda_d \lambda_e f([d,e]) + R$$
where
$$R = \sum_d \sum_e \lambda_d \lambda_e R_{[d,e]}.$$

- Thus
$$S(A,P) \le X \sum_d \sum_e \lambda_d \lambda_e f([d,e]) + R$$
where
$$R = \sum_d \sum_e \lambda_d \lambda_e R_{[d,e]}.$$
- **Example 9** Consider Example 1, $a(n) = 1$ iff $n \in (Y, Y+X]$ with $\mathcal{D}$ as in Example 8. Then
$$|R| \le \left(\sum_d |\lambda_d| \right)^2 \le D^2 \|\lambda\|_\infty^2.$$

- Thus
$$S(A,P) \le X \sum_d \sum_e \lambda_d \lambda_e f([d,e]) + R$$
where
$$R = \sum_d \sum_e \lambda_d \lambda_e R_{[d,e]}.$$
- **Example 9** Consider Example 1, $a(n)=1$ iff $n \in (Y, Y+X]$ with $\mathcal{D}$ as in Example 8. Then
$$|R| \le \left(\sum_d |\lambda_d| \right)^2 \le D^2 \|\lambda\|_\infty^2.$$
- The interesting part is the main term $X\mathcal{F}$ where
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$

- Thus
$$S(A,P) \le X \sum_d \sum_e \lambda_d \lambda_e f([d,e]) + R$$
where
$$R = \sum_d \sum_e \lambda_d \lambda_e R_{[d,e]}.$$
- **Example 9** Consider Example 1, $a(n) = 1$ iff $n \in (Y, Y+X]$ with $\mathcal{D}$ as in Example 8. Then
$$|R| \le \left(\sum_d |\lambda_d| \right)^2 \le D^2 \|\lambda\|_\infty^2.$$
- The interesting part is the main term $X\mathcal{F}$ where
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$
- We want to minimise this subject to the condition $\lambda_1 = 1$.

- We have

$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$

- We have
$$\mathcal{F} = \sum_{d} \sum_{e} \lambda_d \lambda_e f([d,e]).$$
- We want to minimise this subject to the condition $\lambda_1 = 1$.

- We have
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$
- We want to minimise this subject to the condition $\lambda_1 = 1$.
- It is helpful to view $\mathcal{F}$ as a quadratic form in the $\boldsymbol{\lambda}$.

- We have
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$
- We want to minimise this subject to the condition $\lambda_1 = 1$.
- It is helpful to view $\mathcal{F}$ as a quadratic form in the $\boldsymbol{\lambda}$.
- Our first objective is to diagonalise
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]),$$
and this can be done quite easily.

- We have
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$
- We want to minimise this subject to the condition $\lambda_1 = 1$.
- It is helpful to view $\mathcal{F}$ as a quadratic form in the $\boldsymbol{\lambda}$.
- Our first objective is to diagonalise
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]),$$
and this can be done quite easily.
- It is also useful to assume that $\mathcal{D}$ is such that $f(d) \neq 0$ when $d \in \mathcal{D}$.

- We have
$$\mathcal{F} = \sum_{d} \sum_{e} \lambda_d \lambda_e f([d, e]).$$

- We have
$$\mathcal{F} = \sum_d \sum_e \lambda_d \lambda_e f([d,e]).$$
- Write $(d,e) = m$, $d = qm$, $e = rm$, so that $(q,r) = 1$. Since $f \in \mathcal{M}$ and qrm is squarefree we have $f([d,e]) = f(qrm) = f(qm)f(rm)/f(m)$ and
$$\mathcal{F} = \sum_m f(m)^{-1} \sum_q \sum_{\substack{r\\(q,r)=1}} \lambda_{qm}\lambda_{rm} f(qm)f(rm).$$

- We have
$$\mathcal{F} = \sum_{d} \sum_{e} \lambda_d \lambda_e f([d,e]).$$
- Write $(d,e) = m$, $d = qm$, $e = rm$, so that $(q,r) = 1$. Since $f \in \mathcal{M}$ and qrm is squarefree we have $f([d,e]) = f(qrm) = f(qm)f(rm)/f(m)$ and
$$\mathcal{F} = \sum_{m} f(m)^{-1} \sum_{q} \sum_{\substack{r \\ (q,r)=1}} \lambda_{qm}\lambda_{rm} f(qm) f(rm).$$
- Now we use the Möbius function to remove the condition $(q,r) = 1$.

- We have

$$\mathcal{F} = \sum_m f(m)^{-1} \sum_q \sum_{\substack{r \\ (q,r)=1}} \lambda_{qm}\lambda_{rm} f(qm)f(rm).$$

- We have

$$\mathcal{F} = \sum_m f(m)^{-1} \sum_q \sum_{\substack{r \\ (q,r)=1}} \lambda_{qm}\lambda_{rm} f(qm) f(rm).$$

- Thus

$$\mathcal{F} = \sum_m f(m)^{-1} \sum_l \mu(l) \left(\sum_d \lambda_{dlm} f(dlm) \right)^2 .$$

- We have

$$\mathcal{F} = \sum_m f(m)^{-1} \sum_q \sum_{\substack{r \\ (q,r)=1}} \lambda_{qm}\lambda_{rm} f(qm) f(rm).$$

- Thus

$$\mathcal{F} = \sum_m f(m)^{-1} \sum_l \mu(l) \left(\sum_d \lambda_{dlm} f(dlm) \right)^2 .$$

- Now we collect together the terms with $lm = n$ and observe that by multiplicativity we have

$$\sum_{\substack{l,m \\ lm=n}} f(m)^{-1} \mu(l) = \prod_{p|n} \frac{1-f(p)}{f(p)}.$$

- We have

$$\sum_{\substack{l,m\\ lm=n}} f(m)^{-1}\mu(l) = \prod_{p|n} \frac{1-f(p)}{f(p)}.$$

- We have

$$\sum_{\substack{l,m\\lm=n}} f(m)^{-1}\mu(l) = \prod_{p|n} \frac{1-f(p)}{f(p)}.$$

- Denoting this expression by $g(n)^{-1}$ we have

$$\mathcal{F} = \sum_n g(n)^{-1}\left(\sum_d \lambda_{dn} f(dn)\right)^2.$$

where

$$g(n) = \prod_{p|n} \frac{f(p)}{1-f(p)}$$

- We have
$$\sum_{\substack{l,m\\ lm=n}} f(m)^{-1}\mu(l)=\prod_{p|n}\frac{1-f(p)}{f(p)}.$$
- Denoting this expression by $g(n)^{-1}$ we have
$$\mathcal{F}=\sum_n g(n)^{-1}\left(\sum_d \lambda_{dn}f(dn)\right)^2.$$
where
$$g(n)=\prod_{p|n}\frac{f(p)}{1-f(p)}.$$
- Let
$$\omega_n=\sum_d \lambda_{dn}f(dn)\quad (n\in\mathcal{D}).$$

- We have

$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2,$$

$$\omega_n = \sum_d \lambda_{dn} f(dn) \quad (n \in \mathcal{D})$$

- We have
$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2,$$
$$\omega_n = \sum_d \lambda_{dn} f(dn) \quad (n \in \mathcal{D})$$
- There is a bijection between the $\boldsymbol{\lambda}$ and the $\boldsymbol{\omega}$. We could view the transformation from the one to the other as being by an upper triangular matrix, which is obviously invertible.

- We have
$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2,$$
$$\omega_n = \sum_d \lambda_{dn} f(dn) \quad (n \in \mathcal{D})$$
- There is a bijection between the $\boldsymbol{\lambda}$ and the $\boldsymbol{\omega}$. We could view the transformation from the one to the other as being by an upper triangular matrix, which is obviously invertible.
- There is a standard number theoretic way of expressing the inversion. Consider
$$\sum_n \omega_{nm}\mu(n) = \sum_n \sum_d \lambda_{dn} f(dnm)\mu(n)$$

- We have
$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2,$$
$$\omega_n = \sum_d \lambda_{dn} f(dn) \quad (n \in \mathcal{D})$$
- There is a bijection between the $\boldsymbol{\lambda}$ and the $\boldsymbol{\omega}$. We could view the transformation from the one to the other as being by an upper triangular matrix, which is obviously invertible.
- There is a standard number theoretic way of expressing the inversion. Consider
$$\sum_n \omega_{nm}\mu(n) = \sum_n \sum_d \lambda_{dn} f(dnm)\mu(n)$$
- Collecting together the terms with $nd = q$ this becomes, for $m \in \mathcal{D}$,
$$\sum_q \lambda_{qm} f(qm) \sum_{n|q} \mu(n) = \lambda_m f(m).$$

- Hence

$$g(n)=\prod_{p|n}\frac{f(p)}{1-f(p)}$$

$$\mathcal{F}=\sum_n g(n)^{-1}\left(\sum_d \lambda_{dn}f(dn)\right)^2$$

$$\omega_n=\sum_d \lambda_{dn}f(dn)\quad (n\in\mathcal{D})$$

$$\lambda_m f(m)=\sum_n \omega_{nm}\mu(n)\quad (m\in\mathcal{D})$$

- Hence

$$g(n) = \prod_{p|n} \frac{f(p)}{1-f(p)}$$

$$\mathcal{F} = \sum_n g(n)^{-1} \left(\sum_d \lambda_{dn} f(dn) \right)^2$$

$$\omega_n = \sum_d \lambda_{dn} f(dn) \quad (n \in \mathcal{D})$$

$$\lambda_m f(m) = \sum_n \omega_{nm} \mu(n) \quad (m \in \mathcal{D})$$

- Thus we are seeking to minimise

$$\mathcal{F} = \sum_n g(n)^{-1} \omega_n^2 \text{ under } \sum_n \omega_n \mu(n) = \lambda_1 = 1.$$

- Thus we are seeking to minimise

$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2 \text{ under } \sum_n \omega_n \mu(n) = \lambda_1 = 1.$$

- Thus we are seeking to minimise

$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2 \text{ under } \sum_n \omega_n \mu(n) = \lambda_1 = 1.$$

- Let $\theta = 1/\sum_{n\in\mathcal{D}} g(n)$. Then $\mathcal{F} =$

$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + 2\theta \sum_n \omega_n\mu(n) - \theta^2 \sum_n g(n)$$

$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + \theta.$$

- Thus we are seeking to minimise

$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2 \text{ under } \sum_n \omega_n\mu(n) = \lambda_1 = 1.$$

- Let $\theta = 1/\sum_{n\in\mathcal{D}} g(n)$. Then $\mathcal{F} =$

$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + 2\theta\sum_n \omega_n\mu(n) - \theta^2\sum_n g(n)$$

$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + \theta.$$

- Obviously $\mathcal{F} \ge \theta$

- Thus we are seeking to minimise

$$\mathcal{F} = \sum_n g(n)^{-1}\omega_n^2 \text{ under } \sum_n \omega_n \mu(n) = \lambda_1 = 1.$$

- Let $\theta = 1/\sum_{n\in\mathcal{D}} g(n)$. Then $\mathcal{F} =$

$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + 2\theta \sum_n \omega_n\mu(n) - \theta^2 \sum_n g(n)$$
$$= \sum_{n\in\mathcal{D}} \frac{(\omega_n - \theta\mu(n)g(n))^2}{g(n)} + \theta.$$

- Obviously $\mathcal{F} \geq \theta$
- and the choice

$$\omega_n = \theta\mu(n)g(n)$$

gives

$$\sum_n \omega_n\mu(n) = 1 \text{ and } \mathcal{F} = \theta.$$

- We have just shown that the minimum of $\mathcal{F}$ is θ and the minimum is attained when

$$\omega_n = \theta\mu(n)g(n)$$

- We have just shown that the minimum of $\mathcal{F}$ is θ and the minimum is attained when

$$\omega_n = \theta\mu(n)g(n)$$

- We can now invert the transform to deduce the minimising λ_m

- We have just shown that the minimum of $\mathcal{F}$ is θ and the minimum is attained when

$$\omega_n = \theta\mu(n)g(n)$$

- We can now invert the transform to deduce the minimising λ_m
- Recall that

$$\lambda_m f(m) = \sum_n \omega_{nm}\mu(n) \quad (m \in \mathcal{D}).$$

- We have just shown that the minimum of $\mathcal{F}$ is θ and the minimum is attained when

$$\omega_n = \theta\mu(n)g(n)$$

- We can now invert the transform to deduce the minimising λ_m
- Recall that

$$\lambda_m f(m) = \sum_n \omega_{nm}\mu(n) \quad (m \in \mathcal{D}).$$

- Thus the minimising λ_m are given by

$$\begin{aligned}\lambda_m &= \frac{\theta}{f(m)} \sum_n g(mn)\mu(mn)\mu(n)\\ &= \theta\mu(m)\frac{g(m)}{f(m)} \sum_{\substack{n\\ nm\in\mathcal{D}}} g(n).\end{aligned}$$

- We have $\lambda_m = \theta\mu(m)\dfrac{g(m)}{f(m)}\displaystyle\sum_{\substack{n\\ nm\in\mathcal{D}}} g(n).$

- We have $\lambda_m = \theta\mu(m)\dfrac{g(m)}{f(m)}\displaystyle\sum_{\substack{n\\ nm\in\mathcal{D}}} g(n).$
- You might not think there is any reason to care about the actual values of the λ_m, since the minimum seems to be the crucial thing.

- We have $\lambda_m = \theta\mu(m)\dfrac{g(m)}{f(m)}\displaystyle\sum_{\substack{n\\ nm\in\mathcal{D}}} g(n).$
- You might not think there is any reason to care about the actual values of the λ_m, since the minimum seems to be the crucial thing.
- However the λ_m also occur in the error term.

- We have $\lambda_m = \theta\mu(m)\dfrac{g(m)}{f(m)} \displaystyle\sum_{\substack{n \\ nm\in\mathcal{D}}} g(n)$.
- You might not think there is any reason to care about the actual values of the λ_m, since the minimum seems to be the crucial thing.
- However the λ_m also occur in the error term.
- Write $\dfrac{g(m)}{f(m)} = \displaystyle\prod_{p|m} \dfrac{1}{1-f(p)} = \prod_{p|m}(1+g(p)) = \sum_{d|m} g(d)$.

- We have $\lambda_m = \theta\mu(m)\frac{g(m)}{f(m)}\sum_{\substack{n\\ nm\in\mathcal{D}}} g(n)$.
- You might not think there is any reason to care about the actual values of the λ_m, since the minimum seems to be the crucial thing.
- However the λ_m also occur in the error term.
- Write $\frac{g(m)}{f(m)} = \prod_{p|m}\frac{1}{1-f(p)} = \prod_{p|m}(1+g(p)) = \sum_{d|m} g(d)$.
- Thus

$$|\lambda_m| \le \theta\sum_{d|m} g(d) \sum_{\substack{n\\ nd\in\mathcal{D}\\ (n,m/d)=1}} g(n)$$
$$= \theta\sum_{d|m}\sum_{\substack{k\\ (k,m)=d}} g(k) = 1.$$

- We have

$$\lambda_m = \theta\mu(m)\frac{g(m)}{f(m)}\sum_{\substack{n\\ nm\in\mathcal{D}}} g(n)$$

Theorem 1 (Selberg)

Suppose that $a : \mathbb{Z} \to \mathbb{R}^+$, $A_d = \sum_n a(dn)$ and that $A_d = f(d)X + R_d$ where $f \in \mathcal{M}$ and $0 \le f(p) < 1$. Let $P \in \mathbb{N}$ be squarefree and $\mathcal{D}$ be a divisor closed subset of the divisors of P. Then

$$S(A,P) \le \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \sum_{d\in\mathcal{D}}\sum_{e\in\mathcal{D}} \lambda_d\lambda_e R_{[d,e]}$$

where $g(n) = \prod_{p|n} \frac{f(p)}{1-f(p)}$. Moreover

$$|\lambda_d| \le 1.$$

- Armed with Selberg's theorem we revisit the various examples.

- Armed with Selberg's theorem we revisit the various examples.
- **Example 10 Primes in arithmetic progressions**
  Following example 3, when $(q,r)=1$

  $$\pi(x+y;q,r)-\pi(y;q,r)\le\pi(\sqrt{X})+S(A,P)$$

  where $a(n)=1$ when $y<n\le x+y$, $n\equiv r \pmod q$ and $a(n)=0$ otherwise, $X=x/q$, $P=P_q=\prod_{\substack{p\le\sqrt{X}\\ p\nmid q}} p$,

  $A_d=\frac{X}{d}+R_d$, $|R_d|\le 1$.

- Armed with Selberg's theorem we revisit the various examples.
- **Example 10 Primes in arithmetic progressions** Following example 3, when $(q,r)=1$
$$\pi(x+y;q,r)-\pi(y;q,r)\le\pi(\sqrt{X})+S(A,P)$$
where $a(n)=1$ when $y<n\le x+y$, $n\equiv r\pmod q$ and $a(n)=0$ otherwise, $X=x/q$, $P=P_q=\prod_{\substack{p\le\sqrt{X}\\ p\nmid q}}p$, $A_d=\frac{X}{d}+R_d$, $|R_d|\le 1$.
- Thus $f(d)=1/d$. Let $\mathcal{D}=\{d|P : d\le D\}$ with $D\le\sqrt{X}$.

- Armed with Selberg's theorem we revisit the various examples.
- **Example 10 Primes in arithmetic progressions**
  Following example 3, when $(q,r)=1$

  $$\pi(x+y;q,r)-\pi(y;q,r)\le\pi(\sqrt{X})+S(A,P)$$

  where $a(n)=1$ when $y<n\le x+y$, $n\equiv r \pmod q$ and $a(n)=0$ otherwise, $X=x/q$, $P=P_q=\prod_{\substack{p\le\sqrt{X}\\ p\nmid q}}p$, $A_d=\frac{X}{d}+R_d$, $|R_d|\le 1$.
- Thus $f(d)=1/d$. Let $\mathcal{D}=\{d|P:d\le D\}$ with $D\le\sqrt{X}$.
- Then, for $d\in\mathcal{D}$,

  $$g(d)=\prod_{p|d}\frac{1/p}{1-1/p}=\frac{1}{\phi(d)}$$

  $$\sum_{d\in\mathcal{D}}g(d)=\sum_{d\le D,(d,q)=1}\frac{\mu(d)^2}{\phi(d)}$$

  and $|\lambda_d|\le 1$.

- **Example 10 continued**

$$\pi(x+y)-\pi(y)\le \pi(\sqrt{X})+S(A,P)$$

$$S(A,P)\le \frac{X}{\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}}+D^2$$

- **Example 10 continued**

$$\pi(x+y)-\pi(y)\le\pi(\sqrt{X})+S(A,P)$$

$$S(A,P)\le\frac{X}{\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}}+D^2$$

- We need a lower bound for

$$\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}$$

- We need a lower bound for $\displaystyle\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}$.

- We need a lower bound for $\displaystyle\sum_{\substack{d\le D\\(d,q)=1}} \frac{\mu(d)^2}{\phi(d)}$.
- First we remove the condition $(d,q)=1$. We have

$$\frac{q}{\phi(q)} = \prod_{p|q}\left(1+\frac{1}{p-1}\right) = \sum_{m|q}\frac{\mu(m)^2}{\phi(m)}$$

- We need a lower bound for $\displaystyle\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}$.
- First we remove the condition $(d,q)=1$. We have

$$\frac{q}{\phi(q)}=\prod_{p|q}\left(1+\frac{1}{p-1}\right)=\sum_{m|q}\frac{\mu(m)^2}{\phi(m)}$$

- Thus

$$\frac{q}{\phi(q)}\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}\ge\sum_{\substack{d\le D\\(d,q)=1}}\frac{\mu(d)^2}{\phi(d)}\sum_{\substack{m\le D/d\\m|q}}\frac{\mu(m)^2}{\phi(m)}$$
$$=\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}$$

- Now we need a lower bound for

$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{n\le D}\frac{\mu(n)^2}{n}\prod_{p|n}\frac{p}{p-1}.$$

- Now we need a lower bound for
$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{n\le D}\frac{\mu(n)^2}{n}\prod_{p|n}\frac{p}{p-1}.$$
- Here we have for the general term, when n is squarefree,
$$\frac{1}{n}\prod_{p|n}\left(\sum_{j=0}^{\infty}\frac{1}{p^j}\right)=\sum_{\substack{m\\ s(m)=n}}\frac{1}{m}$$
where $s(m)$ is the squarefree kernel of m, $s(m)=\prod_{p|m}p$.

- Now we need a lower bound for
$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{n\le D}\frac{\mu(n)^2}{n}\prod_{p|n}\frac{p}{p-1}.$$
- Here we have for the general term, when n is squarefree,
$$\frac{1}{n}\prod_{p|n}\left(\sum_{j=0}^{\infty}\frac{1}{p^j}\right)=\sum_{\substack{m\\ s(m)=n}}\frac{1}{m}$$
where $s(m)$ is the squarefree kernel of m, $s(m)=\prod_{p|m}p$.
- Thus, by Euler,
$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{\substack{m\\ s(m)\le D}}\frac{1}{m}\ge\sum_{m\le D}\frac{1}{m}\ge\log D.$$

- Now we need a lower bound for
$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{n\le D}\frac{\mu(n)^2}{n}\prod_{p|n}\frac{p}{p-1}.$$
- Here we have for the general term, when n is squarefree,
$$\frac{1}{n}\prod_{p|n}\left(\sum_{j=0}^{\infty}\frac{1}{p^j}\right)=\sum_{\substack{m\\ s(m)=n}}\frac{1}{m}$$
where $s(m)$ is the squarefree kernel of m, $s(m)=\prod_{p|m}p$.
- Thus, by Euler,
$$\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}=\sum_{\substack{m\\ s(m)\le D}}\frac{1}{m}\ge\sum_{m\le D}\frac{1}{m}\ge\log D.$$
- Hence we just showed that $\displaystyle\sum_{\substack{d\le D\\ (d,q)=1}}\frac{\mu(d)^2}{\phi(d)}\ge\frac{\phi(q)}{q}\log D.$

- We just showed the very neat upper bound

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{x}{\phi(q)\log D}+\pi(\sqrt{x/q})+D^2.$$

- We just showed the very neat upper bound

$$\pi(x+y;q,r)-\pi(y;q,r)\le \frac{x}{\phi(q)\log D}+\pi(\sqrt{x/q})+D^2.$$

- When $x\ge e^2q$, a close to optimal choice for D is

$$D=\frac{\sqrt{x/q}}{\log\sqrt{x/q}}$$

- We just showed the very neat upper bound

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{x}{\phi(q)\log D}+\pi(\sqrt{x/q})+D^2.$$

- When $x\ge e^2q$, a close to optimal choice for D is

$$D=\frac{\sqrt{x/q}}{\log\sqrt{x/q}}$$

- and this leads to

Theorem 2 (Brun-Titchmarsh)

Suppose that $(q,r)=1$ *and* $y\ge e^2q$. *Then*

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}+O\left(\frac{x\log\log\frac{x}{q}}{\phi(q)\log^2\frac{x}{q}}\right).$$

- We have

$$\pi(x+y;q,r)-\pi(y;q,r) \le \frac{2x}{\phi(q)\log\frac{x}{q}} + O\left(\frac{x\log\log\frac{x}{q}}{\phi(q)\log^2\frac{x}{q}}\right).$$

- We have
$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}+O\left(\frac{x\log\log\frac{x}{q}}{\phi(q)\log^2\frac{x}{q}}\right).$$

- The Generalised Riemann Hypothesis gives nothing when $x<\phi(q)\sqrt{x}$, so this is a really useful result.

- We have

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}+O\left(\frac{x\log\log\frac{x}{q}}{\phi(q)\log^2\frac{x}{q}}\right).$$

- The Generalised Riemann Hypothesis gives nothing when $x<\phi(q)\sqrt{x}$, so this is a really useful result.
- The best general result we know is, for $(q,r)=1$,

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}$$

(Montgomery & RCV 1972).

- We have

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}+O\left(\frac{x\log\log\frac{x}{q}}{\phi(q)\log^2\frac{x}{q}}\right).$$

- The Generalised Riemann Hypothesis gives nothing when $x<\phi(q)\sqrt{x}$, so this is a really useful result.
- The best general result we know is, for $(q,r)=1$,

$$\pi(x+y;q,r)-\pi(y;q,r)\le\frac{2x}{\phi(q)\log\frac{x}{q}}$$

(Montgomery & RCV 1972).

- If one could prove this with the 2 replaced by any smaller constant, then one could establish something very profound about zeros of L-functions, namely that Siegel zeros do not exist.

- When $q = 1$, the optimising choice of λ_m in the proof of the Brun–Titchmarsh theorem is

$$\lambda_m = \mu(m)m\phi(m)^{-1}\frac{\sum_{\substack{n\le D/m\\(n,m)=1}}\frac{\mu(n)^2}{\phi(n)}}{\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}}.$$

- When $q=1$, the optimising choice of λ_m in the proof of the Brun–Titchmarsh theorem is

$$\lambda_m = \mu(m)m\phi(m)^{-1}\frac{\sum_{\substack{n\le D/m\\(n,m)=1}}\frac{\mu(n)^2}{\phi(n)}}{\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}}.$$

- The sum in the denominator is asymptotically $\log D$ and, at least when m is not too close to D, the sum in the numerator ought to be asymptotically $\phi(m)m^{-1}\log(D/m)$.

- When $q = 1$, the optimising choice of λ_m in the proof of the Brun–Titchmarsh theorem is

$$\lambda_m = \mu(m)m\phi(m)^{-1}\frac{\sum_{\substack{n\le D/m\\(n,m)=1}}\frac{\mu(n)^2}{\phi(n)}}{\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}}.$$

- The sum in the denominator is asymptotically $\log D$ and, at least when m is not too close to D, the sum in the numerator ought to be asymptotically $\phi(m)m^{-1}\log(D/m)$.
- Thus λ_m should be close to

$$\lambda_m^* = \mu(m)\frac{\log D/m}{\log D}.$$

- When $q = 1$, the optimising choice of λ_m in the proof of the Brun–Titchmarsh theorem is

$$\lambda_m = \mu(m)m\phi(m)^{-1}\frac{\sum_{\substack{n\le D/m\\(n,m)=1}} \frac{\mu(n)^2}{\phi(n)}}{\sum_{n\le D} \frac{\mu(n)^2}{\phi(n)}}.$$

- The sum in the denominator is asymptotically $\log D$ and, at least when m is not too close to D, the sum in the numerator ought to be asymptotically $\phi(m)m^{-1}\log(D/m)$.
- Thus λ_m should be close to

$$\lambda_m^* = \mu(m)\frac{\log D/m}{\log D}.$$

- Indeed λ_m^* can be used instead of the optimal choice, although there is more work involved in the analysis to push things through.

- When $q = 1$, the optimising choice of λ_m in the proof of the Brun–Titchmarsh theorem is
$$\lambda_m = \mu(m)m\phi(m)^{-1}\frac{\sum_{\substack{n\le D/m\\(n,m)=1}}\frac{\mu(n)^2}{\phi(n)}}{\sum_{n\le D}\frac{\mu(n)^2}{\phi(n)}}.$$
- The sum in the denominator is asymptotically $\log D$ and, at least when m is not too close to D, the sum in the numerator ought to be asymptotically $\phi(m)m^{-1}\log(D/m)$.
- Thus λ_m should be close to
$$\lambda_m^* = \mu(m)\frac{\log D/m}{\log D}.$$
- Indeed λ_m^* can be used instead of the optimal choice, although there is more work involved in the analysis to push things through.
- Later, we will see situations where the optimal choice is not known but a choice of this kind is still effective.

- Let me turn now to the twin prime problem, which we looked at in Example 4.

- Let me turn now to the twin prime problem, which we looked at in Example 4.
- Now we have

$$\pi_2(X) \le \pi(\sqrt{X}) + \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \rho([d,e])$$

where $\rho, f, g \in \mathcal{M}$, $f(d) = \rho(d)/d$, $g(p) = f(p)/(p - f(p))$, $\rho(2) = 1$, $\rho(p) = 2$ $(p > 2)$,

- Let me turn now to the twin prime problem, which we looked at in Example 4.
- Now we have

$$\pi_2(X) \le \pi(\sqrt{X}) + \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \sum_{d\in\mathcal{D}}\sum_{e\in\mathcal{D}} \rho([d,e])$$

where $\rho, f, g \in \mathcal{M}$, $f(d) = \rho(d)/d$,
$g(p) = f(p)/(p - f(p))$, $\rho(2) = 1$, $\rho(p) = 2$ $(p > 2)$,

- so that

$$g(2) = 1, \quad g(p) = \frac{2}{p-2} \quad (p > 2).$$

- Let me turn now to the twin prime problem, which we looked at in Example 4.
- Now we have

$$\pi_2(X) \le \pi(\sqrt{X}) + \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \rho([d,e])$$

  where $\rho, f, g \in \mathcal{M}$, $f(d) = \rho(d)/d$, $g(p) = f(p)/(p - f(p))$, $\rho(2) = 1$, $\rho(p) = 2$ $(p > 2)$,
- so that

$$g(2) = 1, \quad g(p) = \frac{2}{p-2} \quad (p > 2).$$

- Thus our first task is to understand

$$\sum_{n\le D} \mu(n)^2 \prod_{\substack{p|n\\ p>2}} \frac{2}{p-2}.$$

- We want to understand $\displaystyle\sum_{n\le D}\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$.

- We want to understand $\sum_{n\le D}\mu(n)^2\prod_{\substack{p|n\\ p>2}}\frac{2}{p-2}$.
- The general term behaves a bit like $\frac{d(n)}{n}$.

- We want to understand $\displaystyle\sum_{n\le D}\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$.
- The general term behaves a bit like $\frac{d(n)}{n}$.
- The sum $\displaystyle\sum_{n\le x}\frac{d(n)}{n}$ can be dealt with by various methods.

- We want to understand $\displaystyle\sum_{n\le D}\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$.
- The general term behaves a bit like $\frac{d(n)}{n}$.
- The sum $\displaystyle\sum_{n\le x}\frac{d(n)}{n}$ can be dealt with by various methods.
- For one write, with $E(t)\ll\sqrt{t}$, $\displaystyle\sum_{n\le x}d(n)\left(\frac{1}{x}+\int_n^x\frac{dt}{t^2}\right)$

$$=\frac{\sum_{n\le x}d(n)}{x}+\int_1^x\frac{\sum_{n\le t}d(n)}{t^2}dt$$

$$=\log x+2\gamma-1+\frac{E(x)}{x}$$

$$+\int_1^x\left(\log t+2\gamma-1+\frac{E(t)}{t}\right)\frac{dt}{t}.$$

- Thus $\displaystyle\sum_{n\le x}\frac{d(n)}{n}=\frac{(\log x)^2}{2}+2\gamma\log x+c_1+O\left(\frac{1}{\sqrt{x}}\right)$.

- We want to understand $\sum_{n\le D} g(d)$ where $g(n)=\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$. We know that

$$\sum_{n\le x}\frac{d(n)}{n}=\frac{1}{2}(\log x)^2+2\gamma\log x+c_1+O(x^{-1/2}).$$

- We want to understand $\sum_{n\le D} g(d)$ where $g(n)=\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$. We know that

$$\sum_{n\le x}\frac{d(n)}{n}=\frac{1}{2}(\log x)^2+2\gamma\log x+c_1+O(x^{-1/2}).$$

- How to get from one sum to the other. Dirichlet convolution!

- We want to understand $\sum_{n\le D} g(d)$ where $g(n)=\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$. We know that
$$\sum_{n\le x}\frac{d(n)}{n}=\frac{1}{2}(\log x)^2+2\gamma\log x+c_1+O(x^{-1/2}).$$
- How to get from one sum to the other. Dirichlet convolution!
- We want to find a function h so that $Ng = d*h$.

- We want to understand $\sum_{n\le D} g(d)$ where $g(n)=\mu(n)^2\prod_{\substack{p|n\\p>2}}\frac{2}{p-2}$. We know that

$$\sum_{n\le x}\frac{d(n)}{n}=\frac{1}{2}(\log x)^2+2\gamma\log x+c_1+O(x^{-1/2}).$$

- How to get from one sum to the other. Dirichlet convolution!
- We want to find a function h so that $Ng=d*h$.
- Recall that $d=1*1$ and $1*\mu=e$. Hence $h=(Ng)*\mu_2$ where we have written $\mu_2=\mu*\mu$.

- Recall that $d = 1 * 1$ and $1 * \mu = e$. Hence $h = (Ng) * \mu_2$ where we have written $\mu_2 = \mu * \mu$.

- Recall that $d = 1 * 1$ and $1 * \mu = e$. Hence $h = (Ng) * \mu_2$ where we have written $\mu_2 = \mu * \mu$.
- We find that μ_2, $h \in \mathcal{M}$ and

$$\mu_2(p) = -2, \mu_2(p^2) = 1, \mu(p^k) = 0 \quad (k > 2),$$

$$h(2) = 0, h(4) = -3, h(8) = 2, h(p^k) = 0 \quad (k > 3)$$

and for $p > 2$,

$$h(p) = \frac{4}{p-2}, h(p^2) = -\frac{3p+2}{p-2},$$

$$h(p^3) = \frac{2p}{p-2}, h(p^k) = 0(k > 3).$$

- Recall that $d = 1 * 1$ and $1 * \mu = e$. Hence $h = (Ng) * \mu_2$ where we have written $\mu_2 = \mu * \mu$.
- We find that μ_2, $h \in \mathcal{M}$ and

$$\mu_2(p) = -2, \mu_2(p^2) = 1, \mu(p^k) = 0 \quad (k > 2),$$

$$h(2) = 0, h(4) = -3, h(8) = 2, h(p^k) = 0 \quad (k > 3)$$

and for $p > 2$,

$$h(p) = \frac{4}{p-2}, h(p^2) = -\frac{3p+2}{p-2},$$

$$h(p^3) = \frac{2p}{p-2}, h(p^k) = 0(k > 3).$$

- We have

$$g(n) = \sum_{m|n} \frac{d(m)}{m} \cdot \frac{h(n/m)}{n/m}$$

- Recall that $d = 1 * 1$ and $1 * \mu = e$. Hence $h = (Ng) * \mu_2$ where we have written $\mu_2 = \mu * \mu$.
- We find that μ_2, $h \in \mathcal{M}$ and

$$\mu_2(p) = -2, \mu_2(p^2) = 1, \mu(p^k) = 0 \quad (k > 2),$$

$$h(2) = 0, h(4) = -3, h(8) = 2, h(p^k) = 0 \quad (k > 3)$$

and for $p > 2$,

$$h(p) = \frac{4}{p-2}, h(p^2) = -\frac{3p+2}{p-2},$$

$$h(p^3) = \frac{2p}{p-2}, h(p^k) = 0 (k > 3).$$

- We have

$$g(n) = \sum_{m|n} \frac{d(m)}{m} \cdot \frac{h(n/m)}{n/m}$$

- Thus

$$\sum_{n \le D} g(n) = \sum_{l \le D} \frac{h(l)}{l} \sum_{m \le D/l} \frac{d(m)}{m}.$$

- Thus
$$\sum_{n\le D} g(n) = \sum_{l\le D} \frac{h(l)}{l} \sum_{m\le D/l} \frac{d(m)}{m}.$$

- Thus
$$\sum_{n\le D} g(n) = \sum_{l\le D} \frac{h(l)}{l} \sum_{m\le D/l} \frac{d(m)}{m}.$$
- Now we substitute in our approximation
$$\frac{1}{2}(\log D/l)^2 + 2\gamma\log(D/l) + c_1 + O(l^{1/2}D^{-1/2})$$
for the inner sum.

- Thus
$$\sum_{n\le D} g(n) = \sum_{l\le D} \frac{h(l)}{l} \sum_{m\le D/l} \frac{d(m)}{m}.$$
- Now we substitute in our approximation
$$\frac{1}{2}(\log D/l)^2 + 2\gamma\log(D/l) + c_1 + O(l^{1/2}D^{-1/2})$$
for the inner sum.
- It turns out that the various sums over l which occur are nicely convergent and we obtain
$$\sum_{n\le D} g(n) = \frac{1}{2}(\log D)^2 \sum_{l=1}^{\infty} \frac{h(l)}{l} + O(\log D).$$

- Thus
$$\sum_{n\le D} g(n) = \sum_{l\le D} \frac{h(l)}{l} \sum_{m\le D/l} \frac{d(m)}{m}.$$
- Now we substitute in our approximation
$$\frac{1}{2}(\log D/l)^2 + 2\gamma \log(D/l) + c_1 + O(l^{1/2}D^{-1/2})$$
for the inner sum.
- It turns out that the various sums over l which occur are nicely convergent and we obtain
$$\sum_{n\le D} g(n) = \frac{1}{2}(\log D)^2 \sum_{l=1}^{\infty} \frac{h(l)}{l} + O(\log D).$$
- The infinite sum here is
$$\prod_p \left(1 + \sum_{j=1}^{3} \frac{h(p^j)}{p^j}\right) = \left(1 - \frac{3}{4} + \frac{1}{4}\right) \prod_{p>2} \left(\frac{(p-1)^2}{p(p-2)}\right).$$

- We have just proved that
$$\frac{x}{\sum_{n\le D} g(n)} = \frac{2C_2x}{(\log D)^2} + O(x(\log D)^{-3})$$
where
$$C_2 = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right).$$

- We have just proved that

$$\frac{x}{\sum_{n\le D} g(n)} = \frac{2C_2x}{(\log D)^2} + O(x(\log D)^{-3})$$

where

$$C_2 = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right).$$

- and so

$$\pi_2(x) \le \frac{2C_2x}{(\log D)^2} + \pi(\sqrt{x}) + \sum_{d,e\le D}\mu(d)^2\mu(e)^2\rho([d,e]) + O(x(\log D)^{-3}).$$

- We have just proved that

$$\frac{x}{\sum_{n\le D} g(n)} = \frac{2C_2 x}{(\log D)^2} + O(x(\log D)^{-3})$$

where

$$C_2 = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right).$$

- and so

$$\pi_2(x) \le \frac{2C_2 x}{(\log D)^2} + \pi(\sqrt{x}) + \sum_{d,e\le D} \mu(d)^2\mu(e)^2\rho([d,e]) + O(x(\log D)^{-3}).$$

- We know that $\rho(m)\le d(m)$ and it is easily seen that $d([d,e])\le d(d)d(e)$ and so the sum here is

$$\ll (D\log D)^2.$$

- We have

$$\pi_2(x) \le \frac{2C_2x}{(\log D)^2} + O(\sqrt{x} + (D\log D)^2 + x(\log D)^{-3}).$$

- We have

$$\pi_2(x) \le \frac{2C_2x}{(\log D)^2} + O(\sqrt{x} + (D\log D)^2 + x(\log D)^{-3}).$$

- Let $D = x^{1/2}(\log x)^{-3}$. Then we have

$$\pi_2(x) \le \frac{8C_2x}{(\log x)^2} + O\left(\frac{x\log\log x}{(\log x)^3}\right).$$

- Hardy and Littlewood (1923) conjectured that

$$\pi_2(x) \sim \frac{C_2x}{(\log x)^2}$$

- We have

$$\pi_2(x) \le \frac{2C_2x}{(\log D)^2} + O(\sqrt{x} + (D\log D)^2 + x(\log D)^{-3}).$$

- Let $D = x^{1/2}(\log x)^{-3}$. Then we have

$$\pi_2(x) \le \frac{8C_2x}{(\log x)^2} + O\left(\frac{x\log\log x}{(\log x)^3}\right).$$

- Hardy and Littlewood (1923) conjectured that

$$\pi_2(x) \sim \frac{C_2x}{(\log x)^2}$$

- The constant

$$C_2 = 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right)$$

is known as the twin prime constant.

- Selberg's theorem applied to Example 6 gives

$$Q(X) \le \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \pi(\sqrt{X}) + \sum_{d,e\in\mathcal{D}} \rho([d,e])$$

where $Q(X) = \text{card}\{m \le X : m^2+1 \text{ prime}\}$, $\rho(p) = 1 + \chi_1(p)$, $g(p) = \rho(p)/\big(p - \rho(p)\big)$.

- Selberg's theorem applied to Example 6 gives

$$Q(X) \le \frac{X}{\sum_{n\in\mathcal{D}} g(n)} + \pi(\sqrt{X}) + \sum_{d,e\in\mathcal{D}} \rho([d,e])$$

where $Q(X) = \mathrm{card}\{m \le X : m^2+1 \text{ prime}\}$, $\rho(p) = 1+\chi_1(p)$, $g(p) = \rho(p)/\big(p-\rho(p)\big)$.

- A somewhat more complex analysis to the previous ones, gives

$$\sum_{n\in\mathcal{D}} g(n) = C_1^{-1}\log D + O(1)$$

where

$$C_1 = \frac{\pi}{4}\prod_p\left(1+\frac{\chi_1(p)}{p\big(p-1-\chi_1(p)\big)}\right)$$

$$= \frac{\pi}{4}\prod_p\left(1-\frac{\rho(p)-1}{p\big(p-\rho(p)\big)}\right).$$

- We have
$$\sum_{n\in\mathcal{D}} g(n) = C_1^{-1}\log D + O(1)$$

- We have
$$\sum_{n\in\mathcal{D}} g(n) = C_1^{-1}\log D + O(1)$$
- Thus
$$Q(X) \le \frac{2C_1X}{\log X} + O\big(X(\log\log X)(\log X)^{-2}\big).$$
- Bateman and Horn (1962) (cf Hardy & Littlewood 1923) have conjectured that
$$Q(X) \sim \frac{C_1X}{\log X}.$$

- We have
$$\sum_{n\in\mathcal{D}} g(n) = C_1^{-1}\log D + O(1)$$
- Thus
$$Q(X) \le \frac{2C_1X}{\log X} + O\big(X(\log\log X)(\log X)^{-2}\big).$$
- Bateman and Horn (1962) (cf Hardy & Littlewoood 1923) have conjectured that
$$Q(X) \sim \frac{C_1X}{\log X}.$$
- The exponent of $\log x$ in these results is often called the **Dimension** of the sieve. An alternative definition is given by
$$\lim_{D\to\infty}\frac{1}{\log\log D}\sum_{p\le D}\frac{g(p)}{p}.$$

- We have
$$\sum_{n\in\mathcal{D}} g(n) = C_1^{-1}\log D + O(1)$$
- Thus
$$Q(X) \le \frac{2C_1X}{\log X} + O\big(X(\log\log X)(\log X)^{-2}\big).$$
- Bateman and Horn (1962) (cf Hardy & Littlewoood 1923) have conjectured that
$$Q(X) \sim \frac{C_1X}{\log X}.$$
- The exponent of $\log x$ in these results is often called the **Dimension** of the sieve. An alternative definition is given by
$$\lim_{D\to\infty}\frac{1}{\log\log D}\sum_{p\le D}\frac{g(p)}{p}.$$
- Thus primes in an interval, or n^2+1 have dimension 1. The twin prime conjecture has dimension 2.

- If we choose $D \approx X^\theta$, then θ is sometimes called the **Level of the Distribution**, but this does not necessarily depend on the nature of the sieve, but rather how clever we are.

- If we choose $D \approx X^\theta$, then θ is sometimes called the **Level of the Distribution**, but this does not necessarily depend on the nature of the sieve, but rather how clever we are.
- Recall that we can set up the Goldbach binary conjecture by considering $\big(n(N-n), P\big) = 1$ and the analysis is similar to the twin prime conjecture.

- If we choose $D \approx X^{\theta}$, then θ is sometimes called the **Level of the Distribution**, but this does not necessarily depend on the nature of the sieve, but rather how clever we are.
- Recall that we can set up the Goldbach binary conjecture by considering $\big(n(N-n), P\big) = 1$ and the analysis is similar to the twin prime conjecture.
- Thus we obtain that for even N

$$R(N) \le 8C_2 \frac{N}{(\log N)^2} \prod_{\substack{p|N\\ p>2}} \frac{p-1}{p-2} + O\big(N(\log\log N)(\log N)^{-3}\big)$$

where $R(N)$ is the number of ordered pairs of primes p_1, p_2 such that $p_1 + p_2 = N$ and C_2 is the twin prime conjecture.

- Sometimes we write

$$C_2 \prod_{\substack{p|N\\ p>2}} \frac{p-1}{p-2}$$

$$= \prod_{p|N}\left(1+\frac{1}{p-1}\right)\prod_{p\nmid N}\left(1-\frac{1}{(p-1)^2}\right) = \mathfrak{S}(N).$$

- Sometimes we write

$$C_2 \prod_{\substack{p|N\\p>2}} \frac{p-1}{p-2}$$
$$= \prod_{p|N}\left(1+\frac{1}{p-1}\right)\prod_{p\nmid N}\left(1-\frac{1}{(p-1)^2}\right) = \mathfrak{S}(N).$$

- Hardy and Littlewood (1923) conjectured that

$$R(N) \sim \mathfrak{S}(N)\frac{N}{(\log N)^2}$$

and deduced, on the assumption of GRH, that this holds for almost all even N. Chudakov, Estermann and van der Corput independently proved this without GRH in 1937 by using Vinogradov's method.

- Sometimes we write

$$C_2 \prod_{\substack{p|N\\p>2}} \frac{p-1}{p-2}$$

$$= \prod_{p|N}\left(1+\frac{1}{p-1}\right)\prod_{p\nmid N}\left(1-\frac{1}{(p-1)^2}\right) = \mathfrak{S}(N).$$

- Hardy and Littlewood (1923) conjectured that

$$R(N) \sim \mathfrak{S}(N)\frac{N}{(\log N)^2}$$

and deduced, on the assumption of GRH, that this holds for almost all even N. Chudakov, Estermann and van der Corput independently proved this without GRH in 1937 by using Vinogradov's method.

- Montgomery & RCV (1975) showed that there is a positive number δ such that the number $E(X)$ of even $N \le X$ such that $R(N) = 0$ satisfies $E(X) \ll X^{1-\delta}$.

- So here is a clever idea. Suppose we know GRH(!).

- So here is a clever idea. Suppose we know GRH(!).
- **Example 11 Recast Twin Primes.** We can consider

$$\pi_2(x) \le \operatorname{card}\{p \le x : (p+2, P) = 1\} + \pi(\sqrt{x}).$$

- Thus, for $d|P$, where now $P = \prod_{2<p\le\sqrt{x}} p$

$$A_d = \pi(x; d, -2) = f(d)X + R_d$$

$f(d) = \frac{1}{\phi(d)}$, $X = \operatorname{li}(x)$, $|R_d| \ll x^{1/2}(\log x)^2$.

- So here is a clever idea. Suppose we know GRH(!).
- **Example 11 Recast Twin Primes.** We can consider

$$\pi_2(x) \le \operatorname{card}\{p \le x : (p+2, P) = 1\} + \pi(\sqrt{x}).$$

- Thus, for $d|P$, where now $P = \prod_{2<p\le\sqrt{x}} p$

$$A_d = \pi(x; d, -2) = f(d)X + R_d$$

$f(d) = \frac{1}{\phi(d)}$, $X = \operatorname{li}(x)$, $|R_d| \ll x^{1/2}(\log x)^2$.

- Then $g(2) = 0$, $g(p) = 1/(p-2)$ $(p > 2)$, and

$$\sum_{d\in\mathcal{D}} g(d) = \sum_{\substack{n\le D \\ 2\nmid n}} \frac{\mu(n)^2}{\prod_{p|n}(p-2)}.$$

- So here is a clever idea. Suppose we know GRH(!).
- **Example 11 Recast Twin Primes.** We can consider

$$\pi_2(x) \le \operatorname{card}\{p \le x : (p+2, P) = 1\} + \pi(\sqrt{x}).$$

- Thus, for $d|P$, where now $P = \prod_{2<p\le\sqrt{x}} p$

$$A_d = \pi(x; d, -2) = f(d)X + R_d$$

$f(d) = \frac{1}{\phi(d)}$, $X = \mathrm{li}(x)$, $|R_d| \ll x^{1/2}(\log x)^2$.

- Then $g(2) = 0$, $g(p) = 1/(p-2)$ $(p > 2)$, and

$$\sum_{d\in\mathcal{D}} g(d) = \sum_{\substack{n\le D\\ 2\nmid n}} \frac{\mu(n)^2}{\prod_{p|n}(p-2)}.$$

- The methods we have for approximating such sums give

$$\sum_{d\in\mathcal{D}} g(d) = \frac{1}{2}\left(\prod_{p>2}\frac{(p-1)^2}{p(p-2)}\right)\log D + O(1).$$

- Thus
$$\pi_2(x) \le \frac{C_2\,\mathrm{li}(x)}{\log D} + O\big(D^2 x^{1/2}(\log x)^2\big)$$

- Thus
$$\pi_2(x) \le \frac{C_2 \operatorname{li}(x)}{\log D} + O\big(D^2 x^{1/2} (\log x)^2\big)$$
- Wait a minute, now the sieve has dimension 1!

- Thus
$$\pi_2(x) \le \frac{C_2 \operatorname{li}(x)}{\log D} + O\big(D^2 x^{1/2} (\log x)^2\big)$$
- Wait a minute, now the sieve has dimension 1!
- But we have to take D smaller, say $D = x^{1/4}(\log x)^{-2}$.
- Hence
$$\pi_2(x) \le \frac{4C_2 x}{(\log x)^2} + O\big(x(\log\log x)(\log x)^{-3}\big).$$

- Thus
$$\pi_2(x) \le \frac{C_2 \operatorname{li}(x)}{\log D} + O\big(D^2 x^{1/2} (\log x)^2\big)$$
- Wait a minute, now the sieve has dimension 1!
- But we have to take D smaller, say $D = x^{1/4}(\log x)^{-2}$.
- Hence
$$\pi_2(x) \le \frac{4C_2 x}{(\log x)^2} + O\big(x(\log\log x)(\log x)^{-3}\big).$$
- So we nevertheless gained a factor of 2 on GRH.

- Thus
$$\pi_2(x) \le \frac{C_2\,\mathrm{li}(x)}{\log D} + O\big(D^2x^{1/2}(\log x)^2\big)$$
- Wait a minute, now the sieve has dimension 1!
- But we have to take D smaller, say $D = x^{1/4}(\log x)^{-2}$.
- Hence
$$\pi_2(x) \le \frac{4C_2x}{(\log x)^2} + O\big(x(\log\log x)(\log x)^{-3}\big).$$
- So we nevertheless gained a factor of 2 on GRH.
- Oh, but wait another minute. In 1965 Bombieri and A. I. Vinogradov proved that GRH holds on average, and that is all that we need!!

- Bombieri's version can be stated: Given any fixed $A > 0$ there is a $B = B(A)$ such that if $Q = x^{1/2}(\log x)^{-B}$, then

$$\sum_{q\le Q} \max_{(a,q)=1} \sup_{y\le x} \left| \pi(y; q, a) - \frac{\mathrm{li}(y)}{\phi(q)} \right| \ll x(\log x)^{-A}.$$

- Bombieri's version can be stated: Given any fixed $A > 0$ there is a $B = B(A)$ such that if $Q = x^{1/2}(\log x)^{-B}$, then

$$\sum_{q\le Q} \max_{(a,q)=1} \sup_{y\le x} \left| \pi(y;q,a) - \frac{\mathrm{li}(y)}{\phi(q)} \right| \ll x(\log x)^{-A}.$$

- I plan to show you a relatively simple proof that we can take $B(A) = A + 4$.

- Bombieri's version can be stated: Given any fixed $A > 0$ there is a $B = B(A)$ such that if $Q = x^{1/2}(\log x)^{-B}$, then

$$\sum_{q\le Q} \max_{(a,q)=1} \sup_{y\le x} \left| \pi(y;q,a) - \frac{\mathrm{li}(y)}{\phi(q)} \right| \ll x(\log x)^{-A}.$$

- I plan to show you a relatively simple proof that we can take $B(A) = A + 4$.
- There is one other example which I want to show you.

- Bombieri's version can be stated: Given any fixed $A > 0$ there is a $B = B(A)$ such that if $Q = x^{1/2}(\log x)^{-B}$, then

$$\sum_{q \le Q} \max_{(a,q)=1} \sup_{y \le x} \left| \pi(y; q, a) - \frac{\mathrm{li}(y)}{\phi(q)} \right| \ll x(\log x)^{-A}.$$

- I plan to show you a relatively simple proof that we can take $B(A) = A + 4$.
- There is one other example which I want to show you.
- This is the ultimate generalisation of the twin prime conjecture, at least for linear polynomials.

- **prime** k**–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X;\mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \text{card}\{m \le X : (m+h_1)\ldots(m+h_k) = n\}$. Then we have
$$\pi_k(X;\mathbf{h}) \le \pi(\sqrt{X}) + S(A, P).$$

- **prime k–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X;\mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \text{card}\{m \le X : (m+h_1)\ldots(m+h_k) = n\}$. Then we have
$$\pi_k(X;\mathbf{h}) \le \pi(\sqrt{X}) + S(A,P).$$
- Now $A_d = f(d)X + R_d$ where $f(d) = \rho(d)/d$, $|R_d| \le \rho(d)$ and $\rho(d)$ is the number of solutions of $(x+h_1)\ldots(x+h_k) \equiv 0 \pmod{d}$.

- **prime k–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X;\mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \operatorname{card}\{m \le X : (m+h_1)\ldots(m+h_k) = n\}$. Then we have
$$\pi_k(X;\mathbf{h}) \le \pi(\sqrt{X}) + S(A,P).$$
- Now $A_d = f(d)X + R_d$ where $f(d) = \rho(d)/d$, $|R_d| \le \rho(d)$ and $\rho(d)$ is the number of solutions of $(x+h_1)\ldots(x+h_k) \equiv 0 \pmod d$.
- Then $\rho \in \mathcal{M}$, $\rho(p) \le k$, and $\rho(p) = k$ when
$$p \nmid \Delta := \prod_{1\le i<j\le k} |h_j - h_i|.$$

- **prime k–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X;\mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \operatorname{card}\{m \le X : (m+h_1)\ldots(m+h_k) = n\}$. Then we have
$$\pi_k(X;\mathbf{h}) \le \pi(\sqrt{X}) + S(A,P).$$
- Now $A_d = f(d)X + R_d$ where $f(d) = \rho(d)/d$, $|R_d| \le \rho(d)$ and $\rho(d)$ is the number of solutions of $(x+h_1)\ldots(x+h_k) \equiv 0 \pmod{d}$.
- Then $\rho \in \mathcal{M}$, $\rho(p) \le k$, and $\rho(p) = k$ when
$$p \nmid \Delta := \prod_{1\le i<j\le k} |h_j - h_i|.$$
- This is an example of a k–dimensional sieve.

- **prime k–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X; \mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \operatorname{card}\{m \le X : (m+h_1)\ldots(m+h_k) = n\}$. Then we have
$$\pi_k(X; \mathbf{h}) \le \pi(\sqrt{X}) + S(A, P).$$
- Now $A_d = f(d)X + R_d$ where $f(d) = \rho(d)/d$, $|R_d| \le \rho(d)$ and $\rho(d)$ is the number of solutions of $(x+h_1)\ldots(x+h_k) \equiv 0 \pmod{d}$.
- Then $\rho \in \mathcal{M}$, $\rho(p) \le k$, and $\rho(p) = k$ when
$$p \nmid \Delta := \prod_{1\le i<j\le k} |h_j - h_i|.$$
- This is an example of a k–dimensional sieve.
- If the $\mathbf{h}$ give a complete set of residues modulo p for some p, then there are not many k-tuples which are simultaneously prime!

- **prime k–tuples.** Let $\mathbf{h} = h_1, h_2, \ldots, h_k$ be k distinct positive integers and $\pi_k(X; \mathbf{h})$ be the number of $m \le X$ such that the $m + h_j$ are simultaneously prime. Let $a(n) = \text{card}\{m \le X : (m + h_1) \ldots (m + h_k) = n\}$. Then we have
$$\pi_k(X; \mathbf{h}) \le \pi(\sqrt{X}) + S(A, P).$$
- Now $A_d = f(d)X + R_d$ where $f(d) = \rho(d)/d$, $|R_d| \le \rho(d)$ and $\rho(d)$ is the number of solutions of $(x + h_1) \ldots (x + h_k) \equiv 0 \pmod{d}$.
- Then $\rho \in \mathcal{M}$, $\rho(p) \le k$, and $\rho(p) = k$ when
$$p \nmid \Delta := \prod_{1 \le i < j \le k} |h_j - h_i|.$$
- This is an example of a k–dimensional sieve.
- If the $\mathbf{h}$ give a complete set of residues modulo p for some p, then there are not many k-tuples which are simultaneously prime!
- Thus a natural condition for the existence of many prime k-tuples is that $\rho(p) < p$ for all p, i.e $f(p) < 1$.

- $\rho(d)$ is the number of solutions of $(x+h_1)\dots(x+h_k)\equiv 0 \pmod{d}$.

- $\rho(d)$ is the number of solutions of $(x+h_1)\dots(x+h_k)\equiv 0 \pmod{d}$.
- Then $\rho\in\mathcal{M}$, $\rho(p)\le k$, and $\rho(p)=k$ when

$$p\nmid\Delta:=\prod_{1\le i<j\le k}|h_j-h_i|.$$

- $\rho(d)$ is the number of solutions of $(x+h_1)\ldots(x+h_k)\equiv 0 \pmod{d}$.
- Then $\rho\in\mathcal{M}$, $\rho(p)\le k$, and $\rho(p)=k$ when

$$p\nmid \Delta := \prod_{1\le i<j\le k} |h_j-h_i|.$$

- Hardy and Littlewood (1923) conjectured that

$$\pi_k(X;\mathbf{h})\sim \mathfrak{S}(\mathbf{h})\frac{x}{(\log x)^k}$$

where

$$\mathfrak{S}(\mathbf{h})=\prod_p\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-k}.$$

- Although not relevant for the rest of the course I want to say something briefly about upper and lower sieve estimates. Hopefully it will give you some idea of where the mainstream of the subject lies.

- Although not relevant for the rest of the course I want to say something briefly about upper and lower sieve estimates. Hopefully it will give you some idea of where the mainstream of the subject lies.
- First let me recall some definitions.

- Although not relevant for the rest of the course I want to say something briefly about upper and lower sieve estimates. Hopefully it will give you some idea of where the mainstream of the subject lies.
- First let me recall some definitions.
- We supposed that $a(n)$ are non-negative real numbers defined on $\mathbb{Z}$, and defined

$$A_d = \sum_n a(dn).$$

- Although not relevant for the rest of the course I want to say something briefly about upper and lower sieve estimates. Hopefully it will give you some idea of where the mainstream of the subject lies.
- First let me recall some definitions.
- We supposed that $a(n)$ are non-negative real numbers defined on $\mathbb{Z}$, and defined

$$A_d = \sum_n a(dn).$$

- We then supposed that there is a multiplicative function $f(d)$ and a positive real number X so that, at least when d is squarefree, we can write

$$A_d = f(d)X + R_d$$

with some expectation that, at least for smaller d, the $f(d)X$ dominate the R_d, albeit on average.

- We also defined the "dimension" κ of a sieve by

$$\kappa = \lim_{D\to\infty} \frac{\sum_{p\le D} f(p)}{\log\log D}$$

when the limit exists.

- We also defined the "dimension" κ of a sieve by

$$\kappa = \lim_{D\to\infty} \frac{\sum_{p\le D} f(p)}{\log\log D}$$

  when the limit exists.
- For the purposes of this exposition I will suppose that $\kappa = 1$.

- We also defined the“dimension” κ of a sieve by

$$\kappa = \lim_{D\to\infty} \frac{\sum_{p\le D} f(p)}{\log\log D}$$

  when the limit exists.
- For the purposes of this exposition I will suppose that $\kappa = 1$.
- Another term which is sometimes used is the “level” θ of a sieve. That is a “good” value of θ for which we can show that

$$\sum_{d\le X^\theta} \mu(d)^2 |R_d| \ll X(\log X)^{-B}$$

  for some suitably large enough value of B. Hopefully $\theta = 1$, but this is not always possible.

- We also defined the"dimension" κ of a sieve by
$$\kappa = \lim_{D\to\infty} \frac{\sum_{p\le D} f(p)}{\log\log D}$$
when the limit exists.
- For the purposes of this exposition I will suppose that $\kappa = 1$.
- Another term which is sometimes used is the "level" θ of a sieve. That is a "good" value of θ for which we can show that
$$\sum_{d\le X^\theta} \mu(d)^2 |R_d| \ll X(\log X)^{-B}$$
for some suitably large enough value of B. Hopefully $\theta = 1$, but this is not always possible.
- My comments are very hand-wavy, but they do fit in with the known facts in lots of interesting examples, as we have seen.

- Within these parameters we are looking for real numbers $\lambda_d^{\pm}$ such that for every n, or at least for every squarefree n is a suitable divisor closed set, we have

$$\sum_{d|n} \lambda_d^- \le \sum_{d|n} \mu(d) = e(n) \le \sum_{d|n} \lambda_d^+.$$

- There are two points of view, each of which in the one dimensional case leads to essentially the same conclusion.

- Within these parameters we are looking for real numbers $\lambda_d^{\pm}$ such that for every n, or at least for every squarefree n is a suitable divisor closed set, we have

$$\sum_{d|n} \lambda_d^- \le \sum_{d|n} \mu(d) = e(n) \le \sum_{d|n} \lambda_d^+.$$

- There are two points of view, each of which in the one dimensional case leads to essentially the same conclusion.
- One is to gain insight by looking at the d formed from exactly k prime divisors and using this to construct optimal $\lambda_d^{\pm}$. This can be considered the ultimate version of Sylvester's inclusion-exclusion principle and the Brun-Merlin sieve, and was developed independently by Rosser (c1945 but unpublished) and Iwaniec (c1975). It is also sometimes called the combinatorial sieve.

- Within these parameters we are looking for real numbers $\lambda_d^{\pm}$ such that for every n, or at least for every squarefree n is a suitable divisor closed set, we have

$$\sum_{d|n} \lambda_d^- \le \sum_{d|n} \mu(d) = e(n) \le \sum_{d|n} \lambda_d^+.$$

- There are two points of view, each of which in the one dimensional case leads to essentially the same conclusion.
- One is to gain insight by looking at the d formed from exactly k prime divisors and using this to construct optimal $\lambda_d^{\pm}$. This can be considered the ultimate version of Sylvester's inclusion-exclusion principle and the Brun-Merlin sieve, and was developed independently by Rosser (c1945 but unpublished) and Iwaniec (c1975). It is also sometimes called the combinatorial sieve.
- The other is to work backwards iteratively from the Selberg upper bound, *via* the Buchstab identity.

- To describe the results in either case we also need to define

$$W(D) = \prod_{p<D} \left(1 - \frac{f(p)}{p}\right)$$

which in general we would expect to be roughly $(\log D)^{-\kappa}$, and so here $(\log D)^{-1}$.

- To describe the results in either case we also need to define
$$W(D)=\prod_{p<D}\left(1-\frac{f(p)}{p}\right)$$
which in general we would expect to be roughly $(\log D)^{-\kappa}$, and so here $(\log D)^{-1}$.
- Then the conclusions are
$$\begin{aligned}\sigma_-\big(\theta(\log X)/(\log D)\big)XW(D)+E_-\le\\ S\big(\mathcal{A},P(D)\big)\\ \le\sigma_+\big(\theta(\log X)/(\log D)\big)XW(D)+E_+\end{aligned}$$
for suitable error terms $E_\pm$.

- To describe the results in either case we also need to define

$$W(D)=\prod_{p<D}\left(1-\frac{f(p)}{p}\right)$$

which in general we would expect to be roughly $(\log D)^{-\kappa}$, and so here $(\log D)^{-1}$.

- Then the conclusions are

$$\begin{aligned}\sigma_-\big(\theta(\log X)/(\log D)\big)XW(D)+E_-\le\\ S\big(\mathcal{A},P(D)\big)\\ \le\sigma_+\big(\theta(\log X)/(\log D)\big)XW(D)+E_+\end{aligned}$$

for suitable error terms $E_\pm$.

- Here the functions $\sigma_\pm$ are more complicated versions of Dickman's function ρ, which occurs on homework 5.

- Here the functions $\sigma_\pm$ are more complicated versions of Dickman's function ρ, which occurs on homework 5.

- Here the functions $\sigma_\pm$ are more complicated versions of Dickman's function ρ, which occurs on homework 5.
- They are continuous for $u > 0$ and differentiable for $u > 0$, $u \neq 2$, and satisfy

$$\begin{cases} \sigma_-(u) = 0 & (0 < u \le 2), \\ \sigma_+(u) = 2e^\gamma u^{-1} & (0 < u \le 2), \\ \big(u\sigma_\pm(u)\big)' = \sigma_\mp(u-1) & (u > 2). \end{cases}$$

- Here the functions $\sigma_\pm$ are more complicated versions of Dickman's function ρ, which occurs on homework 5.
- They are continuous for $u > 0$ and differentiable for $u > 0$, $u \neq 2$, and satisfy

$$\begin{cases} \sigma_-(u) = 0 & (0 < u \leq 2), \\ \sigma_+(u) = 2e^\gamma u^{-1} & (0 < u \leq 2), \\ \big(u\sigma_\pm(u)\big)' = \sigma_\mp(u-1) & (u > 2). \end{cases}$$

- See how the definitions in the range $0 < u \leq 2$ fit Selberg's example.

- An alternative is to define

$$\begin{cases} \eta(u) = \frac{1}{u}, & \xi(u) = 1 \qquad (0 < u \le 2), \\ u\eta'(u) = \eta(u-1) & (u > 2), \\ (u-1)\xi'(u) = -\xi(u-1) & (u > 2) \end{cases}$$

and put

$$\sigma_\pm(u) = e^\gamma \left(\eta(u) \pm \frac{\xi(u)}{u} \right).$$

- An alternative is to define

$$\begin{cases} \eta(u) = \frac{1}{u}, \qquad \xi(u) = 1 & (0 < u \le 2), \\ u\eta'(u) = \eta(u-1) & (u > 2), \\ (u-1)\xi'(u) = -\xi(u-1) & (u > 2) \end{cases}$$

  and put

$$\sigma_\pm(u) = e^\gamma \left(\eta(u) \pm \frac{\xi(u)}{u} \right).$$

- By the way, one can check that $\xi(u+1) = \rho(u)$, the function that corresponds to $\psi(X, Y)$ in homework 5.

- An alternative is to define

$$\begin{cases} \eta(u) = \frac{1}{u}, \qquad \xi(u) = 1 & (0 < u \le 2), \\ u\eta'(u) = \eta(u-1) & (u > 2), \\ (u-1)\xi'(u) = -\xi(u-1) & (u > 2) \end{cases}$$

  and put

$$\sigma_\pm(u) = e^\gamma \left(\eta(u) \pm \frac{\xi(u)}{u} \right).$$

- By the way, one can check that $\xi(u+1) = \rho(u)$, the function that corresponds to $\psi(X, Y)$ in homework 5.
- Let me conclude by observing that the case of dimension $\kappa < 1$ is largely understood, but in the case $\kappa > 1$ we have less precise results.