We now come to the crux of the matter. Since the x_i are all 0 or 1, all of the $2x_i - 1$ values are ± 1, so the vector $\boldsymbol{t}$ is quite short, $\|\boldsymbol{t}\| = \sqrt{n}$. On the other hand, we have seen that $m_i = \mathcal{O}(2^{2n})$ and $S = \mathcal{O}(2^{2n})$, so the vectors generating L all have lengths $\|\boldsymbol{v}_i\| = \mathcal{O}(2^{2n})$. Thus it is unlikely that L contains any nonzero vectors, other than $\boldsymbol{t}$, whose length is as small as $\sqrt{n}$. If we postulate that Eve knows an algorithm that can find small nonzero vectors in lattices, then she will be able to find $\boldsymbol{t}$, and hence to recover the plaintext $\boldsymbol{x}$.

Algorithms that find short vectors in lattices are called *lattice reduction algorithms*. The most famous of these is the LLL algorithm, to which we alluded earlier, and its variants such as LLL-BKZ. The remainder of this chapter is devoted to describing lattices, cryptosystems based on lattices, the LLL algorithm, and cryptographic applications of LLL. A more detailed analysis of knapsack cryptosystems is given in Sect. 7.14.2; see also Example 7.33.

7.3 A Brief Review of Vector Spaces

Before starting our discussion of lattices, we pause to remind the reader of some important definitions and ideas from linear algebra. Vector spaces can be defined in vast generality,[3] but for our purposes in this chapter, it is enough to consider vector spaces that are contained in $\mathbb{R}^m$ for some positive integer m.

We start with the basic definitions that are essential for studying vector spaces.

Vector Spaces. A *vector space* V is a subset of $\mathbb{R}^m$ with the property that

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 \in V \quad \text{for all } \boldsymbol{v}_1, \boldsymbol{v}_2 \in V \text{ and all } \alpha_1, \alpha_2 \in \mathbb{R}.$$

Equivalently, a vector space is a subset of $\mathbb{R}^m$ that is closed under addition and under scalar multiplication by elements of $\mathbb{R}$.

Linear Combinations. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in V$. A *linear combination* of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in V$ is any vector of the form

$$\boldsymbol{w} = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k \quad \text{with } \alpha_1, \ldots, \alpha_k \in \mathbb{R}.$$

The collection of all such linear combinations,

$$\{\alpha_1\boldsymbol{v}_1 + \cdots + \alpha_k\boldsymbol{v}_k : \alpha_1, \ldots, \alpha_k \in \mathbb{R}\},$$

is called the *span* of $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.

Independence. A set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in V$ is (*linearly*) *independent* if the only way to get

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k = \boldsymbol{0} \tag{7.5}$$

is to have $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. The set is (*linearly*) *dependent* if we can make (7.5) true with at least one α_i nonzero.

[3] For example, we saw in Sect. 3.6 a nice application of vector spaces over the field $\mathbb{F}_2$.

Bases. A *basis* for V is a set of linearly independent vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ that span V. This is equivalent to saying that every vector $\boldsymbol{w} \in V$ can be written in the form

$$\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n$$

for a *unique* choice of $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$.

We next describe the relationship between different bases and the important concept of dimension.

Proposition 7.11. *Let $V \subset \mathbb{R}^m$ be a vector space.*

(a) *There exists a basis for V.*

(b) *Any two bases for V have the same number of elements. The number of elements in a basis for V is called the* dimension of V.

(c) *Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be a basis for V and let $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ be another set of n vectors in V. Write each $\boldsymbol{w}_j$ as a linear combination of the $\boldsymbol{v}_i$,*

$$\begin{aligned} \boldsymbol{w}_1 &= \alpha_{11}\boldsymbol{v}_1 + \alpha_{12}\boldsymbol{v}_2 + \cdots + \alpha_{1n}\boldsymbol{v}_n, \\ \boldsymbol{w}_2 &= \alpha_{21}\boldsymbol{v}_1 + \alpha_{22}\boldsymbol{v}_2 + \cdots + \alpha_{2n}\boldsymbol{v}_n, \\ \vdots & \qquad\qquad \vdots \\ \boldsymbol{w}_n &= \alpha_{n1}\boldsymbol{v}_1 + \alpha_{n2}\boldsymbol{v}_2 + \cdots + \alpha_{nn}\boldsymbol{v}_n. \end{aligned}$$

Then $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ is also a basis for V if and only if the determinant of the matrix

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

is not equal to 0.

We next explain how to measure lengths of vectors in $\mathbb{R}^n$ and the angles between pairs of vectors. These important concepts are tied up with the notion of dot product and the Euclidean norm.

Definition. Let $\boldsymbol{v}, \boldsymbol{w} \in V \subset \mathbb{R}^m$ and write $\boldsymbol{v}$ and $\boldsymbol{w}$ using coordinates as

$$\boldsymbol{v} = (x_1, x_2, \ldots, x_m) \quad \text{and} \quad \boldsymbol{w} = (y_1, y_2, \ldots, y_m).$$

The *dot product of $\boldsymbol{v}$ and $\boldsymbol{w}$* is the quantity

$$\boldsymbol{v} \cdot \boldsymbol{w} = x_1 y_1 + x_2 y_2 + \cdots + x_m y_m.$$

We say that $\boldsymbol{v}$ and $\boldsymbol{w}$ are *orthogonal* to one another if $\boldsymbol{v} \cdot \boldsymbol{w} = 0$.

The *length*, or *Euclidean norm*, of $\boldsymbol{v}$ is the quantity

$$\|\boldsymbol{v}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}.$$

Notice that dot products and norms are related by the formula

$$\boldsymbol{v} \cdot \boldsymbol{v} = \|\boldsymbol{v}\|^2.$$

Proposition 7.12. *Let $\boldsymbol{v}, \boldsymbol{w} \in V \subset \mathbb{R}^m$.*

(a) *Let θ be the angle between the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, where we place the starting points of $\boldsymbol{v}$ and $\boldsymbol{w}$ at the origin $\mathbf{0}$. Then*

$$\boldsymbol{v} \cdot \boldsymbol{w} = \|\boldsymbol{v}\| \, \|\boldsymbol{w}\| \cos(\theta), \tag{7.6}$$

(b) (**Cauchy–Schwarz inequality**)

$$|\boldsymbol{v} \cdot \boldsymbol{w}| \le \|\boldsymbol{v}\| \, \|\boldsymbol{w}\|. \tag{7.7}$$

Proof. For (a), see any standard linear algebra textbook. We observe that the Cauchy–Schwarz inequality (b) follows immediately from (a), but we feel that it is of sufficient importance to warrant a direct proof. If $\boldsymbol{w} = \mathbf{0}$, there is nothing to prove, so we may assume that $\boldsymbol{w} \ne \mathbf{0}$. We consider the function

$$\begin{aligned} f(t) = \|\boldsymbol{v} - t\boldsymbol{w}\|^2 &= (\boldsymbol{v} - t\boldsymbol{w}) \cdot (\boldsymbol{v} - t\boldsymbol{w}) \\ &= \boldsymbol{v} \cdot \boldsymbol{v} - 2t\boldsymbol{v} \cdot \boldsymbol{w} + t^2 \boldsymbol{w} \cdot \boldsymbol{w} \\ &= \|\boldsymbol{v}\|^2 - 2t\boldsymbol{v} \cdot \boldsymbol{w} + t^2 \|\boldsymbol{w}\|^2. \end{aligned}$$

We know that $f(t) \ge 0$ for all $t \in \mathbb{R}$, so we choose the value of t that minimizes $f(t)$ and see what it gives. This minimizing value is $t = \boldsymbol{v} \cdot \boldsymbol{w} / \|\boldsymbol{w}\|^2$. Hence

$$0 \le f\left(\frac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{w}\|^2}\right) = \|\boldsymbol{v}\|^2 - \frac{(\boldsymbol{v} \cdot \boldsymbol{w})^2}{\|\boldsymbol{w}\|^2}.$$

Simplifying this expression and taking square roots gives the desired result. □

Definition. An *orthogonal basis* for a vector space V is a basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ with the property that

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \quad \text{for all } i \ne j.$$

The basis is *orthonormal* if in addition, $\|\boldsymbol{v}_i\| = 1$ for all i.

There are many formulas that become much simpler using an orthogonal or orthonormal basis. In particular, if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthogonal basis and if $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n$ is a linear combination of the basis vectors, then

$$\begin{aligned} \|\boldsymbol{v}\|^2 &= \|a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n\|^2 \\ &= (a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) \cdot (a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j (\boldsymbol{v}_i \cdot \boldsymbol{v}_j) \\
&= \sum_{i=1}^{n} a_i^2 \|\boldsymbol{v}_i\|^2 \quad \text{since } \boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \text{ for } i \neq j.
\end{aligned}$$

If the basis is orthonormal, then this further simplifies to $\|\boldsymbol{v}\|^2 = \sum a_i^2$.

There is a standard method, called the Gram–Schmidt algorithm, for creating an orthonormal basis. We describe a variant of the usual algorithm that gives an orthogonal basis, since it is this version that is most relevant for our later applications.

Theorem 7.13 (Gram–Schmidt Algorithm). *Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be a basis for a vector space $V \subset \mathbb{R}^m$. The following algorithm creates an orthogonal basis $\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*$ for V:*

```
Set $\boldsymbol{v}_1^* = \boldsymbol{v}_1$.
Loop $i = 2, 3, \ldots, n$.
    Compute  $\mu_{ij} = \boldsymbol{v}_i \cdot \boldsymbol{v}_j^* / \|\boldsymbol{v}_j^*\|^2$  for $1 \leq j < i$.
    Set  $\boldsymbol{v}_i^* = \boldsymbol{v}_i - \sum_{j=1}^{i-1} \mu_{ij} \boldsymbol{v}_j^*$.
End Loop
```

The two bases have the property that

$$\mathrm{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i\} = \mathrm{Span}\{\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_i^*\} \quad \textit{for all } i = 1, 2, \ldots, n.$$

Proof. The proof of orthogonality is by induction, so we suppose that the vectors $\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_{i-1}^*$ are pairwise orthogonal and we need to prove that $\boldsymbol{v}_i^*$ is orthogonal to all of the previous starred vectors. To do this, we take any $k < i$ and compute

$$\begin{aligned}
\boldsymbol{v}_i^* \cdot \boldsymbol{v}_k^* &= \left(\boldsymbol{v}_i - \sum_{j=1}^{i-1} \mu_{ij} \boldsymbol{v}_j^* \right) \cdot \boldsymbol{v}_k^* \\
&= \boldsymbol{v}_i \cdot \boldsymbol{v}_k^* - \mu_{ik} \|\boldsymbol{v}_k^*\|^2 \quad \text{since } \boldsymbol{v}_k^* \cdot \boldsymbol{v}_j^* = 0 \text{ for } j \neq k, \\
&= 0 \quad \text{from the definition of } \mu_{ik}.
\end{aligned}$$

To prove the final statement about the spans, we note first that it is clear from the definition of $\boldsymbol{v}_i^*$ that $\boldsymbol{v}_i$ is in the span of $\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_i^*$. We prove the other inclusion by induction, so we suppose that $\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_{i-1}^*$ are in the span of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}$ and we need to prove that $\boldsymbol{v}_i^*$ is in the span of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i$. But from the definition of $\boldsymbol{v}_i^*$, we see that it is in the span of $\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_{i-1}^*, \boldsymbol{v}_i$, so we are done by the induction hypothesis. □