Модулярные формы

1. Автоморфные функции на верхней полуплоскости.

Группа $\mathrm{GL}_2^+(\mathbf{R})$ действует на верхней полуплоскости H по следующему правилу: если $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, то $\alpha z = \frac{az+b}{cz+d}$. Скалярные матрицы при этом действуют тривиально.

Пусть $k \in \mathbf{Z}$. Определим действие $\mathrm{GL}_2^+(\mathbf{R})$ справа на функциях на H по правилу $f[\alpha]_k(z) \stackrel{\text{def}}{=} (\det \alpha)^{k/2}(cz+d)^{-k} f(\alpha z)$. Будем называть функцию f автоморфной веса k для подгруппы $G \subset \mathrm{GL}_2^+(\mathbf{R})$, если $f[\alpha]_k = f$ при всех $\alpha \in G$.

В дальнейшем будут обсуждаться автоморфные функции только для группы $\mathrm{SL}_2(\mathbf{Z}) \stackrel{\text{def}}{=} \Gamma$ и некоторых её подгрупп G конечного индекса. Условие автоморфности веса k в этом случае выглядит как $\forall \alpha \in G \;\; f(\alpha z) = (cz+d)^k f(z)$.

2. Комплексная структура на $G \setminus \overline{\mathrm{H}}$.

Положим $\overline{\mathrm{H}} \stackrel{\text{def}}{=} \mathrm{H} \cup \mathbf{Q} \cup i\infty$. Группа Γ, а значит и группа G действует также и на $\overline{\mathrm{H}}$ (не смешивая при этом H и $\mathbf{Q} \cup i\infty$). Будем считать, что $\overline{G} \stackrel{\text{def}}{=} G$, если матрица $(-1)_2 \stackrel{\text{def}}{=} \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \notin G$, и $\overline{G} \stackrel{\text{def}}{=} G/(\pm 1)_2$ в противном случае. Группа $\overline{G}$ действует на H свободно везде, кроме конечного числа орбит. Точки H, лежащие на этих орбитах, а иногда и сами орбиты (в зависимости от контекста) называются эллиптическими точками группы G.
Эллиптические точки группы Γ суть Γ-орбиты точки $z = i \stackrel{\text{def}}{=} \sqrt{-1}$ (индекс $\stackrel{\text{def}}{=}$ порядок стационарной подгруппы в $\overline{\Gamma}$ равен 2, сама эта подгруппа для $z = i$ порождена образом в $\overline{\Gamma}$ матрицы $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$), а стационарные подгруппы остальных точек орбиты с ней сопряжены) и точки $z = \rho \stackrel{\text{def}}{=} \frac{1+\sqrt{-3}}{2}$ (индекс равен 3, стационарная подгруппа для $z = \rho$ порождена образом матрицы $TS = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$). Все эллиптические точки группы G суть прообразы эллиптических точек Γ при отображении $G \setminus \mathrm{H} \to \Gamma \setminus \mathrm{H}$. Прообраз эллиптической точки может не быть эллиптической точкой, но если он ею остается, то индекс не меняется.

Параболической точкой группы G называется орбита G в $\mathbf{Q} \cup i\infty$. У группы Γ одна параболическая точка, поскольку все элементы $\mathbf{Q} \cup i\infty$ Γ- эквивалентны. Стационарная подгруппа $\overline{\Gamma}_{i\infty} \subset \overline{\Gamma}$ изоморфна $\mathbf{Z}$ и порождается образом матрицы $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, стационарные подгруппы остальных точек орбиты с ней сопряжены. Все параболические точки подгруппы конечного индекса $G \in \Gamma$ лежат над единственной параболической точкой Γ. Если $a \in \mathbf{Q} \cup i\infty$ - представитель какой-нибудь параболической точки G, то его стационарная подгруппа $\overline{G}_a$ имеет конечный индекс в $\overline{\Gamma}_a$. Последняя сопряжена с $\overline{\Gamma}_{i\infty}$ и поэтому изоморфна $\mathbf{Z}$, поэтому и $\overline{G}_a$ изоморфна $\mathbf{Z}$. В частности, $\overline{G}_{i\infty}$ порождается образом матрицы T^e для некоторого положительного $e \in \mathbf{Z}$.

Комплексная структура на $G \setminus \overline{\mathrm{H}} \stackrel{\mathrm{def}}{=} X_G$ определяется так. Для орбит свободного действия $\overline{G}$ на H локальным параметром можно считать локальный параметр на H в окрестности любого представителя орбиты. Для параболической точки, являющейся G - орбитой $i\infty$, в качестве локального параметра можно использовать $q_e(z) = e^{2\pi iz/e}$, а для остальных параболических точек выбрать представителя $a \in \mathbf{Q}$ и элемент $\gamma \in \Gamma$ такой, что $\gamma a = i\infty$, после чего использовать $q_e(\gamma a)$ в качестве локального параметра. Наконец, для эллиптической точки с представителем $z_0 \in \mathrm{H}$ можно использовать локальный параметр $(\frac{z-z_0}{z-\overline{z}_0})^e$, где e - индекс эллиптической точки.

Несложно проверить, что X_G с такой комплексной структурой является компактной римановой поверхностью с выделенными конечными наборами, соответственно, эллиптических точек (их может не быть) и параболических точек (хотя бы одна обязательно есть). Если $G_1 \subset G$, то определено каноническое голоморфное накрытие $X_{G_1} \to X_G$, которое может ветвиться только над эллиптическими и параболическими точками X_G. Если прообраз эллиптической точки для G есть эллиптическая точка для G_1, то в этом прообразе ветвления нет, а если в этом прообразе $\overline{G}_1$ действует свободно, то индекс ветвления равен индексу образа как эллиптической точки для G (т.е. 2 или 3). Индекс ветвления в параболической точке группы G_1 равен индексу $(\overline{G}_a : (\overline{G_1})_a)$ для какого-нибудь её представителя a.

3. Условия аналитичности.

Начиная с этого момента, мы будем изучать сравнительно обозримый класс автоморфных функций, которые являются мероморфными на H, а также мероморфны во всех параболических точках группы G. Будем называть такие функции модулярными. Ограничимся модулярными функциями четного веса (для групп G, содержащих матрицу $(-1)_2$, в частности, для самой группы Γ, автоморфных функций нечетного веса не существует), поскольку для них теория технически проще.

В параболической точке, которую представляет $i\infty$, функцию будем считать мероморфной, если она разлагается в ряд Лорана по степеням локального параметра q_e, определенного в предыдущем разделе. Если $a \in \mathbf{Q}$ - представитель другой параболической точки, то выберем, как и прежде, матрицу $\gamma \in \Gamma$ такую, что $\gamma a = i\infty$, и рассмотрим функцию $f[\gamma]_k$. Операторы $[\alpha]_k$ для $\alpha \in \gamma^{-1} G \gamma$ действуют на эту функцию тождественно, в частности это верно при $\alpha \in \gamma^{-1} G_a \gamma$. Образ последней группы в $\overline{\Gamma}$ порождается матрицей $\binom{1\,h}{0\,1}$ при некотором $h \in \mathbf{Z}_{>0}$, а оператор $[(-1)_2]_k$ действует тождественно в силу четности веса, так что $f[\gamma]_k$ инвариантна относительно сдвига $z \mapsto z+h$, а следовательно, зависит только от q_h. Условие мероморфности f в параболической точке, представленной a, состоит в том, что, что $f[\gamma]_k$ должна разлагаться в ряд Лорана по q_h.

Произведение автоморфных функций веса k и веса l для G является, очевидно, автоморфной функцией веса $k+l$ для G, так что пространство (линейное над $\mathbf{C}$) модулярных функций веса k является не более чем одномерным векторным пространством над полем модулярных функций веса 0 для G, которое, в свою очередь, совпадает с полем мероморфных функций на римановой поверхности X_G.

На самом деле это ровно одномерное пространство, поскольку непостоянная модулярная функция веса 0, очевидно, существует, а её производная, как легко проверить, является модулярной функцией веса 2, и k-я степень последней имеет вес $2k$. Однако при $k > 0$ неверно, что производная модулярной функции веса k является модулярной функцией.

Будем называть модулярную функцию веса k для G модулярной формой, если она голоморфна на H и во всех параболических точках, и параболической формой, если во всех параболических точках она обращается в нуль.

4. Теорема Римана-Роха.

Имеется взаимнооднозначное соответствие межу модулярными функциями веса k на G и мероморфными дифференциалами степени $k/2$ на X_G (напомним, что мы считаем вес k четным). Действительно, дифференциал $f(z)(dz)^{k/2}$, очевидно, G-инвариантен.
Пусть $f(z)$ соответствует дифференциалу ω на X_G. Тогда порядок $\mathrm{v}_Q(\omega)$ в точке $Q \in X_G$ можно вычислить, сравнивая локальные параметры. Выберем какой-нибудь представитель $P \in \overline{\mathrm{H}}$ орбиты, отвечающей Q. Ответ такой:
- если $P \in \mathrm{H}$ и точка Q не эллиптическая, то $\mathrm{v}_P(f) = \mathrm{v}_Q(\omega)$;
- если Q - эллиптическая индекса 2, то $\mathrm{v}_P(f) = 2\mathrm{v}_Q(\omega) + k/2$;
- если Q - эллиптическая индекса 3, то $\mathrm{v}_P(f) = 3\mathrm{v}_Q(\omega) + k$;

- если Q - параболическая точка, то $\mathrm{v}_P(f) = \mathrm{v}_Q(\omega) + k/2$.

Действительно, локальные параметры на H и на X_G связаны соотношением $t_Q = \left(\frac{t_P}{z-\overline{P}}\right)^e$, где e равно 1 в обычных точках и индексу в эллиптических. $f(z) \sim t_P^{\mathrm{v}_P(f)} dz^{k/2}$ в окрестности P, а $\omega \sim t_Q^{\mathrm{v}_Q(\omega)} dt_Q^{k/2}$ в окрестности Q. Поскольку $t_Q \sim t_P^e$, приравнивая эти выражения, получаем, что $\mathrm{v}_P(f) = e\mathrm{v}_Q(\omega) + (e-1)\frac{k}{2}$. Если Q - параболическая точка, отвечающая орбите $P = i\infty$, то локальным параметром в P по нашему определению мероморфной на $\overline{\mathrm{H}}$ функции является $t_Q = e^{2\pi i z/e}$, и он же определяет комплексную структуру на X_G в окрестности Q, следовательно, $(dt_Q)^{k/2} \sim (t_Q)^{k/2} dz$, что завершает доказательство, ибо в остальных параболических точках локальнные параметры определяются через посредство локальных параметров в $i\infty$.

Пусть D_2, D_3 и D_∞ - дивизоры на X_G, являющиеся формальной суммой, соответственно, эллиптических индекса 2, эллиптических индекса 3 и параболических точек. Из приведенных формул следует, что модулярные формы (см.определение выше) веса k с точностью до умножения на константу соответствуют дивизорам из класса (ω), превращающимся в эффективные после прибавления дивизора $D_0 \stackrel{\text{def}}{=} [k/4]D_2 + [k/3]D_3 + (k/2)D_\infty$. Пусть D_ω - какой-нибудь дивизор из класса (ω). Обозначим $M_k(G)$ пространство (векторное над $\mathbf{C}$) модулярных форм веса k, тогда $\dim M_k(G) = l(D_0 + D_\omega)$.

Модулярных форм отрицательного веса не бывает (это следует из того, что степень дивизора $D_0 + D_\omega$ при $k < 0$ отрицательна, что очевидно при $g > 0$, ибо $\deg \omega = \frac{k}{2}(2g-2) = k(g-1)$), и будет проверено позже при $g = 0$). Если $k = 0$, то $D_0 = 0$ и можно считать, что $D_\omega = 0$, так что пространство $M_0(G)$ одномерно и состоит из констант. Далее мы считаем, что $k > 0$.

Степень дивизора $D_0 + D_\omega$ равна $[k/4]\nu_2 + [k/3]\nu_3 + (k/2)\nu_\infty + k(g-1)$, где g - род X_G, а ν_j - количества точек на X_G из вышеприведенного списка. При $g > 0$ эта степень превосходит $2g-2$ и, тем самым, дивизор не является специальным и теорема Римана-Роха дает $\dim M_k(G) = l(D_0 + D_\omega) = 1 - g + \deg(D_0 + D_\omega) = [k/4]\nu_2 + [k/3]\nu_3 + (k/2)\nu_\infty + (k-1)(g-1)$. Если же $g = 0$, то $\dim M_k(G) = \max(0, 1 + \deg(D_0 + D_\omega)) = \max(0, [k/4]\nu_2 + [k/3]\nu_3 + (k/2)\nu_\infty + 1 - k)$.

Аналогично вычисляется размерность пространства $S_k(G)$ параболических форм, только дивизор D_0 надо заменить на дивизор $D_1 \stackrel{\text{def}}{=} [k/4]D_2 + [k/3]D_3 + (k/2-1)D_\infty$.

Константы не параболичны, так что $\dim S_0(G) = 0$. При $k > 2$ и $g > 0$ дивизор $D_1 + D_\omega$ тоже неспециален и $\dim S_k(G) = \dim M_k(G) - \nu_\infty$. При $k = 2$ и $g > 0$ дивизор $D_1 = 0$, следовательно, $D_1 + D_\omega$ лежит в каноническом классе, поэтому $\dim S_k(G) = g$. Случай $g = 0$ охватывается прежней формулой (с заменой в ней степени D_0 на степень D_1).

5. Модулярные формы для $\Gamma = \mathrm{SL}_2(\mathbf{Z})$.

Γ порождается матрицами $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right))$ и $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Фундаментальная область для действия Γ на H нарисована на скрижалях. Уже из её вида следует, что X_Γ - это риманова сфера (т.е. что $g(X_\Gamma) = 0$), но ниже будет определена модулярная функция $j(z)$ веса 0 для Γ, осуществляющая изоморфизм $X_\Gamma \to \mathbf{P}^1(\mathbf{C})$.
На X_Γ две эллиптические точки (орбита i индекса 2 и орбита ρ индекса 3) и одна параболическая (орбита $i\infty$).

Пусть f - модулярная функция веса k. Тогда справедлива формула:
$\mathrm{v}_{i\infty}(f) + \frac{1}{2}\mathrm{v}_i(f) + \frac{1}{3}\mathrm{v}_\rho(f) + \sum\limits_{P \neq i,\rho,i\infty} \mathrm{v}_P(f) = \frac{k}{12}$,
где суммирование ведется по орбитам Γ в $\overline{\mathrm{H}}$.
Эта формула немедлено следует из рассмотрений предыдущего раздела и того, что $g(X_\Gamma) = 0$, а значит, степень дивизора, отвечающего дифференциалу степени $k/2$ на X_Γ. равна $(-k)$. Впрочем, можно доказать её и непосредственно, интегрируя $d\log f$ по границе фундаментальной области.

Ряд Эйзенштейна $G_k(z) \stackrel{\mathrm{def}}{=} \sum\limits_{(m,n)\neq(0,0)} (mz+n)^{-k}$.
При $k \geq 4$ сходится абсолютно и определяет модулярную форму для Γ веса k.
При $k = 2$ можно определить $G_2(z) \stackrel{\mathrm{def}}{=} \sum\limits_m \sum\limits_{n,(m,n)\neq(0,0)} (mz+n)^{-k}$
Это уже не автоморфная функция, поскольку, хоть G_2 инвариантен относительно замены $z \mapsto z+1$ (т.е. матрицы T), но действие оператора $[S]_2$ меняет порядок суммирования (а сходимость при $k = 2$ не абсолютная).

$G_k(z) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum\limits_{n=1}^{\infty} \sigma_{k-1}(n)q^n$, где $q = e^{2\pi iz}$ и $\sigma_{k-1}(n) = \sum\limits_{d|n} d^{k-1}$.
Эта формула справедлива при всех четных $k \geq 2$ (следует из разложения в ряд Фурье $\sum\limits_{n\in\mathbf{Z}} \frac{1}{(z+n)^k} = \frac{(-2\pi i)^k}{(k-1)!}\sum\limits_{d=1}^{\infty} d^{k-1}e^{2\pi idz}$).

Положим $E_k \stackrel{\text{def}}{=} G_k/2\zeta(k) = 1 - \frac{2k}{B_k}\sum\limits_{n=1}^{\infty}\sigma_{k-1}(n)q^n$

Напомним, что $\zeta(2k) = (-1)^{k+1}\frac{(2\pi)^k}{2(2k)!}B_{2k}$, где B_k - числа Бернулли, заданные производящей функцией $\frac{t}{e^t-1} = \sum\limits_{k=0}^{\infty}\frac{B_k}{k!}t^k$. Это можно проверить, вычисляя контурный интеграл $\int\frac{z^{-2k}}{e^z-1}dz$ по контуру, обходящему положительную полуось, методом вычетов, а затем применяя функциональное уравнение для $\zeta(s)$, с учетом того, что $\Gamma(s)\zeta(s)$ есть преобразование Меллина функции $(e^t-1)^{-1}$.

В частности, $E_4 = 1 + 240\sum\limits_{n=1}^{\infty}\sigma_3(n)q^n$, а $E_6 = 1 - 504\sum\limits_{n=1}^{\infty}\sigma_5(n)q^n$. Из формулы в начале этого раздела следует, что E_4 имеет простой нуль на Γ-орбите $z = \rho$, а E_6 - на Γ-орбите $z = i$, и больше у них никаких нулей на H нет. Кратный дифференциал на X_Γ, отвечающий E_4, имеет простые полюса в эллиптических точках и двойной в параболической, а кратный дифференциал, отвечающий E_6 - простой полюс в эллиптической точке индекса 2, двойной в эллиптической точке индекса 3 и тройной в параболической точке. Нулей ни тот, ни другой не имеют.

Определим $\Delta \stackrel{\text{def}}{=} \frac{1}{1728}(E_4^3 - E_6^2)$ - это параболическая форма веса 12. Из формулы в начале раздела следует, что у неё нет нулей на H. На самом деле $\Delta = q\prod\limits_{n=1}^{\infty}(1-q^n)^{24}$ (ни сама формула, ни её доказательство нам не понадобятся).

Если $f_1 \in M_k(\Gamma)$ и $f_2 \in M_k(\Gamma)$ не пропорциональны, то некоторая их ненулевая линейная комбинация является параболической формой (ибо параболическая точка одна), и её можно поделить на Δ (так как Δ не имеет нулей на H и имеет простой нуль в параболической точке), поэтому $M_k(\Gamma) = \Delta M_{k-12}(\Gamma) + \langle E_k\rangle$ при $k \geq 12$. Используя формулы для размерности пространств модулярных форм из предыдущего раздела, легко вычислить все $M_k(\Gamma)$ и убедиться, что кольцо модулярных форм всех весов для Γ порождено E_4 и E_6.

Положим $j(z) \stackrel{\text{def}}{=} \frac{E_4^3(z)}{\Delta(z)}$. Тогда j - модулярная функция для Γ веса 0, голоморфная на H и имеющая простой полюс в параболической точке. Применяя формулу из начала раздела к функциям $j(z) - c$ при различных $c \in \mathbf{C}$. нетрудно убедиться, что $j(z)$ определяет изоморфизм X_Γ и $\mathbf{P}^1(\mathbf{C})$.

6. Конгруэнцподгруппы в $\mathrm{SL}_2(\mathbf{Z})$.

Пусть $G \subset \Gamma$ - подгруппа конечного индекса, и $(\overline{\Gamma} : \overline{G}) = \mu$ Тогда $X_G \to X_\Gamma$ - накрытие степени μ, и род X_G можно вычислить по формуле Гурвица. Накрытие может ветвиться только над эллиптическими или параболической точками X_Γ, и формула Гурвица дает: $2g(X_G) - 2 = -2\mu + \sum\limits_P (e_P - 1)$. Каждый слой накрытия содержит μ точек, если считать каждую с кратностью e_P. В слое над эллиптической точкой индекс ветвления e_P может быть равен только её индексу (2 или 3) или 1, при этом эллиптическими на X_G будут как раз те точки, в которых ветвления нет. Элементарная арифметика дает: $g(X_G) = 1 + \frac{\mu}{12} - \frac{\nu_2}{4} - \frac{\nu_3}{3} - \frac{\nu_\infty}{2}$.
Используя эту формулу, нетрудно завершить доказательство того, что модулярных форм отрицательного веса не бывает, даже когда $g(X_G) = 0$.

Определим нормальную подгруппу $\Gamma(N) \subset \Gamma$ как ядро естественного гомоморфизма $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))$. Подгруппа в Γ, содержащая $\Gamma(N)$, называется конгруэнцподгруппой уровня N. Определим конгруэнцподгруппу $\Gamma_1(N)$ как состоящую из матриц, чей образ в $\mathrm{SL}_2(\mathbf{Z}/(N))$ имеет вид $\left(\begin{smallmatrix}1 & * \\ 0 & 1\end{smallmatrix}\right)$, а подгруппу $\Gamma_0(N)$ как состоящую из матриц, чей образ в $\mathrm{SL}_2(\mathbf{Z}/(N))$ имеет вид $\left(\begin{smallmatrix}* & * \\ 0 & *\end{smallmatrix}\right)$. Легко видеть, что $\Gamma_1(N)/\Gamma(N) \simeq \mathbf{Z}/(N)$. Подгруппа $\Gamma_1(N)$ нормальна в $\Gamma_0(N)$, и $\Gamma_0(N)/\Gamma_1(N) \simeq (Z/(N))^*$.

Подгруппа $\Gamma_0(N)$ не является нормальной в Γ. Если определить правое действие Γ на $\mathbf{P}^1(\mathbf{Z}/(N))$ по формуле $(x : y)\left(\begin{smallmatrix}a & b \\ c & d\end{smallmatrix}\right) \stackrel{\text{def}}{=} ax + cy : bx + dy$, то это действие будет транзитивно и стационарная подгруппа точки $0 : 1$ совпадет с $\Gamma_0(N)$, так что множество правых смежных классов $\Gamma_0(N)\backslash\Gamma$ отождествится с $\mathbf{P}^1(\mathbf{Z}/(N))$. Стоит напомнить, что точки $\mathbf{P}^1(\mathbf{Z}/(N))$ суть орбиты действия $(\mathbf{Z}/(N))^*$ на множестве пар $x : y$ вычетов $\mod N$, таких что x, y и N взаимно просты в совокупности.

Подгруппа $\Gamma_0(N)$ достаточно мала для того, чтобы теория модулярных форм для неё была нетривиальной (согласно работам Уайлза и его соратников, дзета-функция любой эллиптической кривой, определенной над $\mathbf{Q}$, является преобразованием Меллина модулярной формы веса 2 для какой-нибудь $\Gamma_0(N)$) и в то же время достаточно велика для того, чтобы пространства модулярных форм удобно было описывать.

Входящие в формулу, приведенную в начале этого раздела, параметры для группы $\Gamma_0(N)$ принимают такие значения:

$\mu = N \prod\limits_{p|N}(1+\frac{1}{p})$
$\nu_2 = \prod\limits_{p|n}(1+\left(\frac{-1}{p}\right))$ при $4 \nmid N$, и 0 иначе
$\nu_3 = \prod\limits_{p|n}(1+\left(\frac{-3}{p}\right))$ при $9 \nmid N$, и 0 иначе
$\nu_\infty = \sum\limits_{d|N} \phi(\gcd(d, N/d))$
(мы формально считаем, что $\left(\frac{-1}{2}\right) = \left(\frac{-3}{3}\right) = 0$).

Параболические точки описываются следующим образом. Пусть $a = \frac{x}{y} \in \mathbf{Q}$, причем $y > 0$ и дробь несократима. Поставим в соответствие $\Gamma_0(N)$-орбите точки a пару $(d = \gcd(y, N), \frac{xy}{d} \mod \gcd(d, N/d))$. Орбите $i\infty$ сопоставим пару $(N, 1)$. Проверка взаимной однозначности этого соответствия и формул для ν_2 и ν_3 - несложное, но полезное упражнение.
В частности, если N просто, то $\Gamma_0(N)$ имеет ровно две параболические точки: орбиту $i\infty$, состоящую из самой $i\infty$ и всех $a \in \mathbf{Q}$, в знаменателе которых есть N, и орбиту нуля, состоящую из N - целых рациональных чисел.

Род кривой $X_{\Gamma_0(N)}$, которую часто обозначают $X_0(N)$, равен нулю при $2 \leq N \leq 10$ и при $N = 12, 13, 16, 18, 25$, а для остальных N положителен (в частности, $X_0(11)$ - кривая рода 1).

7. Операторы Гекке.

Будем использовать обозначения $z = x + iy$.

На пространстве $S_k(G)$ можно определить эрмитово скалярное произведение ("произведение Петерссона") по формуле $\langle f, g\rangle = \frac{1}{(\Gamma:G)} \int\limits_E f(z)\overline{g(z)} y^k \frac{dxdy}{y^2}$.
Подынтегральная функция G - инвариантна, а мера $\frac{dxdy}{y^2}$ инвариантна относительно любой матрицы из $\mathrm{GL}_2^+(\mathbf{R})$. Отсюда легко следует независимость от выбора фундаментальной области E группы G, по которой ведется интегрирование. Множитель в формуле подобран так, чтобы результат не зависел и от выбора самой группы G (параболическая форма f является также параболической формой для любой подгруппы $H \subset G$).
Для проверки (абсолютной) сходимости используется легко проверяемая оценка: для любой матрицы $\gamma \in \Gamma \quad |f(\gamma z) << e^{-cy}$, где $c > 0$ - не зависящая от γ константа.
Из этой оценки следует, что на самом деле интеграл, определяющий скалярное произведение, сходится, даже если только одна из модулярных форм f и g параболична.

Для работы с операторами Гекке нам понадобится распространить определение скалярного произведения на модулярные формы для групп вида $G_\alpha = \alpha G \alpha^{-1}$, где $\alpha \in \mathrm{GL}_2^+(\mathbf{Q})$ (можно считать, что α - целочисленная матрица). В отличие от G, такие группы не обязаны содержаться в Γ. Однако всегда существует подгруппа $G'_\alpha \subset G_\alpha$, являющаяся подгруппой Γ конечного индекса. Действительно, если α целочисленна и $\det \alpha = N$, то элементарно проверяется, что $\Gamma(N)_\alpha = \alpha \Gamma(N) \alpha^{-1} \subset \Gamma$. За G'_α можно взять пересечение G_α с $\Gamma(N)_\alpha$. Поскольку пересечение двух подгрупп конечного индекса также имеет конечный индекс, то теперь, пользуясь независимостью определения скалярного произведения от выбора подгруппы, можно определить $\langle f, g \rangle$ для любой пары модулярных форм веса k для групп вида G_α (возможно, с разными G и α).

Пусть $\alpha \in \mathrm{GL}_2^+(\mathbf{Q})$. Рутинная проверка позволяет установить следующие свойства скалярного произведения:
1) $\langle f[\alpha]_k, g[\alpha]_k \rangle = \langle f, g \rangle$ (в частности, если α нормализует G, то $[\alpha]_k$ унитарен).
2) $\langle f[\alpha]_k, g \rangle = \langle f, g[\alpha']_k \rangle$, где $\alpha' \stackrel{\mathrm{def}}{=} \det \alpha (\alpha^{-1})$
3) Если f и g - модулярные формы для группы G, то скалярное произведение из предыдущей строчки зависит только от двойного класса смежности $G \alpha G$.

Теперь мы определим пронумерованную натуральными числами серию коммутирующих друг с другом "усредняющих" операторов на множестве автоморфных функций веса k. Предположим сначала, что $G = \Gamma$.

Рассмотрим множество всех решеток $L \subset \mathbf{C}$. Комплекснозначная функция на решётках имеет вес k, если $F(\lambda L) = \lambda^{-k} F(L)$. Далее, рассмотрим множество векторов - столбцов $\binom{w_1}{w_2}$, где w_1 и w_2 - комплексные числа, такие, что $w_1/w_2 \in \mathrm{H}$. Любой функции F веса k на решетках соответствует функция $\tilde{F}$ веса k (определение такое же, как для решёток) на векторах-столбцах $\tilde{F}\binom{w_1}{w_2} \stackrel{\mathrm{def}}{=} F(L = \langle w_1, w_2 \rangle)$, инвариантная относительно действия группы $\mathrm{SL}_2(\mathbf{Z})$, определяемого естественным действием последней на векторах-столбцах.. Функции $\tilde{F}$, в свою очередь, соответствует автоморфная функция $f(z) \stackrel{\mathrm{def}}{=} \tilde{F}\binom{z}{1}$ веса k. Обратные отображения выглялят как $\tilde{F}\binom{w_1}{w_2} = w_2^{-k} f(w_1/w_2)$ и $F(L) = \tilde{F}($любого правильно ориентированного базиса $L)$.

Рассмотрим векторное пространство LT над $\mathbf{C}$, элементы базиса в котором пронумерованы символами $[L]$, где L - решётка. Положим $\mathrm{T}_n[L] \stackrel{\mathrm{def}}{=} \sum\limits_{(L:L')=n} [L']$ и $R_n[L] \stackrel{\mathrm{def}}{=} [nL]$.

Нетрудно проверить, что имеют место соотношения:
$\mathrm{T}_n\mathrm{T}_m = \mathrm{T}_{nm}$, если $\gcd(n,m) = 1$
$\mathrm{T}_{p^n}\mathrm{T}_p = \mathrm{T}_{p^{n+1}} + pR_p\mathrm{T}_{p^{n-1}}$ при $n \geq 1$.

Первая формула очевидна (любая подрешетка $L' \subset L$ индекса mn содержится в единственной подрешетке индекса n). Для доказательства второй достаточно сравнить кратности, с которыми каждая решетка индекса p^{n+1} входит в левую и в правую части равенства. Эти кратности равны 1, если $L' \subset pL$ и $p+1$, если $L' \not\subset pL$.

Функции F веса k можно продолжить на LT по линейности и введенные операции переносятся на них ($\mathrm{T}F(L) \stackrel{\mathrm{def}}{=} F(\mathrm{T}L)$). Имеют место соотношения:
$\mathrm{T}_n\mathrm{T}_m = \mathrm{T}_{nm}$, если $\gcd(n,m) = 1$
$\mathrm{T}_{p^n}\mathrm{T}_p = \mathrm{T}_{p^{n+1}} + p^{1-k}\mathrm{T}_{p^{n-1}}$ при $n \geq 1$.

Используя установленное выше взаимнооднозначное соответствие, перенесем наши операторы (они называются операторами Гекке) на автоморфные функции веса k, для удобного вида формул умножив каждый на подходящую константу. А именно, положим $\mathrm{T}_n f(z) = n^{k-1}\sum\limits_{a,b,d} d^{-k} f(\frac{az+b}{d}) = n^{k/2-1}\sum\limits_{a,b,d} f[\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)]_k$, где $ad = n$, $d > 0$, и $0 \leq b \leq d-1$, так что $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ - представители орбит левого действия Γ на множестве $\mathrm{Mat}_2(n)$ целочисленных 2×2-матриц с детерминантом n. Стоит также отметить, что $\mathrm{Mat}_2(n)$ может быть представлено и как объединение двойных смежных классов $\Gamma\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)\Gamma$ по наборам $(a, d > 0; ad = n; a|d)$.

Полезно выписать формулу для операторов T_p с простыми номерами отдельно:
$\mathrm{T}_p f(z) = p^{k/2-1}(f[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k + \sum\limits_{0 \leq b \leq p-1} f[\left(\begin{smallmatrix} 1 & b \\ 0 & p \end{smallmatrix}\right)]_k)$

Отличие в формуле коммутации вызвано другой нормировкой:
$\mathrm{T}_n\mathrm{T}_m = \mathrm{T}_{nm}$, если $\gcd(n,m) = 1$
$\mathrm{T}_{p^n}\mathrm{T}_p f = \mathrm{T}_{p^{n+1}} f + p^{k-1}\mathrm{T}_{p^{n-1}} f$ при $n \geq 1$.

Поскольку операторы T_{p^n} являются полиномами от T_p (по индукции из второй формулы), алгебра операторов Гекке коммутативна и порождается операторами T_p.

Все предыдущее касалось произвольных автоморфных функций веса k, а теперь мы перейдем к модулярным формам.
Чуть ниже мы увидим, что операторы Гекке переводят параболические формы в

параболические. Убедимся, что они эрмитовы по отношению к введенному выше скалярному произведению. Достаточно проверить это для операторов T_p. Множество $\mathrm{Mat}_2(p)$ состоит, как видно из приведенного выше описания, из единственного двойного смежного класса. По свойству 3) скалярного произведения $\langle f[\alpha]_k, g\rangle$ не зависит от того, какую матрицу $\alpha \in \mathrm{Mat}_2(p)$ мы выберем, а отображение $\alpha \to \alpha' = p\alpha^{-1}$ просто переставляет эти матрицы.

Модулярные формы обладают q-разложениями, действие операторов Гекке на которых легко описать, поскольку в определении T_n мы использовали только матрицы α с нулем в левом нижнем углу, а для таких матриц действие $[\alpha]_k$ на $q = e^{2\pi i z}$ выражается просто. Ответ выглядит так:
Пусть $f(z) = \sum\limits_{m\geq 0} c_m q^m$. Тогда $\mathrm{T}_n f = \sum\limits_{m\geq 0} \gamma_{m,n} q^m$, где $\gamma_{m,n} = \sum\limits_{d|\,\gcd(m.n)} d^{k-1} c_{mn/d^2}$.
В частности, $\gamma_{0,n} = \sigma_{k-1}(n)c_0$ (поэтому T_n переводит параболические формы в параболические), а $\gamma_{1,n} = c_n$.

Ограничимся проверкой формулы для простых p (общий случай мало отличается).

По определению $\mathrm{T}_p f(z) = p^{k/2-1}(f[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k + \sum\limits_{0\leq b\leq p-1} f[\left(\begin{smallmatrix} 1 & b \\ 0 & p \end{smallmatrix}\right)]_k) = p^{k-1} f(pz) + \frac{1}{p} \sum\limits_{0\leq b\leq p-1} f(\frac{z+b}{p}) =$ $p^{k-1} f(pz) + \frac{1}{p} \sum\limits_{0\leq b\leq p-1} \sum\limits_{m\geq 0} c_m e^{2\pi i m \frac{z+b}{p}} = p^{k-1} f(pz) + \frac{1}{p} \sum\limits_{m\geq 0} c_m e^{2\pi i m\frac{z}{p}} \sum\limits_{0\leq b\leq p-1} e^{2\pi i m \frac{b}{p}} =$ (поскольку последняя сумма равна 0 при $p \nmid m$ и p при $p|m$) $= p^{k-1} f(pz) + \sum\limits_{m\geq 0,\, p|m} c_m e^{2\pi i m \frac{z}{p}}$.

Итак, мы видим, что в этом случае $\gamma_{m,p} = c_{pm}$, если $p \nmid m$, и $\gamma_{m,p} = c_{pm} + p^{k-1} c_{m/p}$, если $p|m$.

Поскольку все операторы T_n эрмитовы и коммутируют друг с другом, в пространстве $S_k(\Gamma)$, согласно спектральной теореме, существует ортогональный базис, состоящий из форм, являющихся собственными для всех T_n. Такова, например, форма $\Delta(z)$ в силу одномерности $S_{12}(\Gamma)$.

Пусть f - не обязательно параболическая модулярная форма (кроме случая $k = 0$ и $f = c$) и $\forall n \;\; \mathrm{T}_n f = \lambda_n f$. Тогда $c_1 \neq 0$, так как, с одной стороны, $\gamma_{1,n} = c_n$, а с другой стороны, $\gamma_{1.n} = \lambda_n c_1$, так что $\forall n \;\; c_n = \lambda_n c_1$. Поделив f на c_1 (получившаяся форма называется нормализованной), получим, что $\forall n \geq 1 \;\; c_n = \lambda_n$. Коэффициент c_0 при этом может быть как нулевым, так и ненулевым (то, что форма не параболическая, не исключает того, что она может оказаться собственной для всех операторов Гекке). Наконец, для коэффициентов нормализованной собственной формы выполнены соотношения:

$c_m c_n = c_{mn}$ при $\gcd(m,n) = 1$
$c_p c_{p^n} = c_{p^{n+1}} + p^{k-1} c_{p^{n-1}}$ при $n \geq 1$.

По формуле из п.4 при любом четном $k \geq 4$ $\dim M_k(\Gamma) = \dim S_k(\Gamma) + 1$ (при $k = 2$ оба пространства нулевые). Тем самым $M_k(\Gamma)$ является суммой подпространства параболических форм и одномерного пространства, натянутого на ряд Эйзенштейна $G_k(z) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!} \sum\limits_{m=1}^{\infty} \sigma_{k-1}(m) q^m$, где $q = e^{2\pi i z}$. Поделив на коэффициент при q (а не на свободный член, как мы это делали, когда определяли E_k), получим нормализованный ряд Эйзенштейна $G_k^*(z) = -\frac{B_k}{2k} + \sum\limits_{m=1}^{\infty} \sigma_{k-1}(m) q^m$. При $m \geq 1$ его коэффициенты Фурье $\sigma_{k-1}(m)$, как легко непосредственно проверить, удовлетворяют только что выписанным соотношениям, а следовательно, действие оператора T_n умножает их на $\sigma_{k-1}(n)$. Что касаеся свободного члена, то T_n, как мы видели выше, умножает его на $\sigma_{k-1}(n)$ вообще для любой модулярной формы. Следовательно, $G_k^*(z)$ - нормализованная собственная форма.

Отметим, что разложение $M_k(\Gamma)$ в сумму $S_k(\Gamma)$ и ряда Эйзенштейна в некотором смысле является ортогональным. А именно, если $f(z) = \sum\limits_{m \geq 1} c_m q^m$ - параболическая собственная форма, то $\langle f, G_k^* \rangle = 0$. "В некотором смысле", потому что на всем пространстве $M_k(\Gamma)$ скалярное произведение не определено - хотя бы одна из форм должна быть параболической. Действительно, пусть f нормализована. Тогда $\forall n \geq 1$ $T_n f = c_n F$ и $T_n G_k^* = \sigma_{k-1}(n) G_k^*$. Поскольку T_n эрмитовы, $\langle T_n f, G_k^* \rangle = \langle f, T_n G_k^* \rangle$, а значит, $\forall n \geq 1 c_n \langle f, G_k^* \rangle = \sigma_{k-1}(n) \langle f, G_k^* \rangle$. Если бы скалярное произведение не равнялось нулю, то $f(z) - G_k^*(z)$ было бы константой, чего не может быть, поскольку константа не является модулярной функцией веса k.

Отступление: формула следа Эйхлера - Сельберга.

Коэффициенты Фурье ряда Эйзенштейна задаются простой формулой, в то время как природа коэффициентов Фурье параболических форм (которые для нормализованных собственных форм совпадают с собственными значениями T_n) на этом этапе изучения остается не слишком понятной. Можно дать суммарную оценку: сравнительно несложное (но длинное и техническое) вычисление приводит к следующей формуле для следа

оператора T_n на пространстве $S_k(\Gamma)$:
$\mathrm{Tr} T_n = -\frac{1}{2}(\sum\limits_{t=-\infty}^{\infty} P_k(t,n)H(4n-t^2) + \sum\limits_{d|n} \min(d, \frac{n}{d})^{k-1})$.
Здесь $P_k(t,n)$ - коэффициент при x^{k-2} в степенном ряду $(1-tx+nx^2)^{-1}$, а $H(\Delta)$- число классов Γ-эквивалентности положительно определенных квадратичных форм $ax^2+bxy+cy^2$ дискриминанта $b^2-4ac=\Delta$. При этом классы эквивалентности форм x^2+y^2 и x^2+xy+y^2 считаются с кратностями, соответственно, $1/2$ и $1/3$, а кроме того, мы полагаем $H(\Delta)=0$ при $\Delta<0$ и $H(0)=-1/12$ (конец отступления).

Для любой формы $f(z) = \sum\limits_{m\geq 1} c_m q^m \in M_k(\Gamma)$ положим формально $L(f,s) \stackrel{\mathrm{def}}{=} \sum\limits_{n=1}^{\infty} \frac{c_n}{n^s}$ (смысл этого определения будет пояснен в следующем разделе; от c_0 $L(f,s)$ не зависит). Тогда, если $f(z)$ - нормализованная собственная форма, то из приведенных соотношений между коэффициентами c_m сразу следует, что
$L(f,s) = \prod\limits_p (1-c_p p^{-s} + p^{k-1-2s})^{-1}$.

8. Снова преобразование Меллина.

Повернем полуось $\mathbf{R}_{>0}$ на 90 градусов и определим преобразование Меллина модулярной формы веса k (не обязательно параболической или собственной для операторов Гекке) формулой $\mathcal{M}f(s) \stackrel{\mathrm{def}}{=} \int\limits_0^{\infty} f(iy)y^s \frac{dy}{y}$. Если положить $f(iy)=\phi(y)$, то инвариатность веса k относительно действия матрицы $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ приводит к соотношению
$f(\frac{-1}{z}) = z^k f(z)$, откуда $\phi(\frac{-1}{y}) = i^k y^k \phi(y)$.

Вместе с очевидной оценкой $f(z)-f(i\infty) << e^{-2\pi y}$ это позволяет применить основную теорему о преобразовании Меллина. Следовательно, функция $\mathcal{M}(f-f(i\infty),s)$ аналитически продолжается на всю комплексную плоскость с полюсами в $s=0$ и $s=k$ и удовлетворяет функциональному уравнению $\mathcal{M}(f-f(i\infty),s) = i^k\mathcal{M}(f-f(i\infty),k-s)$ При этом на самом деле если $f(i\infty)=0$ (т.е. форма f параболическая), то полюсов нет.
С другой стороны, элементарный подсчет показывает, что $\mathcal{M}(f-f(i\infty),s) = (2\pi)^{-s}\Gamma(s)L(f,s)$, где $L(f,s)$ - ряд, рассмотренный в конце предыдущего раздела.

Эта конструкция не обобщается впрямую на подгруппы, поскольку матрица S, меняющая местами z и $\frac{-1}{z}$ в большинстве из них не лежит. На самом деле для получения функционального уравнения достаточно более слабых свойств инвариантности. Например, если $G=\Gamma_0(N)$, то можно рассмотреть матрицу $w_N = \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. Эта матрица нормализует подгруппу $\Gamma_0(N)$ и значит, инволюция $[w_N]_k$ действует на модулярных формах веса k

для $\Gamma_0(N)$. L - ряд любой формы, инвариантной (или антиинвариантной) относительно этого действия, будет удовлетворять функциональному уравнению такого же типа.

Именно таким образом из параболических форм веса 2 для $\Gamma_0(N)$ получаются L - ряды, представляющие собой основную часть дзета-функций Хассе-Вейля всех (как мы теперь знаем) эллиптических кривых над $\mathbf{Q}$. Эти ряды, по определению, разлагаются также в эйлерово произведение, что, в согласии с предыдущим разделом, указывает на то, что они соответствуют формам, собственным относительно алгебры Гекке. Определение последней для конгруэнцподгрупп $G \in \Gamma$ будет дано в следующем разделе.

9. Теория Гекке для конгруэнцподгрупп.

Основным объектом рассмотрения будет группа $G = \Gamma_1(N)$. Напомним, что имеют место вложения $\Gamma_1(N) \triangleleft \Gamma_0(N) \subset \Gamma$, при этом $\Gamma_0(N)/\Gamma_1(N) \simeq (\mathbf{Z}/(N))^*$, а однородное Γ-пространство смежных классов $\Gamma_0(N)\backslash\Gamma$ отождествляется с $\mathbf{P}^1(\mathbf{Z}/(N))$.

Если $a \in (\mathbf{Z}/(n))^*$, то на пространстве автоморфных функций веса k для G определен оператор $\langle a\rangle_k$. Пусть $\sigma_a \in \Gamma_0(N)$ - матрица, такая, что $\sigma_a \equiv \left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right) \mod N$. Положим $\langle a\rangle_k f \stackrel{\text{def}}{=} f[\sigma_a]_k$. Поскольку G нормальна в $\Gamma_0(N)$, оператор $\langle a\rangle_k$ переводит автоморфные функции в автоморфные, а действие матрицы σ_a переставляет параболические точки группы G, лежащие над данной параболической точкой $\Gamma_0(N)$, и все индексы ветвления над ней (а значит, и над параболической точкой $i\infty$ группы Γ) одинаковы. Поэтому набор условий, обеспечивающих мероморфность (соответственно, голоморфность либо обращение в нуль) одновременно во всех параболических точках не меняется при переходе от f к $\langle a\rangle_k f$, так что оператор $\langle a\rangle_k$ переводит каждое из пространств $M_k(G)$ и $S_k(G)$ в себя.

Ясно, что $\langle a\rangle_k\langle b\rangle_k = \langle b\rangle_k\langle a\rangle_k = \langle ab\rangle_k$. Если $\epsilon : (\mathbf{Z}/(N))^* \to \mathbf{C}$ - характер Дирихле, то можно определить подпространство $M_k(G,\epsilon) \subset M_k(G)$, которое задается условием $f[\gamma]_k = \epsilon(d)$ для любой матрицы $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$, и аналогично для $S_k(G)$. Ясно, что $M_k(G) = \oplus_\epsilon M_k(G,\epsilon)$ и $S_k(G) = \oplus_\epsilon S_k(G,\epsilon)$,где ϵ пробегает все характеры Дирихле по модулю N. Тривиальному характеру соответствуют подпространства модулярных форм для $\Gamma_0(N)$. Как было отмечено выше, на $S_k(G)$ сопряженным оператором к оператору $\langle a\rangle_k$ является оператор $\langle a^{-1}\rangle_k$. Если $f \in S_k(G,\epsilon_1)$, а $g \in S_k(G,\epsilon_2)$, то $\epsilon_1(a)\langle f,g\rangle = \langle\langle a\rangle_k f, g\rangle = \langle f, \langle a^{-1}\rangle_k g\rangle = \overline{\epsilon_2(a^{-1})}\langle f,g\rangle = \epsilon_2(a)\langle f,g\rangle$, поэтому разложение пространства параболических форм по характерам является ортогональным.

Определение операторов Гекке, которое было дано в п.7 для группы Γ, необходимо будет модифицировать. А именно, вместо решеток L будем рассматривать пары $(L, t \in \mathbf{C}/L)$, где t имеет порядок в точности N. Функция F на таких парах имеет вес k, если $F(\lambda L, \lambda t) = \lambda^{-k} F(L,t)$. Такие функции взаимнооднозначно соответствуют функкциям $\tilde{F}$ веса k на векторах столбцах $\binom{w_1}{w_2}$, где w_1 и w_2 - комплексные числа, такие, что $w_1/w_2 \in$ H, инвариантным относительно правого действия группы $G = \Gamma_1(N)$, отвечающего естественному левому действию на векторах-столбцах (соответствие задается формулами $\tilde{F}\binom{w_1}{w_2} = F(L = [w_1, w_2], t \equiv \frac{w_2}{N} \mod L)$; $F(L,t) = \tilde{F}\binom{w_1}{w_2}$, где (w_1, w_2) - какой-нибудь правильно ориентированный базис L такой, что $\frac{w_2}{N} \equiv t \mod L)$. Как и раньше, функции $\tilde{F}$ взаимнооднозначно соответствуют автоморфным функциям веса k, на этот раз для группы G, по формуле $w_2^k \tilde{F}\binom{w_1}{w_2} = f(\frac{w_1}{w_2})$.

Положим $T_n(L,t) \overset{\text{def}}{=} \sum (L', nt)$, где $L' \subset L$ пробегает не все подрешётки индекса n, а только такие, по модулю которых nt имеет порядок ровно N (если $\gcd(n,N) = 1$, то это все подрешетки индекса n). Базис $\binom{w'_1}{w'_2}$ решетки L' получается из базиса $\binom{w_1}{w_2}$ решетки L применением целочисленной матрицы с детерминантом n и нужно, чтобы по модулю Ł$'$ выполнялось условие $\frac{w'_2}{N} \equiv \frac{nw_2}{N} \mod$ Ł$'$. Тем самым, множество пар, по которым ведется суммирование, отождествляется со множеством орбит левого действия G на множестве $\mathrm{Mat}_2(n,N)$ целочисленных матриц с детерминантом n, сравнимых с $\binom{1\,*}{0\,n}$ по модулю N.

Пусть p - простое число. Нетрудно проверить, что при $p|N$ имеется p орбит, которые представлены матрицами $\binom{1\,b}{0\,p}, 0 \le b \le p-1$ (Γ-орбита матрицы $\binom{p\,0}{0\,1}$ с $\mathrm{Mat}_2(n,N)$ не пересекается). Если же $p \nmid N$, то орбит, как и в случае группы Γ, $p+1$. Первые p из них представлены теми же матрицами, что и выше, а оставшаяся - произведением матриц $\binom{a\,b}{N\,p}\binom{p\,0}{0\,1}$, где $aN - bp = 1$.

Перенесем операторы T_n на автоморфные функции по образцу п.7, как и прежде, умножив по дороге на n^{k-1}. Для простых p, согласно предыдущему абзацу,

$\mathrm{T}_p f(z) = p^{k/2-1}(\sum\limits_{0 \le b \le p-1} f[\binom{1\,b}{0\,p}]_k)$, если $p|N$, и
$\mathrm{T}_p f(z) = p^{k/2-1}(\sum\limits_{0 \le b \le p-1} f[\binom{1\,b}{0\,p}]_k + f[\binom{a\,b}{N\,p}]_k[\binom{p\,0}{0\,1}]_k)$, если $p \nmid N$.

Используя определение через решётки, легко убедиться, что операторы T_n коммутируют друг с другом и с операторами $\langle a \rangle_k$ (последний переводит пару (L,t) в пару (L, at)).

Кроме того, несложно проверить (примерно так же, как проверялось в случае группы Γ), что $T_m T_n = T_{mn}$ при $\gcd(n,m)=1$, что $T_{p^r} = T_p^r$ при $p|N$ и что
$T_{p^{r+1}} = T_p T_{p^r} - p^{k-1}\langle p\rangle_k T_{p^{r-1}}$ при $r \geq 1$ и при $p \nmid N$.
Последнюю формулу можно считать справедливой и при $p|N$, полагая (что мы в дальнейшем будем делать), что в этом случае $\langle p\rangle_k = 0$.

Действие операторов Гекке на коэффициентах Фурье модулярных форм вычисляется так же, как и для группы Γ (единственная матрица без нуля в левом нижнем углу, фигурирующая в формулах для T_p - это матрица $\left(\begin{smallmatrix} a & b \\ N & p \end{smallmatrix}\right)$, которая лежит в $\Gamma_0(N)$ и поэтому действует как $\langle p\rangle_k$). Параболическая точка $i\infty$ группы G неразветвлена над параболической точкой $i\infty$ группы Γ (матрица $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, переводящая $f(z)$ в $f(z+1)$, содержится в G), поэтому модулярная форма f для группы G обладает
q - разложением $f(z) = \sum\limits_0^\infty c_m(f) q^m$, где $q = e^{2\pi i z}$.
Для коэффициентов Фурье $T_p f$ получается формула: $c_m(T_p f) = c_{mp}(f) + p^{k-1} c_{m/p}(\langle p\rangle_k f)$.
Здесь мы считаем, что $c_{m/p} = 0$ при $p \nmid m$.
В частности, для формы $f \in M_k(G,\epsilon)$, собственной для характера Дирихле,
$c_m(T_p f) = c_{mp}(f) + \epsilon(p) p^{k-1} c_{m/p}(f)$.

Для полноты картины приведём формулы для произвольного n:
$c_m(T_n f) = \sum\limits_{a|\gcd(m,n)} a^{k-1} c_{mn/a^2}(\langle a\rangle_k f)$, в частности, при $f \in M_k(G,\epsilon)$
$c_m(T_n f) = \sum\limits_{a|\gcd(m,n)} \epsilon(a) a^{k-1} c_{mn/a^2}(f)$.

Изучим действие алгебры Гекке на эрмитовом пространстве параболических форм $S_k(G)$. Для начала предположим, что $p \nmid N$. Сопряженные операторы вычисляются так:
$\langle p\rangle_k^* = \langle p^{-1}\rangle_k \quad T_p^* = \langle p^{-1}\rangle_k T_p$.

Это сразу следует из свойства 2) скалярного произведения из п.7, поскольку умножение на матрицу $\sigma_p = \left(\begin{smallmatrix} a & b \\ N & p \end{smallmatrix}\right)$ превpщает слагаемые T_p^* в представителей тех же смежных классов, которые соответствуют слагаемым T_p, меняя только их порядок.

Алгебра, порожденная операторами $\langle a\rangle_k$, а также операторами T_p для $p \nmid N$, называется анемичной алгеброй Гекке. Формула для сопряженных операторов показывает, что все элементы анемичной алгебры Гекке суть нормальные операторы (то есть коммутируют со своими сопряженными), а значит к её действию на $S_k(G)$ применима спектральная теорена: $S_k(G) = \oplus_\chi S_k(G,\chi)$, где χ пробегает характеры анемичной алгебры Гекке, каждый из которых соответствует некоторому набору собственных значений λ_p операторов T_p, а на операторах $\langle a\rangle_k$ является характером Дирихле ϵ, причем это разложение ортогонально по отношению к скалярному произведению. Тем самым, $S_k(G)$ обладает ортогональным базисом, состоящим из параболических форм, собственных для анемичной алгебры Гекке. Одному и тому же характеру могут при этом, вообще говоря, отвечать несколько элементов базиса.

Следует отметить, что можно определить ряды Эйзенштейна для группы G в количестве, равном количеству параболических точек (кроме веса 2, когда на единицу меньше), задающиеся несложными формулами, которые мы здесь не станем выписывать, так что эти ряды также будут собственными для анемичной алгебры Гекке и ортогональными к пространству параболических форм и $M_k(G)$ превратится в сумму $S_k(G)$ и пространства, порожденного указанными рядами.

10. Полная алгебра Гекке.

Осталось изучить вопрос о действии на модулярных формах полной алгебры Гекке, то есть операторов T_n при $\gcd(n,N)\neq 1$. Эти операторы уже не обязаны быть нормальными. Сопряженные к ним могут быть вычислены по общей формуле $T^* = [w_N]_k T[w_N]_k$ (матрица $w_N = \left(\begin{smallmatrix}0 & -1\\ N & 0\end{smallmatrix}\right)$, рассматривавшаяся в п.8, группу G тоже нормализует и определяет инволюцию на $S_k(G)$), справедливой для всех операторов Гекке, но это мало что даёт. Таким образом, базиса, состоящего из форм, собственных для полной алгебры Гекке, в $S_k(G)$ может не быть.

Однако такие формы представляют немалый интерес, прежде всего потому, что если $f = \sum\limits_{m=0}^{\infty} c_m(f)q^m$ - нормализованная (т.е. $c_1(f)=1$) такая форма, то её коэффициенты Фурье удовлетворяют соотношениям $c_m c_n = c_{mn}$ при $\gcd(m,n)=1$, и $c_{p^{r+1}} = c_p c_{p^r} - \epsilon(p)p^{k-1}c_{p^{r-1}}$, автоматически следующему из вышеприведенных правил коммутации для операторов Гекке, а из этого следует, что L-ряд, соответствующий f, обладает эйлеровым произведением $L(f,s) = \prod\limits_p (1 - c_p(f)p^{-s} + \epsilon(p)p^{k-1-2s})^{-1}$, где p пробегает все простые числа, в том числе и делящие N (напомним, что для последних $\epsilon(p)=0$).

Оказывается, что в $S_k(G)$ можно естественным образом выделить подпространство, обладающее базисом из форм, собственных для полной алгебры Гекке, - это подпространство примитивных (или новых) форм.

Пусть $d|N$. Рассмотрим линейное вложение $i_d = i_1(d) + i_2(d) : \quad S_k(\Gamma_1(N/d)^2 \to S_k(\Gamma_1(N) = G)$, на первой координате задающееся формулой $i_1(f) = f$, а на второй - формулой $i_2(g) = g[\alpha_d]_k$, где $\alpha_d = \binom{d\,0}{0\,1}$ (i_2 переводит $g(z)$ в $d^{k/2}g(dz)$). Назовем сумму образов i_d для всех $d|N$ (можно ограничиться простыми делителями N) подпространством $S_k^{old}(G)$ старых форм, а ортогональное дополнение к нему относительно скалярного произведения - подпространством $S_k^{new}(G)$ примитивных или новых форм.

Почти очевидно, что каждая из компонент отображения i_d коммутирует с операторами из анемичной алгебры Гекке (мы отождествляем анемичный оператор для G и одноименный оператор (тоже анемичный) для $\Gamma_1(N/d)$), и легко проверить, что $i_1(d)[w_N]_k = [w_{N/d}]_k i_2(d)$. Следовательно, как анемичная алгебра Гекке, так и инволюция $[w_N]_k$ переводят $S_k^{old}(G) \subset S_k(G)$ в себя, а значит, и $S_k^{new}(G)$ тоже, так как анемичная алгебра Гекке замкнута относительно сопряжения, а инволюция самосопряжена. Поскольку $[w_N]_k\langle a\rangle_k = \langle a^{-1}\rangle_k[w_N]_k$, действие инволюции меняет местами подпространства, отвечающие характерам Дирихле ϵ и $\bar{\epsilon}$.

За неимением места мы опустим требующее довольно длинных вычислений доказательство следующей важной теоремы (Аткин и Ленер доказали её для $\Gamma_0(N)$, а Ли распространил доказательство на $\Gamma_1(N)$).

Теорема. 1.Пусть $f = \sum_{m=1}^{\infty} c_m q^m \in S_k(G)$ - параболическая форма, собственная для анемичной алгебры Гекке. Если $c_1 = 0$, то $f \in S_k^{old}(G)$.
2. Пусть f и g - параболические формы, собственные для анемичной алгебры Гекке с одним и тем же характером χ. Если $f \neq 0$ и $f \in S_k^{new}(G)$, то g пропорциональна f.
3. В подпространстве $f \in S_k^{old}(G)$ любое ненулевое собственное подпространство для какого-либо характера анемичной алгебры Гекке, имеет размерность больше 1.

Поскольку полная алгебра Гекке коммутативна, то из п.2. этой теоремы следует, что любая примитивная форма, собственная относительно анемичной алгебры Гекке, собственна и относительно полной алгебры Гекке.

11. Модулярный символ.

В предыдущем пункте было установлено, что пространство $S_k(G)$ имеет ортогональный базис из форм, собственных для анемичнной алгебры Гекке. Однако вычисление этого базиса представляет собой непростую задачу. Теория модулярного символа позволяет решить эту задачу для форм веса 2, которые находятся во взаимнооднозначном соответствии с регулярными дифференциалами на римановой поверхности $X_G = G \setminus \overline{\mathrm{H}}$ (порядок нуля модулярной формы веса 2 в представителе параболической точки превосходит на 1 порядок нуля в образе этой точки на X_G дифференциала, соответствующего ей по общей теории - см. п.4).

Комплексное сопряжение действует на верхней полуплоскости, переводя z в $-\overline{z}$. Оно нормализует группы $\Gamma_0(N)$, $G = \Gamma_1(N)$ и $\Gamma(N)$, поэтому римановы поверхности, отвечающие этим группам, как алгебраические кривые, определены над $\mathbf{R}$ (в случае G или $\Gamma_0(N)$ на самом деле даже над $\mathbf{Q}$).

Будем предполагать, что род $g(X_G) > 0$ (иначе $S_2(G) = 0$). Рассмотрим группу вещественных когомологий де Рама $H^1_{dR}(X_G, \mathbf{R})$. Интегрирование форм по замкнутым циклам по линейности продолжается до двойственности между $H^1_{dR}(X_G, \mathbf{R})$ и группой сингулярных гомологий $H_1(X_G, \mathbf{R})$, то есть $\mathrm{Hom}(H^1_{dR}(X_G, \mathbf{R}), \mathbf{R}) = H_1(X_G, \mathbf{R})$.
С другой стороны, имеет место разложение Ходжа $H^1_{dR}(X_G), \mathbf{R}) \otimes_{\mathbf{R}} \mathbf{C} = H^1(X_G, \mathbf{C}) = H^{1,0} \oplus H^{0,1}$, где $H^{1,0} = H^0(X_G, \Omega_1)$ есть в точности интересующее нас пространство регулярных дифференциалов, а $H^{0,1} = \overline{H^{1,0}}$. Несмотря на то, что $H^{1,0}$ над $\mathbf{R}$ не определено, его размерность над $\mathbf{R}$ равна размерности $H^1_{dR}(X_G), \mathbf{R})$, то есть $2g$, и интегрирование определяет изоморфизм векторных пространств над $\mathbf{R}$:
$\mathrm{Hom}_{\mathbf{C}}(H^0(X_G, \Omega_1), \mathbf{C}) = H_1(X_G, \mathbf{R})$.

Любая пара точек (z_1, z_2) на $\overline{\mathrm{H}}$ определяет комплекснозначный линейный функционал $\{z_1, z_2\}_G$ на $S_2(X_G)$ - интеграл формы по пути от z_1 до z_2 (если какая-то из точек параболическая, то надо считать, что путь удовлетворяет некоторым аналитическим условиям). Это справедливо для любой подгруппы конечного индекса в Г. Согласно предыдущему абзацу, $\{z_1, z_2\}_G \in H_1(X_G, \mathbf{R})$. Этот функционал и называется модулярным символом. Его простейшие свойства:
$\{z_1, z_2\}_G + \{z_2, z_3\}_G + \{z_3, z_1\}_G = 0$.
$\{gz_1, gz_2\}_G = \{z_1, z_2\}_G$ при $g \in G$ (это частный случай очевидной формулы $\int_{\alpha z_1}^{\alpha z_2} \omega = \int_{z_1}^{z_2} \omega[\alpha]_2$, которая справедлива для любого дифференциала на H и любой матрицы $\alpha \in \mathrm{GL}_2^+(\mathbf{Q})$).

Теорема Манина-Дринфельда. Пусть $G \in \Gamma$ - любая конгруэнцподгруппа, а x и y - представители параболических точек для G. Тогда $\{x, y\}_G \in H_1(X_G, \mathbf{Q})$.
Достаточно доказать для $G = \Gamma(N)$. Определение оператора Гекке T_p на модулярных формах аналогично таковому для $\Gamma_1(N)$. А именно, $T_p = \sum\limits_{i=1}^{s} [\alpha_i]_2$, где α_i - система

представителей орбит левого действаия G на множестве целочисленных матриц с определителем p, сравнимых с единичной матрицей по модулю N.
Лемма. Оператор $T_p - s$ обратим на $S_2(G)$.
Действительно, все операторы $[\alpha_i]_2$ переводят формы $f \in S_2(G)$ в параболические формы (для других групп), не меняя скалярный квадрат, а это означает, что если $T_p f = sf$, то все $f[\alpha_i]_2$ однонаправлены, то есть совпадают с f. Однако действие матрицы $\binom{p\,0}{0\,1}$ переводит форму $\sum c_m q^m$ в форму $\sum c_m q^{pm}$, поэтому номера всех ненулевых коэффициенто Фурье инвариантной относительно действия этой матрицы формы должны делиться на p. Продолжая применять ту же матрицу, получаем противоречие/
Перенесем определение T_p по двойственности с $S_2(G)$ на $H_1(X_G, \mathbf{R})$. Тогда оператор $T_p - s$ и там будет обратим. Кроме того, T_p отображает $H_1(X_G, \mathbf{Z})$ в себя (если $z_2 = gz_1$ при некотором $g \in G$, то $T_p\{z_1, z_2\}_G = \sum\{\alpha_i(z_1), \alpha_i g(z_1\}_G$, и класс гомологий остается целочисленным, поскольку отображение $G\alpha_i \mapsto Gg\alpha_i$ просто переставляет орбиты).
Теперь выберем $p \equiv 1 \mod N$. Тогда если x - представитель параболической точки, то его образ под действием любой матрицы α_i (напомним, что все они сравнимы с единичной по модулю N) переводит x в представителя той же самой параболической точки, а значит, если x и y - два таких представителя, то $(T_p - s)\{x, y\}_G = \sum_{i=1}^{s}(\{\alpha_i x, \alpha_i y\}_G - \sum_{i=1}^{s}(\{x, y\}_G = \sum_{i=1}^{s}(\{\alpha_i x, x\}_G - \{\alpha_i y, y\}_G)$ является целочисленным классом гомологий.
Применяя обратный опепратор, получаем, что $\{x, y\}$ - рациональный класс ■

Если выкинуть из X_G образы параболических и эллиптических точек, то над тем, что осталось (обозначим его X_G^0), остаток верхней полуплоскости H_G^0 будет неразветвленным накрытием с группой Галуа $\overline{G}$. Поэтому любая точка $\alpha \in X_G^0$ определяет гомоморфизм фундаментальной группы $\pi_1(X_G^0, \alpha) \to \overline{G}$. Отображение $\phi_\alpha : G \to H_1(X_G^0, \mathbf{Z}), \quad g \mapsto \{a, ga\}_G$, пропускается через $\overline{G}$ и, очевидно, не зависит от выбора представителя a точки α, а его композиция с указанным гомоморфизмом совпадает со стандартным гомоморфизмом $\pi_1(X_G^0, \alpha) \to H_1(X_G^0, \mathbf{Z})$. Ядро последнего есть коммутант фундаментальной группы, а если спроектировать $H_1(X_G^0, \mathbf{Z})$ на $H_1(X_G, \mathbf{Z})$, то к ядру добавятся обходы вокруг образов выкинутых точек. Из сказанного ясно, что отображене $G \to H_1(X_G, \mathbf{Z})$ сюръективно и не зависит от выбора самой точки α, и несложный анализ покзывает, что его ядро порождено коммутантом группы G и её эллиптическими и параболическими элементами.

Теперь мы построим конечный набор элементов пространства $H_1(X_G, \mathbf{Q})$, чья линейная оболочка порождает всё пространство, такой, что действие операторов Гекке на этих образующих будет чрезвычайно просто описываться. Установив, какие между этими элементами имеются линейные соотношения, мы сможем выразить через них базис пространства и затем вычислить спектр алгебры Гекке на $H_1(X_G, \mathbf{R})$, а следовательно и на пространстве $S_2(G)$, что и было основной задачей, поставленной в начале этого пункта (наборы собственных значений, отвечающих примитивным формам, выделятся на фоне остальных по признаку однократности).

Пусть $J = G\backslash\Gamma$. Если $j \in J$, то назовем отмеченным классом $\xi(j) = \{\gamma(0), \gamma(i\infty)\}_G \in H_1(X_g, \mathbf{Q})$, где $\gamma \in \Gamma$ - какая-нибудь матрица, представляющая j (от её выбора $\xi(j)$ не зависит, поскольку модулярный символ не меняется при действии на оба конца одного и того же элемента $g \in G$). Отмеченный класс не обязан лежать в $H_1(X_G, \mathbf{Z})$, даже если g - единичная матрица.

Любой элемент $H_1(X_G, \mathbf{Z})$ представляется в виде суммы отмеченных классов, причем можно выбрать эти классы так, что алгебраическая сумма проекций их концов (левые со знаком $+$, а правые со знаком -) равнялась нулю как цикл на X_G. Из описания $H_1(X_G, \mathbf{Z})$, данного двумя абзацами выше, видно, что любой элемент этой группы имеет вид $\{0, g(0)\}$, где $g \in G$. Если $g(0) = i\infty$ (при $G \subset \Gamma_0(N)$ такого не бывает), то искомое представление построено. Если $g(0) = \frac{b}{a}$ (можно считать, что $a > 0$), то разложим $\frac{b}{a}$ в непрерывную дробь с подходящими дробями $\frac{b}{a} = \frac{b_n}{a_n}, \frac{b_{n-1}}{a_{n-1}}, \dots, \frac{b_0}{a_0} = \frac{b_0}{1}, \frac{1}{0}, \frac{0}{1}$. Тогда матрицы $\gamma_i = \binom{b_i \ (-1)^{i-1}b_{i-1}}{b_i \ (-1)^{i-1}b_{i-1}}$ по свойству подходящих дробей лежат в Γ, так что классы $\{\frac{b_{i-1}}{a_{i-1}}, \frac{b_i}{a_i}\} = \{\gamma_i(0), \gamma_i(i\infty)\}_G$ являются отмеченными.

Осталось указать соотношения между отмеченными классами. Поскольку матрицы $S = \binom{0 -1}{1 \ 0}$ и $ST = \binom{1 -1}{1 \ 0}$ имеют порядки 2 и 3 соответственно, то отмеченные классы удовлетворяют соотношениям:
$\xi(j) + \xi(jS) = 0$; $\xi(j) = 0$, если $j = jS$
$\xi(j) + \xi(jTS) + \xi(j(TS)^2) = 0$; $\xi(j) = 0$, если $j = jTS$.
Вычисление, основанное на построении специального клеточного разбиения X_G, показывает, что это полный набор соотношений. Иначе говоря, если рассмотреть подгруппу в свободной группе, натянутой на элементы $j \in J$, состоящую из таких линейных комбинаций образующих, что применение к ним отображения ξ дает целочисленные классы гомологий (критерий указан в предыдущем абзаце) и профакторизовать эту подгруппу по всем перечисленным соотношениям, то получится $H_1(X_G, \mathbf{Z})$.