CHAPTER 1

The Elementary Theory of Partitions

1.1 Introduction

In this book we shall study in depth the fundamental additive decomposition process: the representation of positive integers by sums of other positive integers.

DEFINITION 1.1. A *partition* of a positive integer n is a finite nonincreasing sequence of positive integers $\lambda_1, \lambda_2, \ldots, \lambda_r$ such that $\sum_{i=1}^{r} \lambda_i = n$. The λ_i are called the *parts* of the partition.

Many times the partition $(\lambda_1, \lambda_2, \ldots, \lambda_r)$ will be denoted by λ, and we shall write $\lambda \vdash n$ to denote "λ is a partition of n." Sometimes it is useful to use a notation that makes explicit the number of times that a particular integer occurs as a part. Thus if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r) \vdash n$, we sometimes write

$$\lambda = (1^{f_1} 2^{f_2} 3^{f_3} \cdots)$$

where exactly f_i of the λ_j are equal to i. Note now that $\sum_{i \geqslant 1} f_i i = n$.

Numerous types of partition problems will concern us in this book; however, among the most important and fundamental is the question of enumerating various sets of partitions.

DEFINITION 1.2. The partition function $p(n)$ is the number of partitions of n.

Remark. Obviously $p(n) = 0$ when n is negative. We shall set $p(0) = 1$ with the observation that the empty sequence forms the only partition of zero. The following list presents the next six values of $p(n)$ and tabulates the actual partitions.

$p(1) = 1$: $1 = (1)$;
$p(2) = 2$: $2 = (2)$, $1 + 1 = (1^2)$;
$p(3) = 3$: $3 = (3)$, $2 + 1 = (12)$, $1 + 1 + 1 = (1^3)$;
$p(4) = 5$: $4 = (4)$, $3 + 1 = (13)$, $2 + 2 = (2^2)$,
$2 + 1 + 1 = (1^2 2)$, $1 + 1 + 1 + 1 = (1^4)$;

ENCYCLOPEDIA OF MATHEMATICS and Its Applications, Gian-Carlo Rota (ed.). 2, George E. Andrews, The Theory of Partitions

$$
\begin{aligned}
p(5) = 7:\quad & 5 = (5),\quad 4 + 1 = (14),\quad 3 + 2 = (23),\\
& 3 + 1 + 1 = (1^3 3),\quad 2 + 2 + 1 = (12^2),\\
& 2 + 1 + 1 + 1 = (1^3 2),\quad 1 + 1 + 1 + 1 + 1 = (1^5);\\
p(6) = 11:\quad & 6 = (6),\quad 5 + 1 = (15),\quad 4 + 2 = (24),\\
& 4 + 1 + 1 = (1^2 4),\quad 3 + 3 = (3^2),\quad 3 + 2 + 1 = (123),\\
& 3 + 1 + 1 + 1 = (1^3 3),\quad 2 + 2 + 2 = (2^3),\\
& 2 + 2 + 1 + 1 = (1^2 2^2),\quad 2 + 1 + 1 + 1 + 1 = (1^4 2),\\
& 1 + 1 + 1 + 1 + 1 + 1 = (1^6).
\end{aligned}
$$

The partition function increases quite rapidly with n. For example, $p(10) = 42$, $p(20) = 627$, $p(50) = 204226$, $p(100) = 190569292$, and $p(200) = 3972999029388$.

Many times we are interested in problems in which our concern does not extend to all partitions of n but only to a particular subset of the partitions of n.

DEFINITION 1.3. Let $\mathscr{S}$ denote the set of all partitions.

DEFINITION 1.4. Let $p(S, n)$ denote the number of partitions of n that belong to a subset S of the set $\mathscr{S}$ of all partitions.

For example, we might consider $\mathscr{O}$ the set of all partitions with odd parts and $\mathscr{D}$ the set of all partitions with distinct parts. Below we tabulate partitions related to $\mathscr{O}$ and to $\mathscr{D}$.

$$
\begin{aligned}
p(\mathscr{O}, 1) = 1:\quad & 1 = (1),\\
p(\mathscr{O}, 2) = 1:\quad & 1 + 1 = (1^2),\\
p(\mathscr{O}, 3) = 2:\quad & 3 = (3),\quad 1 + 1 + 1 = (1^3),\\
p(\mathscr{O}, 4) = 2:\quad & 3 + 1 = (13),\quad 1 + 1 + 1 + 1 = (1^4),\\
p(\mathscr{O}, 5) = 3:\quad & 5 = (5),\quad 3 + 1 + 1 = (1^2 3),\\
& 1 + 1 + 1 + 1 + 1 = (1^5),\\
p(\mathscr{O}, 6) = 4:\quad & 5 + 1 = (15),\quad 3 + 3 = (3^2),\\
& 3 + 1 + 1 + 1 = (1^3 3),\\
& 1 + 1 + 1 + 1 + 1 + 1 = (1^6),\\
p(\mathscr{O}, 7) = 5:\quad & 7 = (7),\quad 5 + 1 + 1 = (1^2 5),\quad 3 + 3 + 1 = (13^2),\\
& 3 + 1 + 1 + 1 + 1 = (1^4 3),\\
& 1 + 1 + 1 + 1 + 1 + 1 + 1 = (1^7).
\end{aligned}
$$

$$
\begin{aligned}
p(\mathscr{D}, 1) = 1:\quad & 1 = (1),\\
p(\mathscr{D}, 2) = 1:\quad & 2 = (2),\\
p(\mathscr{D}, 3) = 2:\quad & 3 = (3),\quad 2 + 1 = (12),\\
p(\mathscr{D}, 4) = 2:\quad & 4 = (4),\quad 3 + 1 = (13),\\
p(\mathscr{D}, 5) = 3:\quad & 5 = (5),\quad 4 + 1 = (14),\quad 3 + 2 = (23),\\
p(\mathscr{D}, 6) = 4:\quad & 6 = (6),\quad 5 + 1 = (15),\quad 4 + 2 = (24),\\
& 3 + 2 + 1 = (123),\\
p(\mathscr{D}, 7) = 5:\quad & 7 = (7),\quad 6 + 1 = (16),\quad 5 + 2 = (25),\\
& 4 + 3 = (34),\quad 4 + 2 + 1 = (124).
\end{aligned}
$$

We point out the rather curious fact that $p(\mathcal{O}, n) = p(\mathcal{D}, n)$ for $n \leqslant 7$, although there is little apparent relationship between the various partitions listed (see Corollary 1.2).

In this chapter, we shall present two of the most elemental tools for treating partitions: (1) infinite product generating functions; (2) graphical representation of partitions.

1.2 Infinite Product Generating Functions of One Variable

DEFINITION 1.5. The generating function $f(q)$ for the sequence $a_0, a_1, a_2, a_3, \ldots$ is the power series $f(q) = \sum_{n \geqslant 0} a_n q^n$.

Remark. For many of the problems we shall encounter, it suffices to consider $f(q)$ as a "formal power series" in q. With such an approach many of the manipulations of series and products in what follows may be justified almost trivially. On the other hand, much asymptotic work (see Chapter 6) requires that the generating functions be analytic functions of the complex variable q. In actual fact, both approaches have their special merits (recently, E. Bender (1974) has discussed the circumstances in which we may pass from one to the other). Generally we shall state our theorems on generating functions with explicit convergence conditions. For the most part we shall be dealing with absolutely convergent infinite series and infinite products; consequently, various rearrangements of series and interchanges of summation will be justified analytically from this simple fact.

DEFINITION 1.6. Let H be a set of positive integers. We let "H" denote the set of all partitions whose parts lie in H. Consequently, $p(\text{“}H\text{”}, n)$ is the number of partitions of n that have all their parts in H.

Thus if H_0 is the set of all odd positive integers, then "H_0" $= \mathcal{O}$.

$$p(\text{“}H_0\text{”}, n) = p(\mathcal{O}, n).$$

DEFINITION 1.7. Let H be a set of positive integers. We let "H"$(\leqslant d)$ denote the set of all partitions in which no part appears more than d times and each part is in H.

Thus if N is the set of all positive integers, then $p(\text{“}N\text{”}(\leqslant 1), n) = p(\mathcal{D}, n)$.

THEOREM 1.1. *Let H be a set of positive integers, and let*

$$f(q) = \sum_{n \geqslant 0} p(\text{“}H\text{”}, n) q^n, \tag{1.2.1}$$

$$f_d(q) = \sum_{n \geqslant 0} p(\text{“}H\text{”}(\leqslant d), n) q^n. \tag{1.2.2}$$

Then for $|q| < 1$

$$f(q) = \prod_{n \in H} (1 - q^n)^{-1}, \tag{1.2.3}$$

$$\begin{aligned} f_d(q) &= \prod_{n \in H} (1 + q^n + \cdots + q^{dn}) \\ &= \prod_{n \in H} (1 - q^{(d+1)n})(1 - q^n)^{-1}. \end{aligned} \tag{1.2.4}$$

Remark. The equivalence of the two forms for $f_d(q)$ follows from the simple formula for the sum of a finite geometric series:

$$1 + x + x^2 + \cdots + x^r = \frac{1 - x^{r+1}}{1 - x}.$$

Proof. We shall proceed in a formal manner to prove (1.2.3) and (1.2.4); at the conclusion of our proof we shall sketch how to justify our steps analytically. Let us index the elements of H, so that $H = \{h_1, h_2, h_3, h_4, \ldots\}$. Then

$$\begin{aligned} \prod_{n \in H} (1 - q^n)^{-1} &= \prod_{n \in H} (1 + q^n + q^{2n} + q^{3n} + \cdots) \\ &= (1 + q^{h_1} + q^{2h_1} + q^{3h_1} + \cdots) \\ &\quad \times (1 + q^{h_2} + q^{2h_2} + q^{3h_2} + \cdots) \\ &\quad \times (1 + q^{h_3} + q^{2h_3} + q^{3h_3} + \cdots) \\ &\quad \cdots \\ &= \sum_{a_1 \geqslant 0} \sum_{a_2 \geqslant 0} \sum_{a_3 \geqslant 0} \cdots q^{a_1 h_1 + a_2 h_2 + a_3 h_3 + \cdots} \end{aligned}$$

and we observe that the exponent of q is just the partition $(h_1^{a_1} h_2^{a_2} h_3^{a_3} \cdots)$. Hence q^N will occur in the foregoing summation once for each partition of n into parts taken from H. Therefore

$$\prod_{n \in H} (1 - q^n)^{-1} = \sum_{n \geqslant 0} p(\text{“}H\text{”}, n) q^n.$$

The proof of (1.2.4) is identical with that of (1.2.3) except that the infinite geometric series is replaced by the finite geometric series:

$$\begin{aligned} &\prod_{n \in H} (1 + q^n + q^{2n} + \cdots + q^{dn}) \\ &\quad = \sum_{d \geqslant a_1 \geqslant 0} \sum_{d \geqslant a_2 \geqslant 0} \sum_{d \geqslant a_3 \geqslant 0} \cdots q^{a_1 h_1 + a_2 h_2 + a_3 h_3 + \cdots} \\ &\quad = \sum_{n \geqslant 0} p(\text{“}H\text{”}(\leqslant d), n) q^n. \end{aligned}$$

If we are to view the foregoing procedures as operations with convergent infinite products, then the multiplication of infinitely many series together requires some justification. The simplest procedure is to truncate the infinite product to $\prod_{i=1}^{n}(1-q^{h_i})^{-1}$. This truncated product will generate those partitions whose parts are among $h_1, h_2, \ldots, h_n$. The multiplication is now perfectly valid since only a finite number of absolutely convergent series are involved. Now assume q is real and $0 < q < 1$; then if $M = h_n$,

$$\sum_{j=0}^{M} p(\text{“}H\text{”}, j)q^j \leqslant \prod_{i=1}^{n}(1-q^{h_i})^{-1} \leqslant \prod_{i=1}^{\infty}(1-q^{h_i})^{-1} < \infty.$$

Thus the sequence of partial sums $\sum_{j=0}^{M} p(\text{“}H\text{”}, j)q^j$ is a bounded increasing sequence and must therefore converge. On the other hand

$$\sum_{j=0}^{\infty} p(\text{“}H\text{”}, j)q^j \geqslant \prod_{i=1}^{n}(1-q^{h_i})^{-1} \to \prod_{i=1}^{\infty}(1-q^{h_i})^{-1} \qquad \text{as} \quad n \to \infty.$$

Therefore

$$\sum_{j=0}^{\infty} p(\text{“}H\text{”}, j)q^j = \prod_{i=1}^{\infty}(1-q^{h_i})^{-1} = \prod_{n\in H}(1-q^n)^{-1}.$$

Similar justification can be given for the proof of (1.2.4). ∎

Corollary 1.2 (Euler). *$p(\mathcal{O}, n) = p(\mathcal{D}, n)$ for all n.*

Proof. By Theorem 1.1,

$$\sum_{n\geqslant 0} p(\mathcal{O}, n)q^n = \prod_{n=1}^{\infty}(1-q^{2n-1})^{-1}$$

and

$$\sum_{n\geqslant 0} p(\mathcal{D}, n)q^n = \prod_{n=1}^{\infty}(1+q^n).$$

Now

$$\prod_{n=1}^{\infty}(1+q^n) = \prod_{n=1}^{\infty}\frac{(1-q^{2n})}{(1-q^n)} = \prod_{n=1}^{\infty}\frac{1}{1-q^{2n-1}}. \tag{1.2.5}$$

Hence

$$\sum_{n\geqslant 0} p(\mathcal{O}, n)q^n = \sum_{n\geqslant 0} p(\mathcal{D}, n)q^n,$$

(The fact that $L(0,\chi_{-4}) = 2c_1(\chi_{-4}) = \dfrac{1}{2}$ is equivalent via the functional equation of $L(s,\chi_{-4})$ to Leibnitz's famous formula $L(1,\chi_{-4}) = 1 - \dfrac{1}{3} + \dfrac{1}{5} - \cdots = \dfrac{\pi}{4}$.) We will see this function again in §3.1.

♠ Identities Involving Sums of Powers of Divisors

We now have our first explicit examples of modular forms and their Fourier expansions and can immediately deduce non-trivial number-theoretic identities. For instance, each of the spaces $M_4(\Gamma_1)$, $M_6(\Gamma_1)$, $M_8(\Gamma_1)$, $M_{10}(\Gamma_1)$ and $M_{14}(\Gamma_1)$ has dimension exactly 1 by the corollary to Proposition 2, and is therefore spanned by the Eisenstein series $E_k(z)$ with leading coefficient 1, so we immediately get the identities

$$\begin{aligned} E_4(z)^2 \;&=\; E_8(z)\,, \quad E_4(z)E_6(z) \;=\; E_{10}(z)\,, \\ E_6(z)E_8(z) \;&=\; E_4(z)E_{10}(z) \;=\; E_{14}(z)\,. \end{aligned}$$

Each of these can be combined with the Fourier expansion given in Proposition 5 to give an identity involving the sums-of-powers-of-divisors functions $\sigma_{k-1}(n)$, the first and the last of these being

$$\begin{aligned} \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m) \;&=\; \frac{\sigma_7(n) - \sigma_3(n)}{120}\,, \\ \sum_{m=1}^{n-1} \sigma_3(m)\sigma_9(n-m) \;&=\; \frac{\sigma_{13}(n) - 11\sigma_9(n) + 10\sigma_3(n)}{2640}\,. \end{aligned}$$

Of course similar identities can be obtained from modular forms in higher weights, even though the dimension of $M_k(\Gamma_1)$ is no longer equal to 1. For instance, the fact that $M_{12}(\Gamma_1)$ is 2-dimensional and contains the three modular forms E_4E_8, E_6^2 and E_{12} implies that the three functions are linearly dependent, and by looking at the first two terms of the Fourier expansions we find that the relation between them is given by $441E_4E_8 + 250E_6^2 = 691E_{12}$, a formula which the reader can write out explicitly as an identity among sums-of-powers-of-divisors functions if he or she is so inclined. It is not easy to obtain any of these identities by direct number-theoretical reasoning (although in fact it can be done). ♡

2.3 The Eisenstein Series of Weight 2

In §2.1 and §2.2 we restricted ourselves to the case when $k > 2$, since then the series (9) and (10) are absolutely convergent and therefore define modular forms of weight k. But the final formula (13) for the Fourier expansion of $\mathbb{G}_k(z)$ converges rapidly and defines a holomorphic function of z also for $k = 2$, so in this weight we can simply *define* the Eisenstein series $\mathbb{G}_2$, G_2 and E_2 by equations (13), (12), and (11), respectively, i.e.,

$$\begin{aligned} \mathbb{G}_2(z) &= -\frac{1}{24} + \sum_{n=1}^{\infty} \sigma_1(n)\, q^n = -\frac{1}{24} + q + 3q^2 + 4q^3 + 7q^4 + 6q^5 + \cdots\,, \\ G_2(z) &= -4\pi^2\, \mathbb{G}_2(z)\,, \quad E_2(z) = \frac{6}{\pi^2}\, G_2(z) = 1 - 24q - 72q^2 - \cdots\,. \end{aligned} \tag{17}$$

Moreover, the same proof as for Proposition 5 still shows that $G_2(z)$ is given by the expression (10), if we agree to carry out the summation over n first and then over m:

$$G_2(z) = \frac{1}{2}\sum_{n\neq 0}\frac{1}{n^2} + \frac{1}{2}\sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(mz+n)^2}\,. \tag{18}$$

The only difference is that, because of the non-absolute convergence of the double series, we can no longer interchange the order of summation to get the modular transformation equation $G_2(-1/z) = z^2 G_2(z)$. (The equation $G_2(z+1) = G_2(z)$, of course, still holds just as for higher weights.) Nevertheless, the function $G_2(z)$ and its multiples $E_2(z)$ and $\mathbb{G}_2(z)$ do have some modular properties and, as we will see later, these are important for many applications.

Proposition 6. *For $z \in \mathfrak{H}$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$ we have*

$$G_2\left(\frac{az+b}{cz+d}\right) = (cz+d)^2\, G_2(z) - \pi i c(cz+d)\,. \tag{19}$$

Proof. There are many ways to prove this. We sketch one, due to Hecke, since the method is useful in many other situations. The series (10) for $k = 2$ does not converge absolutely, but it is just at the edge of convergence, since $\sum_{m,n} |mz+n|^{-\lambda}$ converges for any real number $\lambda > 2$. We therefore modify the sum slightly by introducing

$$G_{2,\varepsilon}(z) = \frac{1}{2}\sum_{m,\,n}{}' \frac{1}{(mz+n)^2\,|mz+n|^{2\varepsilon}} \qquad (z\in\mathfrak{H},\ \varepsilon > 0)\,. \tag{20}$$

(Here $\sum'$ means that the value $(m,n) = (0,0)$ is to be omitted from the summation.) The new series converges absolutely and transforms by $G_{2,\varepsilon}\left(\frac{az+b}{cz+d}\right) = (cz+d)^2|cz+d|^{2\varepsilon} G_{2,\varepsilon}(z)$. We claim that $\lim\limits_{\varepsilon\to 0} G_{2,\varepsilon}(z)$ exists and equals $G_2(z) - \pi/2y$, where $y = \Im(z)$. It follows that each of the three non-holomorphic functions

$$G_2^*(z) = G_2(z) - \frac{\pi}{2y}\,, \quad E_2^*(z) = E_2(z) - \frac{3}{\pi y}\,, \quad \mathbb{G}_2^*(z) = \mathbb{G}_2(z) + \frac{1}{8\pi y} \tag{21}$$

transforms like a modular form of weight 2, and from this one easily deduces the transformation equation (19) and its analogues for E_2 and $\mathbb{G}_2$. To prove the claim, we define a function I_ε by

$$I_\varepsilon(z) \;=\; \int_{-\infty}^{\infty} \frac{dt}{(z+t)^2\,|z+t|^{2\varepsilon}} \qquad \bigl(z\in\mathfrak{H},\ \varepsilon>-\tfrac12\bigr)\,.$$

Then for $\varepsilon>0$ we can write

$$\begin{aligned} G_{2,\varepsilon} - \sum_{m=1}^{\infty} I_\varepsilon(mz) \;&=\; \sum_{n=1}^{\infty}\frac{1}{n^{2+2\varepsilon}} \\ &+ \sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}\biggl[\frac{1}{(mz+n)^2\,|mz+n|^{2\varepsilon}} - \int_n^{n+1}\frac{dt}{(mz+t)^2|mz+t|^{2\varepsilon}}\biggr]\,. \end{aligned}$$

Both sums on the right converge absolutely and locally uniformly for $\varepsilon>-\frac12$ (the second one because the expression in square brackets is $\mathrm{O}\bigl(|mz+n|^{-3-2\varepsilon}\bigr)$ by the mean-value theorem, which tells us that $f(t)-f(n)$ for any differentiable function f is bounded in $n\le t\le n+1$ by $\max_{n\le u\le n+1}|f'(u)|$), so the limit of the expression on the right as $\varepsilon\to0$ exists and can be obtained simply by putting $\varepsilon=0$ in each term, where it reduces to $G_2(z)$ by (18). On the other hand, for $\varepsilon>-\frac12$ we have

$$\begin{aligned} I_\varepsilon(x+iy) \;&=\; \int_{-\infty}^{\infty}\frac{dt}{(x+t+iy)^2\,((x+t)^2+y^2)^\varepsilon} \\ &=\; \int_{-\infty}^{\infty}\frac{dt}{(t+iy)^2\,(t^2+y^2)^\varepsilon} \;=\; \frac{I(\varepsilon)}{y^{1+2\varepsilon}}\,, \end{aligned}$$

where $I(\varepsilon)=\int_{-\infty}^{\infty}(t+i)^{-2}(t^2+1)^{-\varepsilon}dt$, so $\sum_{m=1}^{\infty}I_\varepsilon(mz)=I(\varepsilon)\zeta(1+2\varepsilon)/y^{1+2\varepsilon}$ for $\varepsilon>0$. Finally, we have $I(0)=0$ (obvious),

$$I'(0) \;=\; -\int_{-\infty}^{\infty}\frac{\log(t^2+1)}{(t+i)^2}\,dt \;=\; \biggl(\frac{1+\log(t^2+1)}{t+i} \;-\; \tan^{-1}t\biggr)\bigg|_{-\infty}^{\infty} \;=\; -\pi\,,$$

and $\zeta(1+2\varepsilon)=\dfrac{1}{2\varepsilon}+\mathrm{O}(1)$, so the product $I(\varepsilon)\zeta(1+2\varepsilon)/y^{1+2\varepsilon}$ tends to $-\pi/2y$ as $\varepsilon\to0$. The claim follows.

Remark. The transformation equation (18) says that G_2 is an example of what is called a *quasimodular* form, while the functions G_2^*, E_2^* and $\mathbb{G}_2^*$ defined in (21) are so-called *almost holomorphic modular forms* of weight 2. We will return to this topic in Section 5.

2.4 The Discriminant Function and Cusp Forms

For $z\in\mathfrak{H}$ we define the *discriminant function* $\Delta(z)$ by the formula

$$\Delta(z) \;=\; e^{2\pi iz}\prod_{n=1}^{\infty}\bigl(1-e^{2\pi inz}\bigr)^{24}\,. \tag{22}$$

(The name comes from the connection with the discriminant of the elliptic curve $E_z = \mathbb{C}/(\mathbb{Z}.z+\mathbb{Z}.1)$, but we will not discuss this here.) Since $|e^{2\pi iz}| < 1$ for $z \in \mathfrak{H}$, the terms of the infinite product are all non-zero and tend exponentially rapidly to 1, so the product converges everywhere and defines a holomorphic and everywhere non-zero function in the upper half-plane. This function turns out to be a modular form and plays a special role in the entire theory.

Proposition 7. *The function* $\Delta(z)$ *is a modular form of weight* 12 *on* $SL(2,\mathbb{Z})$*.*

Proof. Since $\Delta(z) \neq 0$, we can consider its logarithmic derivative. We find

$$\frac{1}{2\pi i}\,\frac{d}{dz}\log\Delta(z) \;=\; 1-24\sum_{n=1}^{\infty}\frac{n\,e^{2\pi inz}}{1-e^{2\pi inz}} \;=\; 1-24\sum_{m=1}^{\infty}\sigma_1(m)\,e^{2\pi imz} \;=\; E_2(z)\,,$$

where the second equality follows by expanding $\dfrac{e^{2\pi inz}}{1-e^{2\pi inz}}$ as a geometric series $\sum_{r=1}^{\infty} e^{2\pi irnz}$ and interchanging the order of summation, and the third equality from the definition of $E_2(z)$ in (17). Now from the transformation equation for E_2 (obtained by comparing (19) and(11)) we find

$$\begin{aligned}\frac{1}{2\pi i}\,\frac{d}{dz}\log\biggl(\frac{\Delta\bigl(\frac{az+b}{cz+d}\bigr)}{(cz+d)^{12}\Delta(z)}\biggr) &= \frac{1}{(cz+d)^2}\,E_2\biggl(\frac{az+b}{cz+d}\biggr) - \frac{12}{2\pi i}\,\frac{c}{cz+d} - E_2(z)\\ &= 0\,.\end{aligned}$$

In other words, $(\Delta|_{12}\gamma)(z) = C(\gamma)\,\Delta(z)$ for all $z \in \mathfrak{H}$ and all $\gamma \in \Gamma_1$, where $C(\gamma)$ is a non-zero complex number depending only on γ, and where $\Delta|_{12}\gamma$ is defined as in (8). It remains to show that $C(\gamma) = 1$ for all γ. But $C : \Gamma_1 \to \mathbb{C}^*$ is a homomorphism because $\Delta \mapsto \Delta|_{12}\gamma$ is a group action, so it suffices to check this for the generators $T = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ of Γ_1. The first is obvious since $\Delta(z)$ is a power series in $e^{2\pi iz}$ and hence periodic of period 1, while the second follows by substituting $z = i$ into the equation $\Delta(-1/z) = C(S)\,z^{12}\Delta(z)$ and noting that $\Delta(i) \neq 0$.

Let us look at this function $\Delta(z)$ more carefully. We know from Corollary 1 to Proposition 2 that the space $M_{12}(\Gamma_1)$ has dimension at most 2, so $\Delta(z)$ must be a linear combination of the two functions $E_4(z)^3$ and $E_6(z)^2$. From the Fourier expansions $E_4^3 = 1 + 720q + \cdots$, $E_6(z)^2 = 1 - 1008q + \cdots$ and $\Delta(z) = q + \cdots$ we see that this relation is given by

$$\Delta(z) \;=\; \frac{1}{1728}\left(E_4(z)^3 \;-\; E_6(z)^2\right). \tag{23}$$

This identity permits us to give another, more explicit, version of the fact that every modular form on Γ_1 is a polynomial in E_4 and E_6 (Proposition 4). Indeed, let $f(z)$ be a modular form of arbitrary even weight $k \geq 4$, with Fourier expansion as in (3). Choose integers $a,\, b \geq 0$ with $4a + 6b = k$ (this is always