MONOGRAFIAS DE MATEMÁTICA Nº 47

FORMAS MODULARES
Uma Introdução

Fernando Quadros Gouvêa

ISBN

85-244-0050-1

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO
INSTITUTO DE MATEMÁTICA PURA E APLICADA
Estrada Dona Castorina, 110
22.460 – Rio de Janeiro – RJ

MONOGRAFIAS DE MATEMÁTICA

1) Alberto Azevedo, Renzo Piccinini - INTRODUÇÃO À TEORIA DOS GRUPOS (reprod)
2) Nathan M. Santos - VETORES E MATRIZES (esgotada)
3) Manfredo P. do Carmo - INTRODUÇÃO À GEOMETRIA DIFERENCIAL GLOBAL (esgot.)
4) Jacob Palis Jr. - SISTEMAS DINÂMICOS (esgotada)
5) João Pitombeira de Carvalho - INTRODUÇÃO À ÁLGEBRA LINEAR (esgotada)
6) Pedro J. Fernandez - INTRODUÇÃO À TEORIA DAS PROBABILIDADES (esgotada)
7) R. C. Robinson - LECTURES ON HAMILTONIAN SYSTEMS (esgotada)
8) Manfredo P. do Carmo - NOTAS DE GEOMETRIA RIEMANNIANA (esgotada)
9) Chaim S. Hönig - ANÁLISE FUNCIONAL E O PROBLEMA DE STURM-LIOUVILLE (esgot)
10) Welington de Melo - ESTABILIDADE ESTRUTURAL EM VARIEDADES DE DIMENSÃO 2 ""
11) Jaime Lesmes - TEORIA DAS DISTRIBUIÇÕES E EQUAÇÕES DIFERENCIAIS (esgotada)
12) Clóvis Vilanova - ELEMENTOS DA TEORIA DOS GRUPOS E DA TEORIA DOS ANÉIS ""
13) Jean Claude Douai - COHOMOLOGIE DES GROUPES (esgotada)
14) H. Blaine Lawson Jr. - LECTURES ON MINIMAL SUBMANIFOLDS, Vol. I (esgotada)
15) Elon L. Lima - VARIEDADES DIFERENCIÁVEIS (esgotada)
16) Pedro Mendes - TEOREMAS DE Ω-ESTABILIDADE E ESTABILIDADE ESTRUTURAL EM VARIEDADES ABERTAS (esgotada)
17) Herbert Amann - LECTURES ON SOME FIXED POINT THEOREMS (esgotada)
18) - EXERCÍCIOS DE MATEMÁTICA / IMPA (esgotada)
19) Djairo G. de Figueiredo - NÚMEROS IRRACIONAIS E TRANSCENDENTES (esgotada)
20) C. [illegible] Zeeman - UMA INTRODUÇÃO INFORMAL À TOPOLOGIA DAS SUPERFÍCIES (esgot)
21) Manfredo P. do Carmo - NOTAS DE UM CURSO DE GRUPOS DE LIE (esgotada)
22) Alexander Prestel - LECTURES ON FORMALLY REAL FIELDS (esgotada)
23) Aron Simis - INTRODUÇÃO À ÁLGEBRA (esgotada)
24) Jaime Lesmes - SEMINÁRIO DE ANÁLISE FUNCIONAL (esgotada)
25) Fred Brauer - SOME STABILITY AND PERTURBATION PROBLEM FOR DIFFERENTIAL AND INTEGRAL EQUATIONS (esgotada)
26) Lúcio Rodriguez - GEOMETRIA DAS SUBVARIEDADES (esgotada)
27) Mário Miranda - FRONTIÈRE MINIME
28) Fernando Cardoso - RESOLUBILIDADE LOCAL DE EQUAÇÕES DIFERENCIAIS PARCIAIS (esgotada)
29) Eberhard Becker - HEREDITARILY-PYTHAGOREAN FIELDS AND ORDERINGS OF HIGHER LEVEL
30) Hyman Bass - PROJECTIVE MODULES AND SYMMETRIC ALGEBRAS
31) J. Neyman - PROBABILIDADE FREQÜENTISTA E ESTATÍSTICA FREQÜENTISTA
32) Freddy Dumortier - SINGULARITIES OF VECTOR FIELDS (esgotada)
33) T. M. Viswanathan - INTRODUÇÃO À ÁLGEBRA E ARITMÉTICA (esgotada)
34) F. Javier Thayer - NOTES ON PARTIAL DIFFERENTIAL EQUATIONS
35) Edward Bierstone - THE STRUCTURE OF ORBIT SPACES AND THE SINGULARITIES OF EQUIVARIANT MAPPINGS
36) F. Javier Thayer - THÉORIE SPECTRALE
37) Manfredo P. do Carmo - FORMAS DIFERENCIAIS E APLICAÇÕES
38) Alexander Prestel, Peter Roquette - LECTURES ON FORMALLY p-ADIC FIELDS
39) Yves Lequain, Arnaldo Garcia - ÁLGEBRA: UMA INTRODUÇÃO (esgotada)
40) J. Lucas Barbosa, A. Gervásio Colares - MINIMAL SURFACES IN R^3
41) Pierre H. Bérard - SPECTRAL GEOMETRY: DIRECT AND INVERSE PROBLEMS
42) Pierre H. Bérard - ANALYSIS ON RIEMANNIAN MANIFOLDS AND GEOMETRIC APPLICATIONS: AN INTRODUCTION
43) Felipe Cano Torres - DESINGULARIZATION STRATEGIES FOR THREE-DIMENSIONAL VECTOR FIELDS
44) Otto Endler - TEORIA DOS CORPOS
45) Winfried Bruns, Udo Vetter - DETERMINANTAL RINGS
46) Abramo Hefez - INTRODUÇÃO À GEOMETRIA PROJETIVA
47) Fernando Quadros Gouvêa - FORMAS MODULARES: UMA INTRODUÇÃO

APRESENTAÇÃO

Este é o têxto de um minicurso ministrado na X ESCOLA DE ÁLGEBRA realizada em Vitória, ES, de 27 de fevereiro a 3 de março de 1989.

Por falta absoluta de recursos não foi possível, como é de praxe, publicar um volume de Atas contendo os minicursos.

Queremos externar os nossos caloroso agradecimentos ao Professor Elon Lages Lima por nos abrir a possibilidade de usar a coleção Monografias de Matemática do IMPA para publicar estes têxtos.

Esperamos que estes minicursos sejam úteis à comunidade matemática pois a sua utilização é a nossa maior recompensa.

Reiteramos aqui os nossos agradecimentos a todos que contribuiram para o sucesso desta Escola.

Rio de Janeiro, abril de 1990
Abramo Hefez
Coordenador da Escola

Introdução

Embora sejam, à primeira vista, objetos mais afins à análise complexa do que à aritmética, as formas modulares têm ocupado, nas últimas décadas, uma posição bastante importante na Teoria de Números. Suas ligações com a teoria de curvas elípticas, com as representações de grupos de Galois, e mais geralmente com a teoria geral de L-funções têm sido alvo de estudo intenso, com resultados profundos e importantes.

O objetivo destas notas é conduzir o leitor nos primeiros passos da teoria, de modo a facilitar seu aprofundamento mais tarde. Para isso, nos concentramos no caso mais simples das formas de nível um, mas nos mantivemos atentos para o caso de nível superior, discutindo-o quando isso complicava pouco e omitindo-o quando as dificuldades técnicas eram maiores. Além disso, mantivemos desde o princípio a posição de que a "interpretação modular" da definição, em termos de classes de isomorfismo de curvas elípticas, é especialmente interessante do ponto de vista aritmético; procuramos, então, tornar transparente a ligação entre a definição clássica e esta interpretação.

Procuramos, como é natural em um texto introdutório, manter ao mínimo os pré-requisitos para a leitura. Todos os fatos sobre curvas elípticas que são usados, por exemplo, são enunciados com cuidado, e poderão ser simplesmente assumidos pelo leitor que não for familiar com essa teoria. (A exceção são alguns exercícios, que visam o leitor que já fez um curso introdutório nessa área.) Algum conhecimento de funções de uma variável complexa e de superfícies de Riemann é pressuposto, mas o uso destas teorias é localizado; o leitor menos familiar, por exemplo, com superfícies de Riemann poderá "ler em torno" delas no Capítulo 3 (que é o momento em que elas aparecem mais), e depois fazer a seqüência de exercícios propostos ao final desse capítulo que permitem evitá-las completamente (ao custo de introduzir um pouco de integração no plano complexo).

As formas modulares são um tema clássico; há portanto, um grande número de livros que tratam do assunto. Mais próximos do espírito destas notas são o tratamento do caso de nível um no *A Course in Arithmetic* de J.-P. Serre e a exposição quase completa mas extremamente

compacta no terceiro capítulo do *Introduction to Elliptic Curves and Modular Forms*, de N. Koblitz, cuja leitura recomendamos para o leitor interessado em saber mais. Outros livros citados com frequência são [Lan76] e [Sil71] (as abreviações se referem à Bibliografia), que são, cada um à sua maneira, abrangentes e insatisfatórios ao mesmo tempo. Devem-se mencionar também as notas [Ogg69] de A. Ogg e os vários artigos nos proceedings da conferência de Antuérpia ([Kui73], [DK73] e [KS73]).

Estas notas representam uma redação preliminar do conteúdo de um curso ministrado na X Escola de Álgebra, realizada de 27 de fevereiro a 3 de março de 1989 em Vitória, ES.

Quaisquer comentários ou correções serão bem-vindos.

São Paulo, 6 de março de 1989

Fernando Q. Gouvêa

Sumário

Capítulo 1

O Semi-plano Superior e o Grupo Modular

O objetivo deste capítulo é estudar a ação de $\mathrm{SL}_2(\mathbf{Z})$ no semi-plano superior, e depois interpretar a teoria em termos do problema da classificação de curvas elípticas (sobre C). No próximo capítulo, vamos definir formas modulares como funções definidas no semi-plano superior que se transformam "bem" sob a ação do grupo modular.

1.1 O grupo modular e sua ação

Queremos considerar o semi-plano complexo superior, isto é,

$$\mathcal{H} = \{\tau = x + iy : y > 0\} \subset \mathbf{C},$$

sob a ação das transformações de Moebius, do tipo

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}.$$

Para começar, notamos que se a, b, c e d são números *reais*, então

$$\Im\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{(ad - bc)\Im(\tau)}{|c\tau + d|^2}. \tag{1.1}$$

Dessa forma, se $\tau \in \mathcal{H}$, teremos que

$$\frac{a\tau + b}{c\tau + d} \in \mathcal{H}$$

se e só se $ad - bc > 0$. A melhor maneira de expressar isto, então, é dizer:

Proposição 1.1.1 *Seja*

$$\mathrm{GL}_2^+(\mathbf{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbf{R}) : ad - bc > 0 \right\},$$

e seja $\mathcal{H}$ o semi-plano superior. Seja

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbf{R});$$

a fórmula

$$\gamma \cdot \tau = \frac{a\tau + b}{c\tau + d}$$

define uma ação de $\mathrm{GL}_2^+(\mathbf{R})$ em $\mathcal{H}$ por transformações holomorfas. As matrizes escalares, e só elas, agem trivialmente.

DEMONSTRAÇÃO: Fácil; fica como exercício para o leitor.□

Como as matrizes escalares agem trivialmente, a ação acima define também uma ação de

$$\mathrm{PSL}_2(\mathbf{R}) = \mathrm{GL}_2^+(\mathbf{R})/\mathbf{R}^\times$$

em $\mathcal{H}$; pode-se provar que esta ação fornece um *isomorfismo* entre $\mathrm{PSL}_2(\mathbf{R})$ e o grupo dos automorfismos holomorfos de $\mathcal{H}$.

Chamaremos de *bordo de $\mathcal{H}$* o conjunto

$$\mathbf{P}^1(\mathbf{R}) = \mathbf{R} \cup \{\infty\} \subset \mathbf{P}^1(\mathbf{C}),$$

com a interpretação usual do ponto no infinito; note que a ação de $\mathrm{GL}_2^+(\mathbf{R})$ preserva este conjunto.

No estudo de formas modulares, estaremos interessados não em todo o grupo $\mathrm{GL}_2^+(\mathbf{R})$, mas sim em seus subgrupos, especialmente os seus subgrupos discretos. Em particular, fazemos as seguintes definições:

Definição 1 *Chamamos de* grupo modular *o grupo $\Gamma = \mathrm{SL}_2(\mathbf{Z})$. Para cada número inteiro positivo* N, *definimos os seguintes subgrupos de Γ, que chamaremos de* subgrupos de congruência de nível N*:*

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma : c \equiv 0 \pmod{N} \right\}$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma : a \equiv b \equiv 1 \pmod{N},\ c \equiv 0 \pmod{N} \right\}$$

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma : a \equiv b \equiv 1 \pmod{N},\ c \equiv d \equiv 0 \pmod{N} \right\}.$$

Uma maneira de entender estes subgrupos é considerar o homomorfismo (sobrejetor— verifique!)

$$r_N : \Gamma = \mathrm{SL}_2(\mathbf{Z}) \longrightarrow \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}),$$

e notar que $\Gamma(N) = \ker(r_N)$, e que $\Gamma_1(N)$ e $\Gamma_0(N)$ são, respectivamente, imagem inversa de um subgrupo unipotente e de um subgrupo de Borel de $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$. Um *subgrupo de congruência de nível* N, em geral, é qualquer subgrupo de Γ que é imagem inversa por r_N de um subgrupo de $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$; os que destacamos são os casos particulares mais importantes. Vai ser relevante notar que $\Gamma(N)$ é normal em Γ (porque é um kernel), e que $\Gamma_1(N)$ é normal em $\Gamma_0(N)$ (verifique, como exercício).

Como os subgrupos $\Gamma_0(N)$, $\Gamma_1(N)$ e $\Gamma(N)$ estão todos contidos em Γ, e portanto em $\mathrm{GL}_2^+(\mathbf{R})$, eles agem no semi-plano superior. Novamente, as matrizes escalares agem trivialmente, e é útil diferenciar entre o grupo $G \subset \mathrm{GL}_2^+(\mathbf{R})$ e sua imagem $\overline{G} \subset \mathrm{PSL}_2(\mathbf{R}) = \mathrm{Aut}(\mathcal{H})$. Note que se $G = \Gamma_0(N)$, então

$$\overline{G} = \Gamma_0(N)/\pm 1,$$

mas que se $N \geq 3$ temos $\overline{\Gamma_1(N)} = \Gamma_1(N)$ e $\overline{\Gamma(N)} = \Gamma(N)$, já que $-1 \notin \Gamma_1(N)$ se $N > 2$.

Seja G um dos subgrupos de $\mathrm{SL}_2(\mathbf{Z})$ indicados acima. Então podemos considerar o espaço de órbitas

$$G\backslash\mathcal{H} = \{G\tau : \tau \in \mathcal{H}\},$$

cujos elementos são as classes de pontos do semi-plano superior equivalentes sob G. Nosso primeiro objetivo é entender um pouco melhor este espaço de órbitas.

1.2 O domínio fundamental para Γ

Começamos obtendo informações sobre o espaço de órbitas no caso em que $G = \Gamma$. Nosso primeiro teorema importante descreve um domínio fundamental para esse caso. Primeiro, lembramos o que isso significa:

Definição 2 *Seja* G *um grupo discreto agindo em* $\mathcal{H}$. *Um domínio fundamental para a ação de* G *é uma região fechada* $D \subset \mathcal{H}$ *tal que*

i) A função natural $D \longrightarrow G\backslash\mathcal{H}$ *é sobrejetora, e*

ii) Sua restrição ao interior de D *é injetora.*

Em outras palavras, D é um domínio fundamental se cada órbita tem pelo menos um representante em D, e este representante é único exceto talvez quando pertencer à fronteira de D.

No caso $G = \Gamma = \mathrm{SL}_2(\mathbf{Z})$, vamos descrever um domínio fundamental; basta, é claro, trabalhar com $\overline{\Gamma} = \mathrm{PSL}_2(\mathbf{Z})$. Sejam

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \overline{\Gamma}$$

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \overline{\Gamma}$$

os elementos cuja ação em $\mathcal{H}$ é dada por

$$S\tau = -\frac{1}{\tau} \qquad T\tau = \tau + 1.$$

É um cálculo imediato verificar que

$$S^2 = 1 \quad \text{e} \quad (ST)^3 = 1;$$

por exemplo,

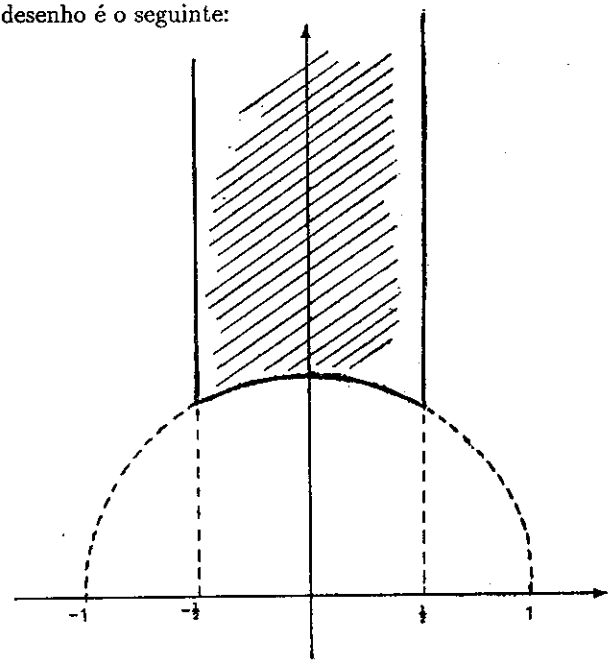
$$S^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = 1$$

em $\overline{\Gamma}$.

Consideremos agora a região $\mathsf{D} \subset \mathcal{H}$ definida da seguinte forma:

$$\mathsf{D} = \{\tau \in \mathcal{H} : |\tau| \geq 1 \text{ e } |\mathrm{Re}(\tau)| \leq 1/2\}.$$

O desenho é o seguinte:



Vamos provar que a região D é um domínio fundamental para a ação de Γ.

Teorema 1.2.1 *Seja* $\mathsf{D} \subset \mathcal{H}$ *a região definida acima, e* $\overline{\Gamma} = \mathrm{PSL}_2(\mathbf{Z})$. *Então:*

i) Para todo $\tau \in \mathcal{H}$ *existe* $\gamma \in \overline{\Gamma}$ *tal que* $\gamma\tau \in \mathsf{D}$.

ii) Se τ *e* τ' *estão em* D *e temos* $\gamma\tau = \tau'$ *para algum* $\gamma \in \overline{\Gamma}$, *então ou* $\Re(\tau) = \pm 1/2$ *e* $\tau = \tau' \pm 1$ *ou* $|\tau| = 1$ *e* $\tau' = -1/\tau$.

iii) Seja $\tau \in \mathsf{D}$, *e defina* $\mathrm{Stab}(\tau) = \{\gamma \in \overline{\Gamma} : \gamma\tau = \tau\}$. *Então* $\mathrm{Stab}(\tau) = 1$ *exceto se:*

(a) $\tau = i$, *quando* $\mathrm{Stab}(\tau) = \langle S \rangle$, *de ordem 2,*

(b) $\tau = \rho = e^{2\pi i/3}$, *quando* $\mathrm{Stab}(\tau) = \langle ST \rangle$, *de ordem 3,*

(c) $\tau = -\overline{\rho} = e^{\pi i/3}$, *quando* $\mathrm{Stab}(\tau) = \langle TS \rangle$, *de ordem 3.*

iv) O grupo $\overline{\Gamma}$ *é gerado por* S *e* T.

DEMONSTRAÇÃO: Seguimos o método de Serre em [Ser73].

i) Seja G o subgrupo de $\overline{\Gamma}$ gerado por S e T, e seja $\tau \in \mathcal{H}$. Vamos provar primeiro que existe $\gamma \in G$ tal que $\gamma\tau \in D$, o que implica, em particular, a primeira asserção do teorema. (O fato de que é possível achar o elemento γ em G vai ser importante, mais adiante, quando formos provar que de fato $G = \overline{\Gamma}$.)

Notemos, primeiro, que se $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \overline{\Gamma}$, temos

$$\Im\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{(ad - bc)\Im(\tau)}{|c\tau + d|^2}. \tag{1.2}$$

Como c e d são inteiros, dado um número M existem apenas um número *finito* de pares (c, d) tais que $|c\tau + d| < M$ (já que $\mathbf{Z}\tau + \mathbf{Z}$ é um conjunto discreto!). Logo, existe $\gamma \in G$ tal que $\mathrm{Im}(\gamma\tau)$ é máximo; fixemos um tal γ. Notando, agora, que $\Im(T^n\gamma\tau) = \Im(\gamma\tau)$ para qualquer $n \in \mathbf{Z}$, vemos que podemos escolher γ de modo que

$$-\frac{1}{2} \leq \Re(\gamma\tau) \leq \frac{1}{2}.$$

(Basta substituir γ por um $T^n\gamma$ apropriado.)

Verifiquemos, então, que $\gamma\tau \in D$; pela construção acima, basta verificar que $|\gamma\tau| \geq 1$. Mas, caso contrário, teríamos que

$$\Im(S\gamma\tau) = \Im(\frac{-1}{\gamma\tau}) > \mathrm{Im}(\gamma\tau),$$

o que contradiz a escolha de γ.

ii) Vamos provar a segunda e a terceira asserções simultaneamente. Seja $\tau \in D$ e tome $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \overline{\Gamma}$ tal que $\gamma\tau \in D$. Trocando, se necessário, o par (τ, γ) pelo par $(\gamma\tau, \gamma^{-1})$, podemos supor que

$$\Im(\gamma\tau) \geq \Im(\tau),$$

ou seja, pela fórmula 1.2, que

$$|c\tau + d| \leq 1.$$

Como $\tau \in D$, $|\tau| \geq 1$, donde a desigualdade só é possível se $c = 0$, 1, ou -1. Vamos examinar caso por caso.

- Se $c = 0$, temos que ter $d = \pm 1$, donde

$$\gamma = \begin{pmatrix} 1 & \pm b \\ 0 & 1 \end{pmatrix}$$

(em $\overline{\Gamma}$, não em Γ). Como a parte real tanto de τ quanto de $\gamma\tau = \tau \pm b$ têm que estar entre $-1/2$ e $1/2$, temos que ter $b = 0$, donde $\gamma = 1$, ou $b = 1$, donde $\gamma = T$ ou $\gamma = T^{-1}$, como queríamos.

- Se $c = 1$, temos $|\tau + d| \leq 1$, o que implica $d = 0$ exceto se $\tau = \rho$ ou $-\overline{\rho}$ e $d \in \{0, 1\}$ ou $d \in \{0, -1\}$, respectivamente. Olhemos os casos separadamente.

- Se $c = 1$ e $d = 0$, temos $|\tau| \leq 1$, donde $|\tau| = 1$; além disso, $ad - bc = 1$, donde $b = -1$, e

$$\gamma\tau = \frac{a\tau - 1}{\tau} = a - \frac{1}{\tau}.$$

Outra vez, segue que $a = 0$ exceto se $\tau = \rho$ ou $-\overline{\rho}$, e neste caso $a \in \{0, 1\}$ ou $\{0, -1\}$, respectivamente.

- Se $\tau = \rho$, $c = d = 1$, o mesmo argumento força $a \in \{0, 1\}$, e analogamente para $\tau = -\overline{\rho}$ e $c = -d = 1$.

- Finalmente, o caso $c = -1$ é tratado trocando o sinal de todos os termos da matriz, o que não muda $\gamma \in \overline{\Gamma}$.

Isto verifica a segunda e terceira asserções (basta escrever ST, TS e seus quadrados matricialmente).

iii) Resta provar que $G = \overline{\Gamma}$. Para isso, tome um elemento $\gamma \in \overline{\Gamma}$ qualquer, escolha τ no interior de D, e considere o ponto $\gamma\tau \in \mathcal{H}$. Por (a), existe $\gamma' \in G$ tal que $\gamma'\gamma\tau \in D$. Os pontos τ e $\gamma'\gamma\tau$ estão ambos em D, e $\tau \in \text{int}(D)$. Por (b), segue que $\tau = \gamma'\gamma\tau$, donde $\gamma'\gamma = 1$, donde $\gamma \in G$, como queríamos demonstrar. □

OBSERVAÇÃO: Pode-se mostrar que $\langle S, T; S^2 = (ST)^3 = 1\rangle$ é uma apresentação do grupo $\overline{\Gamma}$, que é portanto o produto livre de um grupo cíclico de ordem 2 por um grupo cíclico de ordem 3.

Uma vez obtido um domínio fundamental para o grupo Γ, é fácil achar domínios fundamentais para os subgrupos de congruência, já que estes são de índice finito em Γ. Basta, então, achar um conjunto de representantes das classes laterais e considerar a união das imagens de D por cada um destes elementos.

Por exemplo, vamos achar um domínio fundamental para o grupo $\Gamma_0(2)$. Precisamos primeiro calcular o seu índice em Γ. Temos

$$\begin{array}{ccc} \Gamma & \xrightarrow{r_2} & \mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z}) \\ \cup & & \cup \\ \Gamma_0(2) & \xrightarrow{r_2} & H \end{array}$$

onde

$$H = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$$

é o subgrupo de Borel e r_2 é redução módulo 2; então

$$[\Gamma : \Gamma_0(2)] = [\mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z}) : H] = 3.$$

Um sistema de representantes é dado por

$$\alpha_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = S, \quad \alpha_3 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = ST.$$

Temos, então,

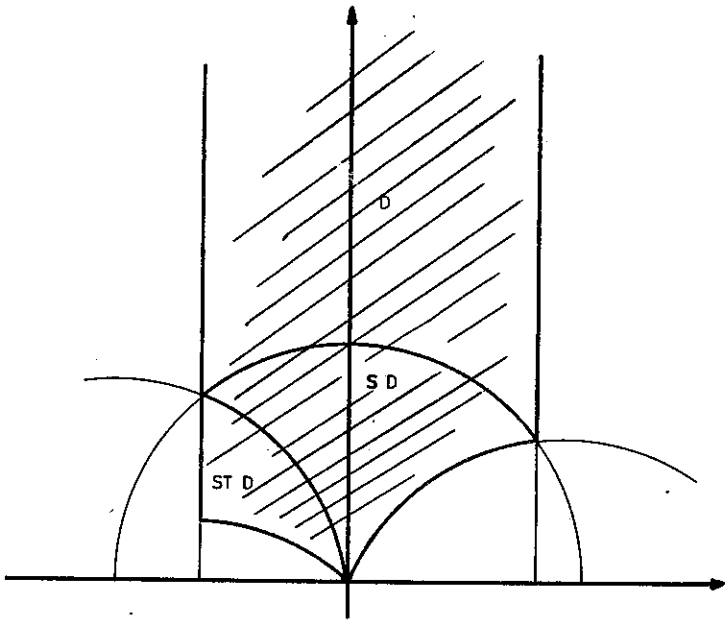
$$\Gamma = \Gamma_0(2)\alpha_1 \cup \Gamma_0(2)\alpha_2 \cup \Gamma_0(2)\alpha_3,$$

e obtemos um domínio fundamental tomando

$$\mathrm{D}_0(2) = \alpha_1\mathrm{D} \cup \alpha_2\mathrm{D} \cup \alpha_3\mathrm{D},$$

já que todo $\gamma \in \Gamma$ pode ser escrito de forma única como $\gamma = \gamma'\alpha_i$, com $\gamma' \in \Gamma_0(2)$. (Veja os exercícios para a verificação de que isto dá certo.)

O desenho é:



Da mesma forma, para $\Gamma(2)$, note que $[\Gamma_0(2) : \Gamma(2)] = 2$, e que T é o representante não-trivial. Logo obtemos um domínio fundamental pondo

$$\mathsf{D}(2) = \mathsf{D}_0(2) \cup T\mathsf{D}_0(2).$$

É claro que há muitas outras escolhas para os sistemas de representantes, e portanto para os domínios fundamentais obtidos. O mesmo processo permite obter domínios fundamentais para outros subgrupos discretos de Γ, com maior ou menor dificuldade.

1.3 As "pontas"

O desenho dos domínios fundamentais acima mostram que, dependendo do grupo em questão, pode haver várias "pontas", isto é, vários lugares em que o domínio fundamental se aproxima do bordo de $\mathcal{H}$. (Vale recordar ao leitor que o bordo de $\mathcal{H}$ foi definido como constando dos pontos da reta real, mais um ponto no infinito.) No caso de Γ, por exemplo, a única ponta é o ponto no infinito (que deve ser pensado como "no eixo y"—às vezes enfatizamos isso chamando o ponto de $i\infty$). No caso de $\Gamma_0(2)$, vê-se do desenho acima que há uma outra ponta, em $\tau = 0$. É fácil ver, a partir do método descrito acima para obter domínios

fundamentais, que qualquer ponta para um subgrupo de Γ tem que ser imagem do ponto no infinito por algum $\gamma \in \Gamma$. Agora, se $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, temos

$$\gamma \cdot \infty = \frac{a}{c} \in \mathbf{Q},$$

de modo que as pontas estarão entre os pontos de uma reta projetiva

$$\mathbf{P}^1(\mathbf{Q}) = \mathbf{Q} \cup \infty,$$

que é estável sob a ação de Γ. Isso leva à seguinte definição:

Definição 3 *Seja $G \subset \Gamma$ um subgrupo de congruência. Uma* ponta *para a ação de G em $\mathcal{H}$ é uma órbita da ação de G em $\mathbf{P}^1(\mathbf{Q})$.*

Note que Γ age transitivamente em $\mathbf{P}^1(\mathbf{Q})$, de modo que toda ponta para um subgrupo G é da forma $\gamma \cdot \infty$, como esperávamos. Outra vez, podemos achar as pontas olhando as imagens de ∞ por um conjunto de representantes das classes laterais. Por exemplo, para $G = \Gamma_0(2)$, usando os representantes acima, temos as pontas $\alpha_1 \cdot \infty = \infty$ e $\alpha_2 \cdot \infty = \alpha_3 \cdot \infty = 0$, que não são equivalentes sob $\Gamma_0(2)$. No caso geral, as pontas estão *entre* as imagens $\alpha_i \cdot \infty$, mas é preciso sempre verificar se duas ou mais destas não são equivalentes sob a ação do subgrupo de congruência em questão.

Uma forma de fazer toda esta discussão numa linha é definir

$$\mathcal{H}^* = \mathcal{H} \cup \mathbf{P}^1(\mathbf{Q}),$$

de modo que, para $G \subset \Gamma$, temos

$$G\backslash\mathcal{H}^* = G\backslash\mathcal{H} \cup \{\text{pontas para G}\}.$$

Como veremos à frente, as pontas têm um papel especial na geometria do espaço quociente, porque fornecem uma compactificação natural.

O ponto no infinito é sempre uma das pontas; o conjunto dos elementos $\gamma \in \Gamma$ que estabilizam este ponto é

$$\text{Stab}(\infty) = \{\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in \mathbf{Z}\} \subset \Gamma.$$

Se agora considerarmos outra ponta $\gamma\cdot\infty$, o estabilizador é simplesmente o conjugado deste:

$$\mathrm{Stab}(\gamma\cdot\infty) = \gamma\,\mathrm{Stab}(\infty)\,\gamma^{-1}.$$

Para qualquer subgrupo de congruência G, definimos a *largura* da ponta $\tau = \infty$ como sendo o menor inteiro $n > 0$ tal que

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in G.$$

Note que se $G \supset \Gamma(\mathrm{N})$, temos certamente que

$$\begin{pmatrix} 1 & \mathrm{N} \\ 0 & 1 \end{pmatrix} \in G,$$

de modo que $n \leq \mathrm{N}$. (É verdade que $n|\mathrm{N}$? A distinção entre G e $\overline{G}$ faz diferença!) Da mesma forma, definimos a largura de uma ponta $\gamma\cdot\infty$ como sendo o menor $n > 0$ tal que

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in \gamma G \gamma^{-1}.$$

Por exemplo:

i) Se $G = \Gamma$, a única ponta é ∞, de largura 1.

ii) Se $G = \Gamma_0(2)$, as pontas são 0 e ∞, de largura 2 e 1, respectivamente. (Verifique!)

iii) Se $G = \Gamma(2)$, as pontas são 0, 1 e ∞, todas de largura 2.

Note que se G for normal em Γ, todas as pontas terão a mesma largura. Note também que a largura de cada ponta é facilmente percebida em um desenho do domínio fundamental.

1.4 A interpretação via curvas elípticas

As construções que fizemos acima podem ser interpretadas em termos da teoria de curvas elípticas. Nesta seção, vamos explicar rapidamente esta interpretação, que é importante em muitas das aplicações aritméticas da teoria. Começamos lembrando os fatos básicos sobre curvas elípticas complexas que são a base da nossa discussão.

Começamos com uma definição simples:

Definição 4 *Um rede $\Lambda \subset \mathbf{C}$ é um subgrupo aditivo de $\mathbf{C}$ que é de posto 2 como $\mathbf{Z}$-módulo.*

É um exercício simples mostrar que uma rede é um subgrupo discreto maximal de $\mathbf{C}$. As redes $\Lambda \subset \mathbf{C}$ estão ligadas de perto às curvas elípticas complexas.

Uma curva elíptica complexa E pode ser definida como uma subvariedade do plano projetivo complexo $\mathbf{P}^2(\mathbf{C})$ dada por uma equação do tipo

$$ZY^2 = 4X^3 + aXZ^2 + bZ^3,$$

com $a, b \in \mathbf{C}$, onde $[X : Y : Z]$ são as coordenadas homogêneas em $\mathbf{P}^2$, e onde a cúbica $x^3 + ax + b$ tem três raízes distintas. É fácil verificar que uma tal equação define uma variedade complexa lisa, de dimensão um sobre $\mathbf{C}$.

Topologicamente, E é uma superfície compacta orientável; para ver que E é de fato um toro, prova-se o seguinte:

Fato 1.4.1 *Dada uma curva elíptica complexa* E *definida por uma equação*

$$ZY^2 = 4X^3 + aXZ^2 + bZ^3,$$

como acima, existe uma função meromorfa $\wp : \mathbf{C} \longrightarrow \mathbf{C}$, que depende de a e b, com as seguintes propriedades:

i) Existem dois números $w_1, w_2 \in \mathbf{C}$, satisfazendo $\Im(w_1/w_2) > 0$, que são períodos de $\wp$, isto é, tais que

$$\wp(z + nw_1 + mw_2) = \wp(z),$$

para quaisquer $n, m \in \mathbf{Z}$;

ii) $\wp$ é holomorfa exceto por polos de ordem 2 com resíduo zero nos pontos da rede $\mathbf{Z}w_1 + \mathbf{Z}w_2 \subset \mathbf{C}$;

iii) A derivada $\wp'$ tem w_1 e w_2 como períodos, e é holomorfa exceto por polos de ordem 3 nos pontos da rede $\mathbf{Z}w_1 + \mathbf{Z}w_2$;

iv) A função $\wp$ é par, e a sua derivada $\wp'$ é ímpar;

v) $\wp$ satisfaz a equação funcional

$$(\wp'(z))^2 = 4\wp(z)^3 + a\wp(z) + b.$$

Juntando tudo isso, é fácil ver que a aplicação

$$\begin{array}{ccc} \mathbf{C}/\mathbf{Z}w_1 + \mathbf{Z}w_2 & \longrightarrow & \mathrm{E} \\ z & \mapsto & [\wp(z) : \wp'(z) : 1] \end{array}$$

determina um isomorfismo entre a curva elíptica E e o toro complexo $\mathbf{C}/(\mathbf{Z}w_1 + \mathbf{Z}w_2)$. Em particular, isto dá a E a estrutura de um *grupo abeliano*, usando a estrutura aditiva em C.

No que segue, queremos estudar um pouco melhor a relação entre a curva E e os períodos w_1 e w_2. Especificamente, queremos tentar classificar as classes de isomorfismo de curvas a partir dos períodos correspondentes. Para isso, precisamos descrever, em termos dos períodos, quando duas curvas são isomorfas. Faremos isto em dois passos: primeiro, relacionamos E à rede $\mathbf{Z}w_1 + \mathbf{Z}w_2$ gerada pelos períodos; depois, estudamos a relação entre os períodos e a rede que eles geram. Em relação a este segundo aspecto, fazemos, por hora, apenas a observação de que se w_1 e w_2 são linearmente independentes sobre $\mathbf{R}$, podemos sempre supor que $\Im(w_1/w_2) > 0$, como no "Fato" acima (basta trocar os índices se necessário). Faremos sempre essa hipótese.

Dada uma curva elíptica E, o "Fato" mostra que existe uma rede Λ tal que $\mathrm{E} = \mathbf{C}/\Lambda$; mas redes diferentes podem definir curvas isomorfas:

Fato 1.4.2 *Sejam* E_1 *e* E_2 *duas curvas elípticas complexas, com*

$$\mathrm{E}_1 \cong \mathbf{C}/\Lambda_1 \quad e \quad \mathrm{E}_2 \cong \mathbf{C}/\Lambda_2.$$

Todo homomorfismo analítico $\varphi : \mathrm{E}_1 \longrightarrow \mathrm{E}_2$ *é induzido por multiplicação por um escalar* $\lambda \in \mathbf{C}$ *que satisfaz* $\lambda\Lambda_1 \subset \Lambda_2$. *Isto é,* φ *se encaixa em um diagrama comutativo:*

$$\begin{array}{ccc} \mathbf{C} & \xrightarrow{z \mapsto \lambda z} & \mathbf{C} \\ \downarrow & & \downarrow \\ \mathbf{C}/\Lambda_1 & \xrightarrow{\varphi} & \mathbf{C}/\Lambda_2 \end{array}$$

Portanto, temos

$$\mathbf{C}/\Lambda_1 \cong \mathbf{C}/\Lambda_2$$

se e só se existe um número $\lambda \in \mathbf{C}^{\times}$ **tal que** $\lambda\Lambda_1 = \Lambda_2$. **Assim, para classificar curvas E a menos de isomorfismo é o mesmo que classificar redes a menos de homotetia.**

A primeira coisa a notar é que se $\Lambda = \mathbf{Z}w_1 + \mathbf{Z}w_2$ com $\Im(w_1/w_2) > 0$, podemos dividir por w_2 e obter a rede homotética

$$w_2^{-1}\Lambda = \mathbb{Z}\frac{w_1}{w_2} + \mathbb{Z} = \mathbb{Z}\tau + \mathbb{Z},$$

onde $\Im(\tau) = \Im(w_1/w_2) > 0$, isto é, onde $\tau \in \mathcal{H}$. Com isto, já provamos:

Proposição 1.4.3 *A cada ponto $\tau \in \mathcal{H}$ do semi-plano superior corresponde uma curva elíptica complexa E_τ satisfazendo*

$$\mathrm{E}_\tau \cong \mathbb{C}/\mathbb{Z}\tau + \mathbb{Z};$$

toda curva elíptica complexa é isomorfa a alguma curva deste tipo.

Resta determinar quando duas curvas deste tipo são isomorfas. Dados τ e σ em $\mathcal{H}$, temos que $\mathrm{E}_\tau \cong \mathrm{E}_\sigma$ se e só se existe um escalar $\lambda \in \mathbb{C}$ tal que

$$\mathbb{Z}\tau + \mathbb{Z} = \mathbb{Z}\lambda\sigma + \mathbb{Z}\lambda,$$

isto é, se e só se existe uma matriz $\gamma \in \mathrm{SL}_2(\mathbf{Z})$, $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, tal que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \tau \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda\sigma \\ \lambda \end{pmatrix},$$

isto é,

$$\begin{pmatrix} a\tau + b \\ c\tau + d \end{pmatrix} = \begin{pmatrix} \lambda\sigma \\ \lambda \end{pmatrix},$$

donde é claro que λ existe se e só se temos

$$\frac{a\tau + b}{c\tau + d} = \sigma,$$

isto é, se e só se τ e σ são equivalentes sob a ação do grupo modular. A moral da história, então, é:

Teorema 1.4.4 *A correspondência*

$$\tau \in \mathcal{H} \mapsto \mathrm{E}_\tau$$

induz uma bijeção entre o conjunto

$$\Gamma\backslash\mathcal{H} = \textit{órbitas de } \mathcal{H} \textit{ sob a ação de } \mathrm{SL}_2(\mathbf{Z})$$

e o conjunto das classes de isomorfismo de curvas elípticas complexas.

Esta correspondência permite dar uma estrutura natural de superfície de Riemann ao conjunto das classes de isomorfismo de curvas elípticas complexas. Mais além, daremos uma descrição explícita dessa superfície de Riemann.

OBSERVAÇÕES:

1) Uma pergunta natural, a esta altura, é como obter τ (ou, equivalentemente, os períodos w_1 e w_2) a partir da curva E. Damos aqui apenas um esboço de como isso é feito. Note, primeiro, que para $\tau \in \mathcal{H}$ dado, a curva E_τ vem munida de uma diferencial holomorfa invariante $\omega = dz$ (isto é, consideramos a imagem de dz módulo a rede $\mathbf{Z}\tau + \mathbf{Z}$). Uma escolha diferente de τ dá uma diferencial diferente (porque muda a rede), e corresponde a uma mudança de variável por um $\gamma \in \mathrm{SL}_2(\mathbf{Z})$. É fácil ver, por outro lado, que integrando ω ao longo de um dos ciclos não triviais de $\mathrm{H}_1(\mathrm{E}_\tau, \mathbf{Z})$ obtemos um elemento da rede $\mathbf{Z}\tau + \mathbf{Z}$ (lembre que E_τ é um toro!); escolhendo dois ciclos que formam uma base da homologia, obtemos uma base da rede $\mathbf{Z}\tau + \mathbf{Z}$, e portanto τ. Assim, a correspondência

$$\tau \in \mathcal{H} \mapsto (\mathrm{E}_\tau, \omega),$$

com $\omega = dz \bmod \mathbf{Z}\tau + \mathbf{Z}$, é uma bijeção.

Ainda, se $\gamma\tau = \sigma$ para $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$, e munirmos E_τ de ω_1 vindo de dz e E_σ de ω_2 vindo de dz, então o isomorfismo $\varphi : \mathrm{E}_\tau \longrightarrow \mathrm{E}_\sigma$ se encaixa no diagrama comutativo:

$$\begin{array}{ccc} \mathrm{E}_\tau & \xrightarrow{\varphi} & \mathrm{E}_\sigma \\ \uparrow & & \uparrow \\ \mathbf{C} & \xrightarrow{z \mapsto \lambda^{-1} z} & \mathbf{C} \end{array},$$

onde $\lambda = c\tau + d$ (veja acima). Assim, temos que $\varphi^* \omega_2 = \lambda^{-1}\omega_1$, de modo que $(\mathrm{E}_\sigma, dz) \cong (\mathrm{E}_\tau, \lambda^{-1} dz)$. Dessa forma, passar de τ para σ equivale a fixar a curva e multiplicar a diferencial por $\lambda^{-1} = (c\tau + d)^{-1}$.

2) Seria muito bom se fosse possível construir uma família analítica de curvas elípticas

$$\mathcal{E} \xrightarrow{\pi} \Gamma\backslash\mathcal{H}$$

tal que $\pi^{-1}(\tau) \cong \mathrm{E}_\tau$. A maneira mais natural de obter tal coisa seria considerar a projeção natural

$$\mathcal{H} \times \mathbf{C} \longrightarrow \mathcal{H},$$

e considerar a relação de equivalência em $\mathcal{H} \times \mathbf{C}$ dada por $(\tau, z) \sim (\tau', z')$ se $\tau = \tau'$ e $z - z' \in \mathbf{Z}\tau + \mathbf{Z}$. Passando ao quociente obtemos

$$\mathrm{X} \xrightarrow{\pi} \mathcal{H}$$

com $\pi^{-1}(\tau) = \{\tau\} \times \mathrm{E}_\tau$. Agora se $\tau = \gamma\sigma$ com $\gamma \in \mathrm{SL}_2(\mathbf{Z})$, basta identificar E_τ e E_σ para podermos passar ao quociente e obtermos a família $\mathcal{E}$. O problema é que o isomorfismo $\mathrm{E}_\tau \cong \mathrm{E}_\sigma$ *não é único*, e pode-se mostrar que não é possível fazer uma escolha coerente. Mais ainda, é possível provar que a "curva universal" $\mathcal{E}$ não existe.

A inexistência da "curva universal" é uma das motivações para o estudo dos subgrupos de congruência. Isto porque podemos evitar o problema da não-unicidade do isomorfismo introduzindo estrutura adicional no problema, isto é, nos restringindo a estudar isomorfismos que respeitem certas condições. Isto se chama "rigidificar o problema". As rigidificações que vamos escolher vão-nos levar aos subgrupos de congruência.

No que segue, fixemos um número inteiro positivo N. Lembremos que uma curva elíptica complexa E tem uma estrutura de grupo; o subgrupo

$$\mathrm{E}[\mathrm{N}] = \{P \in \mathrm{E}\,;\, \mathrm{N}P = 0\}$$

é isomorfo ao produto de dois grupos cíclicos de ordem N (veja acima a estrutura de E como toro complexo). Nesta situação, consideraremos três tipos de estrutura adicional em uma curva elíptica complexa E:

- um subgrupo $C \subset \mathrm{E}$, cíclico de ordem N
- um ponto $P \in \mathrm{E}$ que gera um subgrupo cíclico de ordem N

- dois pontos P, $Q \in \mathrm{E}$ que são uma base do subgrupo E[N], isto é, que dão um isomorfismo

$$\mathrm{E[N]} \cong \mathbf{Z}/\mathrm{N}\mathbf{Z} \times \mathbf{Z}/\mathrm{N}\mathbf{Z}.$$

A cada $\tau \in \mathcal{H}$ fazemos corresponder, em cada caso:

- $(\mathrm{E}_\tau, C_\tau)$, onde C_τ é o subgrupo gerado por $1/\mathrm{N} \bmod \mathbf{Z}\tau + \mathbf{Z}$, isto é

$$C_\tau = \frac{\mathbf{Z}\tau + \mathbf{Z}\frac{1}{\mathrm{N}}}{\mathbf{Z}\tau + \mathbf{Z}}$$

- $(\mathrm{E}_\tau, P_\tau)$, onde $P_\tau = \dfrac{1}{\mathrm{N}} \bmod \mathbf{Z}\tau + \mathbf{Z}$

- $(\mathrm{E}_\tau, P_\tau, Q_\tau)$, onde $P_\tau = \dfrac{1}{\mathrm{N}}$ e $Q_\tau = \dfrac{\tau}{\mathrm{N}} \bmod \mathbf{Z}\tau + \mathbf{Z}$.

Dada uma curva elíptica complexa E existe sempre um $\tau \in \mathcal{H}$ tal que $\mathrm{E} \cong \mathrm{E}_\tau$. Isto continua verdade com estrutura adicional, exceto que no terceiro caso acima (a "estrutura de nível plena") é preciso um cuidado técnico com o qual não nos queremos preocupar muito. Diremos apenas, então, que dada uma base (P, Q) de E[N], é possível definir um "determinante" $\det(P, Q)$, e que, com a definição acima, $\det(P_\tau, Q_\tau) = 1$. Temos, então, o seguinte teorema, cuja demonstração não é difícil se entendermos bem como obter o parâmetro τ a partir da curva E.

Fato 1.4.5 *Seja* E *uma curva elíptica complexa.*

i) Dado um subgrupo cíclico $C \subset \mathrm{E}$ *de ordem* N, *existe* $\tau \in \mathcal{H}$ *tal que* $(\mathrm{E}, C) \cong (\mathrm{E}_\tau, C_\tau)$, *isto é, tal que existe um isomorfismo* $\varphi : \mathrm{E} \longrightarrow \mathrm{E}_\tau$ *com* $\varphi(C) = C_\tau$.

ii) Dado um ponto $P \in \mathrm{E}$ *de ordem* N, *existe* $\tau \in \mathcal{H}$ *tal que* $(\mathrm{E}, P) \cong (\mathrm{E}_\tau, P_\tau)$.

iii) Dada uma base (P, Q) *do subgrupo* E[N] *com* $\det(P, Q) = 1$, *existe* $\tau \in \mathcal{H}$ *tal que* $(\mathrm{E}, P, Q) \cong (\mathrm{E}_\tau, P_\tau, Q_\tau)$.

Agora a ligação com os subgrupos de congruência não é difícil de descrever:

Proposição 1.4.6 *Sejam* $\tau, \sigma \in \mathcal{H}$. *Então:*

i) $(\mathrm{E}_\tau, C_\tau) \cong (\mathrm{E}_\sigma, C_\sigma)$ *se e só se* $\sigma = \gamma \cdot \tau$ *para algum* $\gamma \in \Gamma_0(\mathrm{N})$;

ii) $(E_\tau, P_\tau) \cong (E_\sigma, P_\sigma)$ *se e só se* $\sigma = \gamma \cdot \tau$ *para algum* $\gamma \in \Gamma_1(N)$;

iii) $(E_\tau, P_\tau, Q_\tau) \cong (E_\sigma, P_\sigma, Q_\sigma)$ *se e só se* $\sigma = \gamma \cdot \tau$ *para algum* $\gamma \in \Gamma(N)$;

DEMONSTRAÇÃO: Não é difícil—fica como exercício para o leitor.□

Corolário 1.4.7 *Existem bijeções*

$$\Gamma_0(N)\backslash\mathcal{H} \longleftrightarrow Isom\{(E, C)\}$$

$$\Gamma_1(N)\backslash\mathcal{H} \longleftrightarrow Isom\{(E, P)\}$$

$$\Gamma(N)\backslash\mathcal{H} \longleftrightarrow Isom\{(E, P, Q) : \det(P, Q) = 1\}$$

onde $C \subset E$ *é um subgrupo cíclico de ordem* N, $P \in E$ *é um ponto de ordem* N, *e* (P, Q) *é uma base de* E[N], *e onde* $Isom(\star)$ *denota o conjunto das classes de isomorfismo dos objetos* $\star$.

Que a estrutura adicional que estamos considerando de fato rigidifica o problema é o conteúdo do próximo resultado:

Proposição 1.4.8 *Com as notações acima, temos:*

i) $\mathrm{Aut}(E, C) = \{\pm 1\}$ *se* $N \geq 2$

ii) $\mathrm{Aut}(E, P) = \mathrm{Aut}(E, P, Q) = \{\pm 1\}$ *se* $N = 2$

iii) $\mathrm{Aut}(E, P) = \mathrm{Aut}(E, P, Q) = \{1\}$ *se* $N \geq 3$

DEMONSTRAÇÃO: Trata-se apenas de calcular os estabilizadores dos pontos de um domínio fundamental para cada um dos grupos $\Gamma_0(N)$, $\Gamma_1(N)$ e $\Gamma(N)$. Veja as referências para os detalhes.□

Nos casos rígidos, isto é, quando o grupo de automorfismos é trivial, é possível construir uma "curva universal com estrutura de nível N", e mesmo generalizar tudo para curvas elípticas definidas sobre outros corpos (veja [KM85] para um tratamento quase completo, mas que requer um bom conhecimento de geometria algébrica).

1.5 Exercícios

1) Prove a Proposição 1.1.1.

2) Seja p um primo.

i) Calcule o índice de $\Gamma(p)$ em Γ.

ii) Calcule os índices $[\Gamma : \Gamma_0(p)]$ e $[\Gamma : \Gamma_1(p)]$.

3) Mostre que $\Gamma_1(N)$ é um subgrupo normal de $\Gamma_0(N)$, e que

$$\Gamma_0(N)/\Gamma_1(N) \cong \left(\mathbf{Z}/N\mathbf{Z}\right)^{\times}.$$

4) Seja D um domínio fundamental para a ação de Γ, e seja $G \subset \Gamma$ um subgrupo de índice finito. Se

$$\Gamma = \bigcup_{i=1}^{n} G\alpha_i$$

é uma decomposição de Γ em classes laterais de G, mostre que

$$D' = \bigcup_{i=1}^{n} \alpha_i D$$

é um domínio fundamental para a ação de G.

5) Sejam S e $T \in \Gamma$ como no texto, e seja p um primo. Mostre que se $\gamma \in \Gamma$, $\gamma \notin \Gamma_0(p)$, então existe $\delta \in \Gamma_0(p)$ e um inteiro k, $0 \leq k \leq p-1$, tais que $\gamma = \delta S T^k$.

6) Seja D o domínio fundamental para o grupo modular Γ obtido no texto. Mostre que se p é um primo,

$$D_0(p) = D \cup \bigcup_{k=0}^{p-1} ST^k D$$

é um domínio fundamental para $\Gamma_0(p)$.

7) Desenhe um domínio fundamental para $\Gamma_0(5)$.

8) Mostre que se p é um primo, há duas pontas para a ação de $\Gamma_0(p)$: $\tau = \infty$, de largura 1, e $\tau = 0$, de largura p.

9) Ache um isomorfismo entre $\Gamma(N)$ e um subgrupo de índice $\phi(N)$ de $\Gamma_0(N^2)$, onde ϕ é a função de Euler. Conclua que $\Gamma(2)$ e $\Gamma_0(4)$ são isomorfos.

10) Descreva todos os subgrupos de congruência de nível 2, isto é, todos os grupos G tais que $\Gamma(2) \subset G \subset \Gamma$. Ache um domínio fundamental para cada tal G. Determine as pontas para a ação de cada um destes grupos.

11) Ache um domínio fundamental para $\Gamma_0(4)$. (Sugestão: mostre primeiro que se Γ_1 e Γ_2 são subgrupos de índice finito em Γ, e $\Gamma_1 = \alpha\Gamma_2\alpha^{-1}$ para algum $\alpha \in GL_2^+(\mathbf{Q})$, e se D_2 é um domínio fundamental para Γ_2, então αD_2 é um domínio fundamental para Γ_1.)

12) Neste problema vamos determinar geradores para a imagem em $PSL_2(\mathbf{R})$ de vários dos subgrupos de congruência G obtidos no exercício 10. Lembre que se $G \subset \Gamma$, denotamos essa imagem por $\overline{G}$.

a) Mostre que S e T^2 geram um dos subgrupos de congruência de nível 2. Este grupo foi estudado por Hecke, e costuma ser denotado por $G(2)$.

b) Mostre que T e ST^2S geram $\overline{\Gamma_0(2)}$. (Sugestão: verifique que $\overline{\Gamma_0(2)}$ e $G(2)$ são conjugados em $\overline{\Gamma}$.)

c) Mostre que T^2 e $ST^{-2}S$ geram $\overline{\Gamma(2)}$.

13) Use os exercícios anteriores para provar que T e ST^4S geram $\overline{\Gamma_0(4)}$.

14) Mostre que uma rede $\Lambda \subset \mathbf{C}$ é um subgrupo aditivo discreto maximal de $\mathbf{C}$, e reciprocamente.

15) O grupo dos pontos de N-divisão de uma curva elíptica E é o kernel da multiplicação por N em E:

$$\mathrm{E[N]} = \{P \in \mathrm{E} : \mathrm{N}P = 0\}.$$

Mostre, usando os fatos dados no texto, que

$$\mathrm{E[N]} \cong \mathbf{Z}/\mathrm{N}\mathbf{Z} \times \mathbf{Z}/\mathrm{N}\mathbf{Z}.$$

16) Prove o Fato 1.4.2. (Sugestão: levante o mapa ao recobrimento universal, e expanda em série de potências.)

17) Prove a Proposição 1.4.6.

1.6 Notas

O material neste capítulo é clássico, e pode ser encontrado em todas as referências, exceto a interpretação em termos de curvas elípticas, que é discutida em mais detalhe no livro [Hus87] de Husemoller. Os livros mais antigos possuem mais detalhes; veja, por exemplo, os domínios fundamentais desenhados no livro [Fri22] de Fricke.

Capítulo 2

Formas Modulares

Neste capítulo, introduzimos o conceito de forma modular e estudamos suas propriedades fundamentais. Procuramos dar alguma ênfase às várias interpretações possíveis da definição, especialmente as que se têm provado úteis em geometria algébrica aritmética: a interpretação "modular" (ligada a curvas elípticas, e discutida neste capítulo) e a interpretação geométrica (ligada às "curvas modulares", que são essencialmente os espaços de órbitas introduzidos no primeiro capítulo). Esperamos que esta variedade de enfoques possa facilitar o acesso do leitor à literatura. Concluímos dando os primeiros exemplos.

2.1 Definição

Começamos com a definição clássica.

Definição 5 *Sejam $G \subset \mathrm{SL}_2(\mathbf{Z})$ um subgrupo de congruência, k um inteiro positivo, e $f : \mathcal{H} \longrightarrow \mathbf{C}$ uma função meromorfa. Dizemos que f é uma função quase-modular de peso k para o grupo G se f satisfaz a equação funcional*

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k f(\tau) \tag{2.1}$$

para todo $\tau \in \mathcal{H}$ e toda matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$.

Vale a pena fazer uma série de observações sobre esta definição:

OBSERVAÇÕES:

1) Note, em primeiro lugar, que a equação 2.1 é compatível com a lei de grupo em G. Isto implica, por exemplo, que basta verificá-la para um conjunto de geradores do grupo G.

2) Se $k = 0$, a definição se reduz a dizer que f é invariante sob a ação de G. Este é o caso mais simples, mas, como veremos adiante, funções quase-modulares de peso 0 nunca podem ser holomorfas (a não ser que sejam constantes). É este fato que torna interessante considerar o caso mais geral.

3) Suponha que

$$\gamma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in G$$

(por exemplo, isto vale se $G = \Gamma$ ou $G = \Gamma_0(N)$ para qualquer inteiro positivo N). Então, como $\gamma \cdot \tau = \tau$ para todo $\tau \in \mathcal{H}$, a condição 2.1 diz que

$$f(\tau) = (-1)^k f(\tau);$$

se k for ímpar, isto força $f(\tau) = 0$ para todo τ. Assim, para $G = \Gamma$ ou $G = \Gamma_0(N)$, a teoria só não é vazia quando o peso é par.

4) Se $G = \Gamma$, basta, como observamos acima, verificar a equação funcional para os geradores

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{e} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

isto é, basta verificar que

$$f(-1/\tau) = \tau^k f(\tau)$$
$$f(\tau + 1) = f(\tau),$$

para todo $\tau \in \mathcal{H}$. Além disso, já sabemos que basta considerar o caso em que k é um inteiro par.

O fato de que a equação 2.1 é compatível com a lei de grupo sugere que talvez seja preferível reescrever a definição de modo a enfatizar o fato. Uma forma de fazer isto é

definir, para cada k, uma ação do grupo G tal que f seja quase-modular de peso k se e só se for invariante sob essa ação. Isto não é difícil de fazer, e prova ser extremamente útil. Vamos até mesmo estender a definição para grupos mais gerais.

Definição 6 *Sejam k um inteiro, $f : \mathcal{H} \longrightarrow \mathbf{C}$ uma função meromorfa e*

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbf{R}).$$

Definimos uma outra função $f|[\gamma]_k$ da seguinte forma:

$$(f|[\gamma]_k)(\tau) = \det(\gamma)^{k/2}(c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right),$$

para $\tau \in \mathcal{H}$.

É imediato verificar o seguinte:

Proposição 2.1.1 *Seja k um inteiro positivo e G um subgrupo de $\mathrm{GL}_2^+(\mathbf{R})$. Então*

$$f \mapsto f|[\gamma]_k$$

define uma ação de G no conjunto das funções meromorfas de $\mathcal{H}$ em $\mathbf{C}$. Além disso, se G for um subgrupo de congruência de $\mathrm{SL}_2(\mathbf{Z})$, temos que f é uma função quase-modular de peso k para G se e só se

$$f|[\gamma]_k = f$$

para todo $\gamma \in G$, isto é, se e só se f é invariante sob esta ação.

DEMONSTRAÇÃO: Fica como exercício.□

A extensão da ação "$|[\gamma]_k$" a $\gamma \in \mathrm{GL}_2^+(\mathbf{R})$ não é de todo arbitrária. Ficará claro adiante, por exemplo, que o caso $\gamma \in \mathrm{GL}_2^+(\mathbf{Q})$ é de grande importância.

No caso em que $G = \Gamma$ é o grupo modular todo, já observamos acima que toda função modular é necessariamente periódica: temos que ter $f(\tau + 1) = f(\tau)$. Com um pouco de

cuidado, vê-se que é assim no caso geral. Seja $G \subset \Gamma$ um subgrupo de congruência, e seja n a largura da ponta no infinito, de modo que

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in G.$$

Se f é quase-modular de peso k para G, segue, considerando a ação desta matriz, que

$$f(\tau + n) = f(\tau),$$

para todo $\tau \in \mathcal{H}$. Assim, podemos expandir f em série de Fourier. Se pusermos $q = e^{2\pi i \tau}$ e $q_n = q^{1/n} = e^{2\pi i \tau / n}$, podemos escrever

$$f(\tau) = \sum_{m=-\infty}^{\infty} a_m q_n^m.$$

Esta expressão se chama a *q-expansão de f na ponta no infinito.* Note que se $G = \Gamma_0(\mathrm{N})$ ou $\Gamma_1(\mathrm{N})$, temos $n = 1$, de modo que a q-expansão no infinito é uma série de potências em q.

Uma forma de interpretar a q-expansão no infinito é a seguinte: a transformação $\tau \mapsto q_n = e^{2\pi i \tau / n}$ manda o semi-plano superior $\mathcal{H}$ sobre o disco unitário perfurado

$$\{q_n \in \mathbf{C} : 0 < |q_n| < 1\}.$$

Deste ponto de vista, a q-expansão é a expansão de Laurent de f em torno de $q_n = 0$. Notando que $q_n \to 0$ exatamente quando $\Im(z) \to \infty$, vê-se que a q-expansão no infinito reflete o comportamento de f perto da ponta no infinito.

É comum, no estudo de funções quase-modulares (especialmente formas modulares) confundir a função f com sua q-expansão no infinito, isto é, escrever

$$f = \sum a_m q_n^m.$$

Isto raramente leva a confusões, mas deve ser entendido como um abuso de linguagem.

Definição 7 *Seja $f : \mathcal{H} \longrightarrow \mathbf{C}$ uma função quase-modular de peso k para um subgrupo de congruência $G \subset \Gamma$, e seja*

$$f(\tau) = \sum_{m=-\infty}^{\infty} a_m q_n^m$$

sua q-expansão no infinito. Dizemos que

i) f é meromorfa no infinito se existe M tal que $a_m = 0$ para $m < M$;

ii) f é holomorfa no infinito se $a_m = 0$ para $m < 0$;

iii) f se anula no infinito se $a_m = 0$ para $m \leq 0$.

As outras pontas também precisam ser consideradas. Para isso, seja $\tau_0 = \gamma \cdot \infty$ (com $\gamma \in \Gamma$) uma ponta para a ação de G. Para estudar $f(\tau)$ perto de τ_0, estudamos essencialmente $f(\gamma \cdot \tau)$ perto de ∞. O único cuidado a tomar é que queremos que se τ_0 e ∞ forem a mesma ponta (isto é, se $\gamma \in G$) se obtenha a mesma q-expansão. Para garantir isso, usamos a ação $|[\gamma]_k$ como acima:

Definição 8 *Seja $\tau_0 = \gamma \cdot \infty$ uma ponta para a ação de G, de largura n, e seja f uma função quase-modular de peso k para G. A q-expansão de f na ponta τ_0 é a q-expansão de $f|[\gamma]_k$ na ponta no infinito.*

Dizemos que f é meromorfa, holomorfa ou zero em τ_0 conforme $f|[\gamma]_k$ for meromorfa, holomorfa ou zero no infinito.

Note, por exemplo, o que esta definição diz se $\gamma = S$, de modo que $\gamma \cdot \tau = -1/\tau$: nesse caso, $(f|[\gamma]_k)(\tau) = \tau^{-k} f(-1/\tau)$, de modo que ser holomorfa (ou zero) em $\tau = 0$ *não* é o mesmo que ter um limite quando $\tau \to 0$. Por outro lado, a definição garante que uma função quase-modular para o grupo modular completo tem a mesma q-expansão no zero e no infinito, o que é agradável, já que os dois valores representam a mesma ponta para Γ.

Com estas definições, estamos em condições de definir os objetos básicos do nosso estudo:

Definição 9 *Uma função modular de peso k para um subgrupo de congruência $G \subset \Gamma$ é uma função quase-modular que é meromorfa em todas as pontas para a ação de G. Denotamos o conjunto das funções modulares de peso k para G por $F_k(G, \mathbf{C})$ ou $F_k(G)$.*

Uma forma modular de peso k para G é uma função quase-modular que é holomorfa *em $\mathcal{H}$ e em todas as pontas para a ação de G. Denotamos o conjunto das formas modulares de peso k para G por $M_k(G, \mathbf{C})$ ou $M_k(G)$.*

Uma forma modular se diz parabólica *(em inglês, "cuspform"; em alemão, "spitzenform") se ela se anula em todas as pontas para a ação de G. Denotamos o conjunto das formas modulares parabólicas de peso k para G por $S_k(G, \mathbf{C})$ ou $S_k(G)$.*

É fácil ver que $M_k(G)$ e $S_k(G)$ são espaços vetoriais sobre $\mathbf{C}$; mais adiante, vamos ver que são de dimensão finita.

Feita a definição, queremos entender suas propriedades e construir alguns exemplos. Dedicaremos a maior parte deste capítulo a estudar a definição sob várias óticas. Como resultado, obteremos desde já alguns exemplos interessantes; a maior parte do trabalho de obter exemplos, entretanto, será deixada para o capítulo 4.

Começamos notando que se f_1 é uma forma modular de peso k_1 e f_2 é uma forma modular de peso k_2, então seu produto $f = f_1 f_2$ é uma forma modular de peso $k_1 + k_2$. Isto é,

$$\mathcal{M}(G) = \bigoplus_{k \in \mathbf{Z}} M_k(G)$$

é uma $\mathbf{C}$-álgebra graduada, que contém

$$\mathcal{S}(G) = \bigoplus_{k \in \mathbf{Z}} S_k(G)$$

como um ideal graduado. (Verifique!)

Há relações, é claro, entre os espaços de formas modulares correspondentes aos vários subgrupos de congruência. Por exemplo, temos

$$M_k(\Gamma) \subset M_k(\Gamma_0(\mathrm{N})) \subset M_k(\Gamma_1(\mathrm{N})).$$

No caso da passagem de $\Gamma_0(\mathrm{N})$ para $\Gamma_1(\mathrm{N})$, podemos até dizer um pouco mais, porque $\Gamma_1(\mathrm{N})$ é um subgrupo normal de $\Gamma_0(\mathrm{N})$(veja o exercício 3 do capítulo um).

Proposição 2.1.2 *Sejam $G_1 \supset G_2$ subgrupos de congruência, e suponhamos que $G_2 \lhd G_1$. Se $f \in M_k(G_2)$ e $\gamma \in G_1$, então $f|[\gamma]_k \in M_k(G_2)$, de modo que $f \mapsto f|[\gamma]_k$ define uma ação do quociente G_1/G_2 em $M_k(G_2)$.*

DEMONSTRAÇÃO: Só é preciso verificar que se $\gamma \in G_1$ e $\delta \in G_2$, temos $(f|[\gamma]_k)|[\delta]_k = f|[\gamma]_k$. Mas

$$(f|[\gamma]_k)|[\delta]_k = f|[\gamma\delta]_k = f|[\gamma\delta\gamma^{-1}\gamma]_k = (f|[\gamma\delta\gamma^{-1}]_k)|[\gamma]_k = f|[\gamma]_k,$$

já que f é invariante sob a ação de $\gamma\delta\gamma^{-1} \in G_2$. □

No caso de $\Gamma_0(N) \supset \Gamma_1(N)$, temos

$$\Gamma_0(N)/\Gamma_1(N) \cong \left(\mathbf{Z}/N\mathbf{Z}\right)^\times,$$

via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d.$$

Para cada $x \in (\mathbf{Z}/N\mathbf{Z})^\times$, escolhemos um representante da co-classe que corresponde a x; para isso, podemos tomar $\gamma(x) \in \Gamma_0(N)$ satisfazendo

$$\gamma(x) \equiv \begin{pmatrix} x^{-1} & 0 \\ 0 & x \end{pmatrix} \pmod{N}.$$

(É fácil ver que existem estes elementos.) Então definimos:

Definição 10 *Para cada $x \in (\mathbf{Z}/N\mathbf{Z})^\times$ e cada $f \in M_k(\Gamma_1(N))$, definimos*

$$\langle x \rangle f = f|[\gamma(x)]_k,$$

onde $\gamma(x)$ é escolhido como acima.

Isto dá uma ação de $(\mathbf{Z}/N\mathbf{Z})^\times$ em $M_k(\Gamma_1(N))$ (que é a ação induzida pela de $\Gamma_0(N)$). Podemos então decompor o espaço $M_k(\Gamma_1(N))$ segundo os caracteres de $(\mathbf{Z}/N\mathbf{Z})^\times$:

Definição 11 *Seja $f \in M_k(\Gamma_1(N))$; dizemos que f é uma forma modular de caráter (ou "nebentypus") $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \mathbf{C}$ se temos $\langle x \rangle f = \varepsilon(x) f$.*

Em particular, se $\varepsilon = 1$ é o caráter trivial, temos que f tem caráter ε se e só se $f \in M_k(\Gamma_0(N))$.

Proposição 2.1.3 *Uma forma modular $f \in M_k(\Gamma_1(N))$ é de caráter ε se e só se, para todo $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, temos*

$$f(\gamma \cdot \tau) = \varepsilon(d)(c\tau + d)^k f(z).$$

Se denotarmos o conjunto das formas de caráter ε por $M_k(\Gamma_1(N), \varepsilon) = M_k(N, \varepsilon)$, temos

$$M_k(\Gamma_1(N)) = \bigoplus M_k(\Gamma_1(N), \varepsilon),$$

onde a soma é sobre todos os caracteres de Dirichlet de nível N, isto é, sobre os caracteres de $(\mathbf{Z}/N\mathbf{Z})^\times$.

DEMONSTRAÇÃO: A equação é uma tradução direta da definição. A decomposição nos espaços correspondentes aos caracteres é um resultado geral sobre ações de grupos em espaços vetoriais—veja o exercício 3 adiante. □

2.2 Reinterpretando a definição

Começamos reinterpretando a definição em termos de redes $\Lambda \in \mathbb{C}$; o ponto principal é que temos uma bijeção entre

- redes $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2 \subset \mathbb{C}$ com $\Im(w_1/w_2) > 0$, módulo homotetia, e
- elementos $\tau \in \mathcal{H}$, módulo a ação do grupo modular Γ.

A correspondência era dada por $\tau \mapsto \Lambda(\tau) = \mathbb{Z}\tau + \mathbb{Z}$. Seja $\mathcal{R}$ o conjunto de todas as redes $\Lambda \subset \mathbb{C}$.

Definição 12 *Uma função $F : \mathcal{R} \longrightarrow \mathbb{C}$ se diz de peso k se satisfaz*

$$F(\lambda\Lambda) = \lambda^{-k}F(\Lambda),$$

para toda rede Λ e todo $\lambda \in \mathbb{C}^\times$.

Também podemos pensar em termos de bases para a rede $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2$. Seja $\mathcal{B}$ o conjunto dos pares (w_1, w_2) satisfazendo $\Im(w_1/w_2) > 0$. Podemos definir uma função $\mathcal{F} : \mathcal{B} \longrightarrow \mathbb{C}$ pondo $\mathcal{F}(w_1, w_2) = F(\mathbb{Z}w_1 + \mathbb{Z}w_2)$. Dizer que F é de peso k equivale a dizer que:

- $\mathcal{F}(w_1, w_2)$ é invariante sob a ação natural de $\mathrm{SL}_2(\mathbb{Z})$ em $\mathcal{B}$, e
- $\mathcal{F}(\lambda w_1, \lambda w_2) = \lambda^{-k}\mathcal{F}(w_1, w_2)$, isto é, $\mathcal{F}$ é homogênea de grau $-k$.

Disto, fica claro que $w_2^k\mathcal{F}(w_1, w_2)$ só depende de $\tau = w_1/w_2$; pomos

$$f(\tau) = w_2^k\mathcal{F}(w_1, w_2) = \mathcal{F}(\tau, 1),$$

o que dá

$$\mathcal{F}(w_1, w_2) = w_2^{-k}f(w_1/w_2).$$

Finalmente, $\mathcal{F}$ é $\mathrm{SL}_2(\mathbf{Z})$-invariante se e só se

$$f(\gamma \cdot \tau) = (c\tau + d)^k f(\tau),$$

para todo $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$. Isto mostra:

Proposição 2.2.1 *Existem bijeções naturais entre os seguintes conjuntos:*

i) funções $F : \mathcal{R} \longrightarrow \mathbf{C}$ de peso k,

ii) funções $\mathcal{F} : \mathcal{B} \longrightarrow \mathbf{C}$ invariantes sob $\mathrm{SL}_2(\mathbf{Z})$ e homogêneas de grau $-k$,

iii) funções $f : \mathcal{H} \longrightarrow \mathbf{C}$ que se transformam sob Γ como as funções quase-modulares.

Resta, para completar a reinterpretação, fazer duas coisas: definir estruturas de variedade complexa nos conjuntos $\mathcal{R}$ e $\mathcal{B}$, para podermos falar em holomorficidade, e estender a proposição aos subgrupos de congruência. Vamos deixar estas tarefas ao leitor interessado, especialmente a segunda, cujo ponto de partida é a discussão no final do capítulo um.

2.3 Séries de Eisenstein

Nesta seção, vamos construir os primeiros exemplos não-triviais de formas modulares para o grupo modular Γ. Para isso, usamos a interpretação em termos de redes e tentamos a idéia mais simples possível: para construir uma função de redes que é de peso k, considere uma série do tipo

$$G_k(\Lambda) = \sum_{\substack{w \in \Lambda \\ w \neq 0}} \frac{1}{w^k}.$$

Se esta série for sempre absolutamente convergente, é imediato que ela definirá uma função de redes de peso k. A convergência é fácil:

Lema 2.3.1 *Seja $\Lambda \subset \mathbf{C}$ uma rede, e seja $s \in \mathbf{R}$. A série*

$$\sum_{\substack{w \in \Lambda \\ w \neq 0}} \frac{1}{|w|^s}$$

é convergente desde que $s > 2$.

DEMONSTRAÇÃO: Trata-se de um exercício simples, que deixamos ao leitor. □

Para simplificar a notação, escrevemos $\sum'$ para denotar a soma sobre todos os elementos da rede exceto $w = 0$. Então podemos definir:

Definição 13 *Seja $k \geq 4$ um inteiro par. A série de Eisenstein de peso k é qualquer uma das seguintes funções, que se correspondem pela bijeção descrita acima:*

- $G_k : \mathcal{R} \longrightarrow \mathbf{C}$ *dada por*
$$G_k(\Lambda) = \sum{}' \frac{1}{w^k},$$
- $G_k : \mathcal{B} \longrightarrow \mathbf{C}$ *dada por*
$$G_k(w_1, w_2) = \sum_{\substack{n,m \in \mathbf{Z} \\ (n,m) \neq (0,0)}} \frac{1}{(nw_1 + mw_2)^k},$$
- $G_k : \mathcal{H} \longrightarrow \mathbf{C}$ *dada por*
$$G_k(\tau) = \sum_{\substack{n,m \in \mathbf{Z} \\ (n,m) \neq (0,0)}} \frac{1}{(n\tau + m)^k}.$$

Proposição 2.3.2 *Seja $k \geq 4$ um inteiro par. A série de Eisenstein $G_k(\tau)$ é uma forma modular de peso k para Γ, cuja q-expansão no infinito é da forma*

$$G_k(\tau) = 2\zeta(k) + \sum a_n q^n,$$

onde ζ denota a função zeta de Riemann.

DEMONSTRAÇÃO: Dado o lema e o fato de que G_k é visivelmente uma função de redes de peso k, basta verificar que G_k é holomorfa em $\mathcal{H}$ e na ponta no infinito.

Para verificar a holomoficidade em $\mathcal{H}$, considere primeiro o caso $\tau \in \mathsf{D}$, onde D é o domínio fundamental usual. Neste caso, como temos certamente que $|m\tau + n| \geq |m\rho + n|$ (com $\rho = e^{2\pi i/3}$ como acima), temos

$$\frac{1}{|m\tau + n|^k} \leq \frac{1}{|m\rho + n|^k};$$

como

$$\sum{}' \frac{1}{|m\rho + n|^k}$$

converge (pelo Lema), segue que $G_k(\tau)$ converge uniforme e absolutamente em D, e portanto define uma função holomorfa em D. Se, agora, $\tau \in \gamma\mathsf{D}$ para algum γ tal que $\gamma^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, temos

$$G_k(\tau) = (c\tau + d)^{-k} G_k(\gamma^{-1}\tau),$$

com $\gamma^{-1}\tau \in \mathsf{D}$. Logo, G_k é holomorfa em $\gamma\mathsf{D}$, e portanto em $\mathcal{H}$.

Finalmente, como a única ponta é ∞, basta verificar que existe o limite

$$\lim_{\Im(\tau)\to\infty} G_k(\tau).$$

Podemos supor $\tau \in \mathsf{D}$, e como a convergência é uniforme em D, podemos calcular o limite termo-a-termo:

$$\bullet \text{se } m \neq 0, \quad \lim_{\tau\to i\infty} \frac{1}{(m\tau+n)^k} = 0$$

$$\bullet \text{se } m = 0, \quad \frac{1}{(m\tau+n)^k} = \frac{1}{n^k}.$$

Logo,

$$\lim_{\tau\to i\infty} G_k(\tau) = \sum_{\substack{n\in\mathbb{Z} \\ n\neq 0}} \frac{1}{n^k} = 2\sum_{n\geq 1} \frac{1}{n^k} = 2\zeta(k).$$

Isto prova a proposição; note que a hipótese de que k é par foi essencial para que a soma não fosse zero. □

OBSERVAÇÕES:

1) Se fizermos $g_4(\tau) = 60G_4(\tau)$ e $g_6(\tau) = 140G_6(\tau)$, teremos

$$g_4 = 120\zeta(4) + \sum a_n q^n = \frac{4}{3}\pi^4 + \cdots$$

e também

$$g_6 = 280\zeta(6) + \sum b_n q^n = \frac{8}{27}\pi^6 + \cdots$$

Sejam

$$\mathrm{E}_4(\tau) = \frac{3}{4\pi^4} g_4(\tau) = \frac{1}{2\zeta(4)} G_4(\tau) \quad \text{e} \quad \mathrm{E}_6(\tau) = \frac{27}{8\pi^6} g_6 = \frac{1}{2\zeta(6)} G_6(\tau),$$

de modo que o "termo independente" da q-expansão de E_4 e E_6 é 1. (Introduzir g_4 e g_6 é talvez uma concessão à história; do ponto de vista aritmético, queremos coeficientes racionais, donde o interesse de E_4 e E_6.) Considere agora

$$\Delta(\tau) = (2\pi)^{-12}[g_4^3(\tau) - 27g_6^2(\tau)] = \frac{1}{1728}(E_4^3 - E_6^2).$$

Então uma conta simples mostra que Δ é uma forma modular *parabólica* de peso 12 para Γ (já que ∞ é a única ponta).

2) A racionalidade dos coeficientes da q-expansão será uma questão importante para as aplicações aritméticas. Mais adiante vamos provar que as q-expansões de

$$E_k = \frac{1}{2\zeta(k)}G_k$$

e de Δ têm coeficientes em $\mathbf{Q}$ (no caso de E_4, E_6 e Δ, os coeficientes são mesmo *inteiros*).

3) As formas modulares E_4, E_6 e Δ têm interpretações importantes em termos de curvas elípticas: dado $\tau \in \mathcal{H}$, considere a curva elíptica

$$E(\tau) = \mathbf{C}/\mathbf{Z}(2\pi i\tau) + \mathbf{Z}(2\pi i).$$

(Esta curva é *isomorfa* à E_τ do capítulo anterior, mas a homotetia por $2\pi i$ ajuda na normalização aritmética.) Então:

Fato 2.3.3 *Pode-se escolher coordenadas x e y em $E(\tau)$ de modo que sua equação seja*

$$y^2 = 4x^3 + a(\tau)x + b(\tau),$$

com

$$a(\tau) = -\frac{1}{12}E_4(\tau) \quad e \quad b(\tau) = \frac{1}{216}E_6(\tau).$$

Além disso, o discriminante da curva $E(\tau)$ é $\Delta(\tau)$; em particular, $\Delta(\tau) \neq 0$.

O fato de que $\Delta(\tau) \neq 0$ para $\tau \in \mathcal{H}$ é bastante importante; mais adiante, daremos uma demonstração analítica. Como já vimos, Δ se anula na ponta.

Na próxima seção, exploramos um pouco mais a relação entre formas modulares e curvas elípticas.

2.4 A interpretação "modular"

Já vimos, no primeiro capítulo, que há uma correspondência entre números complexos $\tau \in \mathcal{H}$ e pares (E, ω), onde E é uma curva elíptica e ω é uma diferencial holomorfa invariante. Nesta seção vamos usar a correspondência

$$\tau \in \mathcal{H} \mapsto (E(\tau), \omega(\tau)),$$

dada por

$$E(\tau) = \mathbf{C}/\mathbf{Z}(2\pi i\tau) + \mathbf{Z}(2\pi i),$$

e $\omega(\tau) = dz \bmod 2\pi i(\mathbf{Z}\tau + \mathbf{Z})$. (Como no final da seção anterior, a passagem da E_τ do capítulo anterior para $E(\tau)$ é feita para evitar fatores de $2\pi i$ por toda a parte!)

Lembremos, ainda, que se $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, as curvas correspondentes a τ e a $\gamma \cdot \tau$ são isomorfas, mas as diferenciais diferem por uma constante; de fato,

$$(E(\gamma\tau), \omega(\gamma\tau)) \cong (E(\tau), \lambda\omega(\tau)),$$

onde $\lambda = (c\tau + d)^{-1}$, como já vimos.

Finalmente, lembremos que a passagem do grupo modular Γ para os subgrupos de congruência $\Gamma_0(N)$ ou $\Gamma_1(N)$ equivale, em termos de curvas elípticas, a adicionar mais estrutura: um subgrupo cíclico de ordem N ou um ponto de ordem N. Nesta seção, limitar-nos-emos so caso de $\Gamma_1(N)$.

Definição 14 *Seja* E *uma curva elíptica complexa. Uma estrutura de nível* N *em* E *é uma escolha de um ponto* $P \in E$ *de ordem exatamente igual a* N. *Um objeto-teste de nível* N *é uma tripla* (E, ω, P), *onde* E *é uma curva elíptica,* ω *é uma diferencial holomorfa invariante, e* P *é um ponto de ordem exatamente igual a* N.

Isto tudo sugere que podemos reinterpretar formas modulares como funções de classes de isomorfismo de curvas elípticas munidas de uma estrutura de nível N.

Proposição 2.4.1 *Uma função modular* f *de peso* k *para* $\Gamma_1(N)$ *define uma função* f *dos objetos-teste de nível* N *a valores em* C, *satisfazendo as seguintes condições:*

i) $f(E, \omega, P) \in \mathbf{C}$ *depende só da classe de isomorfismo do objeto-teste* (E, ω, P),

ii) $f(\mathrm{E}, \lambda\omega, P) = \lambda^{-k} f(\mathrm{E}, \omega, P)$, *para todo* $\lambda \in \mathbf{C}^\times$.

DEMONSTRAÇÃO: Dada toda a discussão acima, isto é imediato—basta juntar todas as peças. □

Dada a Proposição, seria interessante *definir* funções modulares nestes termos; o problema é conseguir uma boa noção de holomorficidade. Isto se faz lançando mão de geometria algébrica mais sofisticada para complicar a situação em duas direções. Primeiro, para conseguir garantir que f seja meromorfa, exigimos que seja possível calcular f em objetos-teste *definidos sobre qualquer* $\mathbf{C}$*-álgebra* A, de modo compatível com extensões de escalares. Segundo, para garantir holomorficidade, consideramos o valor de f numa curva especial: a "curva de Tate", que é uma curva elíptica sobre o anel de séries de Laurent $\mathbf{C}((q))$. Uma vez feita essa interpretação, temos uma definição completamente *algébrica*, que faz sentido, por exemplo, sobre qualquer corpo K, o que daria uma teoria de "formas modulares definidas sobre K". Esta versão da teoria é bastante importante, mas requer (como já deve estar claro!) um conhecimento de geometria algébrica mais profundo do que estamos pressupondo nestas notas; assim, referimos o leitor interessado à discussão em [Kat73] e [DR73].

Como exemplo da utilidade da "interpretação modular", vamos reinterpretar a ação de $(\mathbf{Z}/\mathrm{N}\mathbf{Z})^\times$ em $M_k(\Gamma_1(\mathrm{N}))$ em termos modulares. Note, primeiro, que se $x \in (\mathbf{Z}/\mathrm{N}\mathbf{Z})^\times$ e $P \in \mathrm{E}$ é um ponto de ordem exatamente N, então xP é um outro ponto de ordem N (bem-definido porque P é de ordem N).

Proposição 2.4.2 *Sejam* $f \in M_k(\Gamma_1(\mathrm{N}))$ *e* $x \in (\mathbf{Z}/\mathrm{N}\mathbf{Z})^\times$*; como função de objetos-teste de nível* N*, a forma modular* $\langle x \rangle f \in M_k(\Gamma_1(\mathrm{N}))$ *é dada por*

$$(\langle x \rangle f)(\mathrm{E}, \omega, P) = f(\mathrm{E}, \omega, xP).$$

DEMONSTRAÇÃO: Basta notar que se

$$\gamma(x) \equiv \begin{pmatrix} x^{-1} & 0 \\ 0 & x \end{pmatrix} \pmod{\mathrm{N}};$$

$\gamma(x)$ leva (E,ω,P) em $(\mathrm{E},\lambda^{-1}\omega,xP)$, onde, como sempre, $\mathrm{E}=\mathrm{E}(\tau)$, $\gamma(x)=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e $\lambda=(c\tau+d)$. (Verifique!!) Então

$$\begin{aligned}(\langle x\rangle f)(\tau) &= \lambda^{-k} f(\gamma(x)\cdot\tau)\\ &= \lambda^{-k} f(\mathrm{E},\lambda^{-1}\omega,xP)\\ &= \lambda^{-k}\lambda^{k} f(\mathrm{E},\omega,xP)\\ &= f(\mathrm{E},\omega,xP),\end{aligned}$$

como queríamos. □

Com isto, podemos dar uma interpretação algébrica do caráter de uma forma modular para $\Gamma_1(\mathrm{N})$. A decomposição

$$M_k(\Gamma_1(\mathrm{N}))=\bigoplus M_k(\Gamma_1(\mathrm{N}),\varepsilon)$$

se estende, desta forma, a qualquer corpo em que $\phi(\mathrm{N})$ (que é a ordem de $(\mathbf{Z}/\mathrm{N}\mathbf{Z})^{\times}$) for inversível.

No que segue, usaremos a interpretação modular apenas em alguns contextos, mas a teremos sempre como pano de fundo, já que ela tem um papel importante em muitas das aplicações. Para concluir esta seção, vamos tornar explícita a interpretação de E_4, E_6 e Δ em termos de curvas elípticas. O leitor que não tiver grande interesse neste aspecto da teoria deve-se sentir livre para omitir esta parte.

Seja E uma curva elíptica complexa, e ω uma diferencial holomorfa invariante em E. Já observamos acima que existem sempre coordenadas x e y em E de tal forma que a equação de E seja da forma

$$y^2=4x^3+ax+b.$$

A diferencial $dx/2y=dy/(12x^2+a)$ é holomorfa e invariante; se exigirmos que $\omega=dx/2y$, a equação fica univocamente determinada (isto é, a e b ficam determinados). Desta forma, podemos pensar em x, y, a, b e o discriminante Δ como funções de (E,ω). É fácil, então,

verificar que

$$x(E, \lambda\omega) = \lambda^2 x(E, \omega)$$

$$y(E, \lambda\omega) = \lambda^3 y(E, \omega)$$

$$a(E, \lambda\omega) = \lambda^{-4} a(E, \omega)$$

$$b(E, \lambda\omega) = \lambda^{-6} b(E, \omega)$$

$$\Delta(E, \lambda\omega) = \lambda^{-12} \Delta(E, \omega).$$

É possível verificar, então, que a, b e Δ são holomorfas em $\mathcal{H}$ e na ponta (x e y não são holomorfas na ponta), e portanto que são formas modulares; de fato, já observamos que

$$a(\tau) = -\frac{1}{12} E_4 \quad \text{e} \quad b(\tau) = -\frac{1}{216} E_6,$$

e que Δ é a forma modular parabólica definida acima. Isto dá uma definição completamente algébrica destas formas modulares (e implica, por exemplo, a racionalidade dos coeficientes da q-expansão—veja [Kat73]).

2.5 Exercícios

1) Prove a Proposição 2.2.1.

2) Seja $x \in \mathbf{Z}$ com $\mathrm{mdc}(x, N) = 1$, e seja

$$\gamma(x) = \begin{pmatrix} x^{-1} & 0 \\ 0 & x \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Q}).$$

Mostre que $f|[\gamma(x)]_k = \langle x \rangle f$. Esta é a definição mais simples da ação de $(\mathbf{Z}/N\mathbf{Z})^\times$ em $M_k(\Gamma_1(N))$.

3) Seja k um corpo qualquer e V um espaço vetorial de dimensão finita sobre k. Seja G um grupo abeliano cuja ordem é inversível em k, e suponha que exista uma representação $G \longrightarrow \mathrm{End}(V)$. Mostre que V se decompõe na soma direta dos espaços em que G age por cada um dos seus caracteres.

4) Verifique as asserções que provam a Proposição 2.2.1.

5) Seja $\Lambda \subset \mathbb{C}$ uma rede. Um elemento $w \in \mathbb{C}$ se diz de ordem N em relação a Λ se $Nw \in \Lambda$ e $nw \notin \Lambda$ para $n < N$. Mostre que há uma bijeção entre:

- pares (Λ, w) onde Λ é uma rede e w é um ponto de ordem N em relação a Λ, módulo homotetias, e
- elementos $\tau \in \mathcal{H}$, módulo a ação de $\Gamma_1(N)$.

6) Enuncie e prove uma versão para $\Gamma_1(N)$ da Proposição 2.2.1.

7) Prove o Lema 2.3.1. (Sugestão: estime o número de pontos da rede que estão entre o círculo de raio n e o de raio $n+1$.)

8) Mostre que a série de Eisenstein normalizada E_k pode ser definida diretamente pondo

$$E_k(\tau) = \frac{1}{2} \sum_{\substack{m,n \in \mathbb{Z} \\ \text{mdc}(m,n)=1}} \frac{1}{(m\tau + n)^k}.$$

9) Dê uma interpretação modular para formas modulares para $\Gamma_0(N)$.

2.6 Notas

Neste capítulo, seguimos quase sempre o tratamento magistral de Serre em [Ser73], exceto em relação à interpretação modular, que é tratada, em termos muito mais gerais do que aqui, no apêndice e no primeiro capítulo de [Kat73], onde Katz realiza completamente o programa esboçado acima de algebraizar a teoria. O tratamento de Katz depende fundamentalmente da existência da curva modular (como esquema de módulos fino); isto é demonstrado no caso mais geral em [KM85], mas o tratamento de [DR73] também merece ser consultado.

Capítulo 3

O Espaço de Órbitas como Superfície de Riemann

Uma das idéias mais importantes na teoria de formas modulares é considerar o quociente

$$G\backslash\mathcal{H}^* = G\backslash\mathcal{H} \cup \{\text{pontas}\}$$

como superfície de Riemann. Isto pode ser feito para qualquer subgrupo discreto $G \subset \Gamma$, e boa parte da teoria pode ser deduzida da geometria do objeto resultante.

Nesta seção, trataremos o caso $G = \Gamma = \mathrm{SL}_2(\mathbf{Z})$, para obter informações sobre os espaços de formas modulares para Γ. Por exemplo, vamos poder computar a dimensão dos espaços $M_k(\Gamma)$. O resultado fundamental é a fórmula 3.1, que no nosso enfoque é essencialmente o teorema de Riemann-Roch. Esta fórmula pode, é claro, ser obtida de uma forma mais elementar (veja os exercícios); o leitor que assim desejar poderá mesmo admiti-la, e assim evitar a teoria de superfícies de Riemann. Parece-nos instrutivo, entretanto, usar neste caso mais simples as técnicas geométricas que são necessárias no caso mais difícil dos subgrupos de congruência.

Nosso tratamento segue o de Serre em [Bo66]. Um tratamento, no mesmo espírito, do caso dos subgrupos de congruência pode ser encontrado no livro [Shi71] de Shimura. Durante todo este capítulo, nos restringimos ao caso do grupo $\Gamma = \mathrm{SL}_2(\mathbf{Z})$.

3.1 A estrutura de superfície de Riemann

Sejam $\Gamma = \mathrm{SL}_2(\mathbf{Z})$ agindo em $\mathcal{H}$ e seja $X = \Gamma\backslash\mathcal{H}$ o espaço de órbitas. Seja

$$\mathsf{D} = \{\tau \in \mathcal{H} : -\frac{1}{2} \leq \Re(\tau) \leq \frac{1}{2}, |\tau| \geq 1\}$$

o domínio fundamental usual para a ação de Γ. Começamos estudando a topologia de X.

Proposição 3.1.1 *Seja $K \subset \mathcal{H}$ um compacto.*

i) Existe M tal que $\Im(\gamma \cdot \tau) < M$ para todo $\gamma \in \Gamma$ e todo $\tau \in K$.

ii) Para todo outro compacto $K' \subset \mathcal{H}$, o conjunto dos $\gamma \in \Gamma$ tais que $\gamma \cdot K \cap K' \neq \emptyset$ é finito.

DEMONSTRAÇÃO: Como $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, temos a relação 1.2:

$$\Im\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{(ad - bc)\Im(\tau)}{|c\tau + d|^2};$$

como também

$$\inf\{|c\tau + d| : \tau \in K,\ c, d \in \mathbf{Z},\ \mathrm{mdc}(c, d) = 1\} > 0,$$

a primeira afirmação segue imediatamente.

Para provar a segunda afirmação, note que a fórmula 1.2 implica de imediato que o número de pares (c, d) que aparecem como a linha inferior de matrizes γ tais que $\gamma \cdot K \cap K' \neq \emptyset$ é finito. Mas γ e δ têm a mesma linha inferior se e só se $\gamma = T^n \delta$ para algum $n \in \mathbf{Z}$. Como $T^n\delta \cdot K \to \infty$ quando $n \to \pm\infty$, segue que temos $T^n\delta \cdot K \cap K' \neq \emptyset$ só para um número finito de n, e a proposição segue. □

Usa-se freqüentemente a expressão "$\gamma \cdot K \to \infty$" para exprimir o conteúdo da segunda afirmação desta proposição.

Podemos agora obter uma boa descrição da topologia de X: fazemos isto com uma seqüência de corolários.

Corolário 3.1.2 *O espaço quociente $X = \Gamma\backslash\mathcal{H}$, munido da topologia quociente, é Hausdorff, e portanto localmente compacto.*

DEMONSTRAÇÃO: É o argumento usual: sejam $w, w' \in \mathcal{H}$, não equivalentes sob Γ. Como $\gamma \cdot w' \to \infty$, existe uma vizinhança compacta U de w que não contém nenhum ponto da forma $\gamma \cdot w'$ (para $\gamma \in \Gamma$). Seja V uma vizinhança de w'. Pela proposição, o conjunto

$$A = \{\gamma \in \Gamma : \gamma \cdot V \cap U \neq \emptyset\}$$

é finito. Para cada $\gamma \in A$, escolha uma vizinhança de V_γ de $\gamma \cdot w'$ com $V_\gamma \cap U = \emptyset$. Considere então

$$U' = V \cap \bigcap_{\gamma \in A} \gamma^{-1} \cdot V_\gamma.$$

U' é uma vizinhança de w' satisfazendo $\gamma \cdot U' \cap U = \emptyset$ para todo $\gamma \in \Gamma$, de modo que $\Gamma \cdot U$ e $\Gamma \cdot U'$ são vizinhanças compactas saturadas disjuntas das órbitas de w e de w'. □

Corolário 3.1.3 *A projeção canônica* $\mathsf{D} \longrightarrow \Gamma\backslash\mathcal{H}$ *é uma função própria.*

DEMONSTRAÇÃO: Se $K \subset \mathcal{H}$ é compacto, temos que mostrar que

$$B = \left(\bigcup_{\gamma \in \Gamma} \gamma \cdot K \right) \cap \mathsf{D}$$

é compacto. Pela proposição, a união $\bigcup \gamma \cdot K$ é localmente finita, e portanto B é fechado. Por outro lado, como K é compacto, a fórmula 1.2 mostra que $\Im(w)$ é limitado para os $w \in B$, donde segue que B é limitado, portanto compacto. □

Corolário 3.1.4 *Seja* $\backsim$ *a relação de equivalência em* D *induzida pela ação de* Γ, *isto é:*

$$\tau \backsim \tau + 1 \quad se\ \Re(\tau) = -\tfrac{1}{2}$$
$$\tau \backsim -1/\tau \quad se\ |\tau| = 1$$

A projeção canônica $\mathsf{D} \longrightarrow X$ *induz um homeomorfismo*

$$\backsim\backslash\mathsf{D} \xrightarrow{\approx} X.$$

Em particular, X ***é homeomorfo a um plano.***

DEMONSTRAÇÃO: A função é bijetora, contínua, e própria. Como os dois espaços são localmente compactos, é um homeomorfismo. Por outro lado, é claro que $\Gamma\backslash\mathcal{D}$ é homeomorfo a um plano. □

O próximo passo é pôr uma estrutura de variedade complexa no espaço topológico X.

Definição 15 *Seja $p : \mathcal{H} \longrightarrow X$ a projeção canônica. Dizemos que uma função f definida em um aberto $U \subset X$ é holomorfa em U se $f \circ p$ é holomorfa em $p^{-1}(U)$.*

Não é muito difícil verificar que isto dá a X uma estrutura de superfície de Riemann. De fato: lembre que se $\tau \in \mathcal{H}$,

$$\mathrm{Stab}(\tau) = \{\gamma \in \Gamma : \gamma \cdot \tau = \tau\}$$

é finito e mesmo cíclico (de ordem ≤ 3). Se então $\mathrm{Stab}(\tau)$ tem ordem e, podemos achar um parâmetro local z_τ em $\mathcal{H}$ perto de τ tal que o gerador de $\mathrm{Stab}(\tau)$ age em z_τ por $z_\tau \mapsto \zeta z_\tau$, onde ζ é uma raiz e-ésima da unidade. Por exemplo, podemos tomar

$$z_\tau(w) = \frac{w - \tau}{w - \overline{\tau}},$$

para $w \in \mathcal{H}$. Então é fácil ver que z_τ^e é um parâmetro local em X perto de $p(\tau)$. (Deixamos como exercício para o leitor a verificação—que é bastante fácil—destas asserções.)

Seja agora $\hat{X} = \Gamma\backslash\mathcal{H}^* = X \cup \{\infty\}$ com a topologia da compactificação com um ponto de X. Para estender a estrutura de superfície de Riemann a $\hat{X}$, basta dar um parâmetro local no infinito. Usamos para isso uma idéia que já foi descrita acima, quando discutimos a ponta no infinito.

Seja $\mathcal{H}_1 = \{\tau \in \mathcal{H} : \Im(\tau) > 1\}$. Então é claro que

$$\Gamma\backslash\mathcal{H}_1 = \langle T \rangle\backslash\mathcal{H}_1,$$

de modo que, via $\tau \mapsto q = e^{2\pi i \tau}$, temos homeomorfismos

$$p(\mathcal{H}_1) \approx \langle T \rangle\backslash\mathcal{H}_1 \approx \{q \in \mathbf{C} : 0 < |q| < e^{-2\pi}\}.$$

Assim, $p(\mathcal{H}_1)$ é um disco perfurado. Pondo $q(\infty) = 0$, e tomando q como parâmetro local em ∞, obtemos a extensão desejada.

Proposição 3.1.5 *A superfície de Riemann $\hat{X}$ é analiticamente isomorfa a $\mathbf{P}^1(\mathbf{C}) = \mathbf{C} \cup \{\infty\}$.*

DEMONSTRAÇÃO: Pelo Corolário 3.1.4, X é homeomorfo a $\Gamma\backslash\mathcal{D}$, que por sua vez é homeomorfo a um plano. Logo, $\hat{X} \approx S^2$ é uma esfera; como a estrutura complexa na esfera é única, a proposição segue. □

Note que isto implica que X é analiticamente isomorfo ao plano complexo, mas que este fato *não* segue do fato de que X é homeomorfo a um plano.

Seja $\lambda : \hat{X} \xrightarrow{\cong} \mathbf{P}^1(\mathbf{C})$ um isomorfismo analítico, e seja $f : \hat{X} \longrightarrow \mathbf{P}^1(\mathbf{C})$ uma função analítica qualquer. Compondo com a projeção $\mathcal{H}^* \longrightarrow \hat{X}$, é claro que f é simplesmente uma função modular de peso zero para Γ. Como as únicas funções meromorfas em $\mathbf{P}^1(\mathbf{C})$ são as funções racionais, segue que $f \circ \lambda^{-1}$ é uma função racional, isto é, que $f = P(\lambda)/Q(\lambda)$ é uma função racional de λ. Em particular, dois isomorfismos analíticos $\hat{X} \longrightarrow \mathbf{P}^1(\mathbf{C})$ diferem por um automorfismo de $\mathbf{P}^1(\mathbf{C})$; vamos fixar um isomorfismo em particular:

Definição 16 *A função modular $j : \mathcal{H} \longrightarrow \mathbf{C}$ é a função meromorfa que induz o isomorfismo analítico $j : \hat{X} \longrightarrow \mathbf{P}^1(\mathbf{C})$ caracterizado pelas condições:*

i) $j(\infty) = \infty$

ii) $j(\rho) = 0$, *onde* $\rho = e^{2\pi i/3} \in \mathcal{H}$

iii) o resíduo de j em ∞ é 1.

Mais adiante, daremos uma construção explícita da função j. Note que já temos o seguinte:

Corolário 3.1.6 *Seja $f : \mathcal{H} \longrightarrow \mathbf{C}$ uma função modular de peso 0. Então f é uma função racional de j, isto é, existem polinômios $P(X)$ e $Q(X)$ tais que*

$$f = \frac{P(j)}{Q(j)}.$$

Note que se interpretarmos o espaço $X = \Gamma\backslash\mathcal{H}$ como o espaço das classes de isomorfismo de curvas elípticas, a discussão acima diz que é possível associar a cada curva elíptica complexa E um número $j(\mathrm{E})$ que determina a classe de isomorfismo de E. Veremos adiante que este invariante é exatamente o "invariante j" da teoria de curvas elípticas.

3.2 Formas modulares como diferenciais

O objetivo desta seção é traduzir a definição de formas modulares de peso k para Γ em termos da superfície de Riemann $\hat{X} = \Gamma\backslash\mathcal{H}^*$. Começamos notando que se

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma,$$

temos que

$$\frac{d(\gamma\cdot\tau)}{d\tau} = (c\tau + d)^{-2}.$$

Se, então, $f \in M_k(\Gamma)$ e $k = 2n$ (lembre que k tem que ser par), temos

$$\frac{f(\gamma\cdot\tau)}{f(\tau)} = (c\tau + d)^{2n} = \left(\frac{d(\gamma\cdot\tau)}{d\tau}\right)^{-n},$$

que podemos escrever como

$$f(\gamma\cdot\tau)(d(\gamma\cdot\tau))^n = f(\tau)(d\tau)^n.$$

Expressões do tipo $f(\tau)(d\tau)^n$ se chamam *formas diferenciais de peso n*. Se pensarmos em uma forma diferencial usual (i.é, de peso um) como uma seção do fibrado cotangente Ω^1, então uma forma diferencial de peso n é simplesmente uma seção do fibrado $\Omega^n = (\Omega^1)^{\otimes n}$.

Com a discussão acima, uma forma modular de peso $2n$ é o mesmo que uma forma diferencial de peso n em $\mathcal{H}$ que é invariante sob a ação de Γ. Passando ao quociente, obtemos uma forma diferencial de peso n holomorfa em X, que se estende a uma forma diferencial generalizada meromorfa em $\hat{X}$ (f é holomorfa no infinito, mas isto não garante que $f(\tau)d\tau^n$ o seja). O mesmo argumento funciona se f for meromorfa, exceto que nesse caso só podemos afirmar que $f(\tau)d\tau^n$ é meromorfa em X. Resumindo:

Proposição 3.2.1 *Uma função meromorfa $f : \mathcal{H} \longrightarrow \mathbb{C}$ é uma função modular de peso $k = 2n$ para Γ se e só se $f(\tau)d\tau^n$ define uma forma diferencial meromorfa de peso n em $\hat{X}$.*

Queremos agora relacionar os zeros e polos de $f(\tau)$ e de $f(\tau)d\tau^n$. Lembre que se f é uma função meromorfa em uma variedade complexa, a *ordem de f em um ponto x* é a ordem do zero de f em x (onde, como sempre, um "zero de ordem negativa $-n$" é um polo de ordem n).

Proposição 3.2.2 *Seja $f \in F_{2n}(\Gamma)$ uma função modular de peso $2n$ para Γ, e seja $\omega = \omega(f) = f(\tau)d\tau^n$ a forma diferencial em $\hat{X}$ correspondente. Seja $p : \mathcal{H}^* \longrightarrow \hat{X}$ a projeção canônica. Para cada ponto $x \in \mathcal{H}^*$, sejam*

$$\mathsf{w}_x(f) = \textit{ordem de } f \textit{ em } x$$
$$\mathsf{v}_x(\omega) = \textit{ordem de } \omega \textit{ em } p(x) \in \hat{X}.$$

Então temos:

$$\textit{(i)} \quad \mathsf{w}_\infty(f) = \mathsf{v}_\infty(\omega) + n$$
$$\textit{(ii)} \quad \mathsf{w}_x(f) = e\mathsf{v}_x(\omega) + n(e-1),$$

onde $e = \#\mathrm{Stab}(x)$ e $x \neq \infty$.

DEMONSTRAÇÃO: *(i)* Segue de

$$2\pi i d\tau = \frac{dq}{q},$$

já que $q = e^{2\pi i\tau}$, donde

$$(2\pi i)^n (d\tau)^n = \frac{(dq)^n}{q^n},$$

e portanto

$$f(\tau)(d\tau)^n = \frac{1}{(2\pi i)^n}\frac{f(q)}{q^n}(dq)^n.$$

Note que mesmo que f for holomorfa no infinito, a forma diferencial pode não o ser.

(ii) Seja $x \in \mathcal{H}$, e escolha z_x como acima, de modo que $t = z_x^e$ seja um parâmetro local em $p(x)$. Então se $\mathsf{v}_x(\omega) = m$, existe uma função holomorfa u com $u(x) \neq 0$, tal que $\omega = ut^m(dt)^n$. Como $dt = ez_x^{e-1}dz_x$, segue que

$$\begin{aligned} \omega &= ut^m(dt)^n \\ &= uz_x^{em}(ez_x^{e-1}dz_x)^n \\ &= ue^n z_x^{em+n(e-1)}(dz_x)^n, \end{aligned}$$

donde segue a equação *(ii)*. □

Um corolário divertido é:

Corolário 3.2.3 *Na correspondência acima formas modulares parabólicas de peso 2 correspondem a formas diferenciais holomorfas em $\hat{X}$. Logo, não existem formas modulares parabólicas de peso 2 para Γ.*

DEMONSTRAÇÃO: Se $f \in S_2(\Gamma)$, as equações acima implicam de imaediato que $\omega(f)$ é holomorfa. Como $\hat{X} \cong \mathbf{P}^1(\mathbf{C})$, não existem formas diferenciais holomorfas (de peso 1) em $\hat{X}$, donde a conclusão. □

OBSERVAÇÃO: No caso mais geral dos subgrupos de congruência, a primeira frase deste corolário permanece verdadeira, e fornece um instrumento geométrico para estudar as formas parabólicas de peso 2. Em particular, a dimensão do espaço $S_2(G)$ é igual ao gênero da superfície de Riemann compacta $G\backslash\mathcal{H}^*$.

Para obter um resultado geral, lembremos que em uma superfície compacta de gênero g, o grau do divisor de qualquer forma diferencial (de peso 1) é $2g-2$, isto é, para uma diferencial ω de peso 1, a soma das ordens $\mathsf{v}_x(\omega)$ quando x percorre todos os pontos de $\hat{X}$ é $2g-2$. Tensorizando n vezes, obtemos:

Fato 3.2.4 *Se ω é uma forma diferencial meromorfa de peso n em uma superfície de Riemann compacta S de gênero g, então*

$$\sum_{x \in S} \mathsf{v}_x(\omega) = n(2g-2).$$

No nosso caso, temos $\hat{X} \cong \mathbf{P}^1(\mathbf{C})$, de modo que $g=0$, donde

$$\sum_{p(x) \in \hat{X}} \mathsf{v}_x(\omega) = -2n = -k.$$

Vamos traduzir esta fórmula em termos de f. Como estamos interessados nas imagens $p(x)$, basta olhar os pontos $x \in \mathsf{D}$, onde D é o domínio fundamental usual. Os únicos pontos com estabilizador não-trivial são as imagens de $x=i$ e $x=\rho$ (lembre que $p(\rho)=p(-\bar{\rho})$). Destacando estes e o ponto no infinito, a fórmula fica

$$\mathsf{v}_\infty(\omega) + \mathsf{v}_i(\omega) + \mathsf{v}_\rho(\omega) + \sum{}^* \mathsf{v}_x(\omega) = -k,$$

onde $\sum^*$ indica a soma sobre os outros pontos de D. Usando a proposição, isto fica

$$\mathsf{w}_\infty(f) - n + \frac{1}{2}(\mathsf{w}_i(f) - n) + \frac{1}{3}(\mathsf{w}_\rho(f) - 2n) + \sum{}^* \mathsf{w}_x(f) = -2n,$$

isto é,

$$w_\infty(f) + \frac{1}{2}w_i(f) + \frac{1}{3}w_\rho(f) + \sum{}^* w_x(f) = \frac{n}{6} = \frac{k}{12}.$$

Então provamos:

Corolário 3.2.5 *Seja $f \in F_k(\Gamma)$ e seja $w_x(f)$ a ordem de f em $x \in \mathcal{H}$. Então*

$$w_\infty(f) + \frac{1}{2}w_i(f) + \frac{1}{3}w_\rho(f) + \sum{}^* w_x(f) = \frac{n}{6} = \frac{k}{12}, \tag{3.1}$$

onde $\sum^*$ indica a soma sobre um conjunto de representantes das outras órbitas da ação de Γ.

A fórmula 3.1 agora permite obter um grande número de informações sobre os espaços $M_k(\Gamma)$.

Corolário 3.2.6 *i) $M_k(\Gamma) = 0$ se $k \leq 0$, k ímpar, ou $k = 2$.*

ii) $\dim(M_k(\Gamma)) = 1$ *se* $k = 4$, 6, 8, 10,*ou* 14.

iii) $M_4(\Gamma) = \mathbb{C}E_4$; além disso, E_4 tem um zero simples em $\tau = \rho$, e nenhum outro zero.

iv) $M_6(\Gamma) = \mathbb{C}E_6$; além disso, E_6 tem um zero simples em $\tau = i$, e nenhum outro zero.

DEMONSTRAÇÃO: Imediato da fórmula 3.1, mais o fato já conhecido que $E_k \in M_k(\Gamma)$. Note que "nenhum outro zero" significa, é claro, "nenhum outro que não seja equivalente sob Γ." □

Além das séries de Eisenstein E_k, já construímos a forma modular $\Delta \in M_{12}(\Gamma)$, e já verificamos que ela é parabólica, isto é, que $w_\infty(\Delta) \geq 1$. De 3.1, segue que

$$w_\infty(\Delta) = 1 \quad \text{e} \quad w_x(\Delta) = 0 \quad \text{se } x \neq \infty,$$

isto é, que $\Delta(\tau) \neq 0$ para todo $\tau \in \mathcal{H}$ (isto prova este fato sem referência à interpretação modular).

A forma modular parabólica Δ ajuda a entender os espaços $M_k(\Gamma)$ para $k \geq 12$:

Proposição 3.2.7 *Seja $\xi : M_k(\Gamma) \longrightarrow \mathbb{C}$ o homomorfismo*

$$\xi(f) = f(\infty),$$

e seja $\delta : M_{k-12}(\Gamma) \longrightarrow M_k(\Gamma)$ *a multiplicação por* Δ. *Para cada* $k \geq 4$, *temos uma seqüência exata*

$$0 \longrightarrow M_{k-12}(\Gamma) \xrightarrow{\delta} M_k(\Gamma) \xrightarrow{\xi} \mathbf{C} \longrightarrow 0.$$

DEMONSTRAÇÃO: Como $\xi(\mathrm{E}_k) = 1$, é claro que ξ é sobrejetora. Da mesma forma, δ é claramente injetora. Além disso, como $\xi(\Delta) = 0$, temos $\xi(\Delta f) = 0$ para qualquer f, donde $\mathrm{Im}(\delta) \subset \ker(\xi)$.

Por outro lado, se $f(\infty) = 0$, temos $\mathrm{w}_x(f) - \mathrm{w}_x(\Delta) \geq 0$ para todo $x \in \mathcal{H}^*$, donde a função f/Δ é holomorfa em $\mathcal{H} \cup \{\infty\}$ e portanto pertence a $M_{k-12}(\Gamma)$. □

O cálculo das dimensões dos espaços $M_k(\Gamma)$ agora é imediato:

Corolário 3.2.8 *Se* $k \geq 4$, *temos*

$$\dim(M_k(\Gamma)) = \begin{cases} 0 & \text{se } k \text{ é ímpar} \\ \left\lfloor \frac{k}{12} \right\rfloor & \text{se } k = 2n \text{ e } n \equiv 1 \pmod 6 \\ \left\lfloor \frac{k}{12} \right\rfloor + 1 & \text{se } k = 2n \text{ e } n \not\equiv 1 \pmod 6 \end{cases}$$

onde $\lfloor x \rfloor$ *denota a parte inteira (o "chão") de* x.

Corolário 3.2.9 *Toda* $f \in M_k(\Gamma)$ *é um polinômio isobárico em* E_4 *e* E_6.

DEMONSTRAÇÃO: Seja $k \geq 4$ um inteiro par. Existem a e b tais que $4a + 6b = k$, porque $\mathrm{mdc}(4, 6) = 2$. Temos então

$$\mathrm{E}_4^a \mathrm{E}_6^b \in M_k(\Gamma) \quad \text{e} \quad \xi(\mathrm{E}_4^a \mathrm{E}_6^b) = 1.$$

Então, pela Proposição 3.2.7, temos

$$f = \xi(f)\mathrm{E}_4^a \mathrm{E}_6^b + \Delta f_1,$$

com $f_1 \in M_{k-12}(\Gamma)$. Como Δ é um polinômio em E_4 e E_6, o resultado segue por indução. □

Note, em particular, que $E_4^2 = E_8$, $E_6 E_4 = E_{10}$ e $E_4^2 E_6 = E_{14}$, porque os espaços em questão são de dimensão um.

OBSERVAÇÃO: Todos os resultados deste capítulo foram para o caso do grupo modular completo: $\Gamma = SL_2(\mathbf{Z})$. Entretanto, os mesmos métodos funcionam para o caso de subgrupos de congruência $G \subset \Gamma$, com algumas complicações adicionais:

i) As superfícies de Riemann $G\backslash\mathcal{H}^*$ não são necessariamente de gênero 0; o gênero pode ser calculado usando a fórmula de Hurwitz.

ii) O peso k não é necessariamente par, donde a interpretação de formas modulares como formas diferenciais generalizadas é mais complicada (precisamos de "formas diferenciais de peso $k/2$").

Apesar disso, é perfeitamente possível realizar o programa e determinar as dimensões dos espaços $M_k(G)$, exceto para o caso $k = 1$, que é mais sutil. Veja, por exemplo, os cálculos no livro [Shi71] de Shimura.

No que segue, vamos *assumir* que os espaços $M_k(G)$ são de dimensão finita.

3.3 Exercícios

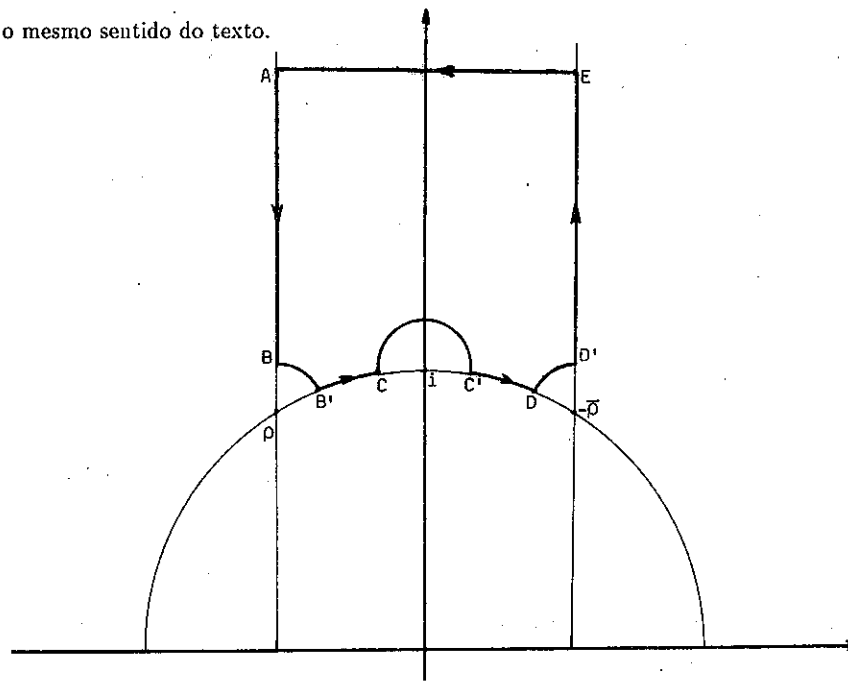
1) Verifique as asserções da página 46 sobre o parâmetro local em X.

2) Este exercício e os seguintes dão uma demonstração elementar da fórmula 3.1. Seja D o domínio fundamental usual para Γ, e seja $f \in M_k(\Gamma)$. Mostre que f tem no máximo um número finito de zeros em D.

3) Seja M um número tal que $f(\tau) \neq 0$ sempre que $\tau \in \mathsf{D}$ e $\Im(\tau) \geq M$, e ponha $\mathsf{D}_M = \{\tau \in \mathsf{D} : \Im(\tau) \leq M\}$. Suponha, para os exercícios (3) a (7), que f não tenha zeros no bordo de D_M, exceto talvez em i, ρ e $-\overline{\rho}$. Seja C o contorno indicado na figura abaixo, isto é, o bordo de D_M modificado por desvios circulares em torno de i, ρ e $-\overline{\rho}$. Mostre que

$$\frac{1}{2\pi i}\int_C \frac{df}{f} = \sum\nolimits^* \mathrm{w}_x(f),$$

onde $\sum^*$ tem o mesmo sentido do texto.



4) Seja agora C_1 o segmento horizontal $AE = \{\tau \in \mathsf{D} : \Im(\tau) = M\}$ (isto é, a parte superior do bordo de D_M), orientado em sentido crescente de $\Re(\tau)$. Use a mudança de variáveis $q = e^{2\pi i \tau}$ para mostrar que

$$\frac{1}{2\pi i}\int_{C_1}\frac{df}{f} = \mathsf{w}_\infty(f).$$

5) Seja C_2 o círculo completo em torno de $\rho = e^{2\pi i/3}$ do qual o arco BB' faz parte. Mostre que

$$\frac{1}{2\pi i}\int_{C_2}\frac{df}{f} = -\mathsf{w}_\rho(f).$$

Conclua que quando o raio de C_2 tende a zero,

$$\frac{1}{2\pi i}\int_{BB'}\frac{df}{f} \longrightarrow -\frac{1}{6}\mathsf{w}_\rho(f).$$

Mostre resultados análogos para os arcos CC' e DD'

6) Use a periodicidade da f para mostrar que

$$\frac{1}{2\pi i}\int_{AB}\frac{df}{f} + \frac{1}{2\pi i}\int_{D'E}\frac{df}{f} = 0.$$

7) Resta lidar com $B'C$ e DC'. Verifique que S manda $B'C$ em DC', e use a modularidade para mostrar que

$$\frac{df(S\tau)}{f(S\tau)} = k\frac{d\tau}{\tau} + \frac{df(\tau)}{f(tau)}.$$

Conclua que

$$\frac{1}{2\pi i}\int_{B'C}\frac{df}{f} + \frac{1}{2\pi i}\int_{C'D}\frac{df}{f} \longrightarrow \frac{k}{12}$$

quando os raios dos três pequenos círculos tendem a zero. Some todos os termos para obter a fórmula 3.1 (sob a hipótese feita no exercício (3).

8) Finalmente, remova a hipótese de que f não tem zeros ou polos nos bordos da região estudada.

3.4 Notas

O tratamento geométrico da demonstração da fórmula 3.1 não é o usual, mas me parece o mais instrutivo em relação à extensão da teoria ao caso de subgrupos de congruência. Seguimos, acima, o artigo de Serre em [Bo66]. Para um tratamento que inclui o caso dos subgrupos de congruência, a melhor referência é sem dúvida o livro [Shi71] de Shimura, mas veja também o livro [Kob84] de Koblitz para alguma informação. As superfícies de Riemann compactas $G\backslash\mathcal{H}^*$ são curvas algébricas, e é possível obter modelos destas curvas sobre $\mathbf{Q}$, e mesmo sobre $\mathbf{Z}$, que preservam sua interpretação "modular" (quando o nível é suficientemente alto). A construção dos modelos sobre $\mathbf{Q}$ pode ser encontrada no livro de Shimura; a extensão a anéis mais gerais é discutida em [DR73] e [KM85].

Capítulo 4

Várias Contas e Exemplos

Neste capítulo faremos vários cálculos que permitirão, entre outras coisas, determinar as q-expansões no infinito de várias das formas modulares que já obtivemos. Além disso, obteremos novos exemplos, inclusive de formas modulares para subgrupos de congruência de nível $N > 1$.

Já definimos, no Capítulo 2, as séries de Eisenstein

$$G_k(\tau) = \sum_{m,n}{}' \frac{1}{(m\tau + n)^k},$$

e suas versões normalizadas

$$E_k(\tau) = \frac{1}{2\zeta(k)} G_k(\tau).$$

Com elas, obtivemos a forma modular parabólica de peso 12

$$\Delta(\tau) = \frac{1}{1728}(E_4^3(\tau) - E_6^2(\tau)).$$

Vamos introduzir mais um exemplo importante:

Definição 17 *Definimos*

$$j(\tau) = \frac{E_4^3(\tau)}{\Delta(\tau)}.$$

Proposição 4.0.1 *A função j é uma função modular de peso zero, holomorfa em $\mathcal{H}$ e com um polo de ordem um no infinito, que define um isomorfismo analítico*

$$\hat{X} = \Gamma\backslash\mathcal{H}^* \xrightarrow{\sim} \mathbf{P}^1(\mathbf{C})$$

que induz um isomorfismo analítico entre $\Gamma\backslash\mathcal{H}$ e $\mathbf{C}$.

DEMONSTRAÇÃO: Como E_4 e Δ são holomorfas em $\mathcal{H}^*$, $\Delta \neq 0$ em $\mathcal{H}$, Δ tem um zero simples no infinito e E_4 não se anula no infinito, resta verificar que j é um isomorfismo. Para isso, basta aplicar a fórmula 3.1 à forma modular $E_4^3 - \lambda\Delta$ para mostrar que há um único zero para cada λ, o que mostra que j é uma bijeção, donde um isomorfismo. □

Já sabemos que $j(\rho) = 0$ (porque $E_4(\rho) = 0$), e que j tem um polo simples no infinito. Para mostrar que j é o isomorfismo analítico considerado no capítulo anterior, resta calcular o resíduo no polo. Para isso, precisamos calcular as q-expansões de E_4 e de Δ. De qualquer forma, já sabemos que:

Corolário 4.0.2 *Seja f uma função meromorfa em $\mathcal{H}$. São equivalentes:*

i) f é uma função modular de peso zero;

ii) f é quociente de duas formas modulares de mesmo peso;

iii) f é função racional de j.

4.1 q-expansões

Vamos agora determinar algumas q-expansões no infinito. Começamos lembrando algumas definições técnicas.

Definição 18 *Os números de Bernouilli B_k são os números racionais definidos por*

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} x^k.$$

Um cálculo tedioso mas simples permite calcular alguns valores:

$$B_0 = 1 \qquad B_1 = -\frac{1}{2} \qquad B_4 = -\frac{1}{30}$$

$$B_6 = \frac{1}{42} \qquad B_8 = -\frac{1}{30} \qquad B_{10} = \frac{5}{66}$$

$$B_{12} = -\frac{691}{2730} \qquad B_{14} = \frac{7}{6}$$

e assim por diante. É imediato verificar que se k é ímpar, $k \neq 1$, temos $B_k = 0$.

Definição 19 *Sejam k e n inteiros positivos. Definimos*

$$\sigma_k(n) = \sum_{d|n} d^k.$$

Proposição 4.1.1 *Seja k um inteiro positivo par. Então*

i) $\zeta(k) = -\dfrac{(2\pi i)^k \mathrm{B}_k}{2k!}$;

ii) se $k \geq 4$, $G_k(\tau) = 2\zeta(k)\left(1 - \dfrac{2k}{\mathrm{B}_k}\displaystyle\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n\right)$;

iii) se $k \geq 4$, $\mathrm{E}_k(\tau) = 1 - \dfrac{2k}{\mathrm{B}_k}\displaystyle\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n$.

DEMONSTRAÇÃO: Trata-se de uma seqüência de truques de análise. Começamos com a identidade

$$\operatorname{sen}(\pi z) = (\pi z)\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right) = (\pi z)\prod_{n=1}^{\infty}\left(1 - \frac{z}{n}\right)\left(1 + \frac{z}{n}\right).$$

Tomando a derivada logarítmica dos dois lados, obtemos

$$\pi\cot(\pi z) = \frac{1}{z} + \sum_{n=1}^{\infty}\left(\frac{1}{z-n} + \frac{1}{z+n}\right).$$

Por outro lado,

$$\pi\cot(\pi z) = \pi i\frac{e^{\pi iz} + e^{-\pi iz}}{e^{\pi iz} - e^{-\pi iz}} = \pi i + \frac{2\pi i}{e^{2\pi iz} - 1}.$$

Multiplicando por z e pondo $x = 2\pi iz$, isto fica

$$\pi z\cot(\pi z) = \frac{x}{2} + \frac{x}{e^x - 1} = 1 + \sum_{k\geq 2}\frac{\mathrm{B}_k}{k!}x^k.$$

Expandindo a primeira fórmula em série de potências, obtemos:

$$\begin{aligned}
\pi z\cot(\pi z) &= 1 + 2\sum_{n=1}^{\infty}\frac{z^2}{z^2 - n^2} \\
&= 1 + 2\sum_{n=1}^{\infty}\frac{\frac{x^2}{(2\pi i)^2}}{\frac{x^2}{(2\pi i)^2} - n^2} \\
&= 1 + 2\sum_{n=1}^{\infty}\frac{x^2}{x^2 - (2\pi i)^2 n^2} \\
&= 1 - 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{x^{2m}}{n^{2m}(2\pi i)^{2m}} \\
&= 1 - 2\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}\frac{1}{n^{2m}}\right)\frac{x^{2m}}{(2\pi i)^{2m}} \\
&= 1 - 2\sum_{m=1}^{\infty}\zeta(2m)\frac{x^{2m}}{(2\pi i)^{2m}}.
\end{aligned}$$

Comparando as duas fórmulas, obtemos, para $k = 2m$ um inteiro par,

$$\frac{B_k}{k!} = -\frac{2\zeta(k)}{(2\pi i)^k},$$

donde segue *(i)*.

Para provar *(ii)* e *(iii)*, voltamos à fórmula

$$\pi \cot(\pi z) = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z+n} + \frac{1}{z-n} \right),$$

e diferenciamos sucessivamente. Notando que, para $q = e^{2\pi i z}$,

$$\pi \cot(\pi z) = \pi i + \frac{2\pi i}{e^{2\pi i z} - 1} = \pi i - 2\pi i \sum_{n=0}^{\infty} q^n,$$

obtemos

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^k} &= \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} n^{k-1} e^{2\pi i n m z} \\ &= -\frac{2k}{B_k} \zeta(k) \sum_{d=1}^{\infty} d^{k-1} q^{dm}. \end{aligned}$$

Somando sobre m,

$$\begin{aligned} G_k(z) &= 2\zeta(k) + \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^k} \\ &= 2\zeta(k) \left(1 - \frac{2k}{B_k} \sum_{d,m} d^{k-1} q^{md} \right) \\ &= 2\zeta(k) \left(1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \right), \end{aligned}$$

conforme desejado. □

Note que em particular nós mostramos que

$$E_k(\tau) = 1 - \frac{2k}{B_k} \sum \sigma_{k-1}(n) q^n$$

tem coeficientes racionais, e que os únicos primos que aparecem no denominador são os que dividem o denominador de $2k/B_k$. Por exemplo:

$$\begin{aligned} E_4 &= 1 - 240 \sum \sigma_3(n) q^n \\ E_6 &= 1 - 504 \sum \sigma_5(n) q^n \\ E_{12} &= 1 + \frac{65520}{691} \sum \sigma_{11}(n) q^n. \end{aligned}$$

Note que, como anunciado, os coeficientes da q-expansão de E_4 e de E_6 são inteiros, mas que isso nem sempre é assim.

Como $1728\Delta = E_4^3 - E_6^2$ e $j = E_4^3/\Delta$, isto permite calcular os primeiros termos das q-expansões de Δ e de j:

$$\Delta = \sum_{n=1}^{\infty} \tau(n) q^n = q - 24q^2 + 252q^3 - 1472q^4 + \cdots$$

$$j = \frac{1}{q} + 744 + \sum_{n=1}^{\infty} c(n) q^n = \frac{1}{q} + 744 + 196884q + \cdots$$

(As notações $\tau(n)$ e $c(n)$ são tradicionais.) Vê-se de imediato que os $\tau(n)$ são inteiros; é possível provar que os $c(n)$ também são. No caso de Δ, obteremos adiante uma outra fórmula, em termos da função eta de Dedekind, que facilita o cálculo dos coeficientes.

Note que a q-expansão de j que obtemos mostra que o resíduo de j no infinito é zero, e portanto prova (finalmente!) que j é exatamente o isomorfismo analítico que destacamos no capítulo anterior.

4.2 E_2 e a função eta

No estudo das séries de Eisenstein, foi necessário excluir o caso $k = 2$. Isto porque a série que definiria $G_2(\tau)$ converge apenas condicionalmente, e por isso não define uma forma modular. Mesmo assim, podemos escolher uma ordem específica para a somatória tal que a série convirja; a função assim obtida não é uma forma modular de peso 2, mas tem propriedades muito próximas.

Definição 20 *Seja $\tau \in \mathcal{H}$; definimos*

$$\begin{aligned} E_2(\tau) &= \frac{1}{2\zeta(2)} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{1}{(m\tau + n)^2} \\ &= 1 + \frac{3}{\pi^2} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{1}{(m\tau + n)^2}. \end{aligned}$$

Para um m fixado, temos

$$\sum_{n=-\infty}^{\infty} \frac{1}{(m\tau + n)^2} = -\frac{4}{B_2} \zeta(2) \sum_{d=1}^{\infty} d q^{dm}$$

(como acima!), donde

$$E_2(\tau) = 1 - 24 \sum_{n=1}^{\infty} \sigma_1(n) q^n,$$

como se esperaria. O fato de que E_2 tem uma q-expansão já mostra que ela é holomorfa em $\mathcal{H}$ e que $E_2(\tau+1) = E_2(\tau)$, mas:

Proposição 4.2.1 $\tau^{-2}E_2(-1/\tau) = E_2(\tau) + \dfrac{12}{2\pi i\tau}$

DEMONSTRAÇÃO: Trata-se de um cálculo delicado com séries, que omitiremos. Veja os exercícios para mais detalhes. □

A função E_2 está ligada de perto a uma outra função, a função eta de Dedekind, que tem um papel muito importante na teoria de formas modulares.

Definição 21 *A função eta de Dedekind é a função* $\eta : \mathcal{H} \longrightarrow \mathbf{C}$ *definida por*

$$\eta(\tau) = e^{\frac{2\pi i\tau}{24}} \prod_{n=1}^{\infty}(1 - e^{2\pi in\tau}) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n),$$

onde, como sempre, $q = e^{2\pi i\tau}$ *e* $q^{1/24} = e^{2\pi i\tau/24}$.

A relação entre η e E_2 é a seguinte:

Lema 4.2.2 $\dfrac{\eta'(\tau)}{\eta(\tau)} = \dfrac{2\pi i}{24}E_2(\tau)$

DEMONSTRAÇÃO: Cálculo direto. □

O lema, junto com a equação funcional de E_2, permite deduzir uma equação funcional para η:

Proposição 4.2.3 *Seja* $\sqrt{\ }$ *o ramo da função raiz quadrada que tem parte real positiva. Então:*

$$\eta(-1/\tau) = \sqrt{\frac{\tau}{i}}\,\eta(z).$$

DEMONSTRAÇÃO: É claro da definição que $\eta(\tau)$ é uma função holomorfa, e que $\eta(\tau) \neq 0$ para todo τ. Como ambos os termos da igualdade que queremos provar são funções holomorfas em $\mathcal{H}$ e como a igualdade é trivialmente verdadeira para $\tau = i$, basta verificar que as derivadas logarítmicas dos dois lados são iguais. Mas

$$\mathrm{dlog}[\eta(-1/\tau)] = \frac{\eta'(\tau)}{\eta(\tau)}\tau^{-2} = \frac{2\pi i}{24}E_2(-1/\tau),$$

e

$$\mathrm{dlog}\left[\sqrt{\frac{\tau}{i}}\eta(\tau)\right] = \frac{1}{2\tau} + \frac{\eta'(\tau)}{\eta(\tau)} = \frac{1}{2\tau} + \frac{2\pi i}{24}\mathrm{E}_2(\tau),$$

e a igualdade segue imediatamente da proposição anterior. □

Uma maneira de entender esta equação funcional é dizer que η é "essencialmente" uma função modular de peso 1/2. A proposição seguinte explora esse fato para relacionar η e Δ, e em particular provar (outra vez, mas independentemente) que $\Delta(\tau) \neq 0$ para todo $\tau \in \mathcal{H}$.

Proposição 4.2.4 $\Delta(\tau) = \eta^{24}(\tau) = q\prod_{n=1}^{\infty}(1-q^n)^{24}$

DEMONSTRAÇÃO: Basta notar que η^{24} é uma forma modular parabólica de peso 12 (pela equação funcional), e portanto é um múltiplo de Δ. Comparando os coeficientes de q, segue que são iguais. □

Segue em particular que se $\Delta = \sum\tau(n)q^n$, temos $\tau(n) \in \mathbf{Z}$. Além disso, esta fórmula dá um meio de calcular os primeiros termos da q-expansão sem muita dificuldade. (Mas não o melhor método de calcular os $\tau(n)$—veja [Nie75].)

4.3 Exemplos de nível N ≠ 1

Até aqui todos os nossos exemplos foram de formas modulares de nível 1, isto é, de formas modulares para $\Gamma = \mathrm{SL}_2(\mathbf{Z})$. Em uma introdução à teoria, é natural que nos concentremos neste caso. Entretanto, para assegurar o leitor de que a teoria para níveis maiores não é vazia, construiremos nesta seção vários exemplos de formas modulares de nível mais alto. Em primeiro lugar, mencionamos um resultado que permite obter formas de nível N a partir de formas de nível menor.

Proposição 4.3.1 *Seja $f \in M_k(\Gamma)$, e seja m um divisor de* N. *A função g_m definida por*

$$g_m(\tau) = m^k f(m\tau)$$

é uma forma modular de peso k para $\Gamma_0(\mathrm{N})$. Além disso, se f for parabólica, então g_m será parabólica.

DEMONSTRAÇÃO: É claro que g_m é uma função holomorfa em $\mathcal{H}$. Vamos verificar primeiro a lei de transformação sob $\Gamma_0(N)$, e depois o comportamento nas pontas.

i) Seja $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, de modo que $N|c$, e portanto também $m|c$. Temos

$$\det \begin{pmatrix} a & mb \\ c/m & d \end{pmatrix} = 1,$$

de modo que

$$\gamma(m) = \begin{pmatrix} a & mb \\ c/m & d \end{pmatrix} \in \Gamma.$$

Então

$$\begin{aligned} g_m(\gamma \cdot \tau) &= m^k f(m\gamma \cdot \tau) = m^k f\left(m\frac{a\tau + b}{c\tau + d}\right) \\ &= m^k f\left(\frac{am\tau + mb}{\frac{c}{m}m\tau + d}\right) = m^k f(\gamma(m) \cdot m\tau) \\ &= m^k (\frac{c}{m}m\tau + d)^k f(m\tau) \\ &= (c\tau + d)^k g_m(\tau). \end{aligned}$$

ii) Seja $f(\tau) = \sum a_n q^n$ a q-expansão de f no infinito. É claro que a q-expansão de g_m no infinito é

$$g_m(\tau) = m^k \sum a_n q^{nm},$$

de modo que g_m é holomorfa no infinito.

Para os outras pontas, queremos achar a q-expansão de $g_m|[\gamma]_k$ no infinito, para cada $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$. Lembrando que

$$f|[\gamma]_k(\tau) = \det(\gamma)^{k/2}(c\tau + d)^{-k} f(\gamma \cdot \tau),$$

note, em primeiro lugar, que se tomarmos

$$\alpha = \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix},$$

temos

$$f|[\alpha]_k(\tau) = m^{k/2} f(\alpha \cdot \tau) = m^{-k/2} g_m(\tau),$$

de modo que $g_m = m^{k/2} f|[\alpha]_k$, e portanto $g_m|[\gamma]_k = m^{k/2} f|[\alpha\gamma]_k$. Agora, é fácil verificar que existem $\delta = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \Gamma$ e x, y e $z \in \mathbf{Z}$ tais que tais que

$$\alpha\gamma = \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & z \end{pmatrix} = \delta \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}.$$

(Basta escolher $x = \mathrm{mdc}(c, m)$ e fazer as contas; note que $xz = m$.) Então

$$\begin{aligned} f|[\alpha\gamma]_k &= (f|[\delta]_k)|\left[\begin{pmatrix} x & y \\ 0 & z \end{pmatrix}\right]_k \\ &= f|\left[\begin{pmatrix} x & y \\ 0 & z \end{pmatrix}\right]_k \\ &= m^{k/2} z^{-k} \sum_{n=0}^{\infty} a_n e^{2\pi i n(x\tau+y)/z} \\ &= \left(\frac{x}{z}\right)^{k/2} \sum_{n=0}^{\infty} a_n e^{2\pi i n x y/m} e^{2\pi i n x \tau/z} \\ &= \left(\frac{x}{z}\right)^{k/2} \sum_{n=0}^{\infty} a_n e^{2\pi i n x y/m} e^{2\pi i n x^2 \tau/m} \\ &= \left(\frac{x}{z}\right)^{k/2} \sum_{n=0}^{\infty} a_n \xi^n q_m^{nx^2}, \end{aligned}$$

onde $\xi = e^{2\pi i x y/m}$ é uma raiz z-ésima de 1. Logo,

$$g_m|[\gamma]_k = x^k \sum_{n=0}^{\infty} (a_n \xi^n) q_m^{nx^2},$$

donde g_m é holomorfa nas pontas, e parabólica se f for. □

Por exemplo, se $m = \mathrm{N} = p$ é um primo e se $f \in M_k(\Gamma)$, então $g_p(\tau) = p^k f(p\tau)$ define uma forma modular $g_p \in M_k(\Gamma_0(p))$. Se a q-expansão de f for

$$f = \sum a_n q^n,$$

então as q-expansões de g_p serão: no infinito,

$$g_p(\tau) = p^k \sum a_n q^{np},$$

e no zero,

$$\tau^{-k} g_p(-1/\tau) = \sum (a_n \xi^n) q_p^n,$$

onde $\xi = e^{2\pi i y/p}$ e y é escolhido como acima, já que no caso em que m é primo podemos supor que $\mathrm{mdc}(c, m) = 1$ na demonstração (porque o item *(i)* já cuida do caso $m|c$). É relevante observar que os fatores m^k na proposição foram escolhidos para preservar integralidade: se a q-expansão de f tem coeficientes inteiros, a q-expansão de g_m também tem. (Veja também os exercícios.)

Esta proposição pode ser facilmente generalizada; veja, por exemplo, a Proposição 17 na página 127 de [Kob84]. Por exemplo, o mesmo argumento mostra que se $f \in M_k(\Gamma_0(\mathrm{N}))$ e $g(\tau) = p^k f(p\tau)$, então $g \in M_k(\Gamma_0(\mathrm{N}p))$.

É importante notar que este teorema ainda não justifica o estudo de formas de nível maior, já que ele só produz formas de nível N que provêm de formas de nível menor. De fato, o que o teorema mostra é que cada forma de peso k para Γ produz um "pacote" de formas modulares de peso k para $\Gamma_0(\mathrm{N})$, uma para cada divisor de N:

$$f \in M_k(\Gamma) \rightsquigarrow \{g_d \in M_k(\Gamma_0(\mathrm{N})) : d|\mathrm{N}\},$$

onde $g_d(\tau) = d^k f(d\tau)$. As formas modulares que pertencem ao subespaço gerado pelas obtidas desta forma se chamam "formas velhas" de $M_k(\Gamma_0(\mathrm{N}))$. É preciso provar que estas não esgotam o espaço $M_k(\Gamma_0(\mathrm{N}))$, isto é, provar que existem "formas novas". (Não vamos definir "formas novas" aqui, porque ainda não temos os pré-requisitos teóricos para fazê-lo; basta dizer que elas formam um subespaço de $M_k(\Gamma_0(\mathrm{N}))$ que é um complemento do espaço gerado pelas formas velhas. Veja [AL70] para mais detalhes—mas primeiro leia sobre operadores de Hecke!)

Há duas maneiras de provar a existência de "formas novas". A mais indireta é simplesmente calcular a dimensão de $M_k(\Gamma_0(\mathrm{N}))$, e verificar que ela é maior que a dimensão do espaço das formas velhas. (Veja [Shi71], por exemplo.) Mais diretamente, pode-se entar construir formas novas explicitamente (o mais difícil é provar que são novas!). Infelizmente, ambas estas estratégias são tecnicamente complicadas, e não poderemos ir muito longe nesse sentido. No que segue, damos apenas uns poucos resultados gerais (seguindo Koblitz).

Em primeiro lugar, observamos que as únicas formas modulares de peso zero, mesmo para um subgrupo de congruência, são as constantes.

Proposição 4.3.2 *Para qualquer subgrupo de congruência $G \subset \Gamma$, temos $M_0(G) = \mathbf{C}$.*

DEMONSTRAÇÃO: O caso $G = \Gamma$ já foi provado; no fundo, a demonstração deste caso se baseou no fato de que as únicas funções holomorfas em todos os pontos de uma superfície de Riemann compacta são as constantes. Caso tivéssemos definido uma estrutura de superfície de Riemann compacta em

$$G\backslash\mathcal{H}^* = G\backslash\mathcal{H} \cup \{\text{pontas}\},$$

o mesmo argumento funcionaria neste caso também. Como não o fizemos, temos que dar uma demonstração mais tortuosa, que é indicada nos exercícios. □

Como aplicação, escolhamos inteiros N e k, com k par, tais que $k(\mathrm{N}+1) = 24$, e consideremos a função

$$f(\tau) = (\eta(\tau)\eta(\mathrm{N}\tau))^k = q\prod_{n=1}^{\infty}(1-q^n)^k(1-q^{\mathrm{N}n})^k.$$

A condição $k(\mathrm{N}+1) = 24$ garante, entre outras coisas, a ausência de potências fracionárias de q. O fato de que η é "essencialmente de peso $1/2$" sugere que f deve ser de peso k.

Proposição 4.3.3 *Seja $f(\tau) = (\eta(\tau)\eta(\mathrm{N}\tau))^k$ onde k e N são inteiros positivos, k é par, e $k(\mathrm{N}+1) = 24$. Seja $g \in S_k(\Gamma_0(\mathrm{N}))$ uma forma modular parabólica de peso k para $\Gamma_0(\mathrm{N})$. Então g é múltiplo de f.*

DEMONSTRAÇÃO: Note que a proposição em não afirma que uma tal g existe. A demonstração é simples, e fica como exercício. □

Note que as possibilidades para k e N são $k = 12, 8, 6, 4, 2$ e $\mathrm{N} = 1, 2, 3, 5, 11$, respectivamente, de modo que em particular N é sempre um primo. Se $\mathrm{N} = 1$ e $k = 12$, o resultado não diz nada de novo, porque já sabemos que $\eta^{24} = \Delta$. É possível provar que em todos os outros casos temos $\dim S_k(\Gamma_0(\mathrm{N})) = 1$, de modo que f de fato é uma forma modular. Veja nos exercícios o caso $k = 8$, $\mathrm{N} = 2$. Repare, também, que o caso $k = 3$, $\mathrm{N} = 7$ só foi excluído porque não há formas modulares de peso ímpar para nenhum dos $\Gamma_0(\mathrm{N})$; a próxima seção mostra como tratar esse caso, estudando mais de perto os espaços de formas modulares para $\Gamma_1(\mathrm{N})$ com caráter ε.

4.4 Formas com Caráter

No Capítulo 2, mencionamos o fato de que o espaço $M_k(\Gamma_1(N))$ se decompõe sob a ação de $\Gamma_0(N)$ segundo os caracteres de Dirichlet de nível N

$$\varepsilon : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \mathbf{C}^\times,$$

e escrevemos

$$M_k(\Gamma_1(N)) = \bigoplus_\varepsilon M_k(\Gamma_1(N), \varepsilon).$$

Considere, agora, a matriz

$$-I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma_0(N).$$

Se $f \in M_k(\Gamma_1(N), \varepsilon)$, temos, por definição, que

$$f(\tau) = f(-I \cdot \tau) = \varepsilon(-1)(-1)^k f(\tau);$$

logo, teremos $f(\tau) = 0$ exceto se $\varepsilon(-1) = (-1)^k$. No caso em que ε é o caráter trivial, isto é, no caso de $\Gamma_0(N)$, isto força k a ser par (como já sabemos). Em geral, isto dá:

Proposição 4.4.1 $M_k(\Gamma_1(N), \varepsilon) = 0$ *exceto se* $\varepsilon(-1) = (-1)^k$.

Considere, por exemplo, N = 4. Neste caso, $(\mathbf{Z}/4\mathbf{Z})^\times$ é cíclico de ordem 2, e há dois caracteres: o trivial, e um caráter χ de ordem 2 dado por $\chi(d) = (-1)^{\frac{d-1}{2}}$ para todo d ímpar.

Corolário 4.4.2 *Temos*

$$M_k(\Gamma_1(4)) = \begin{cases} M_k(\Gamma_0(4)) & \text{se } k \text{ é par} \\ M_k(\Gamma_1(4), \chi) & \text{se } k \text{ é ímpar,} \end{cases}$$

onde χ *denota o caráter não-trivial de* $(\mathbf{Z}/4\mathbf{Z})^\times$.

Vamos exibir um exemplo interessante. Considere a função em $\mathcal{H}$ definida por

$$\Theta(\tau) = \sum_{n \in \mathbf{Z}} q^{n^2} = \sum_{n \in \mathbf{Z}} e^{2\pi i n^2 \tau}.$$

Como $\tau \in \mathcal{H}$, a série é claramente convergente, e define uma função holomorfa em $\mathcal{H}$. Vamos mostrar que o quadrado de Θ é uma forma modular de peso 1.

Proposição 4.4.3 *Seja* $f = \Theta^2$. *Então* $f \in M_1(\Gamma_1(4)) = M_1(\Gamma_1(4), \chi)$, *onde* $\chi(d) = (-1)^{(d-1)/2}$.

DEMONSTRAÇÃO: Vamos verificar primeiro a regra de transformação para $-I$, T e

$$ST^4S = \begin{pmatrix} -1 & 0 \\ 4 & -1 \end{pmatrix},$$

que geram $\Gamma_0(4)$. (Veja o exercício 13 do Capítulo 1.) A lei de transformação sob $-I$ é imediata pela escolha do caráter; sob T, é imediata pela definição de Θ. Basta, então, considerar ST^4S. Seja $\alpha(4) = \begin{pmatrix} 0 & -1 \\ 4 & 0 \end{pmatrix}$. Então

$$\alpha(4)^{-1} = -\frac{1}{4}\alpha(4) = \begin{pmatrix} 0 & 1/4 \\ -1 & 0 \end{pmatrix},$$

e portanto

$$\alpha(4)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\alpha(4)^{-1} = \begin{pmatrix} d & -c/4 \\ -4b & a \end{pmatrix},$$

donde

$$ST^4S = -\alpha(4)T\alpha(4)^{-1} = \frac{1}{4}\alpha(4)T\alpha(4).$$

Como matrizes escalares agem trivialmente, segue que $f|[ST^4S]_1 = f|[\alpha(4)T\alpha(4)]_1$. Para calcular esta última expressão, precisamos determinar $f|[\alpha(4)]_1$. Fazemos isto em um lema que essencialmente determina a equação funcional da função Θ.

Lema 4.4.4 $f|[\alpha(4)]_1 = -if$

DEMONSTRAÇÃO DO LEMA: 1) Considere a função

$$\theta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi t n^2},$$

definida para $\Re(t) > 0$. Vamos verificar primeiro que o lema segue da equação funcional $\theta(1/t) = \sqrt{t}\theta(t)$, onde $\sqrt{\ }$ é, como antes, o ramo da raiz quadrada com parte real positiva.

De fato, temos $\Theta(\tau) = \theta(-2i\tau)$, e a equação funcional fica sendo

$$\Theta(-1/4\tau) = \sqrt{\frac{2z}{i}}\Theta(\tau);$$

elevando ambos os termos ao quadrado e rearranjando, vem

$$\Theta(-1/4\tau)^2 = \frac{2\tau}{i}\Theta(\tau)^2$$

Como $f = \Theta^2$, segue

$$\begin{aligned}(f|[\alpha(4)]_1)(\tau) &= 4^{1/2}(4\tau)^{-1}f(-1/4z) \\ &= \frac{2}{4\tau}\frac{2\tau}{i}f(\tau) = -if(\tau),\end{aligned}$$

como queríamos.

2) Resta mostrar a equação funcional para $\theta(t)$. Como tudo é analítico, basta provar a equação para $t \in R$, $t > 0$. Para isso, aplicamos a fórmula somatória de Poisson à função $g(x) = e^{-\pi t x^2}$, com $t > 0$ fixo. Como

$$g(x) = f(\sqrt{t}x), \quad \text{com } f(x) = e^{-\pi x^2},$$

a transformada de Fourier é

$$\hat{g}(y) = t^{-1/2}e^{-\pi y^2/t}.$$

A fórmula de Poisson dá, então,

$$\sum_{m\in\mathbf{Z}} g(m) = \sum_{m\in\mathbf{Z}} \hat{g}(m),$$

isto é,

$$\theta(t) = \sum_{m\in\mathbf{Z}} e^{-\pi t m^2} = t^{-1/2}\sum_{m\in\mathbf{Z}} e^{\pi m^2/t} = t^{-1/2}\theta(1/t),$$

que é a equação funcional. □

Provado o Lema, é fácil terminar:

$$\begin{aligned}f|[ST^4S]_1 &= f|[\alpha(4)T\alpha(4)]_1 = -if|[T\alpha(4)]_1 \\ &= -if|[\alpha(4)]_1 = -f = \chi(-1)f,\end{aligned}$$

como queríamos.

Resta verificar as condições nas pontas. No infinito, a holomorficidade é evidente pela definição; para as outras pontas, veja os exercícios.□

A forma modular $\Theta^2 \in M_1(\Gamma_1(4), \chi)$ é um exemplo interessante porque os coeficientes da sua q-expansão no infinito têm um significado aritmético importante: pondo

$$\Theta^2 = \left(\sum_{n\in\mathbf{Z}} q^{n^2}\right)^2 = \sum_{n\geq 0} a(n)q^n,$$

fica claro que $a(n)$ é igual ao número de modos de se escrever n como soma de dois quadrados. Por exemplo,

$$\begin{aligned} a(0) &= 1 \quad \text{já que} \quad 0 = 0^2 + 0^2 \\ a(1) &= 4 \quad \text{já que} \quad 1 = (\pm 1)^2 + 0^2 = 0^2 + (\pm 1)^2 \\ a(2) &= 4 \quad \text{já que} \quad 2 = (\pm 1)^2 + (\pm 1)^2 \\ a(3) &= 0 \\ &\textit{etc.} \end{aligned}$$

Em particular, se for possível expressar Θ^2 em termos de outras formas modulares, a igualdade de q-expansões dará uma fórmula para o número de modos em que se pode escrever n como soma de quadrados. Uma estimativa para os $a(n)$ quando $n \to \infty$ se interpreta da mesma forma.

Há muitas outras maneiras de se obter exemplos de formas modulares. Por exemplo, se $f \in M_k(\Gamma_1(N), \epsilon)$ e

$$\chi : (\mathbf{Z}/M\mathbf{Z})^\times \longrightarrow \mathbf{C}^\times$$

é um caráter de Dirichlet, é possível construir uma forma modular $f^\chi \in M_k(\Gamma_1(NM), \epsilon\chi^2)$, chamada o "twist" de f por χ, tal que as q-expansões no infinito são

$$f = \sum a_n q^n \quad \text{e} \quad f^\chi = \sum \chi(n) a_n q^n,$$

onde estendendemos χ a $\mathbf{Z}$ da maneira usual. (Veja os exercícios.)

4.5 Exercícios

1) Verifique os valores dos números de Bernouilli dados no texto. Prove que se k é ímpar, $k \geq 3$, então $B_k = 0$.

2) Mostre que as funções σ_k são multiplicativas, isto é, que se $\mathrm{mdc}(m,n) = 1$, então $\sigma_k(mn) = \sigma_k(m)\sigma_k(n)$.

3) Traduza as identidades $E_4^2 = E_8$ e $E_4 E_6 = E_{10}$ para obter relações entre as funções σ_k.

4) Traduza a igualdade $1728\Delta = E_4^3 - E_6^2$ para obter uma fórmula para $\tau(n)$ em termos de σ_3 e σ_5. Conclua que $\tau(n) \in \mathbf{Z}$, como sugere o texto.

5) Prove a Proposição 4.2.1. (Este exercício é mais para quem gosta de trabalhar com séries. Aqui está uma sugestão (seguindo [Kob84]): defina

$$a_{m,n}(z) = \frac{1}{(mz+n-1)(mz+n)} = \frac{1}{mz+n-1} - \frac{1}{mz+n},$$

e mostre que

$$E_2(z) = \tilde{E}_2(z) = 1 + \frac{3}{\pi^2} \sum_{m \neq 0} \sum_{n=-\infty}^{\infty} \left(\frac{1}{(mz+n)^2} - a_{m,n}(z) \right),$$

e que esta série é absolutamente convergente; inverta a ordem da somatória para concluir que

$$E_2(z) = -z^2 E_2(-1/z) - \frac{3}{\pi^2} \sum_{n=-\infty}^{\infty} \sum_{m \neq 0} a_{m,n}(z),$$

e use a expansão para $\pi \cot(\pi z)$ obtida no texto para calcular esta última soma.)

6) Prove o Lema 4.2.2.

7) Seja G um subgrupo de congruência de Γ, e seja $f \in M_0(G)$ uma forma modular de peso zero para G. Este exercício indica uma demonstração de que f é constante. Fixe $\tau_0 \in \mathcal{H}$ e ponha $a = f(\tau_0)$.

a) Seja $\Gamma = \bigcup \alpha_i \cdot G$ a decomposição de Γ em classes laterais de G, e defina

$$g(\tau) = \prod (f(\alpha_i^{-1} \cdot \tau) - a).$$

Mostre que $g \in M_0(\Gamma)$ e que portanto g é constante.

b) Mostre que $g = 0$.

c) Conclua que f é constante.

8) Sejam $(k, N) = (8,2), (6,3), (4,5)$, ou $(2,11)$, $g \in S_k(\Gamma_0(N))$ e $f(\tau) = (\eta(\tau)\eta(N\tau))^k$. Mostre que f^{N+1} é uma forma modular de peso 24 para $\Gamma_0(N)$ que não se anula em $\mathcal{H}$ e tem zeros de ordem $N+1$ no infinito e no zero. (Note que como N é primo, estas são as únicas pontas.) Conclua que $(g/f)^{N+1}$ é uma forma modular de peso zero, e que portanto g é múltiplo de f.

9) Use os geradores de $\Gamma_0(4)$ obtidos no exercício 13 do Capítulo 1 para mostrar que $f(\tau) = (\eta(\tau)\eta(4\tau))^8$ é uma forma modular parabólica de peso 8 para $\Gamma_0(4)$.

10) Prove as seguintes identidades:

$$\Theta(\tau) + \Theta(\tau + \frac{1}{2}) = 2\Theta(4\tau)$$

$$\eta(\tau + \frac{1}{2}) = e^{\frac{2\pi i}{48}} \frac{\eta^3(2\tau)}{\eta(\tau)\eta(4\tau)}$$

11) Prove que

$$\frac{\eta^8(z)}{\eta^4(2z)} \in M_2(\Gamma_0(4),$$

e determine seu valor em cada ponta.

12) Mostre que $\Theta^4 \in M_2(\Gamma_0(4))$, seguindo o texto, e notando que as pontas são ∞, 0 e 1/2. Use este fato para concluir a demonstração de que $\Theta^2 \in M_1(\Gamma_0(4))$. Qual é a interpretação aritmética dos coeficientes da q-expansão no infinito de Θ^4? Qual é a relação entre Θ^4 e a forma modular do exercício 11?

13) Mostre que se $f \in M_{k_1}(\Gamma_1(N), \varepsilon_1)$ e $g \in M_{k_2}(\Gamma_1(N), \varepsilon_2)$, então $fg \in M_{k_1+k_2}(\Gamma_1(N), \varepsilon_1\varepsilon_2)$.

14) Prove a afirmação do texto sobre a existência do "twist" $f^{\chi} \in M_k(\Gamma_1(NM), \varepsilon\chi^2)$ de uma forma modular $f \in M_k(\Gamma_1(N), \varepsilon)$ por um caráter de Dirichlet $\chi : (\mathbf{Z}/M\mathbf{Z})^{\times} \longrightarrow \mathbf{C}^{\times}$. (Sugestão: use as idéias que apareceram na demonstração da Proposição 4.3.1.)

15) (Para quem sabe um pouco de curvas elípticas.) Mostre que a operação $f \rightsquigarrow g_m$ discutida no texto pode ser interpretada modularmente da seguinte forma: seja (E, C) uma curva elíptica munida de um subgrupo cíclico de ordem N, seja $N = md$, e seja $dC = \{dP : P \in C\}$ (que é um subgrupo cíclico de ordem m); então

$$g_m(E, C, \omega) = f(E/dC, \omega'),$$

onde E/dC é o quociente de E por dC e ω' é a imagem de ω pela isogenia dual da passagem ao quociente.

4.6 Notas

Os vários resultados analíticos que tratamos são todos clássicos, e estão em todas (ou quase todas) as referências, às vezes com demonstrações diferentes. Em relação a formas de nível superior, nós apenas arranhamos a superfície. Em particular, deixamos de construir as séries de Eisenstein de nível superior, e de explorar um pouco mais seriamente o problema de determinar formas parabólicas (este é um problema não-trivial que é bastante estudado—veja [?]). A construção geral de séries de Eisenstein de nível N pode ser encontrada em [Kob84], bem como vários outros resultados sobre formas de nível superior. Da mesma forma, há muitos outros resultados sobre funções theta (parentes da nossa Θ), que são importantes tanto nas teorias de curvas elípticas e variedades abelianas quanto na teoria de formas modulares. Em [Ser73], Serre dá exemplos usando outras formas quadráticas que não $x^2 + y^2$, que em particular dão formas de nível 1; em [Kob84], Koblitz discute a interpretação de Θ como "forma modular de nível 1/2".

Capítulo 5

A L-função de uma Forma Modular

O objetivo central deste capítulo é definir a L-função associada a uma forma modular. A L-função é uma série de Dirichlet, mas pode ser definida, na nossa situação, como a transformada de Mellin da forma modular em questão; isto permite obter facilmente sua equação funcional e continuação analítica. Para simplificar a discussão, vamo-nos restringir ao caso de nível N = 1, dando apenas indicações das alterações necessárias no caso geral.

5.1 Estimativas

Para garantir a convergência das várias operações analíticas a serem discutidas no restante do capítulo, precisamos de estimativas da rapidez de crescimento dos coeficientes da q-expansão de uma forma modular. Começamos com uma forma modular $f \in M_k(\Gamma)$ de peso k para o grupo modular completo Γ. Suponhamos que a q-expansão (no infinito) de f é

$$f = \sum a_n q^n.$$

Proposição 5.1.1 *Se $f = \mathrm{E}_k$ é uma série de Eisenstein, então existem constantes positivas A e B tais que*

$$An^{k-1} \leq |a_n| \leq Bn^{k-1}.$$

DEMONSTRAÇÃO: Por um lado, temos que $a_n = \pm A\sigma_{k-1}(n)$ para uma constante real (racional, até) positiva A, donde

$$|a_n| = A\sigma_{k-1}(n) \geq An^{k-1}.$$

Por outro lado,

$$\frac{|a_n|}{n^{k-1}} = A\sum_{d|n}\frac{1}{d^{k-1}} \leq A\sum_{d=1}^{\infty}\frac{1}{d^{k-1}} = A\zeta(k-1) < \infty,$$

donde é claro que existe B como queremos. □

Como toda forma de nível um é soma de um múltiplo de E_k e uma forma parabólica, resta considerar agora o caso das formas parabólicas. Nesse caso, podemos provar bem mais:

Teorema 5.1.2 (Hecke) *Se $f \in S_k(\Gamma)$ é uma forma modular parabólica de peso k para Γ, então $a_n = O(n^{k/2})$, isto é, existe uma constante B tal que $|a_n| \leq Bn^{k/2}$ quando $n \to \infty$.*

DEMONSTRAÇÃO: Como f é parabólica, podemos escrever

$$f(\tau) = \sum_{n\geq 1} a_n q^n = q\left(\sum_{n\geq 1} a_n q^{n-1}\right).$$

Logo, $|f(\tau)| = O(|q|) = O(e^{-2\pi y})$ quando $q \to 0$. (Aqui, é claro, $y = \Im(\tau)$.)

Considere, então, a função $\phi(\tau) = |f(\tau)|y^{k/2}$. Como f é modular, temos

$$\begin{aligned}\phi(\gamma\cdot\tau) &= |f(\gamma\cdot\tau)|\Im(\gamma\cdot\tau)^{k/2}\\ &= |(c\tau+d)^k f(\tau)|(\Im(\tau)|c\tau+d|^{-2})^{k/2}\\ &= |f(\tau)|\Im(\tau)^{k/2} = \phi(\tau),\end{aligned}$$

de modo que ϕ é invariante sob Γ. Além disso, ϕ é contínua em $\mathcal{H}$ e $\phi(\tau) \to 0$ quando $y \to \infty$, já que $|f(\tau)| = O(e^{-2\pi y})$. Segue que ϕ é limitada, isto é, que existe $M \geq 0$ tal que $\phi(\tau) \leq M$ para todo $\tau \in \mathcal{H}$, o que quer dizer

$$|f(\tau)| \leq My^{-k/2},$$

para todo $\tau \in \mathcal{H}$.

Para passar desta estimativa para f para uma estimativa para os a_n, usamos o método usual para determinar os coeficientes da série de Fourier: fixe y e faça $x = \Re(\tau)$ variar de 0 a 1. O ponto $q = e^{2\pi i(x+iy)}$ percorre um círculo $\mathcal{C}_y$ de centro $q = 0$. Pela fórmula dos resíduos,

$$a_n = \frac{1}{2\pi i}\int_{\mathcal{C}_y} f(\tau)q^{-n-1}dq = \int_0^1 f(x+iy)q^{-n}dx,$$

e de $|f(\tau)| \leq My^{-k/2}$ segue agora que $|a_n| \leq My^{-k/2}e^{2\pi ny}$, para todo $y \geq 0$. Pondo $y = 1/n$, obtemos $|a_n| \leq e^{2\pi}Mn^{k/2}$. □

Corolário 5.1.3 *Se f não for parabólica, temos $|a_n| = O(n^{k-1})$.*

DEMONSTRAÇÃO: Claro, já que $f = \lambda E_k + f_0$, com f_0 parabólica e $\lambda \neq 0$. □

OBSERVAÇÕES: 1) A estimativa do teorema não é a melhor possível: as Conjeturas de Weil, provadas por Deligne, dão, para f parabólica,

$$a_n = O(n^{\frac{k-1}{2}}\sigma_0(n)) = O(n^{\frac{k-1}{2}+\epsilon}),$$

para todo $\epsilon > 0$.

2) O teorema se estende facilmente a formas parabólicas para subgrupos de congruência— veja os exercícios. Quanto às formas não-parabólicas para subgrupos de congruência, é preciso introduzir séries de Eisenstein generalizadas que permitem uma decomposição em soma direta, também neste caso, do tipo $M_k(G) = \mathit{Eisenstein} \oplus S_k(G)$. (Veja-se [Kob84], por exemplo.) Isto feito, uma estimativa semelhante à obtida acima segue sem dificuldade. Do ponto de vista das L-funções, o caso parabólico é nitidamente o mais interessante, e nos concentraremos nele no que segue.

5.2 A L-função

Nesta seção, queremos definir a L-série associada a uma forma modular. Como já dissemos acima, nos restringimos ao caso de formas de nível um. Nosso objetivo é o seguinte: dada uma forma modular $f \in M_k(\Gamma)$ com q-expansão $f = \sum a_n q^n$, queremos estudar a série de Dirichlet

$$\mathrm{L}(f,s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}.$$

As estimativas obtidas acima já mostram que esta série converge em algum semi-plano $\Re(s) > M$, mas não seria fácil obter uma equação funcional ou a continuação analítica a partir desta expressão. Assim, usamos um outro método, que obtém a L-função a partir da transformação de Mellin da forma modular f.

As estimativas da seção anterior mostram que para qualquer $f \in M_k(\Gamma)$ temos

$$|a_n| = O(n^c)$$

com $c = k-1$ se f não for parabólica e $c = k/2$ se f for parabólica. Consideremos a função

$$g(s) = \int_0^{i\infty} f(\tau)\tau^s \frac{d\tau}{\tau}.$$

Lema 5.2.1 *Se f for parabólica, a integral que define $g(s)$ converge desde que $\Re(s) > c+1$.*

DEMONSTRAÇÃO: Formalmente,

$$\begin{aligned}\int_0^{i\infty} f(\tau)\tau^s \frac{d\tau}{\tau} &= \sum_{n=1}^{\infty} a_n \int_0^{i\infty} \tau^s e^{2\pi i n\tau} \frac{d\tau}{\tau} \\ &= \sum_{n=1}^{\infty} a_n \left(\frac{-1}{2\pi i n}\right)^s \int_0^{\infty} t^s e^{-t} \frac{dt}{t} \qquad (t = -2\pi i n\tau) \\ &= (-2\pi i)^{-s}\Gamma(s) \sum_{n=1}^{\infty} a_n n^{-s}.\end{aligned}$$

Agora, como $|a_n n^{-s}| = O(n^{c-\Re(s)})$, a última série é absolutamente convergente se $\Re(s) > c+1$, donde todas as passagens são válidas. □

Se f não for parabólica, basta substituir $f(\tau)$ por $f(\tau) - a_0$, e repetir o argumento acima para definir $g(s)$. Em qualquer caso, obtemos

$$g(s) = (-2\pi i)^{-s}\Gamma(s)\mathrm{L}(f,s),$$

onde

$$\mathrm{L}(f,s) = \sum_{n=1}^{\infty} a_n n^{-s},$$

e onde tudo está definido para $\Re(s) > c+1$.

Para obter a equação funcional e a continuação analítica, o essencial é usar a modularidade de f; a equação crucial é $f(-1/\tau) = \tau^k f(\tau)$. Escrevemos

$$g(s) = \int_0^{i\infty} f(\tau)\tau^s \frac{d\tau}{\tau} = \int_0^{i} f(\tau)\tau^s \frac{d\tau}{\tau} + \int_i^{i\infty} f(\tau)\tau^s \frac{d\tau}{\tau}.$$

Agora, pondo $\tau = -1/w$,

$$\begin{aligned}\int_0^i f(\tau)\tau^s\frac{d\tau}{\tau} &= \int_i^{i\infty} f(-1/w)(-1)^s w^{-s}\frac{dw}{w}\\ &= (-1)^s\int_i^{i\infty} w^k f(w)w^{-s}\frac{dw}{w}\\ &= (-1)^s\int_i^{i\infty} f(w)w^{k-s}\frac{dw}{w}.\end{aligned}$$

Logo, podemos escrever

$$g(s) = \int_i^{i\infty}[f(\tau)\tau^s + (-1)^s f(\tau)\tau^{k-s}]\frac{d\tau}{\tau},$$

donde se vê imediatamente que $g(k-s) = (-1)^s g(s)$. Escolhendo

$$C = (-1)^{k/2} = \begin{cases} 1 & \text{se } k \equiv 0 \pmod 4\\ -1 & \text{se } k \equiv 2 \pmod 4,\end{cases}$$

podemos reescrever isto como

$$(-i)^s g(s) = C(-i)^{k-s}g(k-s).$$

Pondo agora

$$\Lambda(s) = (-i)^s g(s) = (-i)^s(-2\pi i)^{-s}\Gamma(s)\mathrm{L}(f,s) = (2\pi)^{-s}\Gamma(s)\mathrm{L}(f,s),$$

temos a equação funcional

$$\Lambda(s) = C\Lambda(k-s),$$

o que dá também a continuação analítica. Resumindo tudo:

Teorema 5.2.2 *Seja $f(\tau) = \sum a_n q^n$ uma forma modular de peso k para $\Gamma = \mathrm{SL}_2(\mathbf{Z})$. Seja $c = k/2$ se f for parabólica, $c = k-1$ se não. Então:*

i) A série de Dirichlet

$$\mathrm{L}(f,s) = \sum_{n=1}^{\infty} a_n n^{-s}$$

converge se $\Re(s) > c+1$;

ii) A função $\Lambda(f,s) = (2\pi)^{-s}\Gamma(s)\mathrm{L}(f,s)$ possui uma continuação analítica para todo o plano complexo, e satisfaz a equação funcional

$$\Lambda(f,s) = C\Lambda(f,k-s),$$

onde $C = (-1)^{k/2}$.

Definição 22 *A L-função associada à forma modular f é a extensão meromorfa da função* $L(f,s)$ *definida pelo teorema.*

Assim,

$$L(f,s) = \frac{(2\pi)^s \Lambda(f,s)}{\Gamma(s)},$$

onde "$\Lambda(f,s)$" denota a continuação analítica acima. Por abuso de linguagem, escreveremos ocasionalmente $L(f,s) = \sum_{n=1}^{\infty} a_n n^{-s}$, embora esta expressão só valha para $\Re(s) > c+1$.

Note que $L(f,s) = 0$ para $s = 0, -1, -2, \ldots$, onde $\Gamma(s)$ tem polos; estes são os "zeros triviais". Quaisquer outros zeros terão que ocorrer na "faixa crítica" $0 \leq \Re(s) \leq k$. Note, também, que se $C = -1$, teremos forçosamente $L(f, k/2) = 0$. A expansão de Taylor de $L(f,s)$ em torno de $s = k/2$ tem um grande interesse nas aplicações aritméticas da teoria.

A L-função $L(f,s) = \sum a_n n^{-s}$ tem um claro parentesco com a função zeta de Riemann e suas generalizações, as L-séries de Dirichlet. Estas têm em comum o fato de possuirem uma expansão em produto de Euler, do tipo

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1},$$

onde p percorre os números primos. É natural, então, perguntar se podemos escrever $L(f,s)$ na forma

$$L(f,s) = \prod_p \mathcal{E}(f,p,s)$$

para algum conjunto de "fatores locais" $\mathcal{E}(f,p,s)$. A resposta é que nem sempre isto é possível, mas que as formas para as quais uma fórmula deste tipo é válida são exatamente as de interesse central na teoria. Para estudar esta questão, Hecke introduziu uma família de operadores, os "operadores de Hecke", que são o tema do próximo capítulo e que são o centro da teoria aritmética das formas modulares.

5.3 Subgrupos de congruência

Não é muito difícil generalizar a definição de $L(f,s)$ para o caso de formas modulares para subgrupos de congruência. Para o caso de formas parabólicas para $\Gamma_1(N)$, por exemplo,

exatamente o mesmo cálculo de antes dá uma expressão da série como transformação de Mellin. O que fica um pouco mais complicado é obter a equação funcional. Neste texto, preferimos omitir este cálculo; veja, por exemplo, o tratamento no livro [Kob84] de Koblitz. O resultado final é muito parecido com o do caso de nível 1: se $f \in M_k(\Gamma_1(\mathrm{N}), \varepsilon)$, a L-função $\mathrm{L}(f,s)$ possui uma continuação analítica a todo o plano complexo. Se ainda f satisfizer

$$f(-1/\mathrm{N}\tau) = C\mathrm{N}^{-k/2}(i\mathrm{N}z)^k f(\tau),$$

com $C = \pm 1$, temos, pondo

$$\Lambda(f,s) = (\sqrt{\mathrm{N}}/2\pi)^s \Gamma(s)\mathrm{L}(f,s),$$

uma equação funcional

$$\Lambda(f,s) = C\Lambda(f,k-s).$$

5.4 Exercícios

1) O objetivo deste exercício é estender as estimativas obtidas no texto para os coeficientes de uma forma parabólica ao caso dos subgrupos de congruência. Seja $G \subset \Gamma$ um subgrupo de congruência, e seja $\overline{\Gamma} = \bigcup \alpha_j \overline{G}$ uma decomposição em classes laterais. Seja $f \in S_k(G)$ e ponhamos $f_j = f|[\alpha_j^{-1}]_k$. Defina $\phi(\tau) = y^{k/2}|f(\tau)|$ e $\phi_j(\tau) = y^{k/2}|f_j(\tau)|$. Agora siga a demonstração no texto para mostrar que os coeficientes das q-expansões de f em todas as pontas satisfazem a estimativa $a_n = O(n^{k/2})$. (É preciso algum cuidado em observar as larguras das pontas!)

2) Seja $f \in M_k(\Gamma_1(\mathrm{N}), \chi)$, e seja $\alpha(\mathrm{N}) = \begin{pmatrix} 0 & -1 \\ \mathrm{N} & 0 \end{pmatrix}$. Mostre que $f|[\alpha(\mathrm{N})]_k \in M_k(\Gamma_1(\mathrm{N}), \overline{\chi})$, que $f \mapsto f|[\alpha(\mathrm{N})]_k$ é um isomorfismo $M_k(\Gamma_1(\mathrm{N}), \chi) \xrightarrow{\cong} M_k(\Gamma_1(\mathrm{N}), \overline{\chi})$, e que seu quadrado é multiplicação por $(-1)^k$ em $M_k(\Gamma_1(\mathrm{N}), \chi)$. Conclua que se $\chi = \overline{\chi}$ (isto é, se χ é real), toda $f \in M_k(\Gamma_1(\mathrm{N}), \chi)$ pode ser escrita como uma soma $f = f^+ + f^-$, onde $f^+|[\alpha(\mathrm{N})]_k = i^k f^+$ e $f^-|[\alpha(\mathrm{N})]_k = -i^k f^-$, de modo que f^+ e f^- verificam equações do tipo mencionado na discussão da equação funcional no caso de nível $\mathrm{N} > 1$.

3) Suponha $f \in M_k(\Gamma_1(N), \chi)$, com χ real, e suponha que $f|[\alpha(N)]_k = Ci^k f$, com $C = \pm 1$. Prove que $L(f, s)$ satisfaz a equação funcional indicada no texto.

4) Que tipo de alteração será necessária no caso em que o caráter não é real?

5.5 Notas

Muitos dos problemas centrais da teoria dos números estão ligados às propriedades de vários tipos de L-funções. Na maioria dos casos de interesse aritmético, sabe-se ou conjetura-se que a L-função em questão possui uma continuação analítica e uma equação funcional; em alguns casos, isto é bastante difícil de demonstrar (em contraste com a situação neste capítulo!). Há uma conjetura, entretanto, que sugere que *todas* as L-funções de interesse aritmético são também as L-funções associadas a (generalizações de) formas modulares (o que implicaria de imediato a existência da continuação analítica e equação funcional). Por exemplo: é possível definir a L-função de uma curva elíptica. Conjetura-se, então, que se E é uma curva elíptica cuja equação tem coeficientes em $\mathbf{Q}$ com L-função $L(E, s)$, então existe uma forma modular parabólica de peso 2 para $\Gamma_0(N)$ (para um N específico, que depende de E) com coeficientes $a_n \in \mathbf{Q}$ tal que $L(E, s) = L(f, s)$. Esta conjetura é devida a Taniyama, Shimura e Weil, mas é conhecida como "Conjetura de Weil-Taniyama"; foi provado recentemente por Ribet que ela implica o "teorema" de Fermat. O melhor resultado no sentido de uma demosntração da conjetura é devido a Weil; o enunciado e uma demonstração podem ser encontrados em [Ogg69]. Vale observar que a recíproca é verdadeira: a qualquer forma modular de peso 2 para $\Gamma_0(N)$ com coeficientes de Fourier racionais corresponde uma curva elíptica definida sobre $\mathbf{Q}$ tal que as L-funções coincidem.

Capítulo 6

Operadores de Hecke

Para decidir se a L-função $L(f,s)$ de uma forma modular tem uma expansão em produto de Euler, Hecke introduziu uma família de operadores T_n, para $n \in \mathbf{Z}$, $n \geq 1$. Estes operadores provaram ser absolutamente fundamentais para a teoria, e dedicaremos este capítulo a estudar os fatos mais simples a seu respeito.

Para definir os operadores de Hecke, partimos primeiro da interpretação de formas modulares em termos de redes, e depois reinterpretamos. Como a teoria em nível $N > 1$ fica sensivelmente mais complicada, outra vez nos restringimos ao caso de nível um, isto é, de formas modulares para $\Gamma = SL_2(\mathbf{Z})$. Nosso tratamento segue de perto o de Serre em [Ser73]. O caso geral é discutido com cuidado em [Kob84] e [Lan76]; o livro [Shi71] de Shimura contém uma versão diferente da definição, em termos de "double cosets", que não iremos mencionar aqui.

6.1 Operadores de Hecke em redes

Começamos definindo a ação dos operadores de Hecke no espaço das redes $\Lambda \subset \mathbf{C}$. Nossos operadores não vão ser funções nesse espaço, mas sim correspondências, no seguinte sentido:

Definição 23 *Seja X um conjunto e seja $\mathcal{L}$ o grupo abeliano livre gerado pelos elementos de X. Uma* correspondência $T : X \to X$ *é um endomorfismo de grupos $\mathcal{L} \longrightarrow \mathcal{L}$.*

Uma correspondência $T : X \to X$ é então dada pelos valores

$$Tx = \sum_{y \in X} n_x(y) y,$$

para cada $x \in X$. A idéia é que uma correspondência é uma função generalizada onde cada elemento de X tem várias imagens (a "imagem" de x é um conjunto de elementos $y \in X$, com multiplicidades dadas pelos $n_x(y)$).

Uma correspondência $\mathrm{T} : X \to X$ define uma transformação no conjunto das funções $F : X \longrightarrow \mathbf{C}$, pondo

$$\begin{aligned}(\mathrm{T}F)(x) &= F(\mathrm{T}x) \\ &= F(\textstyle\sum n_x(y)y) \\ &= \textstyle\sum n_x(y)F(y).\end{aligned}$$

(É claro que isto funciona para funções com imagem em qualquer grupo abeliano.)

Para definir os operadores T_n, começamos definindo correspondências no conjunto $\mathcal{R}$ das redes $\Lambda \subset \mathbf{C}$.

Definição 24 *Seja $n \in \mathbf{Z}$, $n \geq 1$. Definimos uma correspondência T_n no conjunto $\mathcal{R}$ das redes $\Lambda \subset \mathbf{C}$ pondo, para cada rede Λ,*

$$\mathrm{T}_n\Lambda = \sum_{(\Lambda:\Lambda')=n} \Lambda.$$

Assim, a imagem de uma rede é a soma das sub-redes de índice n. Note que há apenas um número finito destas, já que temos que ter $\Lambda' \supset n\Lambda$, donde seu número é igual ao número de subgrupos de ordem n do grupo finito

$$\Lambda/n\Lambda \cong \left(\mathbf{Z}/n\mathbf{Z}\right)^2.$$

Por exemplo, se $n = p$ é primo, há $p + 1$ tais subgrupos.

Para cada $\lambda \in \mathbf{C}^\times$, definimos também um operador de homotetia de redes:

Definição 25 *Se $\Lambda \subset \mathbf{C}$ é uma rede, pomos $\mathrm{R}_\lambda\Lambda = \lambda\Lambda$.*

Proposição 6.1.1 *Temos:*

i) $\mathrm{R}_\lambda\mathrm{R}_\mu = \mathrm{R}_{\lambda\mu}$

ii) $\mathrm{R}_\lambda\mathrm{T}_n = \mathrm{T}_n\mathrm{R}_\lambda$

iii) Se $\mathrm{mdc}(m,n) = 1$, *então* $\mathrm{T}_n\mathrm{T}_m = \mathrm{T}_{nm}$

iv) Para todo p primo, $T_{p^{n+1}} = T_{p^n} T_p - p T_{p^{n-1}} R_p$

DEMONSTRAÇÃO: é um exercício fácil de grupos abelianos, que fica para o leitor. □

Corolário 6.1.2 *Para cada p primo e cada* $n \geq 1$, T_{p^n} *é um polinômio em* T_p *e* R_p. *Além disso, a álgebra de operadores gerada pelos* R_λ *e os* T_p *para p primo é comutativa, e contém todos os* T_n.

O corolário salienta o fato de que em muitos casos basta considerar os T_p para p primo.

Seja agora $F : \mathcal{R} \longrightarrow \mathbf{C}$ uma função de redes de peso k, isto é, tal que $F(\lambda\Lambda) = \lambda^{-k}\Lambda$. Como acima, podemos considerar operadores T_n e R_λ no conjunto das funções deste tipo:

$$(T_n F)(\Lambda) = \sum_{(\Lambda:\Lambda')=n} F(\Lambda')$$

$$(R_\lambda F)(\Lambda) = F(\lambda\Lambda).$$

Como F é de peso k, temos $R_\lambda F = \lambda^{-k} F$; além disso, como R_λ e T_n comutam,

$$R_\lambda(T_n F) = T_n(R_\lambda F) = T_n(\lambda^{-k} F) = \lambda^{-k} T_n F,$$

donde $T_n F$ também é de peso k. Temos, portanto, uma transformação no espaço das funções de rede de peso k, que vamos transferir para uma transformação no espaço das formas modulares de peso k, usando a correspondência entre os dois espaços que exploramos no Capítulo 2. Antes disso, notemos desde já que temos:

$$T_n T_m F = T_{nm} F \qquad \text{se } \mathrm{mdc}(m,n) = 1$$

$$T_{p^{n+1}} F = T_p T_{p^n} F - p^{1-k} T_{p^{n-1}} F,$$

para p primo, $n, m \in \mathbf{Z}$, $n, m \geq 1$.

6.2 Operadores de Hecke em $M_k(\Gamma)$

Vimos, no Capítulo 2, que uma forma modular de peso k para Γ define uma função de rede de peso k, e que uma função de rede de peso k define uma função *quase-modular* de peso k para Γ. Dada $f \in M_k(\Gamma)$, isto permite definir $T_n f : \mathcal{H} \longrightarrow \mathbf{C}$; introduzimos o fator normalizador óbvio.

Definição 26 *Seja $f \in M_k(\Gamma)$ e $F : \mathcal{R} \longrightarrow \mathbf{C}$ a função de rede associada, isto é,*

$$F(\mathbf{Z}w_1 + \mathbf{Z}w_2) = w_2^{-k} f(w_1/w_2).$$

Então $\mathrm{T}_n f$ é a função quase-modular associada à função de rede $n^{k-1}\mathrm{T}_n F$:

$$(\mathrm{T}_n f)(\tau) = n^{k-1}(\mathrm{T}_n F)(\mathbf{Z}\tau + \mathbf{Z}).$$

Para verificar que de fato $\mathrm{T}_n f$ é uma forma modular, precisamos provar a holomorficidade em $\mathcal{H}$ e na ponta no infinito (como conseqüência, poderemos descrever a ação de T_n na q-expansão, o que também é importante). Para isso, precisamos de uma descrição mais explícita das subredes de índice n de uma rede dada.

Lema 6.2.1 *Sejam $\Lambda = \mathbf{Z}w_1 + \mathbf{Z}w_2$ uma rede e n um inteiro positivo. Seja S_n o conjunto das matrizes*

$$\sigma = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

com $a, b, d \in \mathbf{Z}$, $ad = n$, $a \geq 1$ e $0 \leq b < d$. Para cada $\sigma \in S_n$, seja Λ_σ a subrede de Λ com base

$$w_1' = aw_1 + bw_2 \quad \text{e} \quad w_2' = dw_2,$$

de modo que

$$\sigma \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} w_1' \\ w_2' \end{pmatrix}.$$

Então a correspondência $\sigma \leftrightarrow \Lambda_\sigma$ dá uma bijeção entre S_n e o conjunto das subredes de índice n de Λ.

DEMONSTRAÇÃO: (Seguindo Serre em [Ser73]) Como $\det(\sigma) = n$, é claro que $(\Lambda : \Lambda_\sigma) = n$. Por outro lado, se $\Lambda' \subset \Lambda$ tem índice n, sejam

$$X = \frac{\Lambda}{\Lambda' + \mathbf{Z}w_2} \quad \text{e} \quad Y = \frac{\mathbf{Z}w_2}{\Lambda' \cap \mathbf{Z}w_2}.$$

Então X e Y são grupos cíclicos gerados pelas imagens de w_1 e w_2, respectivamente. Sejam $a = \#X$ e $d = \#Y$. Como a seqüência

$$0 \longrightarrow Y \longrightarrow \Lambda/\Lambda' \longrightarrow X \longrightarrow 0$$

é exata, temos $ad = n$; resta determinar b de modo que $\Lambda' = \Lambda_\sigma$.

Se tomarmos $w_2' = dw_2$, então $w_2' \in \Lambda'$ (porque $dw_2 = 0$ em Y). Da mesma forma,

$$aw_1 \in \Lambda' + \mathbf{Z}w_2,$$

donde existe w_1' tal que

$$aw_1 - w_1' \in \mathbf{Z}w_2.$$

É fácil ver, agora, que

$$\Lambda' = \mathbf{Z}w_1' + \mathbf{Z}w_2',$$

e que

$$w_1' = aw_1 + bw_2,$$

com b univocamente determinado módulo d. Escolhendo $0 \leq b < d$ determina b, e portanto σ, e estabelece a bijeção desejada. □

O exemplo mais simples é o caso em que $n = p$ é um primo, quando

$$S_p = \left\{ \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix} : 0 \leq b \leq p-1 \right\},$$

que tem, como esperado, $p+1$ elementos.

Dado o lema, temos imediatamente que

$$(\mathrm{T}_n f)(\tau) = n^{k-1} \sum_{\substack{a \geq 1 \\ 0 \leq b < d \\ ad = n}} d^{-k} f\left(\frac{a\tau + b}{d}\right),$$

ou, usando a ação $|[\gamma]_k$,

$$\mathrm{T}_n f = \frac{1}{n} \sum n^{k/2} f \Big| \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}_k.$$

Logo:

Corolário 6.2.2 *Se $f \in M_k(\Gamma)$, então $\mathrm{T}_n f$ é holomorfa em $\mathcal{H}$.*

Mais interessante é a questão da q-expansão no infinito:

Proposição 6.2.3 *Seja $f \in M_k(\Gamma)$, e suponha que sua q-expansão no infinito seja*

$$f = \sum_{m \geq 0} \alpha(m) q^m.$$

Então $T_n f \in M_k(\Gamma)$, com q-expansão no infinito

$$T_n f = \sum_{m \geq 0} \beta(m) q^m,$$

onde

$$\beta(m) = \sum_{\substack{a \mid \mathrm{mdc}(m,n) \\ a \geq 1}} a^{k-1} \alpha(mn/a^2).$$

DEMONSTRAÇÃO: Basta fazer o cálculo. Temos:

$$\begin{aligned}(T_n f)(\tau) &= n^{k-1} \sum_{\substack{a \geq 1 \\ 0 \leq b < d \\ ad = n}} d^{-k} f\left(\frac{a\tau + b}{d}\right) \\ &= n^{k-1} \sum_{\substack{a \geq 1 \\ 0 \leq b < d \\ ad = n}} d^{-k} \sum_{m \geq 0} \alpha(m) e^{2\pi i m(a\tau + b)/d}.\end{aligned}$$

A soma

$$\sum_{0 \leq b < d} e^{2\pi i m b/d}$$

vale zero se $d \nmid m$ e d se $d|m$. Pondo $m' = m/d$,

$$(T_n f)(\tau) = n^{k-1} \sum_{\substack{a \geq 1 \\ m' \geq 0 \\ ad = n}} d^{-k+1} \alpha(m'd) q^{am'},$$

donde segue a fórmula anunciada reagrupando os termos.□

Os casos mais interessantes são os seguintes:

Corolário 6.2.4 *Com as hipóteses e notações acima, temos:*

i) $\beta(0) = \sigma_{k-1}(n)\alpha(0)$ *e* $\beta(1) = \alpha(n)$

ii) Se $n = p$ é um primo, então:

$$\begin{aligned}\beta(m) &= \alpha(pm) && \text{se } p \nmid m \\ \beta(m) &= \alpha(pm) + p^{k-1}\alpha(m/p) && \text{se } p|m\end{aligned}$$

iii) Se $f \in S_k(\Gamma)$, então $T_n f \in S_k(\Gamma)$.

6.3 Autoformas

Vamos agora considerar as formas modulares $f \in M_k(\Gamma)$ que são autovetores para a ação dos T_n.

Definição 27 *Uma forma modular $f \in M_k(\Gamma)$ se diz uma* autoforma *se for autovetor para a ação de todos os T_n, isto é, se existirem $\lambda(n) \in \mathbb{C}$ tais que para todo n tenhamos*

$$T_n f = \lambda(n) f.$$

O seguinte resultado é fundamental:

Proposição 6.3.1 *Seja $f \in M_k(\Gamma)$ uma autoforma de peso $k > 0$, com q-expansão $f = \sum \alpha(n) q^n$. Então*

i) $\alpha(1) \neq 0$

ii) Se f for normalizada pela condição $\alpha(1) = 1$, então teremos $\lambda(n) = \alpha(n)$ para todo n, de modo que $T_n f = \alpha(n) f$.

DEMONSTRAÇÃO: Seja $T_n f = \sum \beta(n) q^n$. Pelo Corolário 6.2.4, temos

$$\beta(1) = \alpha(n).$$

Por outro lado, como f é autoforma, temos $T_n f = \lambda(n) f$, donde

$$\beta(1) = \lambda(n)\alpha(1).$$

As duas afirmações seguem imediatamente, já que f não é constante (o peso é positivo). □

Corolário 6.3.2 *Se $f \in M_k(\Gamma)$, $k > 0$, é autoforma e não é parabólica, então $f = \lambda E_k$ é um múltiplo da série de Eisenstein.*

DEMONSTRAÇÃO: Se $T_n f = \sum \beta(n) q^n$, temos, pelo Corolário 6.2.4, $\beta(0) = \sigma_{k-1}(n)\alpha(0)$, donde $\lambda(n) = \sigma_{k-1}(n)$; mas, pela Proposição acima, $\alpha(n) = \lambda(n)\alpha(1)$, donde a conclusão. □

Dizemos que uma autoforma é *normalizada* se o coeficiente de q em sua q-expansão é 1, como na Proposição acima. Então temos:

Corolário 6.3.3 *Duas autoformas normalizadas de mesmo peso $k > 0$ com os mesmos autovetores $\lambda(n)$ são iguais.*

DEMONSTRAÇÃO: Pelo teorema, ambas têm a mesma q-expansão, exceto talvez pelo termo constante. Como a diferença não pode ser uma forma de peso zero (porque $k > 0$), segue que são iguais. □

Este corolário se chama "teorema de multiplicidade um" para formas de nível um; ele é um dos resultados fundamentais da teoria.

Corolário 6.3.4 *Se $f = \sum \alpha(n)q^n$ é uma autoforma normalizada, então*

$$\alpha(n)\alpha(m) = \alpha(nm) \qquad \text{se } \mathrm{mdc}(m,n) = 1$$
$$\alpha(p^{n+1}) = \alpha(p^n)\alpha(p) - p^{k-1}\alpha(p^{n-1}),$$

para p primo e $n \geq 1$.

DEMONSTRAÇÃO: Claro, porque os autovalores dos T_n satisfazem as mesmas identidades que os T_n. □

Corolário 6.3.5 *Se $f = \sum \alpha(n)q^n$ é uma autoforma normalizada e $\mathrm{L}(f,s) = \sum \alpha(n)n^{-s}$ é sua L-função, temos*

$$\mathrm{L}(f,s) = \prod_{p \text{ primo}} (1 - \alpha(p)p^{-s} + p^{k-1-2s})^{-1},$$

para $\Re(s) > c+1$, onde $c = k/2$ se f é parabólica, $c = k-1$ se não.

DEMONSTRAÇÃO: Como a função $n \mapsto \alpha(n)$ é multiplicativa, é imediato ver que

$$\mathrm{L}(f,s) = \prod_{p \text{ primo}} \left(\sum_{n=0}^{\infty} \alpha(p^n)p^{-ns} \right).$$

A questão, então, é provar que esta série tem a soma indicada. Pondo $p^{-s} = X$, trata-se de provar uma identidade formal:

$$\sum_{n=0}^{\infty} \alpha(p^n).X^n = \frac{1}{1 - \alpha(p).X + p^{k-1}T^2},$$

que segue imediatamente da fórmula do corolário anterior (multiplique os dois lados por $1 - a(p)X + p^{k-1}X^2$). □

Temos assim a resposta da pergunta formulada no capítulo anterior: $L(f,s)$ tem uma expansão em produto de Euler quando f é uma autoforma (para a recíproca, veja os exercícios!).

6.4 Exemplos de autoformas

Nesta seção, damos alguns exemplos simples de autoformas (de nível um). Para formas que não são parabólicas, não é difícil:

Proposição 6.4.1 *Se $k \geq 4$ é um inteiro par, a série de Eisenstein E_k é uma autoforma de peso k; os autovalores são $\lambda(n) = \sigma_{k-1}(n)$, e a autoforma normalizada é*

$$f = (-1)^k \frac{B_k}{2k} E_k = (-1)^k \frac{B_k}{2k} + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n.$$

Finalmente, temos

$$L(f,s) = \zeta(s)\zeta(s-k+1).$$

DEMONSTRAÇÃO: Já provamos que uma autoforma não parabólica é necessariamente múltiplo de E_k. Para ver que E_k é de fato uma autoforma, basta verificar que $T_p E_k = \sigma_{k-1}(p) E_k$ para todo primo p. Isto pode ser feito verificando diretamente as identidades acima, ou pensando em E_k como função de rede. Vamos usar este último método; é claro que podemos trabalhar com G_k, cuja interpretação em termos de redes é mais simples.

Seja, então, $\tilde{G}_k$ a função de rede correspondente a G_k, de modo que

$$\tilde{G}_k(\Lambda) = \sum_{\substack{\gamma \in \Lambda \\ \gamma \neq 0}} \gamma^{-k}.$$

Logo, temos

$$(T_p \tilde{G}_k)(\Lambda) = \sum_{(\Lambda:\Lambda')=p} \sum_{\substack{\gamma \in \Lambda' \\ \gamma \neq 0}} \gamma^{-k}.$$

Seja $\gamma \in \Lambda$. Se $\gamma \in p\Lambda$, então γ pertence a todas as $p+1$ subredes Λ', de modo que sua contribuição na soma é $(p+1)\gamma^{-k}$. Se, por outro lado, $\gamma \notin p\Lambda$, γ só pode pertencer a *uma*

das Λ', e contribui γ^{-k} na soma. Assim,

$$\begin{aligned}(\mathrm{T}_p\tilde{G}_k)(\Lambda) &= \tilde{G}_k(\Lambda) + p\sum_{\substack{\gamma\in p\Lambda\\ \gamma\neq 0}}\gamma^{-k}\\ &= \tilde{G}_k(\Lambda) + p\tilde{G}_k(p\Lambda)\\ &= (1+p^{1-k})\tilde{G}_k(\Lambda).\end{aligned}$$

Agora, como

$$\widetilde{\mathrm{T}_p G_k} = p^{k-1}\mathrm{T}_p\tilde{G}_k,$$

segue que

$$\mathrm{T}_p G_k = (1+p^{k-1})G_k = \sigma_{k-1}(p)G_k,$$

como desejado. Isto mostra que G_k é uma autoforma, e a versão normalizada segue da fórmula para sua q-expansão.

Finalmente,

$$\begin{aligned}\mathrm{L}(f,s) &= \sum_{n=1}^{\infty}\sigma_{k-1}(n)n^{-s} = \sum_{a,d\geq 1}a^{k-1}a^{-s}d^{-s}\\ &= \left(\sum_d d^{-s}\right)\left(\sum_a a^{-(s+1-k)}\right)\\ &= \zeta(s)\zeta(s-k+1).\square\end{aligned}$$

Em particular, isto mostra que a decomposição

$$M_k(\Gamma) = \mathrm{CE}_k \oplus S_k(\Gamma)$$

é estável sob a ação dos operadores de Hecke.

Exemplos de autoformas parabólicas são um pouco mais difíceis de obter. O caso mais fácil, é claro, é quando $\dim S_k(\Gamma) = 1$.

Proposição 6.4.2 *Para $k = 12$, 16, 18, 20, 22 ou 26, seja f_k uma forma modular parabólica de peso k para Γ. Então f_k é uma autoforma.*

DEMONSTRAÇÃO: Claro, já que a dimensão é um nesses casos. □

Em particular,

$$\Delta = \sum_{n\geq 1} \tau(n) q^n \in S_{12}(\Gamma)$$

é uma autoforma normalizada, de modo que a função $\tau(n)$ satisfaz as identidades acima. A função

$$\mathrm{L}(\Delta, s) = \sum_{n=1}^{\infty} \tau(n) n^{-s}$$

não se expressa em termos da função zeta ou de L-séries de Dirichlet: é um objeto genuinamente novo, e tem grande interesse aritmético.

6.5 O produto escalar de Petersson

Apesar de não ser de todo fácil exibir exemplos de autoformas parabólicas, é fácil provar que elas existem em abundância. Para isso, definimos um produto escalar em $S_k(\Gamma)$, da seguinte forma: se $f, g \in S_k(\Gamma)$, vê-se facilmente que a medida

$$\mu(f,g) = f(\tau)\overline{g(\tau)} y^k \frac{dxdy}{y^2}$$

(onde, é claro, $\tau = x + iy$) é invariante sob Γ, e que é uma medida *limitada* no espaço quociente $\Gamma\backslash\mathcal{H}$ (é aqui que usamos a hipótese de que f e g são parabólicas).

Definição 28 *Se $f, g \in S_k(\Gamma)$, definimos*

$$\langle f, g \rangle = \int_{\Gamma\backslash\mathcal{H}} \mu(f,g) = \int_{\mathrm{D}} f(\tau)\overline{g(\tau)} y^k \frac{dxdy}{y^2},$$

onde D *é um domínio fundamental para* Γ.

Verifica-se, então, que $\langle\ ,\ \rangle$ é um produto escalar positivo hermitiano e não-degenerado em $S_k(\Gamma)$, e que

$$\langle \mathrm{T}_n f, g \rangle = \langle f, \mathrm{T}_n g \rangle,$$

de modo que os T_n são hermitianos em relação a este produto interno. Como os T_n comutam entre si, segue que existe uma *base* de $S_k(\Gamma)$ formada de autoformas. (Segue também que os $\lambda(n)$ são reais; é possível provar que são números algébricos totalmente reais.) Em muitos casos, é possível determinar essa base explicitamente usando o teorema de multiplicidade um mencionado acima.

6.6 Dualidade

Um dos temas básicos da teoria aritmética das formas modulares é que há uma dualidade entre formas modulares e operadores de Hecke. Nesta seção, damos uma versão elementar de um resultado importante de Miller nesse sentido.

Seja $\mathfrak{h}_k = \mathfrak{h}_k(\Gamma)$ a sub-álgebra de $\mathrm{End}_{\mathbf{C}}(S_k(\Gamma))$ gerada pelos T_n. Vamos definir um pareamento entre $\mathfrak{h}_k$ e $S_k = S_k(\Gamma)$. (Como Γ é fixo em toda esta discussão, vamos omiti-lo da notação.) Nesta seção, escreveremos $a(n, f)$ para denotar o coeficiente de q^n na q-expansão de f, de modo que

$$f = \sum_{n \geq 1} a(n, f) q^n.$$

Definição 29 *O pareamento de Miller é a função*

$$(\, ,) : \mathfrak{h}_k \times S_k \longrightarrow \mathbf{C}$$

definida por $(\mathrm{T}, f) = a(1, \mathrm{T}f)$.

Teorema 6.6.1 *O pareamento de Miller estabelece uma dualidade perfeita entre os* $\mathbf{C}$*-espaços vetoriais* $\mathfrak{h}_k$ *e* S_k. *Em particular, temos*

$$S_k \cong \mathrm{Hom}_{\mathbf{C}}(\mathfrak{h}_k, \mathbf{C}).$$

Sob este isomorfismo, as autoformas normalizadas correspondem exatamente aos homomorfismos de $\mathbf{C}$*-álgebras* $\mathfrak{h}_k \longrightarrow \mathbf{C}$.

DEMONSTRAÇÃO: Seja, primeiro, $f \in S_k$ tal que $(\mathrm{T}, f) = 0$ para todo $\mathrm{T} \in \mathfrak{h}_k$. Então, para todo n,

$$a(n, f) = a(1, \mathrm{T}_n f) = (\mathrm{T}_n, f) = 0,$$

donde $f = 0$.

Reciprocamente, se $\mathrm{T} \in \mathfrak{h}_k$ satisfaz $(\mathrm{T}, f) = 0$ para toda $f \in s_k$, temos, para todo n e todo f,

$$a(n, \mathrm{T}f) = a(1, \mathrm{T}\mathrm{T}_n f) = (\mathrm{T}, \mathrm{T}_n f) = 0,$$

donde $\mathrm{T}f = 0$ para toda f, donde $\mathrm{T} = 0$. Isto estabelece que o pareamento é perfeito. A última afirmação é então imediata. □

Note que o homomorfismo correspondente a uma autoforma normalizada $f = \sum a(n)q^n$ é simplesmente $\mathrm{T}_n \mapsto a(n)$, isto é manda cada T_n no autovalor correspondente.

OBSERVAÇÃO: O pareamento de Miller é especialmente importante porque se generaliza. Se denotarmos por $S_k(\mathbf{Z}) = S_k(\Gamma, \mathbf{Z})$ o $\mathbf{Z}$-submódulo de S_k formado pelas formas modulares parabólicas cujas q-expansões têm coeficientes inteiros,

$$S_k(\mathbf{Z}) = \{f \in S_k : a(n, f) \in \mathbf{Z}, n = 1, 2, \cdots\}.$$

Note que $S_k(\mathbf{Z})$ é estável sob os T_n. Agora, se $\mathsf{h}_k(\mathbf{Z})$ é a $\mathbf{Z}$-subálgebra de $\mathrm{End}_{\mathbf{Z}}(S_k(\mathbf{Z}))$ gerada pelos T_n, temos:

Teorema 6.6.2 (Miller) *O pareamento*

$$\mathsf{h}_k(\mathbf{Z}) \times S_k(\mathbf{Z}) \longrightarrow \mathbf{Z}$$

dado por $(\mathrm{T}, f) = a(1, \mathrm{T}f)$ *estabelece uma dualidade perfeita de* $\mathbf{Z}$*-módulos, e em particular identifica as autoformas normalizadas de* $S_k(\mathbf{Z})$ *com os homomorfismos de* $\mathbf{Z}$*-álgebras de* $\mathsf{h}_k(\mathbf{Z})$ *em* $\mathbf{Z}$.

Mais geralmente, segue que uma autoforma $f = \sum a(n)q^n$ com coeficientes em alguma $\mathbf{Z}$-álgebra $\mathcal{O}$ corresponde a um homomorfismo de anéis $\mathsf{h}_k(\mathbf{Z}) \longrightarrow \mathcal{O}$, e que toda autoforma provém de um tal homomorfismo. Isto mostra que as autoformas parabólicas estão ligadas de perto à estrutura da "álgebra de Hecke" $\mathsf{h}_k(\mathbf{Z})$.

6.7 Exercícios

1) Prove as afirmações feitas no texto sobre a medida

$$\mu(f, g) = f(\tau)\overline{g(\tau)}y^k \frac{dxdy}{y^2}$$

e sobre o produto escalar de Petersson.

2) Sejam $f_1 = \Delta^2$ e $f_2 = \Delta E_{12}$. Já sabemos que f_1 e f_2 são uma base de $S_{24}(\Gamma)$. Verifique que f_1 e f_2 não são autoformas (para f_1 é imediato!). Use uma calculadora para determinar a matriz de T_2 em relação à base $\{f_1, f_2\}$; diagonalize essa matriz para obter uma base de autoformas para $S_{24}(\Gamma)$. (Explique por que basta diagonalizar T_2.)

3) Seja $f = \sum a_n q^n$ e seja $L(f,s) = \sum a_n n^{-s}$ sua L-série. Suponha que temos

$$L(f,s) = \prod (1 - a_p p^{-s} + p^{k-1-2s})^{-1}.$$

Mostre que f é uma autoforma.

6.8 Notas

Neste capítulo, nós nem mesmo tocamos na questão de como as definições precisam ser alteradas no caso de nível mais alto. Há dois casos: se $\text{mdc}(n, N) = 1$, o leitor verificará que exatamente a mesma definição acima dá certo, porque um ponto de ordem N em relação a Λ permanece de ordem N em relação a Λ' se $(\Lambda : \Lambda') = n$. Se, por outro lado, $\text{mdc}(m, n) > 1$, é preciso modificar a definição de T_n para somar apenas sobre aquelas subredes de índice n em relação às quais o ponto de ordem N dado permanece de ordem N. Isto não é muito difícil, e produz uma teoria completamente análoga à exposta acima (mas com complicações: multiplicidade um só vale para formas novas, por exemplo). Veja-se [Kob84] e [Lan76] para maiores detalhes.

Bibliografia

[AL70] A. O. L. Atkin and J. Lehner. "Hecke Operators on $\gamma_0(m)$". *Math. Ann.*, 185:134–160, 1970.

[Bo66] A. Borel et. al. *Seminar on Complex Multiplication.* Volume 21 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, Heidelberg, New York, 1966.

[DK73] P. Deligne and W. Kuijk, editors. *Modular Functions in One Variable II.* Volume 349 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[DR73] P. Deligne et M. Rapoport. "Les Schémas de Modules de Courbes Elliptiques". In P. Delígne and W. Kuijk, editors, *Modular Functions in One Variable II (SLN 349)*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[Fri22] R. Fricke. *Die Elliptischen Funktionen und ihre Anwendungen.* Teubner, Leipzig, Berlin, vol. 1, 1916, vol. 2, 1922.

[Hus87] Dale Husemöller. *Elliptic Curves.* Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[Igu72] J. I. Igusa. *Theta Functions.* Springer-Verlag, Berlin, Heidelberg, New York, 1972.

[Kat73] Nicholas M. Katz: "p-adic Properties of Modular Schemes and Modular Forms". In W. Kuijk and Jean-Pierre Serre, editors, *Modular Functions in One Variable III (SLN 350)*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[KM85] Nicholas M. Katz and Barry Mazur. *Arithmetic Moduli of Elliptic Curves.* Volume 108 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, New Jersey, 1985.

[Kob84] Neal Koblitz. *Introduction to Elliptic Curves and Modular Forms.* Springer-Verlag, Berlin, Heidelberg, New York, 1984.

[KS73] W. Kuijk and J-P. Serre, editors. *Modular Functions in One Variable III.* Volume 350 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[Kui73] W. Kuijk, editor. *Modular Functions in One Variable I.* Volume 320 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[Lan76] Serge Lang. *Introduction to Modular Forms.* Springer-Verlag, Berlin, Heidelberg, New York, 1976.

[Mum19] David Mumford. *Tata Lectures on Theta I, II.* Birkhäuser, Boston, 19??

[Nie75] D. Niebur. "A Formula for Ramanujan's τ-function". *Ill. J. Math.*, 19:448–449, 1975.

[Ogg69] Andrew Ogg. *Modular Forms and Dirichlet Series.* W. A. Benjamin, New York, 1969.

[Ser73] J.-P. Serre. *A Course in Arithmetic.* Springer-Verlag, 1973.

[Shi71] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Forms.* Princeton University Press, 1971.

Impresso na Gráfica do
impa