# 1 Dedekind's Eta Function and Modular Forms

## 1.1 Identities of Euler, Jacobi and Gauss

Throughout this monograph we use the notation

$$e(z) = e^{2\pi i z}$$

where $z$ is a complex number. We define the *Dedekind eta function* by the infinite product

$$\eta(z) = e\left(\tfrac{z}{24}\right) \prod_{n=1}^{\infty} (1 - q^n) \qquad \text{with} \qquad q = e(z). \tag{1.1}$$

The product *converges normally* for $q$ in the unit disc or, equivalently, for $z$ in the *upper half plane* $\mathbb{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$. This means that the product of the absolute values $|1 - q^n|$ converges uniformly for $z$ in every compact subset of $\mathbb{H}$. The normal convergence of the product implies that $\eta$ is a holomorphic function on $\mathbb{H}$ and that $\eta(z) \neq 0$ for all $z \in \mathbb{H}$.

Throughout this monograph, $\left(\frac{c}{d}\right)$ denotes the *Legendre–Jacobi–Kronecker symbol* of quadratic reciprocity. Its definition and properties, especially for an even denominator, can be found in many textbooks on Number Theory, for example [45], §5.3, or [49], §46. For the readers' convenience, we reproduce the definition. First of all, the symbol takes the value 0 whenever $\gcd(c, d) > 1$. If $d \neq 2$ is prime and $d \nmid c$ then $\left(\frac{c}{d}\right) = 1$ or $-1$ as to wether $c$ is or is not a square modulo $d$. (This is the *Legendre symbol*.) For $d = 2$ the definition reads

$$\left(\tfrac{c}{2}\right) = \begin{cases} 1 & \text{if} \quad c \equiv 1 \bmod 8, \\ -1 & \text{if} \quad c \equiv 5 \bmod 8, \end{cases}$$

while $\left(\frac{c}{2}\right)$ remains undefined if $c \equiv 3 \bmod 4$. This is the appropriate procedure in order to validate the decomposition law for primes in quadratic number fields which will be stated in Sect. 5.3. Finally, $\left(\frac{c}{d}\right)$ is totally mutiplicative as

G. Köhler, *Eta Products and Theta Series Identities*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-16152-0_1, 

a function of the denominator $d$, and it follows that it is totally multiplicative also as a function of the numerator $c$. We will frequently and silently use the *law of quadratic reciprocity*; we do not state it here, but refer to the textbooks.

Euler's identity

$$\prod_{n=1}^{\infty}(1-q^n) = \sum_{m=-\infty}^{\infty}(-1)^m q^{\frac{1}{2}m(3m-1)}$$

is easily transformed (see below in this subsection) into the series expansion

$$\eta(z) = \sum_{n=1}^{\infty}\left(\frac{12}{n}\right) e\left(\frac{n^2 z}{24}\right) \tag{1.2}$$

for the eta function. Euler succeeded to prove his identity in 1750. His proof rests on a tricky inductive argument and can be studied in [114], §98. Nowadays the Euler identity is commonly viewed as a special case of a more general identity, which Jacobi published in 1829 in his famous *Fundamenta Nova Theoriae Functionum Ellipticarum*. Proofs of this so-called triple product identity are given in [9], §1.3, [14], §3.1, [36], §2.8.1, [38], §17, [45], §12.4, [70], §3.2, [114], §100, and at other places.

**Theorem 1.1 (Jacobi Triple Product Identity)** *Suppose that $q, w \in \mathbb{C}$ and $|q| < 1$, $w \neq 0$. Then*

$$\prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}w)(1+q^{2n-1}w^{-1}) = \sum_{n=-\infty}^{\infty} q^{n^2} w^n.$$

We will present a proof of this identity because of its fundamental importance, although many proofs are available in textbooks. We join [9] and [70] and give a proof which is due to Andrews [4]. It is based upon another of Euler's identities (Chap. 16 of his *Introductio in Analysin Infinitorum*):

**Lemma 1.2 (Euler)** *For $q, w \in \mathbb{C}$ with $|q| < 1$ we have*

$$\prod_{n=0}^{\infty}(1+q^n w) = \sum_{m=0}^{\infty}\frac{q^{m(m-1)/2}w^m}{(1-q)(1-q^2)\dots(1-q^m)}. \tag{1.3}$$

*If also $|w| < 1$, then*

$$\prod_{n=0}^{\infty}\frac{1}{1+q^n w} = \sum_{m=0}^{\infty}\frac{(-1)^m w^m}{(1-q)(1-q^2)\dots(1-q^m)}. \tag{1.4}$$

*Proof.* The infinite product

$$f(q,w) \;=\; \prod_{n=0}^{\infty}(1+q^n w)$$

converges absolutely for $|q|<1$ and any $w\in\mathbb{C}$ because of the convergence of $\sum_{n=0}^{\infty}|q^n w|$. Therefore for any $q$ with $|q|<1$ there is a power series expansion

$$f(q,w) \;=\; \sum_{m=0}^{\infty} a_m(q) w^m$$

which is valid on the entire $w$-plane. The definition of $f$ clearly implies that $f(q,w)=(1+w)f(q,qw)$, hence

$$\sum_{m=0}^{\infty} a_m(q)w^m \;=\; \sum_{m=0}^{\infty} a_m(q)q^m w^m \;+\; \sum_{m=0}^{\infty} a_m(q)q^m w^{m+1}\,.$$

Comparing coefficients yields $a_m(q)=a_m(q)q^m+a_{m-1}(q)q^{m-1}$ for $m\geq 1$, or

$$a_m(q)=a_{m-1}(q)q^{m-1}(1-q^m)^{-1}\,.$$

Since $a_0(q)=1$, it follows by induction that

$$a_m(q) \;=\; \frac{q^{(m-1)+(m-2)+\ldots+1}}{(1-q)(1-q^2)\ldots(1-q^m)} \;=\; \frac{q^{m(m-1)/2}}{(1-q)(1-q^2)\ldots(1-q^m)}\,.$$

Thus the result (1.3) follows.

Now we consider

$$g(q,w) \;=\; \prod_{n=0}\frac{1}{1+q^n w}\,.$$

For $|q|<1$, $|w|<1$ this product converges absolutely because of the convergence of

$$\sum_{n=0}^{\infty}\left|1-\frac{1}{1+q^n w}\right| \;=\; \sum_{n=0}^{\infty}\left|\frac{q^n w}{1+q^n w}\right| \;\leq\; \frac{|w|}{1-|w|}\sum_{n=0}^{\infty}|q^n w|\,.$$

Therefore for any $q$ with $|q|<1$, $g$ is an analytic function of $w$ with a power series expansion $g(q,w)=\sum_{m=0}^{\infty}b_m(q)w^m$ which is valid for $|w|<1$. The definition of $g$ implies that $g(q,qw)=(1+w)g(q,w)$, and hence

$$\sum_{m=0}^{\infty} b_m(q)q^m w^m \;=\; \sum_{m=0}^{\infty} b_m(q)w^m \;+\; \sum_{m=0}^{\infty} b_m(q)w^{m+1}\,.$$

We conclude that $b_m(q)q^m=b_m(q)+b_{m-1}(q)$, or $b_m(q)=-b_{m-1}(q)/(1-q^m)$ for $m\geq 1$. Since $b_0(q)=1$, we obtain by induction that

$$b_m(q) \;=\; \frac{(-1)^m}{(1-q)(1-q^2)\ldots(1-q^m)}\,,$$

and the result (1.4) follows. □

*Proof* of Theorem 1.1. Assume that $|q| < 1$ and $w \in \mathbb{C}$. From (1.3) we obtain

$$\begin{aligned}
\prod_{n=0}^{\infty}(1+q^{2n+1}w) &= \prod_{n=0}^{\infty}\big(1+(q^2)^n(qw)\big) \\
&= \sum_{m=0}^{\infty}\frac{q^{2m(m-1)/2}q^m w^m}{(1-q^2)(1-q^4)\dots(1-q^{2m})} \\
&= \sum_{m=0}^{\infty}\frac{q^{m^2}w^m}{(1-q^2)(1-q^4)\dots(1-q^{2m})} \\
&= \sum_{m=0}^{\infty}q^{m^2}w^m\prod_{\nu=0}^{\infty}(1-q^{2m+2+2\nu})\Big/\prod_{\nu=0}^{\infty}(1-q^{2\nu+2}) \\
&= \prod_{\nu=0}^{\infty}\frac{1}{1-q^{2\nu+2}}\sum_{m=0}^{\infty}q^{m^2}w^m\prod_{\nu=0}^{\infty}(1-q^{2m+2+2\nu})\,.
\end{aligned}$$

For $m < 0$ the product inside the infinite sum is identically 0 because of the factor with $\nu = -m-1$. Therefore we can write

$$\prod_{n=0}^{\infty}(1+q^{2n+1}w) = \prod_{\nu=0}^{\infty}\frac{1}{1-q^{2\nu+2}}\sum_{m=-\infty}^{\infty}q^{m^2}w^m\prod_{\nu=0}^{\infty}(1-q^{2m+2+2\nu})\,.$$

Applying (1.3) once more, we get

$$\begin{aligned}
\prod_{\nu=0}^{\infty}(1-q^{2m+2+2\nu}) &= \prod_{\nu=0}^{\infty}\big(1+(q^2)^\nu(-q^{2+2m})\big) \\
&= \sum_{k=0}^{\infty}\frac{q^{k(k-1)}(-q^{2+2m})^k}{(1-q^2)(1-q^4)\dots(1-q^{2k})} \\
&= \sum_{k=0}^{\infty}\frac{(-1)^k q^{k^2+k+2mk}}{(1-q^2)(1-q^4)\dots(1-q^{2k})}\,.
\end{aligned}$$

Together with the preceding result this yields

$$\prod_{n=0}^{\infty}(1+q^{2n+1}w) = \prod_{\nu=0}^{\infty}\frac{1}{1-q^{2\nu+2}}\sum_{m=-\infty}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^k q^{m^2+k^2+2mk+k}w^m}{(1-q^2)(1-q^4)\dots(1-q^{2k})}\,.$$

We want to interchange the summation in the double sum, and for this purpose we need absolute convergence. We have convergence for all $w \in \mathbb{C}$. But an estimate of the double sum in reversed order of summation shows that absolute convergence does only hold if $|q| < 1$ and $|w| > |q|$. Under this assumption we get

$$\begin{aligned}
&\prod_{n=0}^{\infty}(1+q^{2n+1}w)\\
&\quad=\prod_{\nu=0}^{\infty}\frac{1}{1-q^{2\nu+2}}\sum_{k=0}^{\infty}\frac{(-1)^k q^k}{(1-q^2)(1-q^4)\dots(1-q^{2k})}\sum_{m=-\infty}^{\infty}q^{(m+k)^2}w^m\\
&\quad=\Bigg(\sum_{m=-\infty}^{\infty}q^{m^2}w^m\Bigg)\prod_{\nu=0}^{\infty}\frac{1}{1-q^{2\nu+2}}\sum_{k=0}^{\infty}\frac{(-1)^k (q/w)^k}{(1-q^2)(1-q^4)\dots(1-q^{2k})}\,.
\end{aligned}$$

Since by assumption $|q/w|<1$, we can apply (1.4) to the inner sum on $k$ and replace it by the product

$$\prod_{n=0}^{\infty}\frac{1}{1+(q^2)^n(q/w)}\,.$$

This yields the Triple Product Identity

$$\sum_{m=-\infty}^{\infty}q^{m^2}w^m=\prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}w)(1+q^{2n-1}w^{-1})$$

under the assumptions that $|q|<1$ and $|w|>|q|$. By the principle of analytic continuation it holds for $|q|<1$ and all $w\neq 0$. □

**Corollary 1.3 (Euler, Gauss)** *For $q\in\mathbb{C}$, $|q|<1$ and $m\in\mathbb{N}$ the following identities hold:*

$$\begin{aligned}
&\prod_{n=1}^{\infty}(1-q^{n(m+1)})(1-q^{n(m+1)-m})(1-q^{n(m+1)-1})\\
&\qquad=\sum_{n=-\infty}^{\infty}(-1)^n q^{\frac{1}{2}n(n(m+1)-m+1)},\\
&\prod_{n=1}^{\infty}(1-q^n)=\sum_{n=-\infty}^{\infty}(-1)^n q^{\frac{1}{2}n(3n-1)},\\
&\prod_{n=1}^{\infty}(1-q^n)^2(1-q^{2n})^{-1}=\sum_{n=-\infty}^{\infty}(-1)^n q^{n^2}.
\end{aligned}$$

*Proof.* In Theorem 1.1 we replace $q$ by $q^{\frac{1}{2}(m+1)}$, and we put $w=-q^{\frac{1}{2}(1-m)}$. This gives the first identity. When we choose $m=2$ then we get the second, which is Euler's identity. Now we choose $m=1$ in the first identity. Then the left hand side is

$$\prod_{n=1}^{\infty}(1-q^{2n})(1-q^{2n-1})(1-q^{2n-1})=\prod_{n=1}^{\infty}(1-q^n)(1-q^{2n-1}),$$

since $2n$ and $2n-1$ together take each positive integer once as a value. We multiply and divide each factor by $1-q^{2n}$. This yields the last identity. □

The third identity in Corollary 1.3 is attributed to Gauss. The right hand side in the triple product identity is the famous *Jacobi theta function* which is traditionally denoted by $\theta(q,w)$, $\theta_3(q,w)$, or by $\theta(z,u)$, $\theta_3(z,u)$ if $q=e(z/2)$, $w=e(u)$.

In order to derive (1.2), we multiply Euler's identity by $q^{1/24}$ and observe that

$$\tfrac{1}{24}+\tfrac{1}{2}n(3n-1)=\tfrac{1}{24}(36n^2-12n+1)=\tfrac{1}{24}(6n-1)^2.$$

We put $6n-1=m$ for $n>0$, $6n-1=-m$ for $n\le 0$. Then $m>0$ for all $n$ and

$$(-1)^n=\chi(m)=\begin{cases} 1 \\ -1 \end{cases} \quad \text{for} \quad m\equiv\begin{cases} \pm 1 \\ \pm 5 \end{cases} \bmod 12.$$

Hence $\chi(m)=\left(\frac{12}{m}\right)$ for $\gcd(m,12)=1$. Since $\left(\frac{12}{m}\right)=0$ for $\gcd(m,12)>1$, we arrive at the series expansion (1.2) for $\eta(z)$.

We put $q=e(z)$ in the third identity in Corollary 1.3. Then we get

$$\frac{\eta^2(z)}{\eta(2z)}=\sum_{n=-\infty}^{\infty}(-1)^n e\left(n^2z\right). \tag{1.5}$$

The coefficient function $\chi(m)=\left(\frac{12}{m}\right)$ in (1.2) is a Dirichlet character modulo 12. In fact, it is the only primitive character among the four characters modulo 12.

We recall that a *Dirichlet character* modulo $N$ is a homomorphism $\chi$ of the group $(\mathbb{Z}/N\mathbb{Z})^\times$ of coprime residues modulo $N$ into the multiplicative group $\mathbb{C}^\times$ of complex numbers. It is lifted to a function $\chi$ on $\mathbb{Z}$ by putting $\chi(m)=\chi(m \bmod N)$ if $\gcd(m,N)=1$ and $\chi(m)=0$ if $\gcd(m,N)>1$. We say that $\chi$ is *induced* by a character $\psi$ modulo a divisor $N_0$ of $N$ if $\chi(m)=\psi(m)$ whenever $\gcd(m,N_0)=1$. The smallest $N_0$ such that $\chi$ is induced by a character modulo $N_0$ is called the *conductor* of $\chi$. If the conductor is $N$ then $\chi$ is called *primitive*; otherwise it is called *imprimitive*.

**Corollary 1.4 (Jacobi)** *For $q\in\mathbb{C}$, $|q|<1$ we have*

$$\prod_{n=1}^{\infty}(1-q^n)^3=\sum_{n=0}^{\infty}(-1)^n(2n+1)q^{\frac{1}{2}n(n+1)}. \tag{1.6}$$

*The third power of the eta function has the expansion*

$$\eta^3(z)=\sum_{n=1}^{\infty}\left(\frac{-1}{n}\right)ne\left(\frac{n^2z}{8}\right). \tag{1.7}$$

*Proof* ([114], §102, or [101]). In Theorem 1.1 we put $q = \sqrt{uv}$ and $w = -\sqrt{u/v}$. This yields

$$\prod_{n=1}^{\infty}(1-u^n v^n)(1-u^n v^{n-1})(1-u^{n-1}v^n) = \sum_{n=-\infty}^{\infty}(-1)^n u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}n(n-1)}, \tag{1.8}$$

valid for $|uv| < 1$, $u \neq 0$, $v \neq 0$. (We start from a small region where holomorphic square roots exist, and then argue by analytic continuation.) In (1.8) we divide by $1-v$. For the left hand side this simply means that we drop the third factor in the term with $n = 1$. On the right hand side we combine, for any $n \geq 0$, the terms with $n$ and $-n-1$, which gives

$$\begin{aligned}(-1)^n &u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}n(n-1)} + (-1)^{n+1} u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}(n+1)(n+2)} \\ &= (-1)^n u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}n(n-1)}(1-v^{2n+1}) \\ &= (-1)^n u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}n(n-1)}(1-v)\sum_{k=0}^{2n} v^k.\end{aligned}$$

Therefore the division yields

$$\begin{aligned}\prod_{n=1}^{\infty}&(1-u^n v^n)(1-u^n v^{n-1})(1-u^n v^{n+1}) \\ &= \sum_{n=0}^{\infty}(-1)^n u^{\frac{1}{2}n(n+1)} v^{\frac{1}{2}n(n-1)} \sum_{k=0}^{2n} v^k.\end{aligned}$$

Here we put $v = 1$ and write $q$ instead of $u$. This gives us the identity (1.6). We multiply (1.6) by $q^{1/8}$, put $q = e(z)$, and observe that $\frac{1}{8} + \frac{1}{2}n(n+1) = \frac{1}{8}(2n+1)^2$. So we arrive at

$$\eta^3(z) = \sum_{n=0}^{\infty}(-1)^n(2n+1)e\left(\frac{(2n+1)^2 z}{8}\right), \tag{1.9}$$

the Jacobi identity for $\eta^3$ in its usual notation. We observe that $(-1)^n = \left(\frac{-1}{2n+1}\right)$ is a quadratic residue symbol. Thus we arrive at (1.7) by the observation that $\left(\frac{-1}{n}\right) = 0$ if $n$ is even. □

*Remarks.* The coefficient function $n \mapsto \left(\frac{-1}{n}\right)$ in (1.7) is the primitive Dirichlet character modulo 4.—When we replace $v$ by 1 in (1.8) then both sides are 0. Thus the replacement gives a useful result only after division by $1-v$. Similarly, one might try to use Theorem 1.1 directly, replacing $q$ by $q^{1/2}$ and $w$ by $-q^{1/2}$. But then too, both sides become 0. Nevertheless, a refinement of this idea yields a proof of (1.6); see [70], §3.2.

## 1.2 The Sign Transform

The map $q \mapsto -q$, applied to a Laurent series or product in the variable $q$, will be called the *sign transform*, after Zucker [142]. For $q = e(z)$ the sign transform corresponds to the translation

$$z \mapsto z + \tfrac{1}{2}$$

of the upper half plane. Zucker succeeded to deduce new identities from known ones in a completely elementary way by means of the sign transform. We give two examples:

**Proposition 1.5** *For $z$ in the upper half plane we have*

$$\eta\big(z+\tfrac{1}{2}\big) = e\big(\tfrac{1}{48}\big)\frac{\eta^3(2z)}{\eta(z)\eta(4z)}\,, \qquad \frac{\eta^3(2z)}{\eta(z)\eta(4z)} = \sum_{n=1}^{\infty}\left(\frac{6}{n}\right) e\left(\frac{n^2 z}{24}\right), \tag{1.10}$$

$$\frac{\eta^5(2z)}{\eta^2(z)\eta^2(4z)} = \sum_{n=-\infty}^{\infty} e\left(n^2 z\right). \tag{1.11}$$

*Proof.* The product expansion for $\eta(z)$ gives

$$\begin{aligned} e\big(-\tfrac{1}{48}\big)\eta\big(z+\tfrac{1}{2}\big) &= e\big(\tfrac{z}{24}\big)\prod_{n=1}^{\infty}\left(1-(-q)^n\right) \\ &= e\big(\tfrac{z}{24}\big)\prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}) \\ &= e\big(\tfrac{z}{24}\big)\prod_{n=1}^{\infty}\frac{(1-q^{2n})(1-q^{4n-2})}{1-q^{2n-1}} \\ &= e\big(\tfrac{z}{24}\big)\prod_{n=1}^{\infty}\frac{(1-q^{2n})^2(1-q^{4n-2})(1-q^{4n})}{(1-q^n)(1-q^{4n})} \\ &= e\big(\tfrac{z}{24}\big)\prod_{n=1}^{\infty}\frac{(1-q^{2n})^3}{(1-q^n)(1-q^{4n})} \\ &= \frac{\eta^3(2z)}{\eta(z)\eta(4z)}\,. \end{aligned}$$

On the other hand, the series expansion yields

$$\begin{aligned} e\big(-\tfrac{1}{48}\big)\eta\big(z+\tfrac{1}{2}\big) &= \sum_{n=1}^{\infty} e\big(-\tfrac{1}{48}\big)\left(\frac{12}{n}\right) e\left(\frac{n^2 z}{24}+\frac{n^2}{48}\right) \\ &= \sum_{n=1}^{\infty} e\left(\frac{n^2-1}{48}\right)\left(\frac{12}{n}\right) e\left(\frac{n^2 z}{24}\right) \\ &= \sum_{n=1}^{\infty}\left(\frac{6}{n}\right) e\left(\frac{n^2 z}{24}\right). \end{aligned}$$

This proves (1.10). In the last line we used

$$\left(\frac{2}{n}\right) = (-1)^{(n^2-1)/8} = e\left(\frac{n^2-1}{16}\right) = e\left(\frac{n^2-1}{48}\right) \quad \text{for} \quad \gcd(n,6)=1.$$

Now we take the sign transform of the Gauss identity (1.5). We plug in (1.10) and observe that the denominator is transformed into $\eta(2z+1) = e\left(\frac{1}{24}\right)\eta(2z)$. So we get

$$\frac{\eta^5(2z)}{\eta^2(z)\eta^2(4z)} = \sum_{n=-\infty}^{\infty} (-1)^n\, e\big(n^2 z + \tfrac{n^2}{2}\big) = \sum_{n=-\infty}^{\infty} e(n^2 z)\,.$$

Thus we have proved (1.11). □

We remark that the right hand side in (1.11) is traditionally called a *Theta-nullwert* and denoted by $\theta(2z)$ or $\theta_3(2z)$. With the notation explained after Corollary 1.3, we have $\theta(2z) = \theta(2z, 0)$. From (1.10) and (1.11) we deduce

$$\theta(z) = \frac{\eta^2\left(\frac{z+1}{2}\right)}{\eta(z+1)} = 1 + 2\sum_{n=1}^{\infty} e^{\pi i n^2 z}.$$

(See also [70], §3.4.)—Another example of a Zucker identity comes from Jacobi's identity for $\eta^3$:

**Proposition 1.6** *For $z$ in the upper half plane we have*

$$\frac{\eta^9(2z)}{\eta^3(z)\eta^3(4z)} = \sum_{n=1}^{\infty}\left(\frac{-2}{n}\right) n\, e\left(\frac{n^2 z}{8}\right). \tag{1.12}$$

*Proof.* In Jacobi's identity (1.7) we take the sign transform, use (1.10), and observe that $\left(\frac{-1}{n}\right) e\big(\frac{n^2-1}{16}\big) = \left(\frac{-2}{n}\right)$. □

The identity (1.12) is contained in Zucker's lists in an equivalent form (items (24) in [141] and (T4.8) in [142]). It was also proved in a more complicated way in [77].

## 1.3 The Multiplier System of $\eta$

The transformation formula

$$\eta(z+1) = e\big(\tfrac{1}{24}\big)\eta(z) \tag{1.13}$$

follows trivially from the definition of the eta function as a product or as a series. (We used it already in the proof of (1.11).) Not at all trivial is the transformation formula

$$\eta\left(-\frac{1}{z}\right) = \sqrt{-iz}\,\eta(z), \tag{1.14}$$

where the square root of $-iz$ is the holomorphic function on the upper half plane which takes positive values for $z = iy$, $y > 0$. There is a rich literature on (1.14) and its proofs. It is partly listed in the references for Appendix D in [110]. Three proofs are given in Apostol [5], §3. Weil [138] reduced (1.14) to a functional equation of a corresponding Dirichlet series; his proof is reproduced in [96], §4.4. In [70], §3.3, Knopp deduces (1.14) from the Poisson summation formula and a theta transformation formula. Here we will sketch Siegel's one-page proof [134] which is based on a skillful application of the calculus of residues:

*Sketch of a proof* for (1.14). By the principle of analytic continuation it suffices to prove (1.14) for $z = yi$ with $y > 0$. The assertion will follow from

$$\log\eta(i/y) \;-\; \log\eta(yi) \;=\; \frac{1}{2}\log y.$$

Taking the logarithm of an infinite product, we obtain

$$\begin{aligned}\log\eta(yi) &= -\frac{\pi y}{12} + \sum_{n=1}^{\infty}\log(1-e^{-2\pi n y}) = -\frac{\pi y}{12} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{e^{-2\pi m n y}}{m} \\ &= -\frac{\pi y}{12} + \sum_{m=1}^{\infty}\frac{1}{m}\,\frac{1}{1-e^{2\pi m y}}\,.\end{aligned}$$

Therefore it suffices to prove that

$$\sum_{m=1}^{\infty}\frac{1}{m}\,\frac{1}{1-e^{2\pi m y}} \;-\; \sum_{m=1}^{\infty}\frac{1}{m}\,\frac{1}{1-e^{2\pi m/y}} \;-\; \frac{\pi}{12}\left(y-\frac{1}{y}\right) \;=\; -\,\frac{1}{2}\,\log y\,. \tag{1.15}$$

For fixed $y > 0$ we consider the sequence of meromorphic functions

$$f_n(w) \;=\; -\,\frac{1}{8w}\,\cot(\pi i N w)\,\cot(\pi N w/y) \qquad \text{with} \qquad n\in\mathbb{N},\;\; N = n+\frac{1}{2}\,.$$

Let $C$ be the contour of the parallelogram with vertices $y$, $i$, $-y$, $-i$ in that order. Inside $C$, the function $f_n$ has simple poles at $w = \frac{mi}{N}$ and at $w = \frac{my}{N}$ for $m \in \mathbb{Z}$, $1 \le |m| \le n$, and there is a triple pole at $w = 0$ with residue $\frac{i}{24}\left(y - y^{-1}\right)$. The residues of $f_n$ at $\frac{mi}{N}$ and at $\frac{my}{N}$ are

$$\frac{1}{8m\pi}\cot(\pi i m/y) = \frac{1}{8m\pi i}\left(1-\frac{2}{1-e^{2\pi m/y}}\right)$$

and

$$-\,\frac{1}{8m\pi}\cot(\pi i m y) = -\,\frac{1}{8m\pi i}\left(1-\frac{2}{1-e^{2\pi m y}}\right),$$

respectively. Using that these expressions are even functions of $m$, we observe that the $2\pi i$-fold sum of the residues of $f_n(w)$ inside $C$ is equal to the left hand side in (1.15), where the summation is restricted to $1 \le m \le n$. On the other hand, by the residue theorem this sum is equal to the contour integral of $f_n$ along $C$. Therefore, in order to complete the proof, it suffices to show that

$$\lim_{n\to\infty} \int_C f_n(w)\,dw = -\frac{1}{2}\log y.$$

On the edges of $C$, except at the vertices, the functions $w\,f_n(w)$ have, as $n \to \infty$, the limit $\frac{1}{8}$ on the edges connecting $y$, $i$ and $-y$, $-i$, and the limit $-\frac{1}{8}$ on the other two edges. A closer inspection shows that the functions $f_n(w)$ are bounded on $C$ uniformly with respect to $n$ (because of $y > 0$ and $N = n + \frac{1}{2}$). Therefore we can use the bounded convergence theorem and interchange integration with taking the limit. We get

$$\begin{aligned}\lim_{n\to\infty} \int_C f_n(w)\,dw &= \int_C \Big(\lim_{n\to\infty} w\,f_n(w)\Big)\frac{dw}{w} = \frac{1}{4}\left(\int_y^i \frac{dw}{w} - \int_{-i}^y \frac{dw}{w}\right)\\ &= \tfrac{1}{4}\Big(\big(\tfrac{\pi i}{2} - \log y\big) - \big(\log y + \tfrac{\pi i}{2}\big)\Big)\\ &= -\tfrac{1}{2}\log y. \quad \square\end{aligned}$$

It is well-known that the matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

generate the (*homogeneous*) *modular group* $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$. Correspondingly, the Möbius transformations $T : z \mapsto z+1$ and $S : z \mapsto -\frac{1}{z}$ of the upper half plane generate the (*inhomogeneous*) *modular group* which we also denote by $\Gamma_1$ and which consists of all transformations $z \mapsto L(z) = \frac{az+b}{cz+d}$ with $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. The relations (1.13) and (1.14) are transformation formulae for $\eta(z)$ with respect to the generators $T$ and $S$ of $\Gamma_1$. They can be written as $\eta(Tz) = e\left(\frac{1}{24}\right)\eta(z)$ and $\eta(Sz) = e\left(-\frac{1}{8}\right)\sqrt{z}\,\eta(z)$, where the holomorphic branch of $\sqrt{z}$ is fixed by $\sqrt{i} = e\left(\frac{1}{8}\right)$. One can verify directly or deduce from the chain rule that the function $J : \left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), z\right) \mapsto cz + d$ satisfies $J(L_1L_2, z) = J(L_1, L_2z)J(L_2, z)$ for all Möbius transformations $L_1, L_2 \in \mathrm{SL}_2(\mathbb{R})$ of the upper half plane. It follows that the eta function satisfies the relations

$$\eta(Lz) = v_\eta(L)(cz+d)^{1/2}\,\eta(z) \qquad \text{for all} \qquad L = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \tag{1.16}$$

with factors $v_\eta(L)$ depending only on $L$ and not on the variable $z$. We will describe them explicitly, but before doing so it is necessary to agree on a convention for square roots and, more generally, for powers with a real exponent.

We fix an *argument* of $z$ for $z \in \mathbb{C}$, $z \neq 0$ by

$$-\pi \leq \arg(z) < \pi.$$

Then for $r \in \mathbb{R}$ we put

$$z^r = |z|^r e^{ir\arg(z)}$$

where, of course, $|z|^r > 0$. In particular we have $\sqrt{z} = \sqrt{|z|}e^{i\arg(z)/2}$. This convention will be used for (1.16). It implies $z^r z^s = z^{r+s}$. But $z^r w^r = (zw)^r$ does not hold in general.

The function $L \mapsto v_\eta(L)$ is called the *multiplier system* of the eta function. Its values $v_\eta(T) = e\left(\frac{1}{24}\right)$, $v_\eta(S) = e\left(-\frac{1}{8}\right)$ for the generators of the modular group are 24th roots of unity. It follows that $v_\eta(L)$ is a 24th root of unity for every $L \in \mathrm{SL}_2(\mathbb{Z})$. The determination of these roots of unity is an important issue in the theory of the eta function. A formula for $v_\eta(L)$ was first given by Rademacher [113] in 1931. He expressed $v_\eta(L)$ in terms of Dedekind sums which can be evaluated recursively; see also Chap. 9 of his book [114]. In 1954, Petersson [109] gave a formula which can be evaluated directly, without a recursive process. It is contained in his book [110], entry (4.14). A similar explicit formula is given by Rademacher in [114], §74. We begin with an example which shows that $v_\eta$ is not a homomorphism on $\mathrm{SL}_2(\mathbb{Z})$: Since $S^2 = -1_2$ is the negative of the $2 \times 2$ unit matrix and operates as the identity on the upper half plane, and since $\sqrt{-1} = e^{-i\pi/2} = -i$ by our convention on roots, we obtain

$$\eta(z) = \eta((-1_2)(z)) = v_\eta(-1_2) \cdot (-i) \cdot \eta(z),$$

and hence $v_\eta(-1_2) = i$. Therefore we get $v_\eta(S^2) = i \neq -i = (v_\eta(S))^2$.— For Petersson's formula we need some notation which extends the symbol of quadratic reciprocity:

**Notation** Let $c$ and $d$ be integers such that $\gcd(c,d) = 1$, $d$ is odd and $c \neq 0$. Let $\mathrm{sgn}(x) = \frac{x}{|x|}$ be the sign of a real number $x \neq 0$. Then we put

$$\left(\frac{c}{d}\right)^* = \left(\frac{c}{|d|}\right) \qquad \text{and} \qquad \left(\frac{c}{d}\right)_* = \left(\frac{c}{|d|}\right) \cdot (-1)^{\frac{1}{4}(\mathrm{sgn}(c)-1)(\mathrm{sgn}(d)-1)}.$$

Furthermore, we put

$$\left(\frac{0}{1}\right)^* = \left(\frac{0}{-1}\right)^* = 1, \qquad \left(\frac{0}{1}\right)_* = 1, \qquad \left(\frac{0}{-1}\right)_* = -1.$$

Now we reproduce Petersson's formula, following Knopp [70], §4.1:

**Theorem 1.7** *For*

$$L = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

*the multiplier system of the eta function is given by*

$$v_\eta(L) = \left(\frac{d}{c}\right)^{*} e\big(\tfrac{1}{24}((a+d)c - bd(c^2-1) - 3c)\big) \qquad \textit{if } c \textit{ is odd,}$$

$$v_\eta(L) = \left(\frac{c}{d}\right)_{*} e\big(\tfrac{1}{24}((a+d)c - bd(c^2-1) + 3d - 3 - 3cd)\big) \qquad \textit{if } c \textit{ is even.}$$

## 1.4 The Concept of Modular Forms

The relations (1.16) say that $\eta(z)$ is a modular form of weight $\frac{1}{2}$ for the modular group $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$. We will use the concept of a modular form mainly for integral weights and for certain congruence subgroups of the modular group. Nevertheless it is necessary to define a more comprehensive concept, since we encountered $\eta(z)$, $\theta(z)$ and $\eta^3(z)$ with half-integral weights, and since we will meet the Fricke groups which are not subgroups of the modular group.

**Definition.** Two subgroups $\Gamma, \widetilde{\Gamma}$ of $\mathrm{SL}_2(\mathbb{R})$ are called *commensurable* if their intersection $\Gamma \cap \widetilde{\Gamma}$ has finite index both in $\Gamma$ and in $\widetilde{\Gamma}$.—Recall that every element $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ acts as a Möbius transformation $z \mapsto Lz = \frac{az+b}{cz+d}$ on the upper half plane $\mathbb{H}$.

**Definition.** Let $\Gamma$ be a subgroup of $\mathrm{SL}_2(\mathbb{R})$ which is commensurable with the modular group $\Gamma_1$, and let $k$ be a real number. A function $f : \mathbb{H} \to \mathbb{C}$ is called a *modular form* of *weight* $k$ and *multiplier system* $v$ for $\Gamma$ if $f$ is holomorphic on $\mathbb{H}$ and has the following two properties:

(1) The relation

$$f(Lz) = f\left(\frac{az+b}{cz+d}\right) = v(L)(cz+d)^k f(z)$$

holds for every $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$. Here, the complex numbers $v(L)$ satisfy $|v(L)| = 1$ and do not depend on the variable $z$, and the powers $(cz+d)^k$ are defined according to the convention in Sect. 1.3.

(2) The function $f$ is holomorphic at all *cusps* $r \in \mathbb{Q} \cup \{\infty\}$.—The meaning of this condition will be explained immediately.

We begin to explain property (2) for the cusp $\infty$. Since $\Gamma$ is commensurable with $\Gamma_1$, there is a positive integer $h$ for which $T^h = \left(\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$. We may assume that $h$ is chosen minimal with this property. From (1) we obtain

$$f(z+h) = v(T^h) f(z).$$

We write $v(T^h) = e(\kappa) = e^{2\pi i\kappa}$ with $0 \le \kappa < 1$. The integer $h$ is called the *width* of $\Gamma$ at the cusp $\infty$, and the number $\kappa$ is called the *cusp parameter* (according to Rankin [117]) or the *Drehrest* (according to Petersson [110]) of $f$ at $\infty$. It follows that $g(z) = e^{-2\pi i\kappa z} f(hz)$ is a holomorphic function with period 1 on the upper half plane. Hence it can be written as a holomorphic function of the variable $q = e(z)$ in the punctured unit disc, which henceforth has a Laurent expansion valid for $0 < |q| < 1$. For $f$ itself we obtain a Fourier expansion of the form

$$f(z) = e^{2\pi i\kappa z/h} \sum_n c(n) e\left(\frac{nz}{h}\right) = \sum_n c(n) e\left(\frac{(n+\kappa)z}{h}\right), \tag{1.17}$$

where the summation is on all $n \in \mathbb{Z}$. The function $f$ is called *holomorphic at* the cusp $\infty$ if powers of $e(z/h)$ with negative exponents do not occur in (1.17), i.e., if $c(n) \ne 0$ implies that $n + \kappa \ge 0$.

Now we consider cusps $r \in \mathbb{Q}$. We write $r = \frac{a}{c}$ with $\gcd(a, c) = 1$. Then

$$r = A(\infty) \qquad \text{with some} \qquad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1 = \mathrm{SL}_2(\mathbb{Z}).$$

Since the conjugate group $A^{-1}\Gamma A$ is commensurable with $A^{-1}\Gamma_1 A = \Gamma_1$, there exists a smallest integer $h > 0$ for which $T^h \in A^{-1}\Gamma A$. The element $L = AT^hA^{-1} \in \Gamma$ fixes the point $r$. We write $L = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ and put $v(L) = e^{2\pi i\kappa}$ with $0 \le \kappa < 1$. As before, $h$ is called the *width* of $\Gamma$ at the cusp $r$, and $\kappa$ is the *cusp parameter* or *Drehrest* of $f$ at $r$. Because of (1) the function

$$\varphi(z) = (z - r)^k f(z)$$

satisfies

$$\varphi(Lz) = (Lz - r)^k f(Lz) = (Lz - r)^k (\gamma z + \delta)^k e^{2\pi i\kappa} f(z).$$

Elementary calculation yields

$$L = \begin{pmatrix} 1 - ach & a^2h \\ -c^2h & 1 + ach \end{pmatrix} \qquad \text{and} \qquad L(z) - r = \frac{z - r}{\gamma z + \delta}\,.$$

Since $L(z) - r$ and $z - r$ both belong to $\mathbb{H}$, their arguments are in the interval from 0 to $\pi$. Hence the difference of the arguments is in the interval from $-\pi$ to $\pi$ where all arguments have to be chosen by the convention from Sect. 1.3. Therefore in this particular situation we get

$$(Lz - r)^k (\gamma z + \delta)^k = ((Lz - r)(\gamma z + \delta))^k = (z - r)^k.$$

It follows that

$$\varphi(Lz) = (z - r)^k e^{2\pi i\kappa} f(z), \qquad \varphi(AT^hA^{-1}z) = e^{2\pi i\kappa}\varphi(z).$$

With $Az$ instead of $z$ we get

$$\varphi(AT^h z) = e^{2\pi i\kappa}\varphi(Az).$$

Now it is easy to verify that the holomorphic function

$$g(z) = e^{-2\pi i\kappa z}\varphi(A(hz))$$

has period 1, and hence can be expanded in a Laurent series in the variable $q = e(z)$ which is valid for $0 < |q| < 1$. Rewriting it for the function $f(z)$, we obtain an expansion of the form

$$f(z) = (z-r)^{-k}\sum_n c(n)e\left(\frac{(n+\kappa)A^{-1}(z)}{h}\right), \tag{1.18}$$

valid for $z \in \mathbb{H}$, with summation over all $n \in \mathbb{Z}$. It is called the *Fourier expansion* of $f$ at the cusp $r$. As before, $f$ is called *holomorphic at* the cusp $r$ if $c(n) \neq 0$ implies that $n + \kappa \geq 0$. It can be shown that this condition is independent of the choice of the matrix $A$ in $\Gamma_1$ which sends $r$ to $\infty$.—So finally, we have explained the meaning of the requirement (2) on modular forms.

At this point a remark on the multiplier system $v$ of a modular form is in order. We use the notation $J(L, z) = cz + d$ for $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ which was introduced in Sect. 1.3. Suppose that there exists a function $f$ which satisfies (1) and is not identically 0. Then it is easy to prove that

$$v(L_1L_2)\, J(L_1L_2, z)^k = v(L_1)v(L_2)\, J(L_1, L_2z)^k\, J(L_2, z)^k$$

for all $L_1, L_2 \in \Gamma$. (See [70], §2.1, for example.) Matters are simplified considerably when we deal with an integral weight $k$. Then we do not have to worry about arguments of complex numbers, and from $J(L_1L_2, z) = J(L_1, L_2z)J(L_2, z)$ we obtain

$$v(L_1L_2) = v(L_1)v(L_2).$$

Thus the multiplier system of a modular form of integral weight on $\Gamma$ is a homomorphism of $\Gamma$ into the complex numbers of absolute value 1.

We continue with some definitions and remarks.

A modular form $f$ is called a *cusp form* if it vanishes at all cusps. This means that for all $r \in \mathbb{Q} \cup \{\infty\}$ we have $c(n) = 0$ whenever $n + \kappa \leq 0$ in the expansions (1.17) and (1.18). Points $z, w$ in $\mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$ are called *equivalent* with respect to the group $\Gamma$ if $w = Lz$ for some $L \in \Gamma$. The set $\Gamma(z)$ of points equivalent to $z$ is called the *orbit* of $z$ under $\Gamma$ or the $\Gamma$-*orbit* of $z$. Let $f$ be a function with property (1). If $f$ is holomorphic or vanishes at a cusp $r$ then it is easy to see that $f$ is holomorphic or vanishes at all cusps in the $\Gamma$-orbit of $r$, respectively. It is well-known that for the groups considered here there exist only finitely many orbits of cusps. Therefore, in order to show that $f$ is a modular form it suffices to verify (2) for a finite set of representatives of cusp orbits.

Clearly, the set of modular forms of weight $k$ and multiplier system $v$ for $\Gamma$ is a complex vector space, and the same is true for cusp forms. We denote these spaces by $\mathcal{M}(\Gamma, k, v)$ and $\mathcal{S}(\Gamma, k, v)$, respectively. (We will rarely need to use these notations.) Compactness arguments show that these spaces are $\{0\}$ whenever $k \leq 0$, except for the equally trivial space $\mathcal{M}(\Gamma, 0, 1) = \mathbb{C}$ where 1 stands for the constant function 1 on $\Gamma$. Moreover, for the groups considered here, all spaces of modular forms have finite dimension. In some cases the dimension can be computed by contour integration with the help of the argument principle; in more cases, the Riemann–Roch theorem yields a dimension formula. We refer to the numerous textbooks for this important topic, but here we will not reproduce dimension formulae.

Frequently the condition of holomorphicity is too strong since it excludes interesting examples. A function $f$ on $\mathbb{H}$ is called a *meromorphic modular form* of *weight* $k$ and *multiplier system* $v$ for $\Gamma$ if it is meromorphic on $\mathbb{H}$, satisfies (1) and is meromorphic at all cusps $r \in \mathbb{Q} \cup \{\infty\}$. This last condition means that in each of the Fourier expansions (1.17) and (1.18) we have $c(n) \neq 0$ for only finitely many $n$ with $n + \kappa < 0$. Also, this condition implies that $f$ is holomorphic in a half plane $\{z \in \mathbb{C} \mid \mathrm{Im}(z) > M\}$ for some sufficiently large $M > 0$. Moreover, now the expansions (1.17) and (1.18) need not hold for all $z \in \mathbb{H}$, but only for $0 < |e(z)| < \varepsilon$ with some sufficiently small $\varepsilon > 0$. An interesting class consists of those meromorphic modular forms whose poles are supported by the cusps, that is, which are holomorphic on $\mathbb{H}$. Eta products belong to this class.

The case of weight $k = 0$ is of foremost importance. A meromorphic modular form $f$ of weight 0 and trivial multiplier system 1 for $\Gamma$ is called a *modular function* for $\Gamma$. It satisfies

$$f(Lz) = f(z) \qquad \text{for all} \qquad L \in \Gamma.$$

Clearly, the set of all modular functions for $\Gamma$ is a field. It can be identified with the field of meromorphic functions on the compact Riemann surface corresponding to $\Gamma$.

Let $f$ be a (holomorphic or) meromorphic modular form of weight $k$ and multiplier system $v$ for $\Gamma$ which is not identically 0, and let $r$ be a cusp. Let $n_0$ be the smallest integer for which $c(n_0) \neq 0$ in the Fourier expansion (1.17) or (1.18). Then we call

$$\mathrm{ord}(f, r) = n_0 + \kappa$$

the *order* of $f$ *at the cusp* $r$.

We give a final remark on products of modular forms. For $j = 1, 2$, let $f_j$ be a (holomorphic or) meromorphic modular form of weight $k_j$ and multiplier system $v_j$ for a group $\widetilde{\Gamma}_j$ commensurable with the modular group. Then, clearly, the product $f_1 f_2$ is a (holomorphic or) meromorphic modular form of weight $k_1 + k_2$ and some multiplier system $v$ for the group $\widetilde{\Gamma}_1 \cap \widetilde{\Gamma}_2$. In the case of integral weights we have $v(L) = v_1(L)v_2(L)$ for $L$ in the intersection of the groups. By this observation one can construct new modular forms from known ones. We will use it when we introduce eta products in Sect. 2.

## 1.5 Eisenstein Series for the Full Modular Group

Part of the fascination in the realm of modular forms comes from the fact that there are several possibilities to construct such functions arithmetically, while on the other hand they form vector spaces of small dimensions. Therefore there are linear relations and other identities among modular forms which encode interesting arithmetical relations among their Fourier coefficients. As for the constructions, we will introduce eta products in Sect. 2, Hecke theta series in Sect. 5, and in the present subsection we introduce a few of the many types of Eisenstein series.

**Definition.** A non-zero modular form is called *normalized* if its first non-zero Fourier coefficient (at the cusp $\infty$) is equal to 1. For an even integer $k \geq 2$, the *normalized Eisenstein series* $E_k$ of *weight* $k$ for the modular group $\Gamma_1$ is defined by

$$E_k(z) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) e(nz) \tag{1.19}$$

for $z \in \mathbb{H}$, where $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}, \ldots$ are the *Bernoulli numbers*, defined by the expansion

$$\frac{w}{e^w - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} w^n \qquad \text{for} \qquad 0 < |w| < 2\pi,$$

and where

$$\sigma_l(n) = \sum_{d|n,\, d>0} d^l$$

for any real $l$. For later use we introduce $\tau(n) = \sigma_0(n)$, the number of positive divisors of $n$, as a special case of the *divisor sums* $\sigma_l(n)$.

It is well-known that $E_k(z)$ is a modular form of weight $k$ and trivial multiplier system for the full modular group $\Gamma_1$ if $k \geq 4$. It is not a cusp form because of the non-zero constant coefficient in (1.19). For $k \geq 4$, $E_k(z)$ is a constant multiple of the (non-normalized) *Eisenstein series*

$$G_k(z) = \sum_{m,n \in \mathbb{Z},\ (m,n) \neq (0,0)} (mz + n)^{-k}$$

for which it is easy to verify that the transformation property $G_k(Lz) = J(L,z)^k G_k(z)$ holds for all $L \in \Gamma_1$. Whereas we have absolute and locally uniform convergence in $\mathbb{H}$ of the series $G_k(z)$ for $k \geq 4$ and of $E_k(z)$ for all $k \geq 2$, the series $G_2(z)$ is only conditionally convergent. By evaluating the difference for two specific orders of summation, one can prove (see Schoeneberg [125], §3.2, or Serre [127], §7.4.4) the important transformation formula

$$E_2\left(-\frac{1}{z}\right) = z^2 E_2(z) - \frac{6i}{\pi} z. \tag{1.20}$$

The relation $E_2(z+1) = E_2(z)$ is obvious. More generally,

$$E_2(Lz) = (cz+d)^2 E_2(z) - \frac{6ic}{\pi}(cz+d) \tag{1.21}$$

holds for all $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1$. Non-zero modular forms of weight 2 and trivial multiplier system for $\Gamma_1$ do not exist.

Non-zero cusp forms with trivial multiplier system for $\Gamma_1$ exist for even weights $k = 12$ and $k \geq 16$, but for no other weights. For $k = 12$ we have the cusp forms $E_4^3 - E_6^2$ and the *discriminant function*

$$\Delta(z) = \eta^{24}(z) = e(z) \prod_{n=1}^{\infty} (1 - e(nz))^{24} = \sum_{n=1}^{\infty} \tau(n) e(nz), \tag{1.22}$$

whose coefficients $\tau(n)$ are called the *Ramanujan numbers*. Since the corresponding space of cusp forms has dimension 1, the two functions are proportional; comparing the first non-zero coefficients yields

$$E_4^3(z) - E_6^2(z) = 12^3 \Delta(z),$$

an instance of the arithmetical relations mentioned at the beginning of this subsection. It is well-known that every modular form with trivial multiplier system for $\Gamma_1$ can uniquely be written as a polynomial in the Eisenstein series $E_4$ and $E_6$.

## 1.6 Eisenstein Series for $\Gamma_0(N)$ and Fricke Groups

In this subsection we introduce Eisenstein series of weights $k \geq 3$ for the subgroups $\Gamma_0(N)$ of the modular group and for the Fricke groups $\Gamma^*(N)$. The relation (1.21) is used to construct an Eisenstein series of weight 2 for $\Gamma^*(N)$. The groups are defined as follows:

For a positive integer $N$ we introduce

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \;\middle|\; c \equiv 0 \pmod{N} \right\}.$$

It is called the *Hecke congruence group* of *level* $N$. The groups are named after Erich Hecke because of his important contributions, although other mathematicians worked on them much earlier. The matrix

$$W_N = \begin{pmatrix} 0 & 1/\sqrt{N} \\ -\sqrt{N} & 0 \end{pmatrix}$$

corresponds to the involution $z \mapsto -\frac{1}{Nz}$ of the upper half plane. It belongs to the normalizer of $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{R})$. The group which is generated by $\Gamma_0(N)$ and $W_N$ is called the *Fricke group* of level $N$ and denoted by $\Gamma^*(N)$. We call $W_N$ a *Fricke involution.* The index of $\Gamma_0(N)$ in $\Gamma^*(N)$ is 2, with cosets represented by the identity and $W_N$. We will not need the full normalizer of $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{R})$ which is generated by $\Gamma_0(N)$ and all the Atkin–Lehner involutions; see [6].

We begin with an observation which is easy to verify but important: Let $M, N, d$ be positive integers such that $M|N$ and $d|(N/M)$. Let $f$ be a modular form of weight $k$ for $\Gamma_0(M)$. Then the function

$$g(z) = f(dz)$$

is a modular form of weight $k$ for $\Gamma_0(N)$. If $f$ has trivial multiplier system then the multiplier system of $g$ is trivial, too. So in particular, for $N, d \in \mathbb{N}$, $d|N$ and even $k \geq 4$ the Eisenstein series $E_k(dz)$ are modular forms of weight $k$ with trivial multiplier system for $\Gamma_0(N)$. A bit more is true:

**Proposition 1.8** *For integers $N \geq 2$, even $k \geq 2$ and $\delta \in \{1, -1\}$, define the Eisenstein series*

$$E_{k,\,N,\delta}(z) = \frac{1}{1 + \delta N^{k/2}} \big(E_k(z) + \delta N^{k/2} E_k(Nz)\big).$$

*Then for $k \geq 4$, $E_{k,\,N,\delta}(z)$ is a modular form of weight $k$ for the Fricke group $\Gamma^*(N)$ whose multiplier system $v$ is given by $v(L) = 1$ for $L \in \Gamma_0(N)$ and $v(L) = \delta$ for $L \notin \Gamma_0(N)$. The function*

$$E_{2,\,N,-1}(z)$$

*is a modular form of weight 2 for $\Gamma^*(N)$ whose multiplier system $v$ is given by $v(L) = 1$ for $L \in \Gamma_0(N)$ and $v(L) = -1$ for $L \notin \Gamma_0(N)$.*

*Proof.* The factor $C = 1/(1+\delta N^{k/2})$ is introduced merely to get a normalized function. We put $f(z) = E_{k,\,N,\delta}(z)$.

Let $k \geq 4$. The introductory remark implies that $f$ is a modular form of weight $k$ for $\Gamma_0(N)$ with trivial multiplier system. For the Fricke involution we obtain

$$\begin{aligned} f(W_N z) &= C\left(E_k\left(-\frac{1}{Nz}\right) + \delta N^{k/2} E_k\left(-\frac{1}{z}\right)\right) \\ &= C\big((Nz)^k E_k(Nz) + \delta N^{k/2} z^k E_k(z)\big) \\ &= \delta\,(\sqrt{N}z)^k f(z). \end{aligned}$$

Thus with respect to $W_N$, $f$ transforms like a modular form of weight $k$ with multiplier $v(W_N) = \delta$. This implies the assertion on $f$.

Now we consider the case $k = 2$, $\delta = -1$. Let $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ be given. From (1.21) we obtain

$$\begin{aligned} f(Lz) &= C\left(E_2\left(\frac{az+b}{cz+d}\right) - NE_2\left(\frac{a\cdot Nz + Nb}{\frac{c}{N}\cdot Nz + d}\right)\right) \\ &= C\left((cz+d)^2\left(E_2(z) - NE_2(Nz)\right) - \frac{6i}{\pi}(cz+d)\left(c - N\cdot\frac{c}{N}\right)\right) \\ &= (cz+d)^2 f(z). \end{aligned}$$

A slightly simpler computation for $W_N$, using (1.20), yields

$$f(W_N z) = -(\sqrt{N}z)^2 f(z).$$

In each case we observe cancellation of the extra terms in (1.20) and (1.21) which indicate the deviation of $E_2(z)$ from a modular form. It follows that $f$ transforms like a modular form of weight 2 for $\Gamma^*(N)$ with multiplier system as stated in the proposition. The correct behavior at cusps follows from the expansion of $E_2(z)$ at $\infty$ and the transformation properties. □

Now we present the Eisenstein series of "Nebentypus" which were introduced by Hecke [53].

**Theorem 1.9 (Hecke [53])** *Let $P$ be an odd prime and let $\chi$ be the Dirichlet character modulo $P$ which is defined by the Legendre symbol $\chi(n) = \left(\frac{n}{P}\right)$. Suppose that $k \geq 3$ and $\chi(-1) = (-1)^k$. Then the Eisenstein series*

$$F_1(z) = \sum_{n=1}^{\infty}\left(\sum_{d>0,\, d|n} \chi\left(\tfrac{n}{d}\right) d^{k-1}\right) e(nz) \tag{1.23}$$

*and*

$$F_2(z) = A_k(P) + \sum_{n=1}^{\infty}\left(\sum_{d>0,\, d|n} \chi(d) d^{k-1}\right) e(nz), \tag{1.24}$$

*with*

$$A_k(P) = (-1)^{\lfloor k/2 \rfloor}\frac{P^{(2k-1)/2}(k-1)!}{(2\pi)^k} L(\chi, k), \qquad L(\chi, k) = \sum_{n=1}^{\infty} \chi(n) n^{-k},$$

*are modular forms of weight $k$ for $\Gamma_0(P)$ with character $\chi$, i.e., they satisfy $F(Lz) = \chi(d)(cz+d)^k F(z)$ for $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(P)$. The transformation $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ interchanges the functions $F_1, F_2$ according to*

$$F_1\left(-\frac{1}{z}\right) = (-i)^k(-1)^{\lfloor k/2 \rfloor} P^{(1-2k)/2} z^k F_2\left(\frac{z}{P}\right), \tag{1.25}$$

$$F_2\left(-\frac{1}{z}\right) = (-i)^k(-1)^{\lfloor k/2 \rfloor} P^{-1/2} z^k F_1\left(\frac{z}{P}\right). \tag{1.26}$$

We use the relations (1.25), (1.26) to define Eisenstein series for the Fricke group $\Gamma^*(P)$ similarly as in Proposition 1.8:

**Definition.** Let $P$, $\chi$, $k$ and $F_1$, $F_2$ be given as in Theorem 1.9. Then we put

$$\begin{aligned} E_{k,P,\,i}(z) &= \frac{1}{A_k(P)}\big(F_2(z) - P^{(k-1)/2}F_1(z)\big) \\ &= 1 + \frac{1}{A_k(P)} \sum_{n=1}^{\infty} \Bigg( \sum_{d|n} \big(\chi(d) - P^{(k-1)/2}\chi\big(\tfrac{n}{d}\big)\big) d^{k-1} \Bigg) e(nz), \end{aligned} \tag{1.27}$$

$$E_{k,P,\,-i}(z) = \frac{1}{A_k(P)}\big(F_2(z) + P^{(k-1)/2}F_1(z)\big). \tag{1.28}$$

Since both $F_1$ and $F_2$ are modular forms of weight $k$ for $\Gamma_0(P)$ with character $\chi$, this holds true also for $E_{k,P,\pm i}$. From (1.25), (1.26) and the definitions one easily deduces

$$E_{k,P,\,\delta i}\big(-\tfrac{1}{Pz}\big) = -\delta(-i)^k(-1)^{\lfloor k/2 \rfloor}(\sqrt{P}z)^k E_{k,P,\,\delta i}(z)$$

for $\delta \in \{1, -1\}$. Hence we have modular forms for the Fricke group:

**Proposition 1.10** *For $P$, $\chi$ and $k$ as in Theorem 1.9, the Eisenstein series $E_{k,P,\,\delta i}$ are modular forms of weight $k$ for the Fricke group $\Gamma^*(P)$. Their multiplier systems $v_\delta$ are given by $v_\delta(L) = \chi(d) = \left(\frac{d}{P}\right)$ for $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(P)$ in both cases, and $v_\delta(W_P) = -\delta(-i)^k(-1)^{\lfloor k/2 \rfloor}$.*

We observe that Theorem 1.9 and Proposition 1.10 yield Eisenstein series of odd weights $k \geq 3$ for prime levels $P \equiv 3 \bmod 4$. The values $L(\chi, k)$ of the $L$-series are explicitly known, and the constant term $A_k(P)$ in $F_2(z)$ is a rational number; see [59], §16.4, [84], §14.2, or [140], §7. For example, for level $P = 3$ we have the weight 3 Eisenstein series

$$E_{3,3,i}(z) = 1 + 18\sum_{n=1}^{\infty}\left(\sum_{d|n}\frac{1}{2}\left(3\left(\tfrac{n/d}{3}\right) - \left(\tfrac{d}{3}\right)\right)d^2\right)e(nz),$$

$$E_{3,3,-i}(z) = 1 - 18\sum_{n=1}^{\infty}\left(\sum_{d|n}\frac{1}{2}\left(3\left(\tfrac{n/d}{3}\right) + \left(\tfrac{d}{3}\right)\right)d^2\right)e(nz).$$

They satisfy

$$E_{3,3,i}\left(-\tfrac{1}{3z}\right) = i(\sqrt{3}z)^3E_{3,3,i}(z), \quad E_{3,3,-i}\left(-\tfrac{1}{3z}\right) = -i(\sqrt{3}z)^3E_{3,3,-i}(z).$$

The signs in these transformation formulae have been the reason for the choice of signs in the notation $E_{k,P,\,\delta i}(z)$. We will meet the functions $E_{3,3,\,\delta i}(z)$ in Sect. 11.2.

There are many more types of Eisenstein series which will not be presented here. We refer to [30], Chap. 4, [96], Chap. 7, and [125], Chap. 7 for a thorough discussion, including the delicate cases of small weights 1 and 2. We will meet several examples in Part II.

## 1.7 Hecke Eigenforms

Spaces of modular forms possess bases of arithmetically distinguished functions: Their Fourier expansions have multiplicative coefficients which, moreover, satisfy simple recursions at powers of each prime. As a consequence, the corresponding Dirichlet series have Euler product expansions of a particularly simple type. The tool for establishing these results is provided by a sequence of linear operators on spaces of modular forms, the Hecke operators, and the basis functions in question are the so-called Hecke eigenforms. For introductions to this body of theory, in complete detail or in a more sketchy form, we can refer to [16], [30], [33], [55], [61], [72], [73], [84], [90], [96], [105], [117], [127], [131]. Here we will reproduce the basic definitions and some of the main results.

Let $f \in \mathcal{M}(\Gamma_1, k, 1)$ be a modular form of integral weight $k$ on the full modular group $\Gamma_1$ with trivial multiplier system. For a positive integer $m$, the action of the $m$th *Hecke operator* $T_m$ on $f$ is given by

$$T_mf(z) = m^{k-1}\sum_{ad=m,\ a>0} d^{-k}\sum_{b \bmod d} f\left(\frac{az+b}{d}\right). \tag{1.29}$$

This definition looks more natural when one interprets modular forms as homogeneous functions on lattices: We consider complex valued functions $F$

on the set of all lattices $\Lambda \subset \mathbb{C}$ which are homogeneous of degree $-k$, that is, which satisfy $F(\alpha\Lambda) = \alpha^{-k}F(\Lambda)$ for all lattices $\Lambda$ and $\alpha \in \mathbb{C}$, $\alpha \neq 0$. Any lattice can be written as $\Lambda = \alpha\Lambda_z$ with $\Lambda_z = \mathbb{Z} + \mathbb{Z}z$ where $z$ in the upper half plane is unique up to a transformation from $\Gamma_1$. Then the assignment $f(z) = F(\Lambda_z)$ yields a bijection from functions $F$ on lattices, homogeneous of degree $-k$, to functions $f$ on the upper half plane satisfying the transformation law (1) in the definition of modular forms in Sect. 1.4 (for $\Gamma = \Gamma_1$, $k$ integral, $v = 1$). The action of the $m$th Hecke operator on degree $-k$ functions $F$ on lattices is simply given by $T_m F(\Lambda) = \sum_{\Lambda'} F(\Lambda')$ where $\Lambda'$ runs over all sublattices of index $m$ in $\Lambda$. Choosing appropriate representatives for sublattices and translating back to modular forms yields the definition (1.29), up to the normalizing factor $m^{k-1}$. In terms of the Fourier expansion (1.17) of $f$, which under our present assumptions simply reads

$$f(z) = \sum_{n=0}^{\infty} c(n)e(nz)\,, \tag{1.30}$$

the action of $T_m$ is given by

$$T_m f(z) = \sum_{n=0}^{\infty} \Bigg( \sum_{d>0\,,\; d|\,\gcd(n,m)} d^{k-1} c\big(\tfrac{mn}{d^2}\big) \Bigg) e(nz)\,. \tag{1.31}$$

The operators $T_m$ map $\mathcal{M}(\Gamma_1, k, 1)$ into itself, they are linear, and they map cusp forms into cusp forms. Any two operators $T_m$, $T_l$ commute and satisfy

$$T_m T_l = \sum_{d>0\,,\; d|\,\gcd(m,l)} d^{k-1} T_{ml/d^2}. \tag{1.32}$$

In particular we have

$$T_p T_{p^r} = T_{p^{r+1}} + p^{k-1} T_{p^{r-1}} \tag{1.33}$$

for primes $p$ and any $r \geq 1$. The subspace $\mathcal{S}(\Gamma_1, k, 1)$ of cusp forms is a Hilbert space with respect to the *Petersson inner product* (whose definition by an integral we are not going to reproduce here), and the Hecke operators are self-adjoint with respect to this inner product. Therefore it follows from linear algebra that the operators $T_m$ can simultaneously be diagonalized on the space of cusp forms. Thus $\mathcal{S}(\Gamma_1, k, 1)$ has a basis of functions $f$ which are eigenvectors for all operators $T_m$ and which are mutually orthogonal with respect to the Petersson inner product. This result extends to $\mathcal{M}(\Gamma_1, k, 1)$ since it is easily seen that the Eisenstein series $E_k$ in (1.19) is an eigenvector. If $f \neq 0$ and $T_m f(z) = \lambda(m) f(z)$ for all $m$ then from (1.31) we obtain (for $n = 1$) that $\lambda(m)c(1) = c(m)$ for all $m$. It follows that $c(1) \neq 0$, and we can achieve that $c(1) = 1$. In this case the eigenvalues coincide with the Fourier coefficients; we have

$$\lambda(m) = c(m), \qquad T_m f(z) = c(m) f(z) \qquad \text{for all} \qquad m,$$

and $f$ is called a *normalized Hecke eigenform*, or simply an *eigenform*. The relations (1.32), (1.33) then imply that

$$c(mn) = c(m)c(n) \qquad \text{for} \qquad \gcd(m,n) = 1, \tag{1.34}$$

$$c(p^{r+1}) = c(p)c(p^r) - p^{k-1}c(p^{r-1}) \tag{1.35}$$

for all primes $p$ and all $r \geq 1$. Thus the Fourier coefficients of an eigenform are multiplicative and satisfy a simple recursion at powers of primes. Moreover, they are totally real algebraic integers. An eigenform is uniquely determined by the eigenvalues.

The dimension of $\mathcal{S}(\Gamma_1, k, 1)$ is equal to 1 for $k = 12, 16, 18, 20, 22, 26$. It is clear then that the normalized modular forms $\Delta$, $\Delta E_4$, $\Delta E_6$, $\Delta E_4^2$, $\Delta E_4 E_6$, $\Delta E_4^2 E_6$ in these spaces are normalized Hecke eigenforms. For the most prominent example of the discriminant function $\Delta(z)$ we obtain that the Ramanujan numbers $\tau(n)$ are multiplicative and satisfy $\tau(p^{r+1}) = \tau(p)\tau(p^r) - p^{11}\tau(p^{r-1})$ for all primes $p$.

For any modular form $f$ with Fourier expansion (1.30), its *Hecke L-series* is defined by

$$L(f,s) = \sum_{n=1}^{\infty} c(n)n^{-s}. \tag{1.36}$$

For an eigenform $f$ the relations (1.34) (1.35) translate into the Euler product expansion

$$L(f,s) = \prod_{p} \left(1 - c(p)p^{-s} + p^{k-1-2s}\right)^{-1}, \tag{1.37}$$

where the product is taken over all primes $p$. We mention in passing that, independently from $f$ being an eigenform or not, the Dirichlet series (1.36) converges for $\mathrm{Re}(s) > k$, has an analytic continuation to the whole complex $s$-plane, and satisfies a functional equation of Riemann type relating the values at $s$ and $k - s$.

In the late 1930's Hecke and Petersson generalized the theory of the operators $T_m$ to spaces of modular forms on congruence subgroups of the modular group, most notably for the groups $\Gamma_0(N)$. But some of the main results, such as the uniqueness of simultaneous eigenforms and the unrestricted Euler product formula (1.37), do not hold true for $N > 1$. Fully satisfactory generalizations were achieved only later by Atkin and Lehner [6], with major contributions by W. Li [87], [88], Pizer [112], and other authors, when the concept of newforms was introduced and elaborated.

We consider the spaces $\mathcal{M}(\Gamma_0(N), k, \chi)$ and their subspaces $\mathcal{S}(\Gamma_0(N), k, \chi)$ of cusp forms $f$ of integral weight $k$ which transform according to

$$f(Lz) = \chi(d)\,(cz+d)^k\, f(z) \qquad \text{for} \qquad L = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N),$$

where $\chi$ is a Dirichlet character modulo $N$. For such a function $f$ and for primes $p$ the action of $T_p$ is defined by

$$T_p f(z) \;=\; p^{k-1}\left(\sum_{b=0}^{p-1} p^{-k} f\big(\tfrac{z+b}{p}\big) + \chi(p) f(pz)\right). \tag{1.38}$$

In terms of the Fourier expansion of $f$, which can also be written as (1.30), this reads

$$T_p f(z) = \sum_{n=0}^{\infty} \big(c(pn) + \chi(p) p^{k-1} c(n/p)\big)\, e(nz)\,, \tag{1.39}$$

where we agree that $c(n/p) = 0$ if $p \nmid n$. More generally, for any positive integer $m$ the action of the *Hecke operator* $T_m$ is given by

$$T_m f(z) = \sum_{n=0}^{\infty} \Bigg( \sum_{d>0\,,\; d|\gcd(n,m)} \chi(d)\, d^{k-1}\, c\big(\tfrac{mn}{d^2}\big) \Bigg) e(nz), \tag{1.40}$$

where we note that $\chi(d) = 0$ whenever $\gcd(d,N) > 1$. Any two of the operators $T_m$ with $\gcd(m,N) = 1$ commute, and they are normal (not necessarily self-adjoint) with respect to the Petersson inner product on $\mathcal{S}(\Gamma_0(N), k, \chi)$. This yields Petersson's result [108]:

*The space $\mathcal{S}(\Gamma_0(N), k, \chi)$ has an orthogonal basis of common eigenfunctions of the operators $T_m$ for all $m$ with* $\gcd(m,N) = 1$.

Generally, and in contrast to the case $N = 1$ handled above, $\mathcal{S}(\Gamma_0(N), k, \chi)$ does not necessarily have a basis of common eigenfunctions for all $T_m$, and subspaces of simultaneous eigenfunctions of the operators $T_m$ with $\gcd(m,N) = 1$ need not be one-dimensional. The reason for this is simple and explained as follows. Suppose that $M$ is a proper divisor of $N$ and that $\chi$ is induced from a character $\chi^{'}$ modulo $M$. (For example, $\chi$ might be trivial and $M$ any proper divisor of $N$.) Let $l$ be a positive integer such that $lM|N$, and let $f \in \mathcal{M}(\Gamma_0(M), k, \chi^{'})$. Then it is easy to see that $g(z) = f(lz)$ belongs to $\mathcal{M}(\Gamma_0(N), k, \chi)$ and that the operators $T_m$ with $\gcd(m,N) = 1$ act on $g$ in exactly the same way as they act on $f$. Thus $\mathcal{M}(\Gamma_0(M), k, \chi^{'})$ sits in at least two different ways (for $l = 1$ and $l = \frac{N}{M}$) in $\mathcal{M}(\Gamma_0(N), k, \chi)$, and the same can be said for cusp forms. Following Atkin and Lehner [6], one denotes by $\mathcal{S}^{\rm old}(\Gamma_0(N), k, \chi)$ the subspace of cusp forms which is spanned by the functions $g(z) = f(lz)$ with cusp forms $f$ when $M$ and $l$ vary as described above. It is called the space of *oldforms*. One concludes that the operators $T_m$ with $\gcd(m,N) = 1$ map $\mathcal{S}^{\rm old}(\Gamma_0(N), k, \chi)$ into itself and that subspaces of common eigenfunctions of these operators have dimensions at least 2.

Let $\mathcal{S}^{\rm new}(\Gamma_0(N), k, \chi)$ be the orthogonal complement of $\mathcal{S}^{\rm old}(\Gamma_0(N), k, \chi)$ in $\mathcal{S}(\Gamma_0(N), k, \chi)$ with respect to the Petersson inner product. It is also invariant under the operators $T_m$ with $\gcd(m,N) = 1$, since these operators are normal, and therefore it also has a basis of common eigenfunctions of the operators $T_m$ with $\gcd(m, N) = 1$. Such an eigenfunction is called a *newform*. We note that $\mathcal{S}^{\text{new}}(\Gamma_0(N), k, \chi) = \mathcal{S}(\Gamma_0(N), k, \chi)$ if $\chi$ is a primitive character modulo $N$.

It turns out that the main assertions of the Hecke theory for $\mathcal{S}(\Gamma_1, k, 1)$ generalize to hold for newforms. In particular, if $f$ is a newform and (1.30) its Fourier expansion, then $c(1) \neq 0$, and we can achieve that $c(1) = 1$, in which case $f$ is called a *normalized newform*. The main results for newforms, embracing the above results for $N = 1$, are summarized as follows:

**Theorem 1.11 (Atkin–Lehner)** *Let $k$, $N$ be positive integers and $\chi$ a Dirichlet character modulo $N$. The following assertions hold.*

(1) *There exists an orthogonal basis of $\mathcal{S}^{\text{new}}(\Gamma_0(N), k, \chi)$ consisting of normalized newforms. Let $f \in \mathcal{S}^{\text{new}}(\Gamma_0(N), k, \chi)$ be a normalized newform and $c(n)$ its Fourier coefficients.*

(2) *For all $m \geq 1$ we have*

$$T_m f = c(m) f.$$

*The eigenvalues $c(m)$ are algebraic integers. For prime divisors $p$ of $N$ we have $|c(p)| = p^{\frac{1}{2}(k-1)}$ if $\chi$ is not induced from a character modulo $\frac{N}{p}$, while otherwise we have $c(p) = 0$ if $p^2 | N$, and $c(p)^2 = \chi(p) p^{k-2}$ if $p^2 \nmid N$.*

(3) *The Dirichlet series associated to $f$ has the Euler product expansion*

$$L(f, s) = \prod_p \left(1 - c(p) p^{-s} + \chi(p) p^{k-1-2s}\right)^{-1}.$$

(*Note that $\chi(p) = 0$ if $p | N$.*)

(4) *If $g(z) = \sum_{n=1}^{\infty} b(n) e(nz)$ is a normalized newform of weight $k$ and some level $M$ and character $\psi$ modulo $M$, and if $b(p) = c(p)$ for all but finitely many primes $p$, then $M = N$, $\psi = \chi$ and $g = f$. The simultaneous eigenspaces of the operators $T_p$ for primes $p \nmid N$ in $\mathcal{S}^{\text{new}}(\Gamma_0(N), k, \chi)$ are one-dimensional, and the normalized newforms constitute the unique orthogonal basis of $\mathcal{S}^{\text{new}}(\Gamma_0(N), k, \chi)$ consisting of normalized common eigenfunctions of the operators $T_p$ for primes $p \nmid N$.*

Part (4) in Theorem 1.11 is called the *multiplicity one theorem*. The eigenvalues $c(p)$ of a normalized newform of weight $k$ satisfy

$$|c(p)| \leq 2 p^{\frac{k-1}{2}}$$

for all primes $p$. This is the celebrated Deligne theorem, formerly the Ramanujan–Petersson conjecture, and a very deep result. We will see in Sect. 5.3 that in the special case of Hecke theta series this inequality follows trivially from the decomposition of prime numbers into prime ideals in quadratic number fields.

## 1.8 Identification of Modular Forms

The dimensions of spaces of modular forms are "small". (We mentioned that in Sect. 1.5.) This follows from the fact that the total number of zeros of a non-zero modular form in a fundamental set of its group is "small". In the simplest case of a modular form $f \neq 0$ of integral weight $k$ and trivial multiplier system on the full modular group $\Gamma_1$, contour integration and the argument principle yield the *valence formula*

$$\operatorname{ord}(f,\infty) + \tfrac{1}{2}\operatorname{ord}(f,i) + \tfrac{1}{3}\operatorname{ord}(f,\omega) + \sum_z \operatorname{ord}(f,z) = \frac{k}{12}\,, \tag{1.41}$$

where $\operatorname{ord}(f,z)$ is the order of $f$ at the point $z$ and the summation is on all $z$ in the standard fundamental domain of $\Gamma_1$ different from the elliptic fixed points $i$ and $\omega = e\left(\frac{1}{6}\right)$. Therefore, if (1.30) is the Fourier expansion of a function $f \in \mathcal{M}(\Gamma_1,k,1)$ and if $c(n) = 0$ for all $n \leq 1 + \frac{k}{12}$, then it follows that $f = 0$, since otherwise the left hand side in (1.41) would be bigger than the right hand side. Equivalently, two modular forms in $\mathcal{M}(\Gamma_1,k,1)$ are identical if their initial segments of $\left\lfloor 1 + \frac{k}{12}\right\rfloor$ Fourier coefficients match. Hence one can prove an identity among modular forms by simply comparing a few of their Fourier coefficients.

This principle generalizes to other spaces of modular forms. In [53] (Math. Werke, p. 811) Hecke gave the following results: If $f \in \mathcal{M}(\Gamma_0(N),k,1)$ with expansion (1.30) satisfies

$$c(n) = 0 \qquad \text{for all} \qquad n \leq 1 + \frac{k}{12}\,\mu_0(N)\,,$$

then $f = 0$. If $f \in \mathcal{M}(\Gamma_0(N),k,\chi)$ with a real character $\chi \neq 1$ satisfies $c(n) = 0$ for all $n \leq 2 + \frac{k}{12}\,\mu_0(N)$, then $f = 0$. Here

$$\mu_0(N) = \left[\,\Gamma_1 \,:\, \Gamma_0(N)\,\right] = N \prod_{p|N}\left(1 + \frac{1}{p}\right)$$

is the index of $\Gamma_0(N)$ in $\Gamma_1$. A similar result is given in [116], Theorem 1. A more general result can be found in Petersson's monograph [110], Satz 3.5, p. 47:

**Theorem 1.12** *Let* $\Gamma$ *be a subgroup with finite index* $\mu(\Gamma) = [\Gamma_1 : \Gamma]$ *in the full modular group* $\Gamma_1$. *For cusp forms* $f, g \in \mathcal{S}(\Gamma, k, v)$ *of weight* $k > 0$ *and multiplier system* $v$ *on* $\Gamma$, *let their Fourier expansions at* $\infty$ *be written as* (1.17) *with coefficients* $c(n)$ *and* $b(n)$, *respectively. Then if*

$$c(n) = b(n) \qquad \textit{for all} \qquad n \leq \frac{k}{12}\,\mu(\Gamma) - \beta(\Gamma, k, v)\,, \tag{1.42}$$

*we have* $f = g$.

We will not reproduce the definition of the entity $\beta(\Gamma, k, v)$ which is concocted from cusp parameters (see Sect. 1.4) and properties of elliptic fixed points. Since $\beta(\Gamma, k, v) \geq 0$, we can simply ignore this term in applying Theorem 1.12 and verify $c(n) = b(n)$ for $n \leq \frac{k}{12}\,\mu(\Gamma)$.

Verifying the identities in Part II provides numerous instances for the application of Theorem 1.12 (or other versions of the same principle). For a simple example, consider the identities for $\eta^2(z)$ in Example 9.1. The function $\eta^2(12z)$ belongs to $\Gamma_0(144)$, and by Theorems 5.1, 5.3 this holds also for the theta series $\Theta_1(3,\, \xi, z)$, $\Theta_1(-4, \chi_\nu, z)$ and $\Theta_1(-3, \psi_\nu, z)$ in this example. Thus for establishing the identities it suffices to compare coefficients for $n \leq \frac{1}{12}\,\mu_0(144) = 24$. This is very easy indeed, since for trivial reasons the coefficients vanish for all $n \not\equiv 1 \bmod 12$. For most of the other examples in Part II the work to be done is lengthier.

In closing this subsection we mention the papers [39], [82], [116], [126] where a quite different, but related problem is discussed: Let $f$ and $g$ be distinct normalized Hecke newforms, not necessarily of the same weights or levels. Find an upper bound for the smallest prime $p$ for which the Hecke operator $T_p$ has distinct eigenvalues at $f$ and at $g$.