Lattices, Linear Codes, and Invariants, Part I

Noam D. Elkies

How should 24-dimensional toy merchants most efficiently store their marbles? This is one rather fanciful statement of the "sphere packing problem" in $\mathbb{R}^{24}$. This problem is not just a plaything of high-dimensional Euclidean geometry: it relates to a surprising range of mathematical disciplines, pure as well as applied, including number theory, finite groups, orthogonal polynomials, and signal transmission. The same is true of the closely related discrete problem of error-correcting codes. This is already true for the important special cases of "lattice" packings and "linear codes".

The present article is a two-part series devoted to lattices, linear codes, and their relations with other branches of mathematics. Even a two-part series does not afford enough space to indicate all the mathematical disciplines relevant to the study of lattices and codes; we have chosen to focus our attention on certain *invariants* attached to lattices and codes. In each case these are invariant in two senses: they can (but do not always) distinguish nonisomorphic lattices or codes, and they can be written as generating functions that are invariant, or at least transform predictably, under certain transformations of the variables. Part I, in this issue, mainly concerns lattices, whose relevant invariants are "theta functions". Linear codes, and their close connections with lattices, will be the theme of Part II.

Noam D. Elkies is professor of mathematics at Harvard University. His e-mail address is elkies@math.harvard.edu.

As is usually the case in expository works, very little in Part I (in fact nothing outside the parenthetical remark on [EOR]) is my own work. I have attributed all results and ideas whose authors are known to me and apologize in advance if I have misattributed anything, or did not give a source for a result whose origin I do not know or wrongly believed to be classical or well known. At any rate, I do not claim any such result as my own.

The Sphere Packing Problem

The *sphere packing problem* is: *What is the maximal density of a sphere packing in a Euclidean space $\mathbb{R}^n$ of given dimension n?* Here a *sphere packing* in $\mathbb{R}^n$ is a collection S of balls of the same radius r whose interiors are disjoint; the *density* $\Delta(S)$ is the proportion of (the volume of) $\mathbb{R}^n$ covered by S.[1] Intuitively, one wants to pack a large container with identical n-dimensional balls as efficiently as possible. Modern mathematical terminology distinguishes between the (closed) "ball" $\overline{B_r}(x_0) := \{x \mid r \geq |x - x_0|\}$ of radius r about x_0 and the "sphere" $\{x \mid d(x, x_0) = r\}$; one would thus like to speak of "ball packing" rather than "sphere packing", but the problem long predates the

[1] *Since $\mathbb{R}^n$ has infinite volume, this proportion must be interpreted as $\lim_U \mathrm{Vol}(\cup_{B\in S} B \cap U)/\mathrm{Vol}(U)$ with U in an increasing sequence of convex subsets of $\mathbb{R}^n$ whose union is all of $\mathbb{R}^n$. For arbitrary sphere packings S, the limit need not exist, or could depend on the choice of U, so it is not immediately obvious that a maximal density exists. Fortunately the existence of a maximal density is not too hard to show, and all S that we shall consider will have an evident density $\Delta(S)$ not depending on the choice of U.*

modern terminology, and one usually writes of S as a family of "spheres" rather than "balls" in this context.

For $n \le 3$, sphere packing in $\mathbb{R}^n$ provides a good model for many familiar packing problems: pennies on a tabletop, atoms of a single element in a crystal, oranges or cannonballs in a crate, etc. The problem is trivial for $n = 1$: the maximal $\Delta = \Delta(S)$ is 1, and is attained for instance by $S = \{[2k-1, 2k+1] \mid k \in \mathbb{Z}\}$. See Figure 1. The case $n = 2$ is less trivial. The densest packing has been known since antiquity but was proved optimal only early in the twentieth century. To obtain it, tile the plane with regular hexagons of side $2r/\sqrt{3}$, and let S consist of the circles inscribed in those hexagons; this is a circle packing with $\Delta(S) = \pi/2\sqrt{3} \approx .9069$. See Figure 2. For $n = 3$ the problem was posed by Kepler. Again, the solution has long been surmised, but a proof was announced only a few years ago; see T. Hales's article in the April 2000 *Notices*.

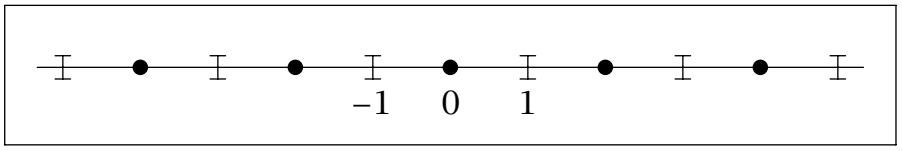

Figure 1. The densest packing in one dimension. The collection S consists of all intervals $[2k-1, 2k+1]$ with k an integer.

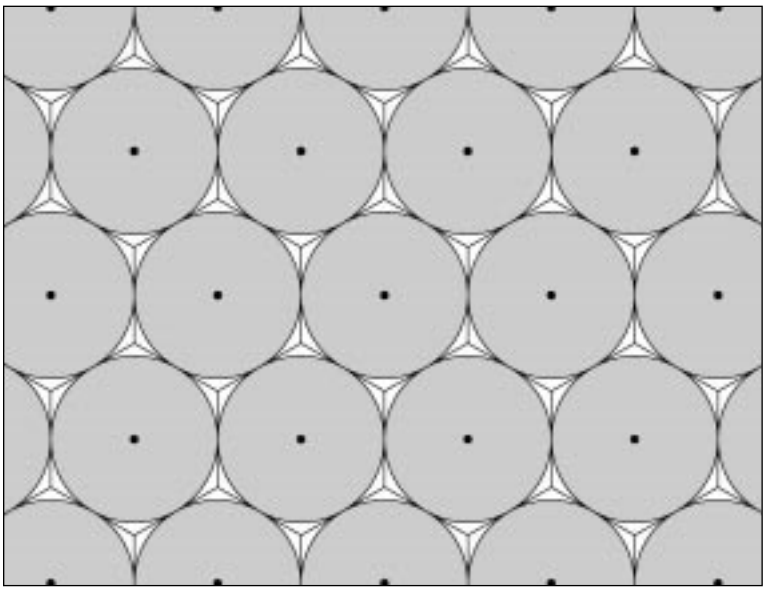

Figure 2. The densest packing in two dimensions. The plane is tiled with regular hexagons, and S consists of the circles inscribed in those hexagons.

Naturally the sphere packing problem in dimension $n \ge 4$ is more recent, though still at least a century old: it is contained in the eighteenth of the famous list of problems posed by Hilbert in 1900. While one does not as often need to pack 24-dimensional marbles as three-dimensional ones, packing spheres in high dimensions still has important applications, as in signal processing, where the space of available signals often has the structure of a bounded but large subset of $\mathbb{R}^n$ for some large n. Of course, the problem has the same mathematical appeal for $n \ge 4$ as for $n < 4$; as noted in the introduction, there are also many specific applications and connections within mathematics.

At present, the sphere packing problem has not been solved for any $n \ge 4$. A few n may be tractable: the approach that solved the case $n = 3$ might eventually, with considerably more effort and computation, handle $n = 4$ as well; other methods may settle the cases $n = 8$ and $n = 24$, which admit remarkably dense packings that we shall describe later. But in general one does not expect to get a closed form for the maximal density as a function of n: one can prove upper bounds and construct or prove the existence of packings that yield lower bounds, but these upper and lower bounds are usually quite far apart.

A simple lower bound is 2^{-n}, obtained by starting from any sphere packing and adding more and more disjoint balls to it until no room is left. One then has a packing $S = \{\overline{B_r}(x_i)\}$ such that $\cup_i \overline{B_{2r}}(x_i) = \mathbb{R}^n$. Since each $\overline{B_{2r}}(x_i)$ has volume $2^n \mathrm{Vol}(\overline{B_r}(x_i))$, it follows that S has $\Delta(S) \ge 2^{-n}$. Exactly the same bound applies to packings of translates of any given centrally symmetric convex body of finite positive volume; in this generality it is known as the *Minkowski-Hlawka bound*, Minkowski having obtained it for spheres with a rather more complicated method. It is rather embarrassing that for large n this simple-minded approach to the sphere packing problem yields a bound that is within a factor $n^{O(1)}$ of the best lower bound known, where $O(1)$ denotes a bounded expression. (For other families of centrally symmetric bodies, notably l_p balls with $p > 2$, one can improve the bound to c^{-n} for some $c < 2$, as shown in [EOR].) By comparison, the best asymptotic upper bound known is $2^{-An+o(n)}$ with $A \approx .599$, requiring a much subtler argument that applies only to spheres and their affine images [CS, Ch. 9, especially pp. 247, 265]; here $o(n)$ denotes an expression that, when divided by n, tends to 0 as n tends to infinity.

Lattices and Lattice Packings of Spheres

An important special case of a sphere packing is a *lattice packing*. Any sphere packing S is completely described by its common radius r and the set C of centers of the spheres in the packing: $S = \{\overline{B_r}(x) \mid x \in C\}$. The spheres do not overlap if and only if $d(x, x') \ge 2r$ for all distinct $x, x' \in C$, i.e., if and only if the *minimal (nonzero) distance* of C is at least $2r$. To find $\Delta(S)$, multiply the density of C by the volume of a ball of radius r in $\mathbb{R}^n$;

here the *density* of a point set is the average number of points per unit volume, and the ball volume is given by the formula $r^n\pi^{n/2}/\Gamma(n/2+1)$. Now S is said to be a "lattice packing" if C is a *lattice* in $\mathbb{R}^n$. This means that C is a discrete additive subgroup of $\mathbb{R}^n$ not contained in any hyperplane; equivalently, C consists of all integer linear combinations $\sum_{j=1}^n c_j v_j$ of some vectors $v_j \in \mathbb{R}^n$ that constitute a basis for $\mathbb{R}^n$. This is true of the optimal packings for $n=1$, where $C=2\mathbb{Z}$, and for $n=2$, where C consists of the centers of the hexagons (as long as the hexagonal tiling is translated so that one of the hexagons is centered on the origin). The maximal density is also attained by a lattice packing for $n=3$ and conjecturally for many other n, including all $n \le 8$ and $n=24$.

These conjecturally maximal packings are known to be optimal at least *among lattice packings* for all $n \le 8$. For instance, the best lattice packing for $n=8$ may be described as follows: C consists of the vectors in $\mathbb{Z}^8 \cup (\mathbb{Z}+\frac{1}{2})^8$ the sum of whose coordinates is even. Equivalently, this is the lattice of integer linear combinations of the eight vectors $2e_1$, $(e_j - e_{j+1})$ for $1 \le j \le 6$, and $\frac{1}{2}\sum_{j=1}^8 e_j$, where $e_1, \dots, e_8$ are the standard unit vectors in $\mathbb{R}^8$. Here $r = 2^{-1/2}$ and C has density 1, so $\Delta(S) = \pi^4/384 \approx .2537$. This $C \subset \mathbb{R}^8$ is nowadays called the E_8 lattice; we shall say much more about it later.

The minimal distance of any lattice C is the minimum of $|x|$ over nonzero $x \in C$, i.e., the *minimal (nonzero) length* of C.[2] For E_8, this minimal length is $\sqrt{2}$, realized by the 112 vectors $\pm e_i \pm e_j$ for $i<j$ and the 128 vectors $\frac{1}{2}\sum_{j=1}^8 a_j e_j$ with $a_j = \pm 1$ and $\prod_j a_j = 1$, for a total of 240.

Let A be the invertible $n \times n$ matrix with column vectors $v_1, \dots, v_n$; then the lattice generated by these vectors is $C = \{Ac \mid c \in \mathbb{Z}^n\}$, with

$$\text{density of } C = 1/|\det A|.$$

For any given C, there are many choices for the generators v_j and thus for the "generator matrix" A: the columns of a matrix B generate C if and only if $B = AM$ for some $n \times n$ integer matrix M of determinant ± 1. Such matrices M constitute a group denoted by $\mathrm{GL}_n(\mathbb{Z})$. A lattice C' is *isometric* with C if it is obtained from C by an orthogonal linear transformation of $\mathbb{R}^n$, and *homothetic* if it is isometric with αC for some $\alpha > 0$; isometry and homothecy are equivalence relations preserving the density of the associated sphere packing (under a homothecy the radius r is also multiplied by α). The lattices generated by matrices A and A' in $\mathrm{GL}_n(\mathbb{R})$ are isometric if and only if $A' = UAM$ for some $M \in \mathrm{GL}_n(\mathbb{Z})$ and U in $\mathrm{O}_n(\mathbb{R})$, the group of orthogonal $n \times n$ matrices; likewise the lattices are homothetic if and only if $A' = \alpha UAM$, for some $\alpha > 0$, U, and M, with U and M as before. Since $|\det A'| = \alpha^n |\det A|$, there is a unique α, namely $|\det A|^{-1/n}$, such that $|\det A'| = 1$. Thus a lattice in $\mathbb{R}^n$ is specified by an $n \times n$ matrix of determinant ± 1 up to multiplication from the left and right by matrices from $\mathrm{O}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{Z})$. We may choose A and U to have determinant +1, i.e., to lie in the "special" linear and orthogonal groups of $n \times n$ matrices; we have thus identified the set $\mathbf{\Lambda}_n$ of homothecy classes of lattices in $\mathbb{R}^n$ with the double coset space $\mathrm{SO}_n(\mathbb{R})\backslash\mathrm{SL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$. This is also the space of isometry classes of lattices of density 1, also known as *unimodular* lattices.

This tells us several things about $\mathbf{\Lambda}_n$. We first obtain its dimension. The Lie groups $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{SO}_n(\mathbb{R})$ have dimension n^2-1 and $\binom{n}{2}$ respectively, while $\mathrm{SL}_n(\mathbb{Z})$ is discrete. Thus $\mathbf{\Lambda}_n$ has dimension $n^2 - 1 - \binom{n}{2} = (n-1)(n+2)/2$. For $n=1$ this dimension is 0 as expected, since $\mathbf{\Lambda}_n$ consists of a single point: all lattices in $\mathbb{R}$ are homothetic. For $n=2$ we find that $\mathbf{\Lambda}_n$ is closely related to the familiar action of $\mathrm{SL}_2(\mathbb{Z})$ on the upper half-plane by fractional linear transformations. Indeed, each of the cosets in $\mathrm{SO}_2(\mathbb{R})\backslash\mathrm{SL}_2(\mathbb{R})$ has a unique representative of the form $y^{-1/2}\left(\begin{smallmatrix} 1 & x \\ 0 & y \end{smallmatrix}\right)$, obtained by rotating the first column vector to a positive multiple of $\binom{1}{0}$; thus $\mathrm{SO}_2(\mathbb{R})\backslash\mathrm{SL}_2(\mathbb{R})$ is isomorphic with the upper half-plane

$$\mathcal{H} := \{\tau = x + iy \mid y > 0\}.$$

If a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$ has determinant +1 Z transformation $\tau \mapsto (a\tau + b)/(c\tau + d)$. This, together with the fact that the involution $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ of determinant -1 is required to act by $x+iy \mapsto -x+iy$, completely describes our action of $\mathrm{GL}_2(\mathbb{Z})$ from the right on $\mathrm{SO}_2(\mathbb{R})\backslash\mathrm{SL}_2(\mathbb{R})$. It is a classical fact that

$$\{(x,y) \mid x^2 + y^2 \ge 1,\ |y| \le 1/2\}$$

is a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$; adding to this the condition $x \ge 0$ yields a fundamental domain for $\mathrm{GL}_2(\mathbb{Z})$. See Figure 3. If $x+iy$ is a point in that fundamental domain, the corresponding unimodular lattice has minimal length $y^{-1/2}$. The point $x+iy = e^{\pi i/3} = (1+\sqrt{-3})/2$ of the domain has minimal y and thus maximal $y^{-1/2}$. We conclude that the density is maximized by the hexagonal lattice, here scaled to density 1. Once $n > 2$ it is of course much harder to visualize $\mathbf{\Lambda}_n$; the spaces $\mathbf{\Lambda}_n$ become still more complicated as n increases, but at least for small n there is a classical "reduction theory" that yields explicit fundamental domains and

[2] *We avoid the word "norm", which has two common and conflicting meanings: in the context of lattices and quadratic forms, a vector x is usually said to have norm $|x|^2$; but elsewhere in mathematics $|x|$ is often called the norm of x.*

identifies the densest lattice packings of spheres in those dimensions.

The identification of $\mathbf{\Lambda}_n$ with

$$\mathrm{SO}_n(\mathbb{R})\backslash \mathrm{SL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$$

also yields a natural choice of *measure* on $\mathbf{\Lambda}_n$, from invariant Haar measures on $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{SO}_n(\mathbb{R})$. For $n=2$ the resulting measure on $\mathcal{H} = \mathrm{SO}_2(\mathbb{R})\backslash \mathrm{SL}_2(\mathbb{R})$, and thus on $\mathbf{\Lambda}_2 = \mathcal{H}/\mathrm{GL}_2(\mathbb{Z})$, is the one coming from the familiar *hyperbolic metric* $dx\,dy/y^2$ invariant under all fractional linear transformations $\tau \mapsto (a\tau+b)/(c\tau+d)$ with $ad-bc>0$. The $n=2$ fundamental region, though not compact, has finite area since $\int_{y_0}^{\infty} dy/y^2 < \infty$; it is known that in fact $\mathbf{\Lambda}_n$ has finite measure for all n. Thus integration yields a well-defined notion of averaging over $\mathbf{\Lambda}_n$—intuitively, of properties of a "random" or "typical" unimodular lattice. For instance, we may ask for the average minimal length of such a lattice or the probability that a random lattice in $\mathbb{R}^n$ has minimal length $\geq 2r$. If that probability is positive, then there exists such a lattice and thus a lattice packing of spheres with $\Delta \geq r^n \pi^{n/2}/\Gamma(n/2+1)$. A lower bound r_0 on r therefore yields a lower bound on an n-dimensional lattice packing of spheres. It is not hard to derive an r_0 that yields lattice packings with $\Delta \geq 2^{-n}$, the same result we found earlier for unrestricted packings. (The Minkowski-Hlawka bound likewise applies even to lattice packings.) This makes it even more embarrassing that we cannot do better than $n^{O(1)}2^{-n}$ for large n: we cannot pack n-dimensional spheres significantly better than is achieved by a *random* lattice!

Of course, it is conceivable that the cases of small n, where the best lattices are much better than random, are misleading and that for high n the random lattices really are essentially best possible; it is an important open question whether there exist sphere packings with $\Delta > 2^{-\theta n}$ for some fixed $\theta < 1$ and all n. But our embarrassment is even more acute than this indicates: not only can we not improve on 2^{-n}, but also we do not even know how to attain this density for large n (say $n > 2000$). It is not even easy to construct any sequence of lattices of dimensions $n \to \infty$ whose densities exceed $2^{-\theta n}$ for any fixed $\theta < \infty$; the smallest θ for which such a sequence is known is ≈ 1.39, using surprisingly sophisticated ideas from number theory and algebraic geometry [TV, p. 585].

This seems paradoxical: surely we could take say $n = 5000$ and just choose a (pseudo)random lattice C, which will almost certainly yield a sphere packing S with $\Delta(S) > 2^{-5010}$. A lattice in $\mathbb{R}^n$ is specified, for instance, by the n^2 entries of a generator matrix; with modern computers it is easy to store and manipulate n^2 entries for $n = 5000$. Generating random lattices according to the natural measure on $\mathbf{\Lambda}_n$ is not entirely trivial, but there are known and efficient ways to do this given a good pseudorandom number generator. The explanation of the apparent paradox is that for large n we know no feasible way to compute the minimal length of a general lattice C given its generators. Much work has been spent on this important problem, but even modern lattice reduction techniques fail long before n reaches 5000. Thus our C almost certainly yields a good sphere packing, but we cannot prove it.

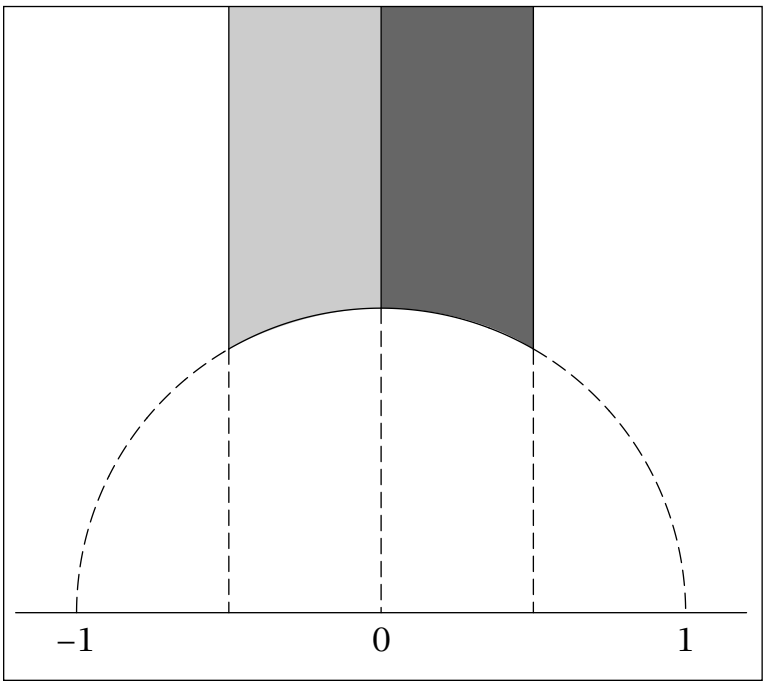

Figure 3. Fundamental domains for the actions of $\mathrm{SL}_2(\mathbb{Z})$ and $\mathrm{GL}_2(\mathbb{Z})$ on the upper half plane. The circle is the unit circle, and the vertical lines are at $x=-1/2$, $x=0$, and $x=1/2$. The total shaded region is a fundamental domain in the case of $\mathrm{SL}_2(\mathbb{Z})$, while the right-hand half is a fundamental domain in the case of $\mathrm{GL}_2(\mathbb{Z})$.

Theta Functions and Modular Forms

Sometimes an appropriate response to a difficult mathematical problem is to pose a much harder problem. Here we find the minimal nonzero length intractable, and thus ask for *all* the lengths of vectors of C and their multiplicities. Equivalently, we ask for the following generating function of the squared lengths, called the *theta function* (or *theta series*) of C:

$$\Theta_C(z) := \sum_{x\in C} z^{(x,x)} = 1 + \sum_{m>0} N_m(C) z^m,$$

where $N_m(C) = \#\{x \in C \mid (x,x) = m\}$ is the number of lattice vectors of length $\sqrt{m}$. For instance,

$$\Theta_{\mathbb{Z}}(z) = \sum_{k=-\infty}^{\infty} z^{k^2} = 1 + 2(z + z^4 + z^9 + \cdots),$$

a specialization of the classical elliptic function ϑ_3. The series defining $\Theta_C(z)$ converges absolutely for $0 \leq z < 1$. As a function of a complex variable, Θ_C

may have a branch point at the origin. But a substitution $z = e^{ci\tau}$ (c a positive constant) yields an analytic function of $\tau \in \mathcal{H}$. Naturally there is no way known to compute $\Theta_C(z)$ for a general lattice $C \in \mathbb{R}^n$ once n is large enough; but theta functions, relations among them, and their properties as analytic functions yield much important information concerning lattices and their vectors' lengths.[3]

An easy identity relates the theta functions of any two lattices $C_1 \subset \mathbb{R}^{n_1}$, $C_2 \subset \mathbb{R}^{n_2}$ with the theta function of their direct sum $C_1 \oplus C_2 \subset \mathbb{R}^{n_1+n_2}$:

$$\Theta_{C_1 \oplus C_2}(z) = \Theta_{C_1}(z)\Theta_{C_2}(z).$$

A more interesting identity arises from the *Poisson summation formula*. Given a suitable function $f : \mathbb{R}^n \to \mathbb{C}$ and a lattice $C \subset \mathbb{R}^n$, Poisson's formula relates the sum $\sum_{x \in C} f(x)$ of f over C with the sum of the Fourier transform

$$\hat{f}(y) := \int_{\mathbb{R}^n} f(x) e^{2\pi i(x,y)} dx$$

of f over the *dual lattice* C^*. The dual lattice is defined by

$$C^* := \{y \in \mathbb{R}^n \mid (x, y) \in \mathbb{Z} \text{ for all } x \in C\};$$

as the name suggests, C^* is again a lattice in $\mathbb{R}^n$, and $(C^*)^* = C$. If A is a generator matrix for C then the transpose of A^{-1} generates C^*. Thus the densities of C and C^* are reciprocals. A function f is "suitable" for the Poisson formula if both f and $\hat{f}$ decrease rapidly enough; in particular, the Gaussian $f(x) = \exp(-c(x, x))$ is suitable for all $c > 0$ (more generally, all c of positive real part). The Fourier transform of a Gaussian is again proportional to a Gaussian, so the Poisson formula will relate Θ_C with Θ_{C^*}. We find:

$$\Theta_C(e^{-\pi y}) = D y^{-n/2} \Theta_{C^*}(e^{-\pi/y})$$

where D is the density of C.

We note in passing the following interpretation of this identity in terms of the geometry of $T_C := \mathbb{R}^n/C$. Topologically, T_C is just an n-torus, but we consider it as a Riemannian manifold, with the metric inherited from $\mathbb{R}^n$. The lengths of vectors in C can then be interpreted as the lengths of closed geodesics on T_C. As to the lengths of $y \in C^*$, those are proportional to square roots of eigenvalues of the Laplacian on T_C, associated to the eigenfunction $x \mapsto e^{2\pi i(x,y)}$. The Poisson formula thus relates for T_C the spectrum of the Laplacian with the geodesic lengths.

The Poisson formula is particularly nice when $C = C^*$, i.e., when C is *self-dual*. This happens if and only if C is unimodular and *integral*, integral meaning that $(x, y) \in \mathbb{Z}$ for all $x, y \in C$. In terms of a generator matrix A, this means that $\det(A) = \pm 1$ and the "Gram matrix" $A^{\mathrm{T}}A$ has integer entries. Examples of self-dual lattices are $\mathbb{Z}$ and the lattice E_8 exhibited above, as well as $C_1 \oplus C_2$ when C_1 and C_2 are self-dual. If C is self-dual then the Poisson formula is a functional equation for Θ_C. Also, the squared lengths of all the lattice vectors are automatically integers, so Θ_C is an analytic function on the unit circle $|z| < 1$. An alternative and even nicer formulation of this is in terms of the analytic function

$$\theta_C(\tau) := \Theta_C(e^{\pi i \tau}) = \sum_{x \in C} e^{\pi i (x,x)\tau}$$

on $\mathcal{H}$, mentioned earlier. The functional equation relates $\theta_C(\tau)$ with $\theta_C(-1/\tau)$; the integrality yields $\theta_C(\tau) = \theta_C(\tau + 2)$. Iterating these two identities gives the ratio between $\theta_C(\tau)$ and $\theta_C(g(\tau))$ where g is any fractional linear transformation in the group Γ generated by the involution $\tau \leftrightarrow -1/\tau$ and the translation $\tau \mapsto \tau + 2$. These are the maps $\tau \mapsto (a\tau + b)/(c\tau + d)$ with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$. It can be shown that this is the same as the index-3 subgroup of $\mathrm{SL}_2(\mathbb{Z})$ consisting of matrices congruent mod 2 to either the identity $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ or the involution $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Thus Γ is an example of a *congruence subgroup* of $\mathrm{SL}_2(\mathbb{Z})$, i.e., a subgroup defined by congruence conditions on the matrix entries. We find that, for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$,

$$\theta_C\left(\frac{a\tau + b}{c\tau + d}\right) = (\epsilon_g \sqrt{c\tau + d})^n \theta_C(\tau),$$

where $\sqrt{\cdot}$ is the principal branch of the square root and ϵ_g is an 8th root of unity depending on (c, d) in an explicit but somewhat complicated way (which incorporates Quadratic Reciprocity!).

This condition on θ_C is quite demanding. An analytic function ϕ on $\mathcal{H}$ satisfying the identity $\phi\left(\frac{a\tau+b}{c\tau+d}\right) = (\epsilon_g \sqrt{c\tau + d})^n \phi(\tau)$ for all $g \in \Gamma$ (and some mild growth conditions that are automatically satisfied by a theta function) is called a *modular form of weight* $(n/2)$ *for* Γ. The subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is one of many congruence subgroups for which the modular forms have been determined completely. For Γ, it is known that any modular form is a weighted homogeneous polynomial in $\Theta_{\mathbb{Z}}$, which has weight $1/2$, and the modular form

$$\begin{aligned}\delta_8(t) &:= e^{\pi i \tau} \prod_{m=1}^{\infty} \left((1 - e^{\pi i m \tau})(1 + e^{2\pi i m \tau})\right)^8 \\ &= e^{\pi i \tau} - 8e^{2\pi i \tau} + 28e^{3\pi i \tau} \\ &\quad - 64e^{4\pi i \tau} + 126e^{5\pi i \tau} \cdots\end{aligned}$$

of weight 4. So, for instance, if $n < 8$, we conclude that $\theta_C = \theta_{\mathbb{Z}}^n$. In particular, C has $2n$ vectors of length 1, from which it readily follows that C is isomorphic with $\mathbb{Z}^n$, the direct sum of n copies of

[3] *Further discussion of some of the topics in this section may be found in [S].*

the lattice $\mathbb{Z}$. For $n = 8$ we have a lattice E_8 that is not isomorphic with $\mathbb{Z}^8$, because E_8 has no vectors of length 1 at all. It is known (see, e.g., [CS])[4] that any self-dual lattice in $\mathbb{R}^8$ is isomorphic with either $\mathbb{Z}^8$ or E_8. Thus θ_{E_8} must be the unique linear combination of $\theta_{\mathbb{Z}}^8$ and δ_8 whose constant and $e^{\pi i t}$ coefficients are respectively 1 and 0. We calculate

$$\begin{aligned}\theta_{E_8}(\tau) &= \theta_{\mathbb{Z}}^8(\tau) - 16\delta_8(\tau)\\ &= 1 + 240e^{2\pi i\tau} + 2160e^{4\pi i\tau}\\ &\quad + 6720e^{6\pi i\tau} + \cdots.\end{aligned}$$

We can now easily compute for each integer m the number $N_m(E_8)$ of vectors of E_8 of length $\sqrt{m}$: it is simply the coefficient of $e^{m\pi i\tau}$ in the expansion of θ_{E_8}. For instance, we confirm our count $N_2(E_8) = 240$ of minimal vectors.

Continuing in this manner past $m = 2$, we find that not only does N_1 vanish but so do $N_3, N_5, N_7, \ldots$. Could it be that $N_m(E_8) = 0$ for all odd m, and thus that the length of every vector of E_8 is the square root of an *even* integer? It transpires that this is true, and not hard to prove. The key is that, for any integral lattice C, the map ν from C to $\mathbb{Z}/2\mathbb{Z}$ defined by $x \mapsto (x, x) \bmod 2$ is a homomorphism. Our claim is then that for E_8, this ν is the zero homomorphism, and it is enough to check it on generators of the lattice. We have already exhibited generators each of whose length is $\sqrt{2}$ or 2, whence our claim follows. A lattice is said to be *even* if the length of every lattice vector is the square root of an even integer. Suppose C is a lattice that, like E_8, is both even and self-dual. Then θ_C is a linear combination of terms $e^{\pi i m\tau}$ with $2|m$, and thus is invariant not only under $\tau \mapsto \tau + 2$ but also under $\tau \mapsto \tau + 1$. Therefore θ_C is a modular form of weight $n/2$ for the group generated by $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$. It turns out that this group is all of $\mathrm{SL}_2(\mathbb{Z})$.

The modular forms for $\mathrm{SL}_2(\mathbb{Z})$ are again known: they are the weighted-homogeneous polynomials in θ_{E_8}, which has weight 4, and the weight 12 form

$$\begin{aligned}\delta_{24}(\tau) &:= e^{2\pi i\tau}\prod_{m=1}^{\infty}(1 - e^{2\pi i m\tau})^{24}\\ &= e^{2\pi i\tau} - 24e^{4\pi i\tau}\\ &\quad + 252e^{6\pi i\tau} - 1472e^{8\pi i\tau}\cdots\end{aligned}$$

whose $e^{2\pi i m\tau}$ coefficient is the value at m of Ramanujan's celebrated multiplicative tau function.[5] In particular, $n/2$ must be a multiple of 4, so *if $\mathbb{R}^n$ contains an even self-dual lattice then n is a multiple of* 8. This necessary condition is also sufficient: if $8|n$ then the direct sum of $n/8$ copies of E_8 is an even self-dual lattice in $\mathbb{R}^n$.

The coefficients of θ_{E_8} can even be given in closed form by identifying that theta function with a normalized *Eisenstein series*. For each $k = 1, 2, 3, \ldots$, the normalized Eisenstein series

$$\mathrm{E}_{4k}(\tau) := \frac{1}{2\zeta(4k)}\sum_{\substack{c,d\in\mathbb{Z}\\(c,d)\neq(0,0)}}\frac{1}{(c\tau + d)^{4k}}$$

is a modular form of weight $4k$ for $\mathrm{SL}_2(\mathbb{Z})$. Here ζ is Riemann's zeta function, defined by $\zeta(s) := \sum_{n=1}^{\infty} n^{-s}$ $(s > 1)$; the normalizing factor $1/2\zeta(4k)$ assures that $\mathrm{E}_{4k}(\tau) \to 1$ as $\tau \to i\infty$. Using the Fourier series for $\sum_{d\in\mathbb{Z}}(x + d)^{-4k}$ to expand $\mathrm{E}_{4k}(\tau)$ in powers of $e^{2\pi i\tau}$, one finds

$$\mathrm{E}_{4k}(\tau) = 1 + \frac{4k}{-B_{4k}}\sum_{m=1}^{\infty}\sigma_{4k-1}(m)e^{2\pi i m\tau}.$$

Here B_{4k} is the $(4k)$-th Bernoulli number $-2(4k)!\zeta(4k)/(2\pi)^{4k}$, known to be rational for each $k = 1, 2, 3, \ldots$; and $\sigma_j(m)$ is the sum of the j-th powers of the positive divisors of m. For instance, $\mathrm{E}_4(t) = 1 + 240\sum_{m=1}^{\infty}\sigma_3(m)e^{2\pi i m\tau}$ is a modular form of weight 4 for $\mathrm{SL}_2(\mathbb{Z})$. Since all such forms are multiples of θ_{E_8}, and E_4, θ_{E_8} have the same constant coefficient 1, we conclude that $\theta_{E_8} = \mathrm{E}_4$. Thus for each positive integer m there are exactly $240\sigma_3(m)$ vectors of E_8 of length $\sqrt{2m}$.

Similarly we conclude that $\theta_{E_8}^2 = \mathrm{E}_8 = 1 + 480\sum_{m=1}^{\infty}\sigma_7(m)e^{2\pi i m\tau}$. Equating the series for E_4^2 and E_8, we find the otherwise mysterious identity

$$\sigma_7(m) = \sigma_3(m) + 120\sum_{m_1+m_2=m}\sigma_3(m_1)\sigma_3(m_2).$$

Returning to theta series, we find that any even self-dual lattice C in $\mathbb{R}^{16}$ must have $\theta_C = \mathrm{E}_8$. While E_8 was the unique such lattice in dimension 8, there are two even self-dual lattices in dimension 16. One, of course, is $E_8 \oplus E_8$. The other is a lattice we might call E_{16}, obtained in the same way we defined E_8: it is generated by $2e_1$, $(e_j - e_{j+1})$ for $1 \le j \le 14$, and $\frac{1}{2}\sum_{j=1}^{16} e_j$, where $e_1, \ldots, e_{16}$ constitute an orthonormal basis for $\mathbb{R}^{16}$. That the lattices are not isomorphic may be seen from the fact that E_{16} is not generated by its 480 vectors $\pm e_i \pm e_j$ of length $\sqrt{2}$, whereas the minimal vectors of $E_8 \oplus E_8$ do generate $E_8 \oplus E_8$. Recalling our earlier discussion of $\mathbb{R}^n/C$, we find that the tori $\mathbb{R}^{16}/E_{16}$ and $\mathbb{R}^{16}/(E_8 \oplus E_8)$ are nonisomorphic compact Riemannian manifolds with the same

[4] *In general, for each n there are only finitely many isomorphism classes of self-dual lattices in $\mathbb{R}^n$, but their number grows superexponentially with n.*

[5] *The reader already conversant with modular forms for $\mathrm{SL}_2(\mathbb{Z})$ may well wonder what became of the Eisenstein series of weight 6. Its absence is explained by the factors ϵ_g^n forced on us by Poisson summation. These factors cannot extend consistently to all of $\mathrm{SL}_2(\mathbb{Z})$ unless $8|n$. To see this, consider $g = \left(\begin{smallmatrix}1 & -1\\ 1 & 0\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$, with $g^3 = -\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$.*

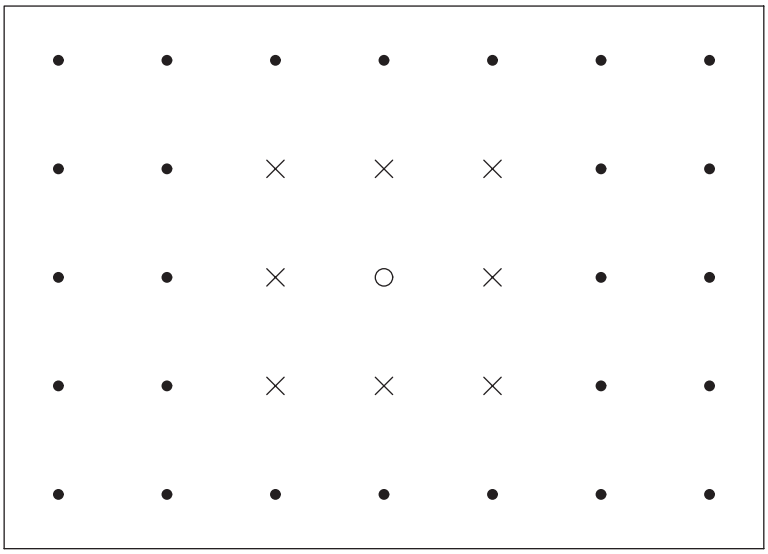

Figure 4. The root lattice of B_2. The origin is marked by a circle, and the roots are marked by crosses. Other lattice points are marked by solid dots. The resulting lattice is just $\mathbb{Z}^2$ in $\mathbb{R}^2$.

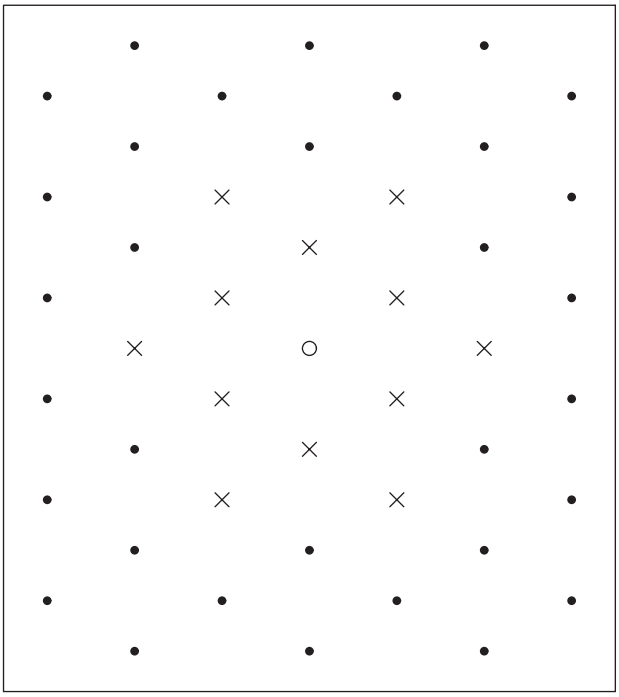

Figure 5. The root lattice of G_2. The origin is marked by a circle, and the roots are marked by crosses. Other lattice points are marked by solid dots. The resulting lattice is homothetic to the hexagonal lattice that appears in Figure 2 if the center of some circle in Figure 2 is regarded as the origin.

Laplacian spectrum; this was the first example of isospectral manifolds discovered (Witt 1941 and Milnor 1964).

In dimension 24 we find for the first time that the theta function of an even self-dual lattice C is not completely determined: the modular-form condition tells us only that $\theta_C = \theta_{E_8}^3 + a\delta_{24}$ for some integer a. To use C for sphere packing, we want its minimal length to be as large as possible; let us try, then, to choose a so that the $e^{2\pi i\tau}$ coefficient of θ_C vanishes, i.e., so that $N_2(C) = 0$ and all nonzero lattice vectors have length at least 2. This condition does completely determine θ_C; we find that $a = -720$ and thus that

$$\theta_C = \theta_{E_8}^3 - 720\delta_{24} = 1 + 196560e^{4\pi i\tau} + 16773120e^{6\pi i\tau} + \cdots$$

Thus any such C would have minimal length 2, attained by 196560 minimal vectors, and also 16773120 vectors of the next-lowest length $\sqrt{6}$, etc. By comparing the first few coefficients with those of the Eisenstein series E_{12}, we find that $N_{2m}(C) = (65520/691)(\sigma_{11}(m) - \tau(m))$ for each $m = 1, 2, 3, \ldots$, where $\tau(m)$ denotes Ramanujan's tau function.

This still begs the question of whether such a lattice exists. J. Leech constructed one in 1966; it was proved unique a few years later in several ways, among them H.-V. Niemeier's classification of all even self-dual lattices in $\mathbb{R}^{24}$ (the last case in which all such lattices are known). We shall describe Leech's remarkable lattice L_{24} explicitly in Part II, using the extended binary Golay code (also to be introduced in Part II). The density $\pi^{12}/12!$ of the resulting sphere packing in $\mathbb{R}^{24}$ is almost certainly the largest possible in that dimension.

In addition to classifying various kinds of lattices up to isomorphism, we may consider the automorphism group of a specific lattice C. The problem of finding the unimodular lattice $C \subset \mathbb{R}^n$ with the largest minimal length has a large symmetry group, namely the group $\mathrm{O}_n(\mathbb{R})$ of orthogonal linear transformations of $\mathbb{R}^n$. Often the solution of a highly symmetric problem will itself have a large and/or interesting symmetry group, and this happens in several cases for our problem. The automorphisms of any lattice $C \subset \mathbb{R}^n$ are those $\phi \in \mathrm{O}_n(\mathbb{R})$ such that $\phi(C) = C$. These constitute a discrete subgroup $\mathrm{Aut}(C)$ of the compact group $\mathrm{O}_n(\mathbb{R})$; thus $\mathrm{Aut}(C)$ is finite. This group always contains the central element -1, taking each $v \in \mathbb{R}^n$ to $-v$. For most lattices C (i.e., in all but a measure-zero subset of $\mathbf{\Lambda}_n$), $\{\pm 1\}$ is all of $\mathrm{Aut}(C)$, but at least in small dimensions the optimal lattices are far from typical in this respect.

One source for nontrivial automorphisms is reflections in hyperplanes. For nonzero $w \in \mathbb{R}^n$, the reflection r_w in the hyperplane orthogonal to w is given by $v \mapsto v - 2\big((v,w)/(w,w)\big)w$. Thus if $w \in C$ and $2(v,w) \in \mathbb{Z}(w,w)$ for all $v \in w$, then $r_w \in \mathrm{Aut}(\mathbb{C})$. Such vectors w are called *roots* of C. For example, if C is integral, then any $w \in C$ with $(w,w) = 1$ or $(w,w) = 2$ is a root. A *root lattice* is a lattice generated by the roots it contains. Roots and

root lattices arise naturally in the classification of Lie groups and algebras; in that context, the group generated by reflections r_w for certain roots $w \in C$ is called the "Weyl group" of the corresponding Lie algebra. For example, the root lattice of the Lie group B_n is isomorphic with $\mathbb{Z}^n$, with roots $\pm e_i$ and $\pm e_i \pm e_j$ for $1 \le i < j \le n$. See Figure 4. The corresponding reflections generate $\mathrm{Aut}(\mathbb{Z}^n)$. This group has a normal subgroup $\{\pm 1\}^n$ and is the semidirect product of that subgroup with the symmetric group Sym_n, acting on $\{\pm 1\}^n$ by permuting its coordinates. This group is also known as the "hyperoctahedral group" because it is the group of symmetries of the hyperoctahedron (cross-polytope) with vertices $\pm e_i$, and as the "signed permutation group" because it consists of the $n \times n$ signed permutation matrices, i.e., matrices with entries in $\{0, \pm 1\}$, each of whose rows and columns contain exactly one nonzero entry. The root lattice of the smallest exceptional Lie group G_2 is homothetic with the hexagonal lattice in $\mathbb{R}^2$. See Figure 5. The roots are the six minimal vectors, of length $\sqrt{2}$, and the six vectors of next smallest length $\sqrt{6}$; again the corresponding reflections generate the automorphism group, which here is the dihedral group of 12 elements. For a final example, the E_8 lattice is a root lattice, associated with the largest of the exceptional simple Lie groups (which is also called E_8). We have already described its 240 roots; the corresponding reflections again generate $\mathrm{Aut}(E_8)$, this time a group of order $2^{14}3^5 5^2 7$, which we describe further in the next paragraph.

The Leech lattice L_{24} has no roots, and indeed $\mathrm{Aut}(L_{24})$ contains no reflections in hyperplanes nor any other orientation-reversing automorphism. In other words, $\mathrm{Aut}(L_{24})$ is its own determinant-one subgroup. Even so, $\mathrm{Aut}(L_{24})$ cannot be simple, because it still contains the central involution -1. But the quotient of $\mathrm{Aut}(L_{24})$ by its center $\{\pm 1\}$ is simple. This quotient group Co_1 was determined by J. H. Conway and bears his name, as does $\mathrm{Aut}(L_{24})$ itself, usually[6] called Co_0. Now it is not unusual for simple groups to arise as quotients of $\mathrm{Aut}(C)$ (or of normal subgroups of $\mathrm{Aut}(C)$) for interesting lattices C; for instance, $\mathrm{Aut}(\mathbb{Z}^n)$ yields Alt_n, and the determinant-one subgroup of $\mathrm{Aut}(E_8)/\{\pm 1\}$ is isomorphic with an orthogonal group $\mathrm{SO}_8^+(\mathbb{Z}/2\mathbb{Z})$. But Co_1 is a "sporadic" simple group: one of the 26 finite simple groups not contained in infinite families of alternating groups or matrix groups of Lie type. A list of these groups may be found in R. Salomon's *Notices* article on the classification of finite simple groups (**42** (1995), pp. 231–239). The Leech lattice and its automorphism group play a central role in the story of the sporadic groups. R. L. Griess used L_{24} and Co_1 to construct the largest of these groups, the "Monster" (named after Griess and B. Fischer, who had independently predicted its existence), from whose subgroups all but six of the sporadic simple groups can be recovered. These include Co_1 itself and eleven further sporadic groups arising naturally from subgroups of Co_1. For instance, $Co_0 = \mathrm{Aut}(L_{24})$ acts transitively on the 196560 vectors of minimal length 2 and on the 16773120 vectors of next-smallest length $\sqrt{6}$; the point stabilizers in these transitive actions are Conway's sporadic simple groups Co_2 and Co_3. Mathieu's highly transitive permutation groups, which were the first five sporadic groups discovered, also occur in Co_0 via the constructions of L_{24} from Golay's error-correcting codes. We turn to error-correcting codes in Part II of this article.

[6] *The other notation for the Conway groups Co_i is $\cdot_i$, pronounced "dot-i"; for instance* $\mathrm{Aut}(L_{24})$ *is called "dotto". This notation was introduced by Conway, who naturally did not call these groups "Co_i" himself.*

References

[CS] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, Springer, New York, 1993.

[EOR] N. D. Elkies, A. M. Odlyzko, and J. A. Rush, On the packing densities of superballs and other bodies, *Invent. Math.* **105** (1991), 613–639.

[S] J.-P. Serre, *A Course in Arithmetic*, Springer, New York, 1973.

[TV] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes*, Kluwer, Dordrecht, 1991.