

A Brief Classical Introduction

A development in graph theory or topology from last July may already be viewed as “ancient history,” but in the theory of quadratic forms, and indeed throughout number theory, “ancient” really means ANCIENT. For example, Babylonian tablets dating back to 1900–1600 BC suggest that it was then known that there are infinitely many **primitive Pythagorean triples**: solutions $(a, b, c) \in \mathbb{Z}^3$ of the equation $x_1^2 + x_2^2 - x_3^2 = 0$, with $\gcd(a, b, c) = 1$; and it was known how to produce them. These results could be said to constitute the first theorem in quadratic forms; and historians suggest that this is the first instance on record in which mathematics was clearly being done for fun, not for commercial purposes or to determine property boundaries.

Historical Remark. Pythagoras actually lived during the period 580–500 BC, long after the above discovery to which we apply his name. See Boyer [Bo] for a thorough treatment of this early history.

We begin with the classical definition of quadratic forms. Later on we will see that it can be advantageous to view a quadratic form as a mapping on a module bearing an inner product.

1.1. Quadratic Forms as Polynomials

Definition 1.1. Let R be an integral domain of characteristic not 2, and suppose F is a field containing R . An n -ary **quadratic form over F** is a

polynomial of the form

$$q(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j \in F[x_1, \dots, x_n].$$

Because multiplication in F is commutative, in considering $q(x_1, \dots, x_n)$ as a function $R^n \rightarrow F$ without loss of generality we can (and will, from now on) suppose that the coefficients of q satisfy the condition $a_{ij} = a_{ji}$; just replace each of these coefficients by $\frac{a_{ij} + a_{ji}}{2}$. (Note that here the underlying assumption that $\text{char } F \neq 2$ is needed.) With this convention in force, q is now given by the symmetric matrix $A = (a_{ij}) \in M_n(F)$. This symmetric matrix is the **Gram matrix** of q .

Let $\alpha \in F$. We say the form q **represents** α **over** R , denoted $\alpha \xrightarrow{R} q$, if there are $\lambda_1, \dots, \lambda_n \in R$ such that $q(\lambda_1, \dots, \lambda_n) = \alpha$. More briefly: viewing q as a function $R^n \rightarrow F$, the statement $\alpha \xrightarrow{R} q$ means $\alpha \in \text{range } q$. (A more general notion of representation will come in Definition 2.8.)

An instance of the **representation problem** is this: Given a form q , for which $\alpha \in F$ is it true that $\alpha \xrightarrow{R} q$?

Example 1.2. Pythagorean triples, mentioned earlier, are precisely the solutions for the representation $0 \xrightarrow{\mathbb{Z}} x_1^2 + x_2^2 - x_3^2$.

Example 1.3. Brahmagupta (598–668) could generate Pythagorean triples, and he studied representations by other quadratic forms as well. For instance, from a given representation $1 \xrightarrow{\mathbb{Z}} x^2 - ay^2$ (with $a \in \mathbb{N}$) he could produce infinitely many.

Example 1.4. (i) Fermat (1601–1665) studied sums of integer squares, and here is one of his typical results, stated by him in 1640 in a private communication—along with hints on his method of proof—and with a formal proof published by Euler in 1754: If $n \in \mathbb{N}$ has prime factorization $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, then $n \xrightarrow{\mathbb{Z}} x^2 + y^2$ if and only if $\alpha_i \equiv 0 \pmod{2}$ for all i such that $p_i \equiv 3 \pmod{4}$.

(ii) In 1967 Leakey [**Le**] returned to Fermat's result, but replacing \mathbb{Z} with a polynomial ring. Let $h(x) \in \mathbb{F}_q[x]$. (Here \mathbb{F}_q denotes a finite field with q elements, with q odd.) Say $h(x) = \varepsilon p_1^{\alpha_1}(x) \cdots p_r^{\alpha_r}(x)$, with the $p_i(x)$ prime (monic and irreducible) polynomials and $\varepsilon \in \mathbb{F}_q^*$. Then $h(x)$ is a sum of two polynomial squares—that is, $h(x) \xrightarrow{\mathbb{F}_q[x]} X^2 + Y^2$ —if and only if $\alpha_i \equiv 0 \pmod{2}$ for all i such that $\partial p_i(x) \equiv 1 \pmod{2}$. (Here we use “ ∂ ” for the degree function.)

Example 1.5. Euler (1707–1783) considered representation of primes by quadratic forms of the form $x^2 \pm Ny^2$ over \mathbb{Z} . A typical result: If p is prime and satisfies $p \equiv 1 \pmod{20}$, then $p \xrightarrow{\mathbb{Z}} x^2 + 5y^2$.

Example 1.6. Lagrange (1736–1813) showed that every nonnegative integer n can be expressed as a sum of at most four squares; that is,

$$n \xrightarrow{\mathbb{Z}} x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

Lagrange also substantially developed the theory of binary (two-variable) quadratic forms over the integers.

Example 1.7. Let p and q be odd primes, with $p \leq q$, and consider representations of the product pq by the quadratic form $Q(x, y) = x^2 - y^2$. Over the field \mathbb{Q} of rational numbers there are infinitely many representations $pq \xrightarrow{\mathbb{Q}} Q$. In fact if β is any rational number, then for all $\lambda \in \mathbb{Q}^*$ we have

$$\beta = \left(\frac{\lambda^2 + \beta}{2\lambda} \right)^2 - \left(\frac{\lambda^2 - \beta}{2\lambda} \right)^2.$$

On the other hand, a straightforward arithmetic argument (Try it!) shows that there are just two nonnegative *integral* representations $pq \xrightarrow{\mathbb{Z}} Q$, given by the pairs

$$(x, y) = \left(\frac{p+q}{2}, \frac{q-p}{2} \right) \quad \text{and} \quad (x, y) = \left(\frac{pq+1}{2}, \frac{pq-1}{2} \right).$$

Now we will move in the opposite direction and *start* with an integer value for x , say $x = n \geq 2$, and ask this question: Must there exist an integer k satisfying $0 \leq k \leq n-2$ and primes p, q such that $n^2 - k^2 = pq$? If the answer is “yes,” then (if $p \leq q$, say) we must have $p = n-k$ and $q = n+k$ by the Fundamental Theorem of Arithmetic, and hence $2n = p+q$. In other words, an affirmative answer to this question on integral representations by the binary quadratic form $Q(x, y) = x^2 - y^2$ would affirmatively settle the celebrated **Goldbach Conjecture: Every even integer ≥ 4 is the sum of two primes.**

Conversely, if the Goldbach Conjecture is true and for each integer $n \geq 2$ there are primes $p \leq q$ such that $2n = p+q$, then $p = n-k$ and $q = n+k$ for some integer k satisfying $0 \leq k \leq n-2$ and hence $n^2 - k^2 = pq$. So the Goldbach Conjecture and our quadratic forms conjecture are actually equivalent. Note that a supplementary conjecture on $Q(x, y) = x^2 - y^2$ that there are infinitely many values of n for which $k = 1$ “works” (i.e., $n^2 - 1 = pq$ for some primes p and q) is equivalent to the **Twin Prime Conjecture: There are infinitely many primes p such that $p+2$ is also prime.**

1.2. Representation and Equivalence; Matrix Connections; Discriminants

The theory of quadratic forms underwent a dramatic change of style and perspective in the Twentieth Century; but before getting on with that (in the next chapter), in keeping with the tenor of this brief look at history we will first consider—without formal proof—an example involving binary quadratic forms in the style of Lagrange. Lagrange observed that if each variable in a given quadratic form is replaced by a linear combination of variables, then the result is again a quadratic form. It may turn out that representations by the transformed form are easier to determine than those of the original form; and then one can carry those representations back and produce representations by the original form. With this general outline, our goal here is to determine what integers n are represented over \mathbb{Z} by the form

$$f = 17x^2 + 94xy + 130y^2.$$

In particular, can we find an explicit representation $5 \xrightarrow{\mathbb{Z}} f$?

Consider the Gram matrix $A = \begin{pmatrix} 17 & 47 \\ 47 & 130 \end{pmatrix}$ for f . First observe that a sequence of elementary row and column operations for \mathbb{Z} -matrices (each time following a row operation with the corresponding column operation to get a symmetric resulting matrix) reduces A to a matrix in which the off-diagonal entry is smaller than the diagonal entries:

$$\begin{aligned} \begin{pmatrix} 17 & 47 \\ 47 & 130 \end{pmatrix} &\mapsto \begin{pmatrix} 17 & 47 \\ 13 & 36 \end{pmatrix} \mapsto \begin{pmatrix} 17 & 13 \\ 13 & 10 \end{pmatrix} \mapsto \begin{pmatrix} 4 & 3 \\ 13 & 10 \end{pmatrix} \\ &\mapsto \begin{pmatrix} 1 & 3 \\ 3 & 10 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Here $T = \begin{pmatrix} 3 & -11 \\ -1 & 4 \end{pmatrix}$ is the product of the elementary matrices—listed from left to right in their order of use—corresponding to the column operations, and we have observed that ${}^tTAT = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Now use the matrix T to dictate a change of variables: in our given form, make the substitutions

$$x \mapsto 3u - 11v \quad \text{and} \quad y \mapsto -u + 4v.$$

Then upon carrying out all the arithmetic we get

$$f(x, y) = g(u, v) = u^2 + v^2.$$

(In other words, we have gotten a new quadratic form with Gram matrix tTAT .) But we know from Fermat's result cited earlier exactly which integers are represented by g . And given explicit values for u and v that solve the equation $n = u^2 + v^2$, we can use the equations relating x and y to

u and v to get a representation $n \xrightarrow{\mathbb{Z}} f$. For instance, the representation $5 = 1^2 + 2^2$ gives

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & -11 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -19 \\ 7 \end{pmatrix}$$

as a solution to $5 = 17x^2 + 94xy + 130y^2$. Also note that since $T \in GL_2(\mathbb{Z})$, we can use the inverse of this transforming matrix to express u, v as integral linear combinations of x, y . This sets up a bijection between the representations

$$n \xrightarrow{\mathbb{Z}} f \quad \text{and} \quad n \xrightarrow{\mathbb{Z}} g.$$

Quadratic forms f and g related in this way via invertible matrices over a ring R are said to be **equivalent** over R . As the example suggests, equivalent forms represent the same elements, and they represent them the same number of times.

Now we will describe the above process in greater generality. Given a quadratic form $q = \sum_{i,j=1}^n a_{ij}x_ix_j$, with $A = (a_{ij}) \in M_n(F)$; if each x_i is replaced by a linear combination of x_1, \dots, x_n with coefficients in R , then the result is a new quadratic form h . More explicitly, given a matrix $T = (t_{ij}) \in M_n(R)$, replacing (in q) each x_i by $\sum_{j=1}^n t_{ij}x_j$ yields a new quadratic form $h = h(x_1, \dots, x_n)$ with associated Gram matrix tTAT . We say h is **represented by q over R** , denoted $h \xrightarrow{R} q$. It is useful to observe that q can be viewed as the result of a matrix product:

$$q(x_1, \dots, x_n) = (x_1, \dots, x_n)A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = {}^txAx, \text{ with } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Thus

$$h(x) = {}^t_x{}^tTATx = {}^t(Tx)A(Tx) = q(Tx).$$

It follows that if $\alpha \xrightarrow{R} h$ and $h \xrightarrow{R} q$ then also $\alpha \xrightarrow{R} q$.

Definition 1.8. Let q, h be n -ary quadratic forms over F , with respective Gram matrices A, C . We say q and h are **equivalent** over R and are in the same **class** over R if there is a matrix $T \in GL_n(R)$ such that $h \xrightarrow{R} q$ via T . That is, $h(x) = q(Tx)$ with T an **invertible** R -matrix. We denote equivalence of quadratic forms h and q over R by $h \cong_R q$. (If R is understood from the context, to avoid visual clutter we will usually denote equivalence by $h \cong q$ and representation by $h \rightarrow q$.)

Remark 1.9. A matrix $T \in GL_n(R)$ is called a **unimodular** R -matrix. It determines an automorphism of R^n via matrix multiplication. So from the equation $h(x) = q(Tx)$ we see that equivalent forms represent the same

elements of F over R ; moreover, for every $\alpha \in F$ represented by these forms, there is a bijection between the subsets $q^{-1}(\alpha)$ and $h^{-1}(\alpha)$ of R^n .

We now have two fundamental problems for quadratic forms over R : Determine a computationally effective set of necessary and sufficient conditions for two given forms to be equivalent or for one to represent the other. Incidentally, notice that the question of whether a given form represents a given element amounts to a special case of the problem of whether one form represents another: $\alpha \xrightarrow[R]{} q(x_1, \dots, x_n)$ if and only if $\alpha x_1^2 \xrightarrow[R]{} q$.

The equivalence problem for quadratic forms over R can be stated as a matrix question: Given symmetric matrices $A, C \in M_n(F)$, is there a matrix $T \in GL_n(R)$ such that $C = {}^tTAT$? If the answer is affirmative we write $A \cong_R C$ and say that A and C are **congruent** over R . (As before, if R is understood we usually write “ $A \cong C$ ”.) The representation problem for quadratic forms has a similar matrix formulation, except that T no longer is required to be unimodular or square.

Example 1.10. Let $q = x_1^2 + x_2^2$ and $h = 2x_1^2 + 2x_2^2$. Then $h(x) = q(Tx)$, where $T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Since $T \in M_2(\mathbb{Z}) \cap GL_2(\mathbb{Q})$, it follows that $h \xrightarrow[\mathbb{Z}]{} q$ and $h \cong_{\mathbb{Q}} q$. But it is NOT the case that $q \xrightarrow[\mathbb{Z}]{} h$, since q represents 1 over \mathbb{Z} while h does not, and hence $h \not\cong_{\mathbb{Z}} q$. In terms of matrix congruence,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cong_{\mathbb{Q}} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{via } T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

while

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \not\cong_{\mathbb{Z}} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Notice that if $q \cong_R h$, say $C = {}^tTAT$ with $T \in GL_n(R)$, then

$$|C| = |A| \cdot |T|^2 \in |A| \cdot U^2 \in F^*/U^2 \cup \{0\}.$$

Let us dissect this latter expression. Here $F^* = F - \{0\}$ is a multiplicative group, $U = R^*$ is the group of units of R , and U^2 is its subgroup of squares. So F^*/U^2 is a multiplicative group, and $F^*/U^2 \cup \{0\}$ then becomes a semi-group if we define $0 \cdot x = 0$ for all x . The **discriminant** of q is the element $dq = |A| \cdot U^2 \in F^*/U^2 \cup \{0\}$. It follows that if $q \cong_R h$ then $dq = dh$; that is, the discriminant is a **class invariant** of quadratic forms. (We will have more to say about discriminants—including some variations—later on. See 2.6 and 2.48.)

Example 1.11. If $R = \mathbb{Z}$ then $U = \{\pm 1\}$, hence $U^2 = \{1\}$ and so $|A|$ is a class invariant, often called the **determinant** of the form. In practice

(even when R is not \mathbb{Z}) it is common to abuse terminology and refer to the determinant of the Gram matrix as the discriminant of the form, with the understanding that two discriminants are to be regarded as equal if one is the square of an R -unit times the other.

Exercises

- (1) Describe the groups F^*/F^{*2} when F is each of the following:
 \mathbb{C} , $\mathbb{C}(x)$, \mathbb{Q} , \mathbb{F}_q , $\mathbb{F}_q(x)$. (Here \mathbb{F}_q is a finite field of odd order q .)
- (2) Consider the quadratic form $f(x, y) = 62x^2 + 176xy + 125y^2$. Is there a representation $4 \xrightarrow{\mathbb{Z}} f$? Is there a representation $11 \xrightarrow{\mathbb{Z}} f$? In each case, if there *is* such a representation, find them all.
- (3) Is the quadratic form $f(x, y) = 3x^2 + 2xy + 2y^2$ equivalent over \mathbb{Z} to a quadratic form $g(x, y) = ax^2 + by^2$ with $a, b \in \mathbb{Z}$?
- (4) Show that over a field F of characteristic not 2 every binary (i.e., two-variable) quadratic form $f(x, y)$ is equivalent to a form $\alpha x^2 + \beta y^2$ with $\alpha, \beta \in F$. If F is an *ordered* field (such as \mathbb{Q} or \mathbb{R}), deduce a necessary and sufficient condition on the discriminant df for f to represent both positive and negative elements of F .

1.3. A Brief Historical Sketch, and Some References to the Literature

Many of the greatest mathematicians have contributed to the theory of quadratic forms, and fortunately much of the early history has been well treated elsewhere, so we will not linger on it here. Particularly recommended are André Weil's book, *Number Theory: An Approach Through History. From Hammurapi to Legendre* [We]; and Winfried Scharlau and Hans Opolka's *From Fermat to Minkowski: Lectures on the Theory of Numbers and its Historical Development* [SO].

Much of the more recent work continues in the spirit of Gauss and the other masters of the past in that it seeks to understand issues of equivalence and representation of quadratic forms through an assortment of **reduction** theories, wherein one attempts to find a hopefully unique canonical form in the equivalence class of a given form. Details of the efforts in this direction—and the degree of success—depend heavily on the underlying coefficient ring of the form. For example, over an algebraically closed field we'll see that there is only one equivalence class of “nondegenerate” quadratic forms in a given number of variables. (A form is **nondegenerate** if it is not equivalent to a form with fewer variables.) So the classification and representation

problems are trivial in this case. But as the coefficient ring becomes more restricted the theory becomes more interesting and the mathematics becomes more challenging.

Sylvester, in the mid-Nineteenth Century gave us his “law of inertia,” which settled matters over \mathbb{R} . Hermite developed some ideas on reduction over \mathbb{Z} , hoping to extend the remarkable earlier work by Gauss and Legendre on reduction and classification of forms in two variables. In particular Hermite obtained a bound on the **minimum** of a positive definite quadratic form over \mathbb{Z} ; that is, the smallest positive integer represented by a quadratic form with integer coefficients that represents only positive integers when its variables are replaced by integers. But Hermite’s work just starts the more general reduction process. Toward the end of the Nineteenth Century Minkowski began what is now called the “local-global” approach, considering the classification of quadratic forms over \mathbb{Z} through the classification over $\mathbb{Z} \pmod{p^k}$, as p ranges through the primes and k grows. In the early Twentieth Century **valuation theory** was developed by Hensel, allowing for the exploration of problems given over the field \mathbb{Q} of rational numbers (and other fields of number-theoretic interest) through the discussion of the same problems over certain extension fields: the fields of **p -adic numbers**. Then Hasse (Hensel’s student) was able to use valuation theory to reformulate and extend Minkowski’s work to a wider number-theoretic context. Roughly speaking, the idea is as follows. A given equation over a given domain may be difficult to solve, and even the question of whether a solution *exists* may be difficult to determine. But if one extends the coefficient ring to a larger and more tractable domain—or to a family of such domains—the solvability “upstairs” may be easy (or at least easier) to determine. Obviously if no solution exists in an extension of the original ring R then there is no solution in R . On the other hand, if there *is* a solution in the extensions then perhaps one can use this information to produce a solution in R , or at least to show that such a solution exists. In order for the process to be successful, R needs to be densely embedded in the extensions, and there must be a kind of coherence or “reciprocity” among the extensions.

The local-global approach developed by Hasse had great success when the coefficient ring was \mathbb{Q} or an algebraic number field. But over \mathbb{Z} (and the other rings of algebraic integers) the success was much more limited, in the sense that without further special information one could not conclude that the existence of **local** solutions—i.e., solutions over the rings \mathbb{Z}_p of **p -adic integers**—forced a **global** solution: a solution over \mathbb{Z} .

In the 1930s C. L. Siegel introduced the methods of complex analysis into the study of quadratic forms in a series of ground-breaking papers. Siegel used his approach to explore the gap—between local integral solutions with

respect to every prime p and global solutions—through the notion of the **genus** of a quadratic form (where he expanded on ideas first developed by Gauss [Ga] in the binary case) and the concept of the **class number**: the number of equivalence classes in a genus.

Also important in the 1930s was a pioneering paper by Witt, who developed a more geometric approach to quadratic forms over fields, for instance transforming the process of equivalence of quadratic forms (via linear changes of variables) into the study of isometries of appropriate associated inner-product spaces of a more general sort than had been considered earlier. Witt's work not only provided a more workable and intuitively fruitful framework for the old questions, but led to what is now called the **algebraic theory of quadratic forms**, which studies the subject over arbitrary fields, not necessarily of a number-theoretic origin. Many of the algebraic issues raised by Witt lay dormant for many years, though a paper by Kaplansky in 1953 brought fresh attention to the subject, and then the remarkable work by Pfister in the 1960s led to an explosion of the algebraic theory. While we will briefly touch on some of that theory in this book, we refer the reader to T. Y. Lam's book, *Introduction to Quadratic Forms over Fields* [Lam], for a full treatment.

Interest in Witt's more geometric approach was heightened by the work of Emil Artin (as represented in his book *Geometric Algebra* [A]) in the 1940s and 1950s, particularly in connection with the developing theory of the **classical groups**. And so by the mid-'50s it was becoming more standard to view quadratic forms not as polynomials to be treated as abstract symbols, but as inner products on modules: **quadratic spaces** if the modules were vector spaces; and **lattices** for appropriate submodules of quadratic spaces. So at this point the old problem of determining whether two quadratic forms were equivalent had become the problem of determining whether two lattices on quadratic spaces were isometric.

The pursuit of the local-global approach to the theory over \mathbb{Z} demanded the solution of the classification problem over \mathbb{Z}_p , and this was the work of Jones and Pall in the 1940s. Their work was extended to the classification of lattices (on quadratic spaces) over the rings of integers of arbitrary **local fields** in the 1950s by O'Meara, who also made major advances on the local **representation** problem: the determination of when one lattice contains an isometric copy of another. In the early 1960s further progress was made by Riehm [Ri] on the representation problem, and recently the local representation problem has apparently been solved in its full generality by Beli [Be].

From the 1950s through the 1970s, Kneser and O'Meara were leaders in exploring the linkage of local and global invariants of lattices, and the impact of that linkage on the study of the classification problem. Particularly important is Kneser's **neighbor lattice** construction [K2], which allowed for a step-by-step transformation of a lattice in one isometry class (in the so-called **unimodular** case) into a lattice in another class. O'Meara's book, *Introduction to Quadratic Forms* [O'M1], which appeared in 1963, is a wonderful introduction to the subject as of that time.

Since the late 1960s there have been important new links forged between the theory of quadratic forms and other areas of mathematics, and corresponding new demands on the subject. In 1968, Conway's discovery [Con1] of a new finite simple group through the study of the **orthogonal group** (the group of inner-product preserving automorphisms) of the **Leech lattice** brought the attention of the group theory community. In topology, work by Milnor, Freedman, and many others reduced certain problems on the classification of manifolds to problems on the structure of associated lattices over \mathbb{Z} . For example, if M is a closed, oriented, simply-connected 4-manifold, then the homology group $H_2(M)$ is a \mathbb{Z} -lattice, and the "intersection pairing" gives $H_2(M)$ a unimodular structure. It turns out that there is an orientation-preserving homotopy equivalence between two such 4-manifolds if and only if the corresponding \mathbb{Z} -lattices are isometric. See *Symmetric Bilinear Forms*, by J. Milnor and D. Husemoller [MH], for more details.

The extraordinary development and proliferation of computing power since the 1970s has of course made possible a broad range of algorithmic approaches to the subject that before were out of the question. And procedures developed earlier (such as Kneser's neighbor-lattice process) that were practical only in very low dimensions were now able to be extended to higher dimensions. In the 1980s came the so-called **LLL-** or **L^3 -algorithm**, due to A. Lenstra, H. Lenstra, and L. Lovasz [LLL], giving a sort of reduced form for a lattice that, while not necessarily canonical, allows for much useful information to be deduced. For instance, L^3 can often be used to find the shortest nonzero vector in a lattice. Finding the length of such a vector is the lattice equivalent of the problem mentioned earlier (pursued by Hermite) of finding the minimum of a quadratic form, and while the inequality on the minimum guaranteed by L^3 is weaker than Hermite's in the worst case, in practice the L^3 inequality is usually stronger than Hermite's; and L^3 actually produces short vectors, whereas Hermite's result does not.

The development of the Internet and other sophisticated electronic means of communication led to the subject of algebraic coding theory and new forms of cryptography. Electronically transmitted messages are usually

encoded as n -tuples of 0's and 1's; that is, as elements of \mathbb{Z}_2^n . Because no medium of communication is perfect, a 0 digit may be received as a 1, or vice versa. Coding theory deals with the practice of introducing redundancy into the transmitted message so that when occasional errors occur they can be recognized and corrected. The greater the likelihood of errors, the more substantial the redundancy needs to be in order to handle the errors. On the other hand, redundancy slows the communication process, so it is best to avoid more redundancy than is absolutely necessary to do the job. Algebraic coding theory deals with the associated problems. If a given code C is a **linear code**, meaning a subspace of \mathbb{Z}_2^n , then one can view C as a subset of \mathbb{Z}^n and define an associated lattice L as the set of all vectors in \mathbb{Z}^n congruent (mod 2) in each coordinate to an element of C . In this way and other ways, problems in coding theory become recast as problems on lattices in \mathbb{R}^n and related problems on **sphere packing**. We refer the reader to the massive *Sphere Packings, Lattices and Groups*, by Conway and Sloane [CS], and Ebeling's *Lattices and Codes* [Eb] for thorough introductions to lattices in coding theory. We note that the Conway–Sloane book has one of the most extensive bibliographies in the mathematical literature.

Cryptography is an ancient subject, but the classical encryption systems—many going back to antiquity—were so-called **private key** systems, in the sense that sender and receiver would agree in advance on an encryption method and key (the “key” being the essential information needed in order to understand the details of how the messages were being encrypted), and anyone knowing the key would be able to decrypt the message. In the 1970s the first of the **public key** cryptosystems (namely RSA encryption) was developed. In such a system, a party wishing to receive a message publicly announces the method of encryption to be used—a method involving some kind of number-theoretic manipulation of the message—but someone intercepting the encrypted message will not be able to decrypt it within a reasonable time period without special additional information known only to the receiver. For example, RSA encryption involves use of a number m that is a product of two large primes; successful decryption requires knowledge of m 's factorization, and factorization is a computationally complex problem. Not long after RSA came **knapsack cryptosystems**, which allowed for much faster encryption than RSA, but it wasn't long before these began to be successfully attacked by a variety of methods, including approaches through lattice reduction based on the L^3 algorithm. In the 1990s an assortment of encryption schemes was developed in which the fundamental difficulty of lattice problems was exploited in the encryption method. Here, though the setting is new, some of the issues go back to the same things that concerned Hermite: the difficulty of finding short vectors in a lattice: the **shortest vector problem (SVP)**; or the more general **closest vector**

problem (CVP): given a lattice L in \mathbb{R}^n and a vector $v \in \mathbb{R}^n$, find a vector $x \in L$ closest to V . A lattice is usually specified by giving a basis for it, and it turns out that questions of the sort just described can be much easier to solve in one basis than another, and the computational complexity of these problems is the basis for some of these cryptosystems. We will have a brief look at lattices in cryptography in the final chapter. Good book sources for more complete treatments of much of this material: *Complexity of Lattice Problems: A Cryptographic Perspective*, by Micciancio and Goldwasser [MG], and *Cryptography and Lattices*, ed. by Silverman [Si].