

0.1 Todo: [The Kuznetsov Trace Formula]

The Kuznetsov trace formula is an analog of the Petersson trace formula for weight zero Maass forms. From ??, $\mathcal{L}(N, \chi)$ admits an orthonormal basis of Maass forms for the point spectrum (these forms are generally not Hecke-Maass eigenforms because they need not be Hecke normalized or even cuspidal in the case of the discrete spectrum). However, by ?? and ?? we make take this orthonormal basis to consist of Hecke-Maass eigenforms and the constant function. Denote this basis by $\{u_j\}_{j \geq 0}$ with $u_0(z) = 1$ and let u_j be of type ν_j for $j \geq 1$. In particular, $\{u_j\}_{j \geq 1}$ is an orthonormal basis of Hecke-Maass eigenforms and each such form admits a Fourier series at the \mathfrak{a} cusp given by

$$(u_j | \sigma_{\mathfrak{a}})(z) = \sum_{n \neq 0} a_{j,\mathfrak{a}}(n) \sqrt{y} K_{\nu_j}(2\pi n y) e^{2\pi i n x}.$$

The Kuznetsov trace formula is an equation relating the Fourier coefficients $a_{j,\mathfrak{a}}(n)$ and $a_{j,\mathfrak{b}}(n)$ of the basis $\{u_j\}_{j \geq 1}$ for two cusps \mathfrak{a} and \mathfrak{b} of $\Gamma_0(N) \backslash \mathbb{H}$ to a sum of integral transforms involving test functions and Salié sums. Similar to the Petersson trace formula, we will compute the inner product of two Poincaré series $P_{n,\chi,\mathfrak{a}}(z, \psi)(z)$ and $P_{m,\chi,\mathfrak{b}}(z, \varphi)(z)$ in two different ways. The first will be geometric in nature while the second will be spectral. We first need to compute the Fourier series of such a Poincaré series. Although we will not need it explicitly, we will work over any congruence subgroup:

Proposition 0.1.1. *Let $m \geq 1$, χ be Dirichlet character with conductor dividing the level, \mathfrak{a} and \mathfrak{b} be cusps of $\Gamma \backslash \mathbb{H}$, and $\psi(y)$ be a smooth function such that $\psi(y) \ll_{\varepsilon} y^{1+\varepsilon}$ as $y \rightarrow 0$. The Fourier series of $P_{m,\chi,\mathfrak{a}}(z, \psi)$ on $\Gamma \backslash \mathbb{H}$ at the \mathfrak{b} cusp is given by*

$$(P_{m,\chi,\mathfrak{a}} | \sigma_{\mathfrak{b}})(z, \psi) = \sum_{t \in \mathbb{Z}} \left(\delta_{\mathfrak{a},\mathfrak{b}} \delta_{m,t} \psi(\text{Im}(z)) + \sum_{c \in \mathcal{C}_{\mathfrak{a},\mathfrak{b}}} \psi(y, m, t, c) S_{\chi,\mathfrak{a},\mathfrak{b}}(m, t, c) \right) e^{2\pi i t z},$$

where $\psi(y, m, t, c)$ is the integral transform given by

$$\psi(y, m, t, c) = \int_{\text{Im}(z)=y} \psi \left(\frac{y}{|cz|^2} \right) e^{-\frac{2\pi i m}{c^2 z} - 2\pi i t z} dz.$$

Proof. From the cocycle condition and ??, we have

$$(P_{m,\chi,\mathfrak{a}} | \sigma_{\mathfrak{b}})(z, \psi) = \delta_{\mathfrak{a},\mathfrak{b}} \psi(\text{Im}(z)) e^{2\pi i m z} + \sum_{\substack{c \in \mathcal{C}_{\mathfrak{a},\mathfrak{b}}, d \in \mathbb{Z} \\ d \pmod{c} \in \mathcal{D}_{\mathfrak{a},\mathfrak{b}}(c)}} \bar{\chi}(d) \psi \left(\frac{\text{Im}(z)}{|cz + d|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + cd} \right)},$$

where a and b are chosen such that $\det \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = 1$ and we have used the fact that

$$\frac{a}{c} - \frac{1}{c^2 z + cd} = \frac{az + b}{cz + d}.$$

Summing over all pairs (c, d) with $c \in \mathcal{C}_{\mathfrak{a},\mathfrak{b}}$, $d \in \mathbb{Z}$, and $d \pmod{c} \in \mathcal{D}_{\mathfrak{a},\mathfrak{b}}(c)$ is the same as summing over all triples (c, ℓ, r) with $c \in \mathcal{C}_{\mathfrak{a},\mathfrak{b}}$, $\ell \in \mathbb{Z}$, and r taken modulo c with $r \in \mathcal{D}_{\mathfrak{a},\mathfrak{b}}(c)$. Indeed, this is seen by writing $d = c\ell + r$. Moreover, since $ad - bc = 1$ we have $a(c\ell + r) - bc = 1$ which further implies that

$ar \equiv 1 \pmod{c}$. So we may take a to be the inverse for r modulo c . Then

$$\begin{aligned}
\sum_{\substack{c \in \mathcal{C}_{a,b}, d \in \mathbb{Z} \\ d \pmod{c} \in \mathcal{D}_{a,b}(c)}} \bar{\chi}(d) \psi \left(\frac{\text{Im}(z)}{|cz + d|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + cd} \right)} &= \sum_{(c, \ell, r)} \bar{\chi}(c\ell + r) \psi \left(\frac{\text{Im}(z)}{|cz + c\ell + r|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + c^2 \ell + cr} \right)} \\
&= \sum_{(c, \ell, r)} \bar{\chi}(r) \psi \left(\frac{\text{Im}(z)}{|cz + c\ell + r|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + c^2 \ell + cr} \right)} \\
&= \sum_{\substack{c \in \mathcal{C}_{a,b} \\ r \in \mathcal{D}_{a,b}(c)}} \sum_{\ell \in \mathbb{Z}} \bar{\chi}(r) \psi \left(\frac{\text{Im}(z)}{|cz + c\ell + r|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + c^2 \ell + cr} \right)} \\
&= \sum_{\substack{c \in \mathcal{C}_{a,b} \\ r \in \mathcal{D}_{a,b}(c)}} \bar{\chi}(r) \sum_{\ell \in \mathbb{Z}} \psi \left(\frac{\text{Im}(z)}{|cz + c\ell + r|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + c^2 \ell + cr} \right)},
\end{aligned}$$

where on the right-hand side it is understood that we are summing over all triples (c, ℓ, r) with the prescribed properties and the second line holds since χ has conductor dividing the level and $d \in \mathcal{D}_{a,b}(c)$ is determined modulo c . Now let

$$I_{c,r}(z, \psi) = \sum_{\ell \in \mathbb{Z}} \psi \left(\frac{\text{Im}(z)}{|cz + c\ell + r|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + c^2 \ell + cr} \right)}.$$

We apply the Poisson summation formula to $I_{c,r}(z, \psi)$. This is allowed since the summands are absolutely integrable by ??, as they exhibit polynomial decay of order $\sigma > 1$ because $\psi(y) \ll_{\varepsilon} y^{1+\varepsilon}$ as $y \rightarrow 0$, and $I_{c,r}(z, \psi)$ is holomorphic because $(P_{m,\chi,a}|\sigma_b)(z, \psi)$ is. By the identity theorem it suffices to apply the Poisson summation formula for $z = iy$ with $y > 0$. So let $f(x)$ be given by

$$f(x) = \psi \left(\frac{y}{|cx + r + icy|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 x + cr + ic^2 y} \right)}.$$

As we have just noted, $f(x)$ is absolutely integrable on \mathbb{R} . We compute the Fourier transform:

$$\hat{f}(t) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i t x} dx = \int_{-\infty}^{\infty} \psi \left(\frac{y}{|cx + r + icy|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 x + cr + ic^2 y} \right)} e^{-2\pi i t x} dx.$$

Complexify the integral to get

$$\int_{\text{Im}(z)=0} \psi \left(\frac{y}{|cz + r + icy|^2} \right) e^{2\pi i m \left(\frac{a}{c} - \frac{1}{c^2 z + cr + ic^2 y} \right)} e^{-2\pi i t z} dz.$$

Now make the change of variables $z \rightarrow z - \frac{r}{c} - iy$ to obtain

$$e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c} - 2\pi i t y} \int_{\text{Im}(z)=y} \psi \left(\frac{y}{|cz|^2} \right) e^{-\frac{2\pi i m}{c^2 z} - 2\pi i t z} dz.$$

As the remaining integral is $\psi(y, m, t, c)$, it follows that

$$\hat{f}(t) = \psi(y, m, t, c) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c} - 2\pi i t y}.$$

By the Poisson summation formula and the identity theorem, we have

$$I_{c,r}(z, \psi) = \sum_{t \in \mathbb{Z}} (\psi(y, m, t, c) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c}}) e^{2\pi i t z},$$

for all $z \in \mathbb{H}$. Substituting this back into the Eisenstein series gives a form of the Fourier series:

$$\begin{aligned}
(P_{m,\chi,a}|\sigma_b)(z, \psi) &= \delta_{a,b} \psi(\text{Im}(z)) e^{2\pi i m z} + \sum_{\substack{c \in \mathcal{C}_{a,b} \\ r \in \mathcal{D}_{a,b}}} \bar{\chi}(r) \sum_{t \in \mathbb{Z}} \psi(y, m, t, c) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c}} e^{2\pi i t z} \\
&= \sum_{t \in \mathbb{Z}} \left(\delta_{a,b} \delta_{m,t} \psi(\text{Im}(z)) + \sum_{\substack{c \in \mathcal{C}_{a,b} \\ r \in \mathcal{D}_{a,b}}} \bar{\chi}(r) \psi(y, m, t, c) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c}} \right) e^{2\pi i t z} \\
&= \sum_{t \in \mathbb{Z}} \left(\delta_{a,b} \delta_{m,t} \psi(\text{Im}(z)) + \sum_{c \in \mathcal{C}_{a,b}} \psi(y, m, t, c) \sum_{r \in \mathcal{D}_{a,b}} \bar{\chi}(r) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c}} \right) e^{2\pi i t z}.
\end{aligned}$$

We will simplify the innermost sum. Since a is the inverse for r modulo c , the innermost sum above becomes

$$\sum_{r \in \mathcal{D}_{a,b}} \bar{\chi}(r) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{r}{c}} = \sum_{r \in \mathcal{D}_{a,b}} \bar{\chi}(\bar{a}) e^{2\pi i m \frac{a}{c} + 2\pi i t \frac{\bar{a}}{c}} = \sum_{r \in \mathcal{D}_{a,b}} \chi(a) e^{\frac{2\pi i (am + \bar{a}t)}{c}} = S_{\chi,a,b}(m, t, c).$$

So at last, we obtain our desired Fourier series:

$$(P_{m,\chi,a}|\sigma_b)(z) = \sum_{t \in \mathbb{Z}} \left(\delta_{a,b} \delta_{m,t} \psi(\text{Im}(z)) + \sum_{c \in \mathcal{C}_{a,b}} \psi(y, m, t, c) S_{\chi,a,b}(m, t, c) \right) e^{2\pi i t z}. \quad \square$$

We can now derive the first half of the Kuznetsov trace formula by computing the inner product between $P_{n,\chi,a}(z, \psi)$ and $P_{m,\chi,b}(z, \varphi)$:

$$\begin{aligned}
\langle P_{n,\chi,a}(\cdot, \psi), P_{m,\chi,b}(\cdot, \varphi) \rangle &= \frac{1}{V_{\Gamma_0(N)}} \int_{\mathcal{F}_{\Gamma_0(N)}} P_{n,\chi,a}(z, \psi) \overline{P_{m,\chi,b}(z, \varphi)} d\mu \\
&= \frac{1}{V_{\Gamma_0(N)}} \int_{\mathcal{F}_{\Gamma_0(N)}} \sum_{\gamma \in \Gamma_b \backslash \Gamma_0(N)} \chi(\gamma) P_{n,\chi,a}(z, \psi) \overline{\varphi(\text{Im}(\sigma_b^{-1} \gamma z))} e^{-2\pi i m \overline{\sigma_b^{-1} \gamma z}} d\mu \\
&= \frac{1}{V_{\Gamma_0(N)}} \int_{\mathcal{F}_{\Gamma_0(N)}} \sum_{\gamma \in \Gamma_b \backslash \Gamma_0(N)} P_{n,\chi,a}(\gamma z, \psi) \overline{\varphi(\text{Im}(\sigma_b^{-1} \gamma z))} e^{-2\pi i m \overline{\sigma_b^{-1} \gamma z}} d\mu \\
&= \frac{1}{V_{\Gamma_0(N)}} \int_{\mathcal{F}_{\sigma_b^{-1} \Gamma_0(N) \sigma_b}} \sum_{\gamma \in \Gamma_b \backslash \Gamma_0(N)} P_{n,\chi,a}(\gamma \sigma_b z, \psi) \overline{\varphi(\text{Im}(\sigma_b^{-1} \gamma \sigma_b z))} e^{-2\pi i m \overline{\sigma_b^{-1} \gamma \sigma_b z}} d\mu \\
&= \frac{1}{V_{\Gamma_0(N)}} \int_{\mathcal{F}_{\sigma_b^{-1} \Gamma_0(N) \sigma_b}} \sum_{\gamma \in \Gamma_\infty \backslash \sigma_b^{-1} \Gamma_0(N) \sigma_b^{-1}} P_{n,\chi,a}(\sigma_b \gamma z, \psi) \overline{\varphi(\text{Im}(\gamma z))} e^{-2\pi i m \overline{\gamma z}} d\mu \\
&= \frac{1}{V_{\Gamma_0(N)}} \int_{\Gamma_\infty \backslash \mathbb{H}} (P_{n,\chi,a}|\sigma_b)(z, \psi) \overline{\varphi(\text{Im}(z))} e^{-2\pi i m \bar{z}} d\mu,
\end{aligned}$$

where in the third line we have used the automorphy of $P_{n,\chi,a}(z, \psi)$, in the forth and fifth lines we have made the change of variables $z \rightarrow \sigma_b z$ and $\gamma \rightarrow \sigma_b \gamma \sigma_b^{-1}$ respectively, and in the sixth line we have unfolded. Now substitute in the Fourier series of $P_{n,\chi,a}(z, \psi)$ at the b cusp to obtain

$$\frac{1}{V_{\Gamma_0(N)}} \int_{\Gamma_\infty \backslash \mathbb{H}} \sum_{t \in \mathbb{Z}} \left(\delta_{a,b} \delta_{n,t} \psi(\text{Im}(z)) + \sum_{c \in \mathcal{C}_{a,b}} \psi(y, n, t, c) S_{\chi,a,b}(n, t, c) \right) \overline{\varphi(\text{Im}(z))} e^{2\pi i t z - 2\pi i m \bar{z}} d\mu,$$

which is equivalent to

$$\frac{1}{V_{\Gamma_0(N)}} \int_0^\infty \int_0^1 \sum_{t \geq 1} \left(\delta_{a,b} \delta_{n,t} \psi(y) + \sum_{c \in \mathcal{C}_{a,b}} \psi(y, n, t, c) S_{\chi,a,b}(n, t, c) \right) \overline{\varphi(y)} e^{2\pi i(t-m)x} e^{-2\pi(t+m)y} \frac{dx dy}{y^2}.$$

By the dominated convergence theorem, we can interchange the sum and the two integrals. Then ?? implies that the inner integral cuts off all of the terms except the diagonal $t = m$. This leaves

$$\frac{1}{V_{\Gamma_0(N)}} \int_0^\infty \left(\delta_{a,b} \delta_{n,m} \psi(y) + \sum_{c \in \mathcal{C}_{a,b}} \psi(y, n, m, c) S_{\chi,a,b}(n, m, c) \right) \overline{\varphi(y)} e^{-4\pi m y} \frac{dy}{y^2}.$$

Interchanging the integral and the remaining sum by the dominated convergence theorem again, we arrive at

$$\langle P_{n,\chi,a}(\cdot, \psi), P_{m,\chi,b}(\cdot, \varphi) \rangle = \delta_{a,b} \delta_{n,m} (\psi, \varphi)_{n,m} + \sum_{c \in \mathcal{C}_{a,b}} S_{\chi,a,b}(n, m, c) V(n, m, c, \psi, \varphi),$$

where we have set

$$(\psi, \varphi)_{n,m} = \frac{1}{V_{\Gamma_0(N)}} \int_0^\infty \psi(y) \overline{\varphi(y)} e^{-2\pi(n+m)y} \frac{dy}{y^2},$$

and

$$V(n, m, c; \psi, \varphi) = \frac{1}{V_{\Gamma_0(N)}} \int_0^\infty \int_{\text{Im}(z)=y} \psi\left(\frac{y}{|cz|^2}\right) \overline{\varphi(y)} e^{-\frac{2\pi i m}{c^2 z} - 2\pi i n z - 4\pi m y} \frac{dz dy}{y^2}.$$

This is the first half of the Kuznetsov trace formula. For the second half, ?? gives

$$P_{n,\chi,a}(\cdot, \psi) = \sum_{j \geq 0} \langle P_{n,\chi,a}(\cdot, \psi), u_j \rangle u_j(z) + \sum_a \frac{1}{4\pi} \int_{-\infty}^\infty \left\langle P_{n,\chi,a}(\cdot, \psi), E_a\left(\cdot, \frac{1}{2} + ir\right) \right\rangle E_a\left(z, \frac{1}{2} + ir\right) dr,$$

and

$$P_{m,\chi,a}(\cdot, \varphi) = \sum_{j \geq 0} \langle P_{m,\chi,a}(\cdot, \varphi), u_j \rangle u_j(z) + \sum_a \frac{1}{4\pi} \int_{-\infty}^\infty \left\langle P_{m,\chi,a}(\cdot, \varphi), E_a\left(\cdot, \frac{1}{2} + ir\right) \right\rangle E_a\left(z, \frac{1}{2} + ir\right) dr.$$

By orthonormality, it follows that

$$\begin{aligned} \langle P_{n,\chi,a}(\cdot, \psi), P_{m,\chi,a}(\cdot, \varphi) \rangle &= \sum_j \langle P_{n,\chi,a}(\cdot, \psi), u_j \rangle \overline{\langle P_{m,\chi,a}(\cdot, \varphi), u_j \rangle} \\ &\quad + \sum_a \frac{1}{4\pi} \int_{-\infty}^\infty \left\langle P_{n,\chi,a}(\cdot, \psi), E_a\left(\cdot, \frac{1}{2} + ir\right) \right\rangle \overline{\left\langle P_{m,\chi,a}(\cdot, \varphi), E_a\left(\cdot, \frac{1}{2} + ir\right) \right\rangle} dr. \end{aligned}$$

Now we must simplify the remaining inner products. Let $f \in \mathcal{L}(N, \chi)$ with Fourier series

$$f(z) = a^+(0) y^{\frac{1}{2}+\nu} + a^-(0) y^{\frac{1}{2}-\nu} + \sum_{n \neq 0} a(n) \sqrt{y} K_\nu(2\pi |n| y) e^{2\pi i n x}.$$

By unfolding the integral in the Petersson inner product and cutting off everything except the diagonal using ?? exactly as in the case for $\langle P_{n,\chi,a}(\cdot, \psi), P_{m,\chi,a}(\cdot, \varphi) \rangle$, we see that

$$\langle P_{n,\chi,a}(\cdot, \psi), f \rangle = \frac{1}{V_\Gamma} \int_0^\infty \overline{a(n) \sqrt{y} K_\nu(2\pi n y)} \psi(y) e^{-4\pi m y} \frac{dy}{y^2}.$$

Now set

$$\omega_\nu(n, \psi) = \frac{1}{V_\Gamma} \int_0^\infty \sqrt{y} K_\nu(2\pi|n|y) \overline{\psi(y)} e^{-4\pi my} \frac{dy}{y^2}.$$

Then it follows from the Fourier series of cusp forms and Eisenstein series that

$$\langle P_{n,\chi,\mathfrak{a}}(\cdot, \psi), u_j \rangle = \overline{a_j(n) \omega_{\nu_j}(n, \psi)},$$

for $j \geq 1$ and

$$\left\langle P_{n,\chi,\mathfrak{a}}(\cdot, \psi), E_{\mathfrak{a}}\left(\cdot, \frac{1}{2} + ir\right) \right\rangle = \overline{\tau_{\mathfrak{a}}\left(n, \frac{1}{2} + ir\right) \omega_{ir}(n, \psi)}.$$

In particular, $\langle P_{n,\chi,\mathfrak{a}}(\cdot, \psi), u_0 \rangle = 0$. So we obtain

$$\begin{aligned} \langle P_{n,\chi,\mathfrak{a}}(\cdot, \psi), P_{m,\chi,\mathfrak{a}}(\cdot, \varphi) \rangle &= \sum_{j \geq 1} \overline{a_j(n) a_j(m) \omega(n, \psi) \omega(m, \varphi)} \\ &\quad + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \overline{\tau_{\mathfrak{a}}\left(n, \frac{1}{2} + ir\right) \tau_{\mathfrak{a}}\left(m, \frac{1}{2} + ir\right) \omega(n, \psi) \omega(m, \varphi)} dr. \end{aligned}$$

This is the second half of the Kuznetsov trace formula. Equating the first and second halves we get the **Kuznetsov trace formula**:

$$\begin{aligned} \delta_{n,m}(\psi, \varphi) + \sum_{\substack{c \geq 1 \\ c \equiv 0 \pmod{N}}} \frac{1}{c} S_\chi(n, m, c) V(n, m, c, \psi, \varphi) &= \sum_{j \geq 1} \overline{a_j(n) a_j(m) \omega(n, \psi) \omega(m, \varphi)} \\ &\quad + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \overline{\tau_{\mathfrak{a}}\left(n, \frac{1}{2} + ir\right) \tau_{\mathfrak{a}}\left(m, \frac{1}{2} + ir\right) \omega(n, \psi) \omega(m, \varphi)} dr. \end{aligned}$$

The left-hand side is called the **geometric side** and the right-hand side is called the **spectral side**. We collect our work as a theorem:

Theorem 0.1.1 (Kuznetsov trace formula). *Let $\{u_j\}_{j \geq 1}$ be an orthonormal basis of Hecke-Maass eigenforms for $\mathcal{L}(N, \chi)$ of types ν_j with Fourier coefficients $a_j(n)$. Then for any positive integers $n, m \geq 1$, we have*

$$\begin{aligned} \delta_{n,m}(\psi, \varphi) + \sum_{\substack{c \geq 1 \\ c \equiv 0 \pmod{N}}} \frac{1}{c} S_\chi(n, m, c) V(n, m, c, \psi, \varphi) &= \sum_{j \geq 1} \overline{a_j(n) a_j(m) \omega(n, \psi) \omega(m, \varphi)} \\ &\quad + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \overline{\tau_{\mathfrak{a}}\left(n, \frac{1}{2} + ir\right) \tau_{\mathfrak{a}}\left(m, \frac{1}{2} + ir\right) \omega(n, \psi) \omega(m, \varphi)} dr. \end{aligned}$$

Chapter 1

The Theory of Sieves

Sieves are an important tool in analytic number theory because they allow for the estimation of the size of a sifted sequence of numbers from some initial sequence. In practice, one is usually sifting out only those indices relatively prime to some fixed integer and up to some prescribed size. In the following, we introduce sieves in some generality.

1.1 The Language of Sieves

Let $\mathcal{A} = (a_n)_{n \geq 1}$ be a nonnegative sequence and \mathcal{P} be a set of primes. We call \mathcal{A} the **sifting sequence**, \mathcal{P} the **sifting range**, and any $p \in \mathcal{P}$ a **sifting prime**. For any integer $d \geq 1$, let \mathcal{A}_d be the subsequence consisting of those terms a_n with $n \equiv 0 \pmod{d}$. Define the sum $S(x; \mathcal{A})$ by

$$S(x; \mathcal{A}) = \sum_{n \leq x} a_n,$$

for $x \geq 0$. Moreover, we define the function $P(z)$ by

$$P(z) = \prod_{\substack{p \in \mathcal{P} \\ p < z}} p,$$

for $z \geq 2$ and call z the **sifting level**. In other words, $P(z)$ is the product of the sifting primes in the sifting range up to z . We define the corresponding **sifting function** $S(x, z; \mathcal{A})$ by

$$S(x, z; \mathcal{A}) = \sum_{\substack{n \leq x \\ (n, P(z)) = 1}} a_n,$$

for $x \geq 0$ and $z \geq 2$. Equivalently, $S(x, z; \mathcal{A})$ consists of only those terms a_n in $S(x; \mathcal{A})$ whose indices n are not divisible by any sifting prime. Thus we have removed those terms a_n with $n \equiv 0 \pmod{p}$ for some sifting prime p up to z (this is equivalent to $p \mid P(z)$). The objective of any particular sieve is to produce upper or lower bounds for the sifting function $S(x, z; \mathcal{A})$. That is, we want to estimate the sum of elements a_n of the sifting sequence \mathcal{A} whose indices n are relatively prime to all sifting primes in the sifting range \mathcal{P} up to the sifting level z . We begin by deriving a useful relation for the sifting function that will motivate the construction of sieves. Applying μ to $S(x, z; \mathcal{A})$ to remove the condition $(n, P(z)) = 1$ gives

$$S(x, z; \mathcal{A}) = \sum_{n \leq x} \sum_{d \mid (n, P(z))} \mu(d) a_n,$$

and upon interchanging the two sums we obtain **Legendre's identity**:

$$S(x, z; \mathcal{A}) = \sum_{d|P(z)} \mu(d) S(x; \mathcal{A}_d).$$

In terms of Legendre's identity, the usefulness of sieves comes from replacing the Möbius function with weights that essentially serve as a truncation of the Möbius function. More precisely, we replace $\mu(d)$ with a weighting factor λ_d coming from a real sequence $\Lambda = (\lambda_d)_{d \geq 1}$ satisfying $\lambda_d = 0$ unless d is square-free, $d | P(z)$, and $d < D$ for some $D \geq 1$. We will also let λ represent the arithmetic function $\lambda(d) = \lambda_d$. We call Λ a **sieve**, λ_d a **sieve weight**, and the minimal such D the **sieve level** of Λ . We also define the **sifting variable** s by

$$s = \frac{\log(D)}{\log(z)}.$$

That is, the sifting variable measures the size of the sieve level relative to the sifting level on a logarithmic scale. Note that the sifting level, sifting variable, and sieve level satisfy the equivalent identity

$$D = z^s.$$

We define the **sieving function** $S^\Lambda(x, z; \mathcal{A})$ by

$$S^\Lambda(x, z; \mathcal{A}) = \sum_{d|P(z)} \lambda_d S(x; \mathcal{A}_d),$$

for $x > 0$ and $z \geq 2$. Expanding $S(x; \mathcal{A}_d)$, we can write

$$S^\Lambda(x, z; \mathcal{A}) = \sum_{n \leq x} a_n \sum_{d|(n, P(z))} \lambda_d.$$

Now set

$$\theta_n^0 = \sum_{d|(n, P(z))} \mu(d) \quad \text{and} \quad \theta_n = \sum_{d|(n, P(z))} \lambda_d.$$

We will also let θ^0 and θ represent the arithmetic functions $\theta^0(n) = \theta_n^0$ and $\theta(n) = \theta_n$ respectively. Equivalently, we can express θ_n^0 and θ_n in terms of the Dirichlet convolutions (see ??) $\theta^0 = \mu * \mathbf{1}$ and $\theta = \lambda * \mathbf{1}$ respectively. Note that θ_n^0 and θ_n (and hence θ^0 and θ as well) both depend on the sifting level z , but we suppress this dependence from the notation. In any case, we have

$$S^\Lambda(x, z; \mathcal{A}) = \sum_{n \leq x} a_n \theta_n.$$

We say that Λ is an **upper sieve** if the sieving function $S^\Lambda(x, z; \mathcal{A})$ is an upper bound for the sifting function $S(x, z; \mathcal{A})$. From the definition of the sifting function, Λ will be an upper sieve if and only if $\theta_n \geq \theta_n^0$. Analogously, we say that Λ is a **lower sieve** if the sieving function $S^\Lambda(x, z; \mathcal{A})$ is a lower bound for the sifting function $S(x, z; \mathcal{A})$. Similarly, Λ will be a lower sieve if and only if $\theta_n \leq \theta_n^0$. We will often optimize the choice of sieve weights such that the lower or upper bound is as tight as possible. When we wish to distinguish upper and lower sieves, we denote them by $\Lambda^\pm = (\lambda_d^\pm)_{d \geq 1}$ respectively and write \pm in place of Λ when denoting dependences upon the sieves. Then we have the upper and lower bounds

$$S^-(x, z; \mathcal{A}) \leq S(x, z; \mathcal{A}) \leq S^+(x, z; \mathcal{A}),$$

provided

$$\theta_n^- \leq \theta_n^0 \leq \theta_n^+. \quad (1.1)$$

We can also compose sieves to form new sieves from old ones. To explain the construction, let $\Lambda' = (\lambda'_d)_{d \geq 1}$ and $\Lambda'' = (\lambda''_d)_{d \geq 1}$ be two sieves of levels D' and D'' respectively. Also let $\theta' = \lambda' * \mathbf{1}$ and $\theta'' = \lambda'' * \mathbf{1}$. Then we define the **composite sieve** $\Lambda = \Lambda' \Lambda''$ of Λ' and Λ'' by

$$\lambda_d = \sum_{\substack{d_1, d_2 | d \\ [d_1, d_2] = d}} \lambda'_{d_1} \lambda''_{d_2},$$

for $d \geq 1$. Clearly Λ is a sieve of level $D'D''$. Moreover, summing over pairs (d_1, d_2) of divisors of n is the same as summing over triples (d_1, d_2, d) of divisors of n with $[d_1, d_2] = d$, we see that

$$\theta = \lambda * \mathbf{1} = (\lambda' * \mathbf{1})(\lambda'' * \mathbf{1}) = \theta' \theta''.$$

From this chain and Equation (1.1) it follows that the composition of an upper and lower sieve is a lower sieve while the composition of two upper or two lower sieves is an upper sieve.

1.2 Estimating the Sifting Function

Ultimately, our aim is to estimate the sifting function in terms of upper or lower sieving functions. In order to achieve this, we require an additional assumption about the sums $S(x; \mathcal{A}_d)$. In particular, we assume that there exists a multiplicative arithmetic function g with $g(1) = 1$, $0 \leq g(d) < 1$ for all $d > 1$ where $g(d) = 0$ unless d is square-free and the primes dividing d belong to \mathcal{P} , satisfies

$$\prod_{w \leq p < z} (1 - g(p))^{-1} \leq K \left(\frac{\log(z)}{\log(w)} \right)^\kappa, \quad (1.2)$$

for constants $K > 1$ and $\kappa \geq 0$ and all w and z with $z > w \geq 2$, and such that

$$S(x; \mathcal{A}_d) = g(d)M(x; \mathcal{A}) + r_d(x; \mathcal{A}), \quad (1.3)$$

for some smooth functions $M(x; \mathcal{A})$ and $r_d(x; \mathcal{A})$ with $M(x; \mathcal{A})$ nonnegative. We call the function g a **density function** and the constant κ the **sieve dimension**. Using Legendre's identity, we can write

$$S(x, z; \mathcal{A}) = M(x; \mathcal{A}) \sum_{d|P(z)} \mu(d)g(d) + \sum_{d|P(z)} \mu(d)r_d(x; \mathcal{A}).$$

Moreover, if we define functions $V(z)$ and $R(x, z; \mathcal{A})$ by

$$V(z) = \sum_{d|P(z)} \mu(d)g(d) \quad \text{and} \quad R(x, z; \mathcal{A}) = \sum_{d|P(z)} \mu(d)r_d(x; \mathcal{A}),$$

for $x > 0$ and $z \geq 2$, then we can further write

$$S(x, z; \mathcal{A}) = M(x; \mathcal{A})V(z) + R(x, z; \mathcal{A}).$$

As g is multiplicative, the definition of the Möbius function allows us to express $V(z)$ as a product:

$$V(z) = \prod_{p|P(z)} (1 - g(p)). \quad (1.4)$$

In particular, since $g(p) = 0$ if $p \notin \mathcal{P}$, Equation (1.2) can be expressed in the form

$$\frac{V(w)}{V(z)} \leq K \left(\frac{\log(z)}{\log(w)} \right)^\kappa. \quad (1.5)$$

Since the Möbius function changes sign, it is difficult to estimate $R(x, z; \mathcal{A})$ beyond trivial bounds. We can do much better with a sieve Λ . In this case, the definition of the sieving function gives

$$S_d^\Lambda(x, z; \mathcal{A}) = M(x; \mathcal{A}) \sum_{d|P(z)} \lambda_d g(d) + \sum_{d|P(z)} \lambda_d r_d(x; \mathcal{A}).$$

Defining functions $V^\Lambda(z)$ and $R^\Lambda(x, z; \mathcal{A})$ by

$$V^\Lambda(z) = \sum_{d|P(z)} \lambda_d g(d) \quad \text{and} \quad R^\Lambda(x, z; \mathcal{A}) = \sum_{d|P(z)} \lambda_d r_d(x; \mathcal{A}),$$

for $x > 0$ and $z \geq 2$, we can further write

$$S^\Lambda(x, z; \mathcal{A}) = M(x; \mathcal{A}) V^\Lambda(z) + R^\Lambda(x, z; \mathcal{A}). \quad (1.6)$$

Therefore, estimates for the sieving function $S^\Lambda(x, z; \mathcal{A})$ reduce to estimates for $M(x; \mathcal{A})$, $V^\Lambda(z)$, and $R^\Lambda(x, z; \mathcal{A})$. In order for Equation (1.3) to be useful, the functions $M(x; \mathcal{A})$ and $r_d(x; \mathcal{A})$ should be such that we know the order of magnitude of $M(x; \mathcal{A})$ and that $r_d(x; \mathcal{A})$ is of smaller order of magnitude. In this case, it suffices to estimate $V^\Lambda(z)$, and $R^\Lambda(x, z; \mathcal{A})$. For $R^\Lambda(x, z; \mathcal{A})$, we will usually discard possible cancellation from the individual terms $\lambda_d r_d(x; \mathcal{A})$ and use the trivial bound

$$R^\Lambda(x, z; \mathcal{A}) \leq |R^\Lambda|(x, z; \mathcal{A}),$$

where

$$|R^\Lambda|(x, z; \mathcal{A}) = \sum_{d|P(z)} |\lambda_d r_d(x; \mathcal{A})|.$$

In order to obtain estimates for $V^\Lambda(z)$, we express it in a more useful form that involves $V(z)$. For this, we define multiplicative arithmetic functions h and j defined by

$$h(p^r) = \frac{g(p^r)}{(1 - g(p^r))} \quad \text{and} \quad j(p^r) = \frac{1}{(1 - g(p^r))},$$

for all primes p and $r \geq 0$. Moreover, these two formulas imply

$$j(p^r) = \frac{h(p^r)}{g(p^r)} = 1 + h(p^r).$$

We call h the **relative density function**. In particular, $j = h * \mathbf{1}$. Now as $\theta = \lambda * \mathbf{1}$, the Möbius inversion formula (see ??) implies $\lambda = \theta * \mu$ so that

$$V^\Lambda(z) = \sum_{d|P(z)} \sum_{e|d} \theta_e \mu\left(\frac{d}{e}\right) g(d).$$

Making the change of variables $d \rightarrow ed$ and noting that $(d, e) = 1$ because $P(z)$ is square-free, we compute

$$\begin{aligned} V^\Lambda(z) &= \sum_{de|P(z)} \theta_e \mu(d) g(de) \\ &= \sum_{e|P(z)} \theta_e g(e) \prod_{p|\frac{P(z)}{e}} (1 - g(p)) \\ &= \sum_{e|P(z)} \theta_e h(e) \prod_{p|P(z)} (1 - g(p)), \end{aligned}$$

where the second line holds by the multiplicativity of g and the definition of the Möbius function. Defining the function $G^\Lambda(z)$ by

$$G^\Lambda(z) = \sum_{d|P(z)} \theta_d g(d),$$

for $z \geq 2$, Equation (1.4) further implies that

$$V^\Lambda(z) = V(z)G^\Lambda(z). \tag{1.7}$$

Therefore estimates for $V^\Lambda(z)$ reduce to estimates for $G^\Lambda(z)$ and $V(z)$. In particular, Equations (1.6) and (1.7) together give the identity

$$S^\Lambda(x, z; \mathcal{A}) = M(x; \mathcal{A})V(z)G^\Lambda(z) + R^\Lambda(x, z; \mathcal{A}).$$

Chapter 2

Types of Sieves

2.1 Todo: [Brun's Sieve]

The modern development of sieves began with Brun (see [?]) in 1915. His idea was to apply a truncation of the inclusion-exclusion principle to the sifting function. To motivate this idea, recall that the sifting function $S(x, z; \mathcal{A})$ is obtained from $S(x; \mathcal{A})$ by removing the terms a_n for which a sifting prime p divides n (maybe more than one). Removing those terms a_n according to the smallest sifting prime p that divides n , we immediately arrive at **Buchstab's identity**:

$$S(x, z; \mathcal{A}) = S(x; \mathcal{A}) - \sum_{p|P(z)} S(x, p; \mathcal{A}_p).$$

Since $S(x, 2; \mathcal{A}) = S(x; \mathcal{A})$, we may apply this formula iteratively r times to obtain

$$S(x, z; \mathcal{A}) = \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d) S(x; \mathcal{A}_d) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d) = r}} S(x, p(d); \mathcal{A}_d), \quad (2.1)$$

where $p(d)$ is the smallest prime divisor of d . This identity can be thought of as an inclusion-exclusion principle for the sifting function. As $S(x, p(d); \mathcal{A}_d)$ is nonnegative, Equation (2.1) produces the upper bound

$$S(x, z; \mathcal{A}) \leq \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d) S(x; \mathcal{A}_d), \quad (2.2)$$

if r is odd, and the lower bound

$$S(x, z; \mathcal{A}) \geq \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d) S(x; \mathcal{A}_d), \quad (2.3)$$

if r is even. Thus Equation (2.1) gives upper and lower bounds for the sifting function depending on the parity of r . Accordingly, for a fixed $r \geq 1$, **Brun's sieve** $\Lambda = (\lambda_d)_{d \geq 1}$ is given by

$$\lambda_d = \begin{cases} \mu(d) & \text{if } \omega(d) < r \text{ and } d \mid P(z), \\ 0 & \text{if } \omega(d) \geq r \text{ or } d \nmid P(z). \end{cases}$$

Clearly the sieve level is $D = P(z)$. Moreover, Equations (2.2) and (2.3) together imply

$$S(x, z; \mathcal{A}) \leq S^\Lambda(x, z; \mathcal{A}) \quad \text{or} \quad S(x, z; \mathcal{A}) \geq S^\Lambda(x, z; \mathcal{A}),$$

according to if r is odd or even. So Brun's sieve is an upper sieve if r is odd and a lower sieve if r is even. If we assume that there exists a density function g for the sums $S(x; \mathcal{A}_d)$, then

$$V^\Lambda(z) = \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d)g(d) \quad \text{and} \quad R^\Lambda(x, z; \mathcal{A}) = \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d)r_d(x; \mathcal{A}).$$

To see the relationship between $V(z)$ and $V^\Lambda(z)$, we will derive an analog of Buchstab's identity. Expanding Equation (1.4) and grouping terms according to the largest sifting prime p such that $g(p)$ divides them, we arrive at the formula

$$V(z) = 1 - \sum_{p|P(z)} g(p)V(p),$$

which is clearly an analog to Buchstab's identity. Since $V(2) = 1$, we may apply this formula iteratively r times and obtain

$$V(z) = \sum_{\substack{d|P(z) \\ \omega(d) < r}} \mu(d)g(d) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d) = r}} g(d)V(p(d)).$$

Note that the first sum on the right-hand side is exactly $V^\Lambda(z)$ and so the difference between $V^\Lambda(z)$ and $V(z)$ is exactly the second sum which alternates in sign according to the parity of r . We now aim to estimate $G^\Lambda(z)$. To this end, let

$$G(z) = \sum_{p|P(z)} g(p).$$

Then $G(z) = -G^\Lambda(z)$ when $r = 1$. Moreover, the Taylor series of the logarithm gives

$$G(z) \leq \sum_{p|P(z)} -\log(1 - g(p)) = -\log V(z).$$

Upon expanding the product $G(z)^r$, it is also clear that $G^\Lambda(z) \leq \frac{G(z)^r}{r!}$ for any $r \geq 1$. These facts with the inequality $r! \geq e \left(\frac{r}{e}\right)^r$ (which easily follows by induction), and the identity $-\log V(z) = |\log V(z)|$, together imply

$$G^\Lambda(z) \leq \frac{G(z)^r}{r!} \leq \frac{1}{e} \left(\frac{eG(z)}{r} \right)^r \leq \frac{1}{e} \left(\frac{e|\log V(z)|}{r} \right)^r. \quad (2.4)$$

We now choose r in order to make this last expression small. Clearly we at least need $r \geq c|\log V(z)|$, for some $c > e$, so that $\frac{e|\log V(z)|}{r} < 1$. To this end, let c be the unique constant satisfying

$$\left(\frac{c}{e}\right)^c = e.$$

Note that $c > e$ because $e > 1$. Moreover, for any $b \geq c$ we have the inequality $b^b \geq e^{2b-c+1}$. Indeed, taking the logarithm and isolating b gives $b(\log(b) - 2) \geq 1 - c$ which is equality when $b = c$ and the left-hand side is an increasing function of b . Letting $r \geq c|\log V(z)|$ and $b = \frac{r}{|\log V(z)|}$, the aforementioned inequality is equivalent to

$$\left(\frac{|\log V(z)|}{r} \right)^r \geq e^{-2r} V(z)^{1-c}. \quad (2.5)$$

Combining Equations (2.4) and (2.5) results in the bound

$$G^\Lambda(z) \leq e^{-(r+1)} V(z)^{1-c}.$$

We choose r to depend on the sifting variable s by setting $r = \lfloor s \rfloor$ provided $s \geq 1 - c \log V(z)$ (recall $-\log V(z) = |\log V(z)|$ so that $s \geq c|\log V(z)|$). Then our bound for $G^\Lambda(z)$ becomes

$$G^\Lambda(z) \leq e^{-s} V(z)^{1-c}.$$

We now establish a relationship between the sifting function $S(x, z; \mathcal{A})$ and the sieving function $S^\Lambda(x, z; \mathcal{A})$ that will allow us to make use of these estimates. **Todo:** [xxx]

2.2 **Todo:** [Beta Sieves]

We now wish to general Brun's sieve by replacing the associated sieving function with one that better approximates the sifting function. Precisely, we will replace the condition $\omega(d) < r$ and $p \mid P(z)$ with the condition $d \in \mathcal{D}$ for set of square-free positive integers \mathcal{D} with a small amount of small prime divisors coming from the sifting range \mathcal{P} . We will be concerned with upper and lower sieves $\Lambda^+ = (\lambda_d^+)_{d \geq 1}$ and $\Lambda^- = (\lambda_d^-)_{d \geq 1}$ where $\lambda_d^\pm = \mu(d)$ only for d belonging to the sets

$$\mathcal{D}^+ = \{d = p_1 p_2 \cdots p_r < D : p_1, \dots, p_r \mid P(z), p_r < p_{r-1} < \cdots < p_1, \text{ and } p_m < y_m \text{ for } m \text{ odd}\},$$

and

$$\mathcal{D}^- = \{d = p_1 p_2 \cdots p_r < D : p_1, \dots, p_r \mid P(z), p_r < p_{r-1} < \cdots < p_1, \text{ and } p_m < y_m \text{ for } m \text{ even}\},$$

respectively, with $1 \in \mathcal{D}^\pm$, and for integer $D \geq 1$ and parameters $y_m > 0$ for all $m \geq 1$. The parameters y_m are chosen to optimize the particular sifting sequence \mathcal{A} . For \mathcal{D}^\pm , the **beta sieve** $\Lambda^\pm = (\lambda_d^\pm)_{d \geq 1}$ is defined by

$$\lambda_d^\pm = \begin{cases} \mu(d) & \text{if } d \in \mathcal{D}^\pm, \\ 0 & \text{if } d \notin \mathcal{D}^\pm. \end{cases}$$

Clearly the sieve level is D . **Todo:** [xxx]

2.3 **Todo:** [Λ^2 Sieves]

2.4 **Todo:** [The Large Sieve]