5.3 QR-Decomposition Using Householder Matrices

First, we state the result geometrically. When translated in terms of Householder matrices, we obtain the fact advertised earlier that every matrix (not necessarily invertible) has a QR-decomposition.

Lemma 5.3.1 *Let E be a nontrivial Euclidean space of dimension n. Given any orthonormal basis $(e_1, \ldots, e_n)$, for any n-tuple of vectors $(v_1, \ldots, v_n)$, there is a sequence of n isometries $h_1, \ldots, h_n$, such that h_i is a hyperplane reflection or the identity, and if $(r_1, \ldots, r_n)$ are the vectors given by*

$$r_j = h_n \circ \cdots \circ h_2 \circ h_1(v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \ldots, e_j)$, $(1 \leq j \leq n)$. Equivalently, the matrix R whose columns are the components of the r_j over the basis $(e_1, \ldots, e_n)$ is an upper triangular matrix. Furthermore, the h_i can be chosen so that the diagonal entries of R are nonnegative.

Remarks. (1) Since every h_i is a hyperplane reflection or the identity,

$$\rho = h_n \circ \cdots \circ h_2 \circ h_1$$

is an isometry.

(2) If we allow negative diagonal entries in R, the last isometry h_n may be omitted.

(3) Instead of picking $r_{k,k} = \|u''_k\|$, which means that

$$w_k = r_{k,k}\, e_k - u''_k,$$

where $1 \leq k \leq n$, it might be preferable to pick $r_{k,k} = -\|u''_k\|$ if this makes $\|w_k\|^2$ larger, in which case

$$w_k = r_{k,k}\, e_k + u''_k.$$

Indeed, since the definition of h_k involves division by $\|w_k\|^2$, it is desirable to avoid division by very small numbers.

Lemma 5.3.1 immediately yields the QR-decomposition in terms of Householder transformations.

Lemma 5.3.2 *For every real $n \times n$-matrix A, there is a sequence $H_1, \ldots, H_n$ of matrices, where each H_i is either a Householder matrix or the identity, and an upper triangular matrix R, such that*

$$R = H_n \cdots H_2 H_1 A.$$

As a corollary, there is a pair of matrices Q, R, where Q is orthogonal and R is upper triangular, such that $A = QR$ (a QR-decomposition of A). Furthermore, R can be chosen so that its diagonal entries are non-negative.

Remarks. (1) Letting

$$A_{k+1} = H_k \cdots H_2 H_1 A,$$

with $A_1 = A$, $1 \leq k \leq n$, the proof of lemma 5.3.1 can be interpreted in terms of the computation of the sequence of matrices $A_1, \ldots, A_{n+1} = R$.

The matrix A_{k+1} has the shape

$$A_{k+1} = \begin{pmatrix} \times & \times & \times & u_1^{k+1} & \times & \times & \times & \times \\ 0 & \times & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \times & u_k^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_{k+1}^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_{k+2}^{k+1} & \times & \times & \times & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & u_{n-1}^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_n^{k+1} & \times & \times & \times & \times \end{pmatrix}$$

where the $(k+1)$th column of the matrix is the vector

$$u_{k+1} = h_k \circ \cdots \circ h_2 \circ h_1(v_{k+1}),$$

and thus

$$u'_{k+1} = (u_1^{k+1}, \ldots, u_k^{k+1}),$$

and

$$u''_{k+1} = (u_{k+1}^{k+1}, u_{k+2}^{k+1}, \ldots, u_n^{k+1}).$$

If the last $n-k-1$ entries in column $k+1$ are all zero, there is nothing to do and we let $H_{k+1} = I$.

Otherwise, we kill these $n-k-1$ entries by multiplying A_{k+1} on the left by the Householder matrix H_{k+1} sending $(0,\ldots,0,u^{k+1}_{k+1},\ldots,u^{k+1}_{n})$ to $(0,\ldots,0,r_{k+1,k+1},0,\ldots,0)$, where

$$r_{k+1,k+1} = \left\|(u^{k+1}_{k+1},\ldots,u^{k+1}_{n})\right\|.$$

(2) If we allow negative diagonal entries in R, the matrix H_n may be omitted ($H_n = I$).

(3) If A is invertible and the diagonal entries of R are positive, it can be shown that Q and R are unique.

(4) The method allows the computation of the determinant of A. We have

$$\det(A) = (-1)^m r_{1,1} \cdots r_{n,n},$$

where m is the number of Householder matrices (not the identity) among the H_i.

(5) The "condition number" of the matrix A is preserved (see Strang [**?**]). This is very good for numerical stability.

We conclude our discussion of isometries with a brief discussion of affine isometries.

5.4 Affine Isometries (Rigid Motions)

Definition 5.4.1 Given any two nontrivial Euclidean affine spaces E and F of the same finite dimension n, a function $f\colon E \to F$ is *an affine isometry (or rigid map)* iff it is an affine map and

$$\|\mathbf{f(a)f(b)}\| = \|\mathbf{ab}\|,$$

for all $a, b \in E$. When $E = F$, an affine isometry $f\colon E \to E$ is also called a *rigid motion*.

Thus, an affine isometry is an affine map that preserves the distance. This is a rather strong requirement.

In fact, we will show that for any function $f\colon E \to F$, the assumption that

$$\|\mathbf{f(a)f(b)}\| = \|\mathbf{ab}\|$$

for all $a, b \in E$, forces f to be an affine map.

Remark: Sometimes, an affine isometry is defined as a *bijective* affine isometry. When E and F are of finite dimension, the definitions are equivalent.

Lemma 5.4.2 *Given any two nontrivial Euclidean affine spaces E and F of the same finite dimension n, an affine map $f: E \to F$ is an affine isometry iff its associated linear map $\overrightarrow{f}: \overrightarrow{E} \to \overrightarrow{F}$ is an isometry. An affine isometry is a bijection.*

Let us now consider affine isometries $f: E \to E$. If $\overrightarrow{f}$ is a rotation, we call f a *proper (or direct) affine isometry*, and if $\overrightarrow{f}$ is a an improper linear isometry, we call f a *an improper (or skew) affine isometry*.

It is easily shown that the set of affine isometries $f: E \to E$ forms a group denoted as $\mathbf{Is}(E)$ (or $\mathbf{Mo}(E)$), and those for which $\overrightarrow{f}$ is a rotation is a subgroup denoted as $\mathbf{SE}(E)$.

The translations are the affine isometries f for which $\overrightarrow{f} = \mathrm{id}$, the identity map on $\overrightarrow{E}$.

The following lemma is the counterpart of lemma 4.3.2 for isometries between Euclidean vector spaces:

Lemma 5.4.3 *Given any two nontrivial Euclidean affine spaces E and F of the same finite dimension n, for every function $f\colon E \to F$, the following properties are equivalent:*

(1) f is an affine map and $\|\mathbf{f}(\mathbf{a})\mathbf{f}(\mathbf{b})\| = \|\mathbf{ab}\|$, for all $a, b \in E$.

(2) $\|\mathbf{f}(\mathbf{a})\mathbf{f}(\mathbf{b})\| = \|\mathbf{ab}\|$, for all $a, b \in E$.

In order to understand the structure of affine isometries, it is important to investigate the fixed points of an affine map.

5.5 Fixed Points of Affine Maps

Recall that $E(1, \overrightarrow{f})$ denotes the eigenspace of the linear map $\overrightarrow{f}$ associated with the scalar 1, that is, the subspace consisting of all vectors $u \in \overrightarrow{E}$ such that $\overrightarrow{f}(u) = u$.

Clearly, $\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = E(1, \overrightarrow{f})$.

Given some origin $\Omega \in E$, since

$$f(a) = f(\Omega + \mathbf{\Omega a}) = f(\Omega) + \overrightarrow{f}(\mathbf{\Omega a}),$$

we get

$$\mathbf{\Omega f(a)} - \mathbf{\Omega a} = \mathbf{\Omega f(\Omega)} + \overrightarrow{f}(\mathbf{\Omega a}) - \mathbf{\Omega a}.$$

Using this, we show the following lemma which holds for arbitrary affine spaces of finite dimension and for arbitrary affine maps.

Lemma 5.5.1 *Let E be any affine space of finite dimension. For every affine map $f\colon E \to E$, let $Fix(f) = \{a \in E \mid f(a) = a\}$ be the set of fixed points of f. The following properties hold.*

(1) If f has some fixed point a, so that $Fix(f) \neq \emptyset$, then $Fix(f)$ is an affine subspace of E such that

$$Fix(f) = a + E(1, \overrightarrow{f}) = a + \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}),$$

where $E(1, \overrightarrow{f})$ is the eigenspace of the linear map $\overrightarrow{f}$ for the eigenvalue 1.

(2) The affine map f has a unique fixed point iff

$$E(1, \overrightarrow{f}) = \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = \{0\}.$$

Remark: The fact that E has finite dimension is only used to prove (2), and (1) holds in general.

If an isometry f leaves some point fixed, we can take such a point Ω as the origin, and then $f(\Omega) = \Omega$ and we can view f as a rotation or an improper orthogonal transformation, depending on the nature of $\overrightarrow{f}$.

Note that it is quite possible that $Fix(f) = \emptyset$. For example, nontrivial translations have no fixed points.

A more interesting example is provided the composition of a plane reflection about a line composed with a a nontrivial translation parallel to this line.

Otherwise, we will see in lemma 5.6.2 that every affine isometry is the (commutative) composition of a translation with an isometry that always has a fixed point.

5.6 Affine Isometries and Fixed Points

Given any two affine subspaces F, G of E such that $\overrightarrow{F}$ and $\overrightarrow{G}$ are orthogonal subspaces of $\overrightarrow{E}$ such that $\overrightarrow{E} = \overrightarrow{F} \oplus \overrightarrow{G}$, for any point $\Omega \in F$, we define $q\colon E \to \overrightarrow{G}$, such that

$$q(a) = p_{\overrightarrow{G}}(\mathbf{\Omega a}).$$

Note that $q(a)$ is independent of the choice of $\Omega \in F$.

Then, the map $g\colon E \to E$ such that $g(a) = a - 2q(a)$, or equivalently

$$\mathbf{ag}(\mathbf{a}) = -2q(a) = -2p_{\overrightarrow{G}}(\mathbf{\Omega a})$$

does not depend on the choice of $\Omega \in F$.

If we identify E to $\overrightarrow{E}$ by choosing any origin Ω in F, we note that g is identified with the symmetry with respect to $\overrightarrow{F}$ and parallel to $\overrightarrow{G}$.

Thus, the map g is an affine isometry, and it is called the *orthogonal symmetry about F*.

Since

$$g(a) = \Omega + \mathbf{\Omega a} - 2p_{\overrightarrow{G}}(\mathbf{\Omega a})$$

for all $\Omega \in F$ and for all $a \in E$, we note that the linear map $\overrightarrow{g}$ associated with g is the (linear) symmetry about the subspace $\overrightarrow{F}$ (the direction of F)

The following amusing lemma shows the extra power afforded by affine orthogonal symmetries: Translations are subsumed!

Lemma 5.6.1 *Given any affine space E, if $f\colon E \to E$ and $g\colon E \to E$ are orthogonal symmetries about parallel affine subspaces F_1 and F_2, then $g \circ f$ is a translation defined by the vector $2\mathbf{ab}$, where $\mathbf{ab}$ is any vector perpendicular to the common direction $\overrightarrow{F}$ of F_1 and F_2 such that $\|\mathbf{ab}\|$ is the distance between F_1 and F_2, with $a \in F_1$ and $b \in F_2$. Conversely, every translation by a vector τ is obtained as the composition of two orthogonal symmetries about parallel affine subspaces F_1 and F_2 whose common direction is orthogonal to $\tau = \mathbf{ab}$, for some $a \in F_1$ and some $b \in F_2$ such that the distance betwen F_1 and F_2 is $\|\mathbf{ab}\|/2$.*

The following result is a generalization of Chasles' theorem about the rigid motions in $\mathbb{R}^3$.

Lemma 5.6.2 *Let E be a Euclidean affine space of finite dimension n. For every affine isometry $f\colon E \to E$, there is a unique isometry $g\colon E \to E$ and a unique translation $t = t_\tau$, with $\overrightarrow{f}(\tau) = \tau$ (i.e., $\tau \in \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})$), such that the set*

$$Fix(g) = \{a \in E \mid g(a) = a\}$$

of fixed points of g is a nonempty affine subspace of E of direction

$$\overrightarrow{G} = \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = E(1, \overrightarrow{f}),$$

and such that

$$f = t \circ g \quad \textit{and} \quad t \circ g = g \circ t.$$

Furthermore, we have the following additional properties:

(a) $f = g$ *and* $\tau = 0$ *iff* f *has some fixed point, i.e., iff* $Fix(f) \neq \emptyset$.

(b) If f *has no fixed points, i.e.,* $Fix(f) = \emptyset$, *then* $\dim(\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})) \geq 1$.

The proof rests on the following two key facts:

(1) If we can find some $x \in E$ such that $\mathbf{xf}(\mathbf{x}) = \tau$ belongs to $\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})$, we get the existence of g and τ.

(2) $\overrightarrow{E} = \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) \oplus \mathrm{Im}\,(\overrightarrow{f} - \mathrm{id})$, and $\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})$ and $\mathrm{Im}\,(\overrightarrow{f} - \mathrm{id})$ are orthogonal. This implies the uniqueness of g and τ.

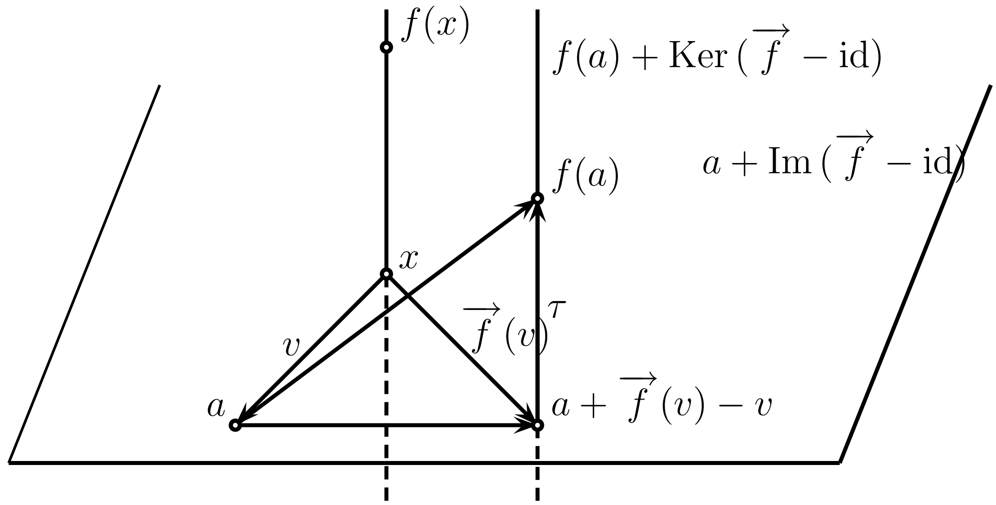

Figure 5.5: Rigid motion as $f = t \circ g$, where g has some fixed point x

Remarks. (1) Note that $\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = \{0\}$ iff $\tau = 0$, iff $Fix(g)$ consists of a single element, which is the unique fixed point of f.

However, even if f is not a translation, f may not have any fixed points.

(2) The fact that E has finite dimension is only used to prove (b).

(3) It is easily checked that $Fix(g)$ consists of the set of points x such that $\|\mathbf{xf(x)}\|$ is minimal.

In the affine Euclidean plane, it is easy to see that the affine isometries are classified as follows.

An isometry f which has a fixed point is a rotation if it is a direct isometry, else a reflection about a line.

If f has no fixed point, then either it is a nontrivial translation or the composition of a reflection about a line with a nontrivial translation parallel to this line.

In an affine space of dimension 3, it is easy to see that the affine isometries are classified as follows.

A proper isometry with a fixed point is a rotation around a line D (its set of fixed points), as illustrated in figure 5.6.

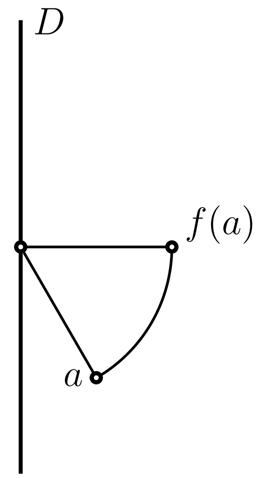

Figure 5.6: $3D$ proper rigid motion with line D of fixed points (rotation)

An improper isometry with a fixed point is either a reflection about a plane H (the set of fixed points), or the composition of a rotation followed by a reflection about a plane H orthogonal to the axis of rotation D, as illustrated in figures 5.7 and 5.8. In the second case, there is a single fixed point $O = D \cap H$.

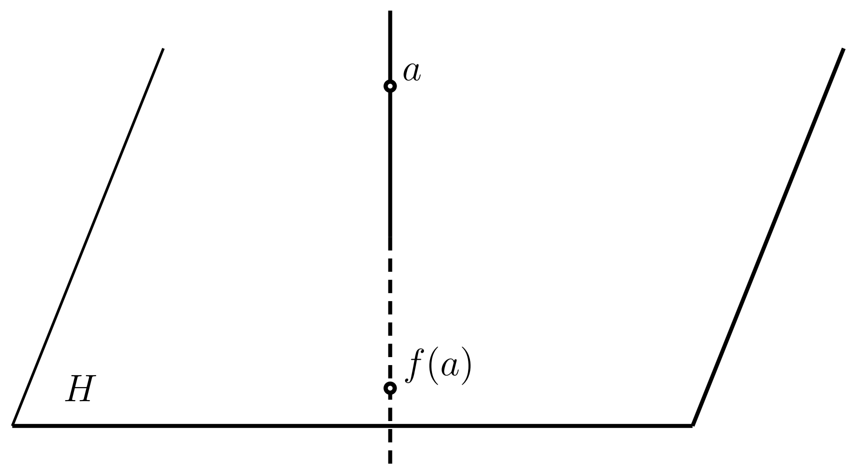

Figure 5.7: $3D$ improper rigid motion with a plane H of fixed points (reflection)

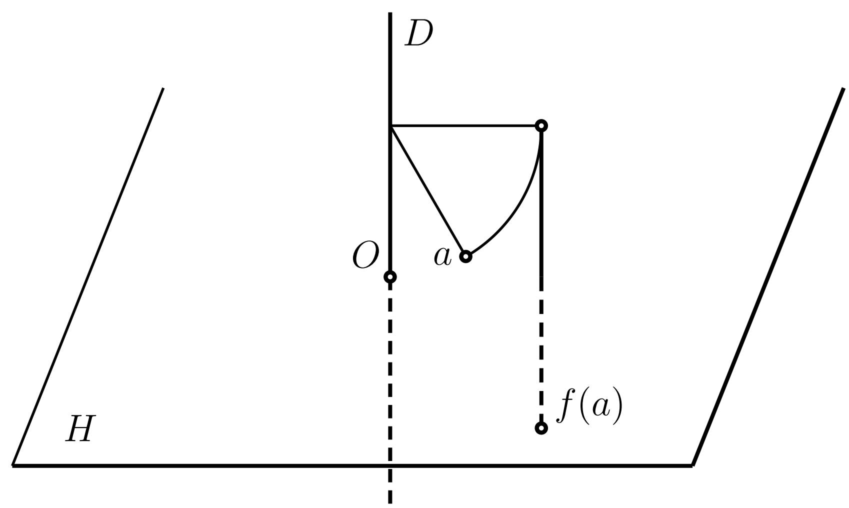

Figure 5.8: $3D$ improper rigid motion with a unique fixed point

There are three types of isometries with no fixed point. The first kind is a nontrivial translation. The second kind is the composition of a rotation followed by a nontrivial translation parallel to the axis of rotation D. Such a rigid motion is proper, and is called a *screw motion*. A screw motion is illustrated in figure 5.9.

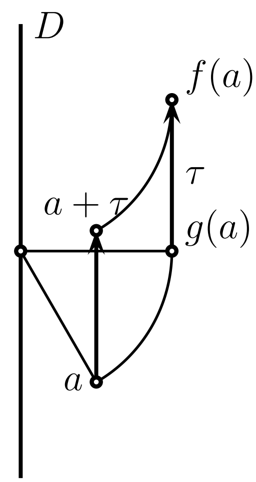

Figure 5.9: $3D$ proper rigid motion with no fixed point (screw motion)

The third kind is the composition of a reflection about a plane followed by a nontrivial translation by a vector parallel to the direction of the plane of the reflection, as illustrated in figure 5.10. It is an improper isometry.

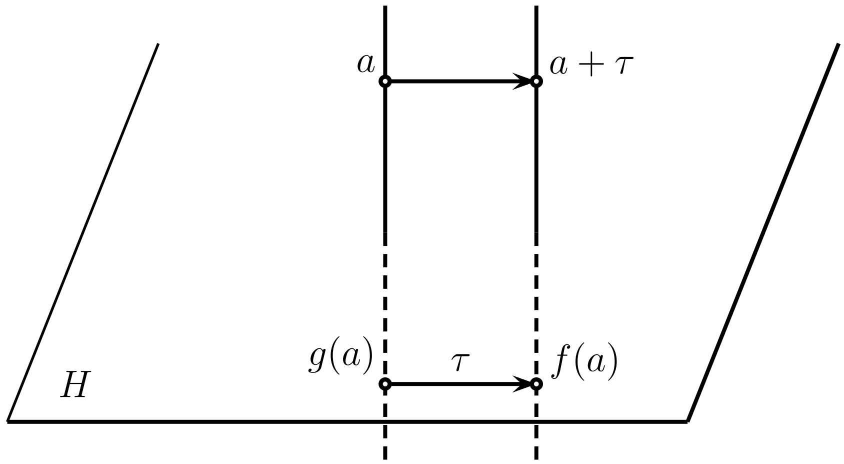

Figure 5.10: $3D$ improper rigid motion with no fixed points

The Cartan-Dieudonné also holds for affine isometries, with a small twist due to translations.

5.7 The Cartan–Dieudonné Theorem for Affine Isometries

Theorem 5.7.1 *Let E be an affine Euclidean space of dimension $n \geq 1$. Every isometry $f \in \mathbf{Is}(E)$ which has a fixed point and is not the identity is the composition of at most n reflections. Every isometry $f \in \mathbf{Is}(E)$ which has no fixed point is the composition of at most $n+2$ reflections. For $n \geq 2$, the identity is the composition of any reflection with itself.*

When $n \geq 3$, we can also characterize the affine isometries in $\mathbf{SE}(n)$ in terms of flips.

Remarkably, not only we can do without translations, but we can even bound the number of flips by n.

Theorem 5.7.2 *Let E be a Euclidean affine space of dimension $n \geq 3$. Every rigid motion $f \in \mathbf{SE}(E)$ is the composition of an even number of flips $f = f_{2k} \circ \cdots \circ f_1$, where $2k \leq n$.*

Remark. It is easy to prove that if f is a screw motion in $\mathbf{SE}(3)$, if D is its axis, θ is its angle of rotation, and τ is the translation along the direction of D, then f is the composition of two flips about lines D_1 and D_2 orthogonal to D, at a distance $\|\tau\|/2$, and making an angle $\theta/2$.

There is one more topic that we would like to cover since it is often useful in practice, the concept of *cross-product of vectors*, also called vector-product. But first, we need to discuss the question of orientation of bases.

5.8 Orientations of a Euclidean Space, Angles

In order to deal with the notion of orientation correctly, it is important to assume that every family $(u_1, \ldots, u_n)$ of vectors is ordered (by the natural ordering on $\{1, 2, \ldots, n\}$).

We will assume that all families $(u_1, \ldots, u_n)$ of vectors, in particular, bases and orthonormal bases are ordered.

Let E be a vector space of finite dimension n over $\mathbb{R}$, and let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be any two bases for E.

Recall that the change of basis matrix from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$ is the matrix P whose columns are the coordinates of the vectors v_j over the basis $(u_1, \ldots, u_n)$.

It is immediately verified that the set of alternating n-linear forms on E is a vector space that we denote as $\Lambda(E)$ (see Lang [**?**]).

It is easy to show that $\Lambda(E)$ has dimension 1.

We now define an equivalence relation on $\Lambda(E) - \{0\}$ (where we let 0 denote the null alternating n-linear form):

φ and ψ are equivalent iff $\psi = \lambda\varphi$ for some $\lambda > 0$.

It is immediately verified that the above relation is an equivalence relation. Furthermore, it has exactly two equivalence classes O_1 and O_2.

The first way of defining an *orientation of E* is to pick one of these two equivalence classes, say O ($O \in \{O_1, O_2\}$).

Given such a choice of a class O, we say that a basis $(w_1, \ldots, w_n)$ has *positive orientation* iff

$$\varphi(w_1, \ldots, w_n) > 0$$

for any alternating n-linear form $\varphi \in O$.

Note that this makes sense, since for any other $\psi \in O$, $\varphi = \lambda\psi$ for some $\lambda > 0$.

According to the previous definition, two bases $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ have the same orientation iff $\varphi(u_1, \ldots, u_n)$ and $\varphi(v_1, \ldots, v_n)$ have the same sign for all $\varphi \in \Lambda(E) - \{0\}$.

From

$$\varphi(v_1, \ldots, v_n) = \det(P)\varphi(u_1, \ldots, u_n),$$

we must have $\det(P) > 0$.

Conversely, if $\det(P) > 0$, the same argument shows that $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ have the same orientation.

This leads us to an equivalent and slightly less contorted definition of the notion of orientation. We define a relation between bases of E as follows:

Two bases $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are related iff $\det(P) > 0$, where P is the change of basis matrix from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$.

Since $\det(PQ) = \det(P)\det(Q)$, and since change of basis matrices are invertible, the relation just defined is indeed an equivalence relation, and it has two equivalence classes.

Furthermore, from the discussion above, any nonnull alternating n-linear form φ will have the same sign on any two equivalent bases.

The above discussion motivates the following definition.

Definition 5.8.1 Given any vector space E of finite dimension over $\mathbb{R}$, we define an *orientation of* E as the choice of one of the two equivalence classes of the equivalence relation on the set of bases defined such that $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ have the same orientation iff $\det(P) > 0$, where P is the change of basis matrix from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$. A basis in the chosen class is said to have *positive orientation, or to be positive*. An *orientation of a Euclidean affine space* E is an orientation of its underlying vector space $\overrightarrow{E}$

In practice, to give an orientation, one simply picks a fixed basis considered as having positive orientation. The orientation of every other basis is determined by the sign of the determinant of the change of basis matrix.

Having the notation of orientation at hand, we wish to go back briefly to the concept of (oriented) angle.

Let E be a Euclidean space of dimension $n = 2$, and assume a given orientation. In any given positive orthonormal basis for E, every rotation r is represented by a matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Actually, we claim that the matrix R representing the rotation r is the same in *all* orthonormal positive bases.

This is because the change of basis matrix from one positive orthonormal basis to another positive orthonormal basis is a rotation represented by some matrix of the form

$$P = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}$$

and that we have

$$P^{-1} = \begin{pmatrix} \cos(-\psi) & -\sin(-\psi) \\ \sin(-\psi) & \cos(-\psi) \end{pmatrix}$$

and after calculations, we find that PRP^{-1} is the rotation matrix associated with $\psi + \theta - \psi = \theta$.

We can choose $\theta \in [0, 2\pi[$, and we call θ the *measure of the angle of rotation of* r *(and* R*)*. If the orientation is changed, the measure changes to $2\pi - \theta$.

We now let E be a Euclidean space of dimension $n = 2$, but we do not assume any orientation.

It is easy to see that given any two unit vectors $u_1, u_2 \in E$ (unit means that $\|u_1\| = \|u_2\| = 1$), there is a unique rotation r such that

$$r(u_1) = u_2.$$

It is also possible to define an equivalence relation of pairs of unit vectors, such that

$$\langle u_1, u_2 \rangle \equiv \langle u_3, u_4 \rangle$$

iff there is some rotation r such that $r(u_1) = u_3$ and $r(u_2) = u_4$.

Then, the equivalence class of $\langle u_1, u_2 \rangle$ can be taken as the definition of the (oriented) *angle of* $\langle u_1, u_2 \rangle$, which is denoted as $\widehat{u_1 u_2}$.

Furthermore, it can be shown that there is a rotation mapping the pair $\langle u_1, u_2 \rangle$ to the pair $\langle u_3, u_4 \rangle$, iff there is a rotation mapping the pair $\langle u_1, u_3 \rangle$ to the pair $\langle u_2, u_4 \rangle$ (all vectors being unit vectors).

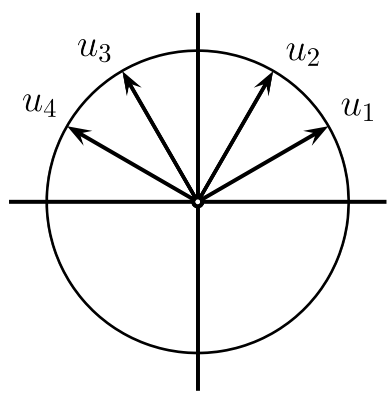

Figure 5.11: Defining Angles

As a consequence of all this, since for any pair $\langle u_1, u_2\rangle$ of unit vectors, there is a unique rotation r mapping u_1 to u_2, the angle $\widehat{u_1 u_2}$ of $\langle u_1, u_2\rangle$ corresponds bijectively to the rotation r, and there is a bijection between the set of angles of pairs of unit vectors and the set of rotations in the plane.

As a matter of fact, the set of angles forms an abelian groups isomorphic to the (abelian) group of rotations in the plane.

Thus, even though we can consider angles as oriented, note that the notion of orientation is not necessary to define angles.

However, to define the *measure of an angle*, the notion of orientation is needed.

If we now assume that an orientation of E (still a Euclidean plane) is given, the unique rotation r associated with an angle $\widehat{u_1u_2}$ corresponds to a unique matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The number θ is defined up to $2k\pi$ (with $k \in \mathbb{Z}$) and is called a *measure of the angle* $\widehat{u_1u_2}$.

There is a unique $\theta \in [0, 2\pi[$ which is a measure of the angle $\widehat{u_1u_2}$.

It is also immediately seen that

$$\cos\theta = u_1 \cdot u_2.$$

In fact, since $\cos\theta = \cos(2\pi - \theta) = \cos(-\theta)$, the quantity $\cos\theta$ does not depend on the orientation.

Now still considering a Euclidean plane, given any pair $\langle u_1, u_2 \rangle$ of nonnull vectors, we define their angle as the angle of the unit vectors $\frac{u_1}{\|u_1\|}$ and $\frac{u_2}{\|u_2\|}$, and if E is oriented, we define the measure θ of this angle as the measure of the angle of these unit vectors.

Note that

$$\cos\theta = \frac{u_1 \cdot u_2}{\|u_1\| \, \|u_2\|},$$

and this independently of the orientation.

Finally if E is a Euclidean space of dimension $n \geq 2$, we define the angle of a pair $\langle u_1, u_2 \rangle$ of nonnull vectors as the angle of this pair in the Euclidean plane spanned by $\langle u_1, u_2 \rangle$ if they are linearly independent, or any Euclidean plane containing u_1 if their are collinear.

If E is an affine Euclidean space of dimension $n \geq 2$, for any two pairs $\langle a_1, b_1 \rangle$ and $\langle a_2, b_2 \rangle$ of points in E, where $a_1 \neq b_1$ and $a_2 \neq b_2$, we define the angle of the pair $\langle \langle a_1, b_1 \rangle, \langle a_2, b_2 \rangle \rangle$ as the angle of the pair $\langle \mathbf{a_1 b_1}, \mathbf{a_2 b_2} \rangle$.

As for the issue of measure of an angle when $n \geq 3$, all we can do is to define the measure of the angle $\widehat{u_1 u_2}$ as either θ or $2\pi - \theta$, where $\theta \in [0, 2\pi[$.

In particular, when $n = 3$, one should note that it is not enough to give a line D through the origin (the axis of rotation) and an angle θ to specify a rotation!

The problem is that depending on the orientation of the plane H (through the origin) orthogonal to D, we get two different rotations: one of angle θ, the other of angle $2\pi - \theta$.

Thus, to specify a rotation, we also need to give an orientation of the plane orthogonal to the axis of rotation.

This can be done by specifying an orientation of the axis of rotation by some unit vector ω, and chosing the basis (e_1, e_2, ω) (where (e_1, e_2) is a basis of H) such that it has positive orientation w.r.t. the chosen orientation of E.

We now return to alternating multilinear forms on a Euclidean space.

When E is a Euclidean space, we have an interesting situation regarding the value of determinants over orthornormal bases described by the following lemma.

Given any basis $B = (u_1, \ldots, u_n)$ for E, for any sequence $(w_1, \ldots, w_n)$ of n vectors, we denote as $\det_B(w_1, \ldots, w_n)$ the determinant of the matrix whose columns are the coordinates of the w_j over the basis $B = (u_1, \ldots, u_n)$.

Lemma 5.8.2 *Let E be a Euclidean space of finite dimension n, and assume that an orientation of E has been chosen. For any sequence $(w_1, \ldots, w_n)$ of n vectors, for any two orthonormal bases $B_1 = (u_1, \ldots, u_n)$ and $B_2 = (v_1, \ldots, v_n)$ of positive orientation, we have*

$$\det_{B_1}(w_1, \ldots, w_n) = \det_{B_2}(w_1, \ldots, w_n).$$

By lemma 5.8.2, the determinant $\det_B(w_1, \ldots, w_n)$ is independent of the base B, provided that B is orthonormal and of positive orientation.

Thus, lemma 5.8.2 suggests the following definition.

5.9 Volume Forms, Cross-Products

Definition 5.9.1 Given any Euclidean space E of finite dimension n over $\mathbb{R}$ and any orientation of E, for any sequence $(w_1, \ldots, w_n)$ of n vectors in E, the common value $\lambda_E(w_1, \ldots, w_n)$ of the determinant $\det_B(w_1, \ldots, w_n)$ over all positive orthonormal bases B of E is called the *mixed product (or volume form) of* $(w_1, \ldots, w_n)$.

The mixed product $\lambda_E(w_1, \ldots, w_n)$ will also be denoted as $(w_1, \ldots, w_n)$, even though the notation is overloaded.

- The mixed product $\lambda_E(w_1, \ldots, w_n)$ changes sign when the orientation changes.
- The mixed product $\lambda_E(w_1, \ldots, w_n)$ is a scalar, and definition 5.9.1 really defines an alternating multilinear form from E^n to $\mathbb{R}$.
- $\lambda_E(w_1, \ldots, w_n) = 0$ iff $(w_1, \ldots, w_n)$ is linearly dependent.

- A basis $(u_1, \ldots, u_n)$ is positive or negative iff $\lambda_E(u_1, \ldots, u_n)$ is positive or negative.
- $\lambda_E(w_1, \ldots, w_n)$ is invariant under every isometry f such that $\det(f) = 1$.

The terminology "volume form" is justified by the fact that $\lambda_E(w_1, \ldots, w_n)$ is indeed the volume of some geometric object.

Indeed, viewing E as an affine space, the *parallelotope defined by* $(w_1, \ldots, w_n)$ is the set of points

$$\{\lambda_1 w_1 + \cdots + \lambda_n w_n \mid 0 \leq \lambda_i \leq 1,\ 1 \leq i \leq n\}.$$

Then, it can be shown (see Berger [**?**], Section 9.12) that the volume of the parallelotope defined by $(w_1, \ldots, w_n)$ is indeed $\lambda_E(w_1, \ldots, w_n)$.

If $(E, \overrightarrow{E})$ is a Euclidean affine space of dimension n, given any $n+1$ affinely independent points $(a_0, \ldots, a_n)$, the set

$$\{a_0 + \lambda_1 \mathbf{a_0 a_1} + \cdots + \lambda_n \mathbf{a_0 a_n} \mid 0 \le \lambda_i \le 1,\ 1 \le i \le n\},$$

is called the *parallelotope spanned by* $(a_0, \ldots, a_n)$.

Then, the volume of the parallelotope spanned by $(a_0, \ldots, a_n)$ is $\lambda_{\overrightarrow{E}}(\mathbf{a_0 a_1}, \ldots, \mathbf{a_0 a_n})$.

It can also be shown that the volume $vol(a_0, \ldots, a_n)$ of the n-simplex $(a_0, \ldots, a_n)$ is

$$vol(a_0, \ldots, a_n) = \frac{1}{n!} \lambda_{\overrightarrow{E}}(\mathbf{a_0 a_1}, \ldots, \mathbf{a_0 a_n}).$$

Now, given a sequence $(w_1, \ldots, w_{n-1})$ of $n-1$ vectors in E, the map

$$x \mapsto \lambda_E(w_1, \ldots, w_{n-1}, x)$$

is a linear form.

Thus, by lemma 4.2.4, there is a unique vector $u \in E$ such that

$$\lambda_E(w_1, \ldots, w_{n-1}, x) = u \cdot x$$

for all $x \in E$.

The vector u has some interesting properties which motivate the next definition.

Definition 5.9.2 Given any Euclidean space E of finite dimension n over $\mathbb{R}$, for any orientation of E, for any sequence $(w_1, \ldots, w_{n-1})$ of $n-1$ vectors in E, the unique vector $w_1 \times \cdots \times w_{n-1}$ such that

$$\lambda_E(w_1, \ldots, w_{n-1}, x) = w_1 \times \cdots \times w_{n-1} \cdot x$$

for all $x \in E$, is called the *cross-product, or vector product, of* $(w_1, \ldots, w_{n-1})$.

The following properties hold.

- The cross-product $w_1 \times \cdots \times w_{n-1}$ changes sign when the orientation changes.
- The cross-product $w_1 \times \cdots \times w_{n-1}$ is a vector, and definition 5.9.2 really defines an alternating multilinear map from E^{n-1} to E.

- $w_1 \times \cdots \times w_{n-1} = 0$ iff $(w_1, \ldots, w_{n-1})$ is linearly dependent. This is because,

  $$w_1 \times \cdots \times w_{n-1} = 0$$

  iff

  $$\lambda_E(w_1, \ldots, w_{n-1}, x) = 0$$

  for all $x \in E$, and thus, if $(w_1, \ldots, w_{n-1})$ was linearly independent, we could find a vector $x \in E$ to complete $(w_1, \ldots, w_{n-1})$ into a basis of E, and we would have

  $$\lambda_E(w_1, \ldots, w_{n-1}, x) \neq 0.$$

- The cross-product $w_1 \times \cdots \times w_{n-1}$ is orthogonal to each of the w_j.

- If $(w_1, \ldots, w_{n-1})$ is linearly independent, then the sequence

  $$(w_1, \ldots, w_{n-1}, w_1 \times \cdots \times w_{n-1})$$

  is a positive basis of E.

We now show how to compute the coordinates of $u_1 \times \cdots \times u_{n-1}$ over an orthonormal basis.

Given an orthonormal basis $(e_1, \ldots, e_n)$, for any sequence $(u_1, \ldots, u_{n-1})$ of $n-1$ vectors in E, if

$$u_j = \sum_{i=1}^{n} u_{i,j} e_i,$$

where $1 \leq j \leq n-1$, for any $x = x_1 e_1 + \cdots + x_n e_n$, consider the determinant

$$\lambda_E(u_1, \ldots, u_{n-1}, x) = \begin{vmatrix} u_{1\,1} & \ldots & u_{1\,n-1} & x_1 \\ u_{2\,1} & \ldots & u_{2\,n-1} & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ u_{n\,1} & \ldots & u_{n\,n-1} & x_n \end{vmatrix}.$$

Calling the underlying matrix above as A, we can expand $\det(A)$ according to the last column, using the Laplace formula (see Strang [**?**]), where $A_{i\,j}$ is the $(n-1)\times(n-1)$-matrix obtained from A by deleting row i and column j, and we get:

$$\begin{vmatrix} u_{1\,1} & \ldots & u_{1\,n-1} & x_1 \\ u_{2\,1} & \ldots & u_{2\,n-1} & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ u_{n\,1} & \ldots & u_{n\,n-1} & x_n \end{vmatrix} =$$

$$(-1)^{n+1}x_1 \det(A_{1\,n}) + \cdots + (-1)^{n+n}x_n \det(A_{n\,n}).$$

Each $(-1)^{i+n}\det(A_{i\,n})$ is called the *cofactor of* x_i.

We note that $\det(A)$ is in fact the inner product

$$\det(A) = \\ ((-1)^{n+1}\det(A_{1\,n})e_1 + \cdots + (-1)^{n+n}\det(A_{n\,n})e_n) \cdot x.$$

Since the cross-product $u_1 \times \cdots \times u_{n-1}$ is the unique vector u such that

$$u \cdot x = \lambda_E(u_1, \ldots, u_{n-1}, x),$$

for all $x \in E$, the coordinates of the cross-product $u_1 \times \cdots \times u_{n-1}$ must be

$$((-1)^{n+1}\det(A_{1\,n}), \ldots, (-1)^{n+n}\det(A_{n\,n})),$$

the sequence of cofactors of the x_i in the determinant $\det(A)$.

For example, when $n = 3$, the coordinates of the cross-product $u \times v$ are given by the cofactors of x_1, x_2, x_3, in the determinant

$$\begin{vmatrix} u_1 & v_1 & x_1 \\ u_2 & v_2 & x_2 \\ u_3 & v_3 & x_3 \end{vmatrix}$$

or more explicitly, by

$$(-1)^{3+1}\begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix}, (-1)^{3+2}\begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix}, (-1)^{3+3}\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix},$$

that is,

$$(u_2v_3 - u_3v_2,\ u_3v_1 - u_1v_3,\ u_1v_2 - u_2v_1).$$

It is also useful to observe that if we let U be the matrix

$$U = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$$

then the coordinates of the cross-product $u \times v$ are given by

$$\begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}$$

We finish our discussion of cross-products by mentioning without proof a few more of their properties, in the case $n = 3$.

Firstly, the following so-called *Lagrange identity* holds:

$$(u \cdot v)^2 + \|u \times v\|^2 = \|u\|^2 \, \|v\|^2 .$$

If u and v are linearly independent, and if θ (or $2\pi - \theta$) is a measure of the angle $\widehat{uv}$, then

$$|\sin \theta| = \frac{\|u \times v\|}{\|u\| \, \|v\|}.$$