Chapter 1

Mappings of Points and Vectors

1.1 Introduction

In this chapter, we explain how vector expressions can be used to describe the effects of various types of transformations that move and deform objects. These types of transformations are usually described by means of matrices, which are inherently coordinate-dependent, of course. We will see in this chapter that coordinate-free expressions can be developed; the usual matrix-based approach is deferred until **??**.

1.2 Mappings

If X and Y are two sets, a **mapping** from X to Y is simply a rule that gives us an element of Y corresponding to each given element of X. Mappings are also known as maps, functions, operators, or transformations. A simple example is the function that returns the square (x^2) of any given number x. The output of a function f corresponding to an input x is denoted by $f(x)$. Sometimes it is convenient to refer to a function anonymously, without giving it a name. So, for example, we might simply say "the function $x \mapsto x^2$", rather than saying "the function f defined by $f(x) = x^2$".

1.3 Linear and Affine Mappings

Let $T : \mathcal{V}_3 \to \mathcal{V}_3$ be a mapping from $\mathcal{V}_3$ to itself. If

$$T(\lambda\mathbf{U} + \mu\mathbf{V}) = \lambda T(\mathbf{U}) + \mu T(\mathbf{V}) \tag{1.1}$$

for all vectors $\mathbf{U}$ and $\mathbf{V}$ in $\mathcal{V}_3$ and all numbers λ and μ, then we say that T is a **linear** mapping, or a linear transformation. We will see later, in chapter **??**, that linear mappings can be defined between any two vector spaces, but, for now, let's just focus on $\mathcal{V}_3$.

If equation (1.1) holds only when $\lambda + \mu = 1$, then T is said to be an **affine** mapping. So, we see that every linear mapping is affine, but not every affine mapping is linear.

We will also be concerned with mappings $T : \mathcal{P}_3 \to \mathcal{P}_3$ that map points to points. Again, we say that a mapping $T : \mathcal{P}_3 \to \mathcal{P}_3$ is affine if equation (1.1) holds whenever $\lambda + \mu = 1$. It does not make sense to discuss whether a mapping from $\mathcal{P}_3$ to $\mathcal{P}_3$ is linear, because the expression $\lambda\mathbf{U} + \mu\mathbf{V}$ is not well defined unless $\lambda + \mu = 1$.

1.4 Translation

Let $\mathbf{D}$ be a fixed vector, and define $T : \mathcal{V}_3 \to \mathcal{V}_3$ by

$$T(\mathbf{V}) = \mathbf{V} + \mathbf{D}$$

If $\lambda + \mu = 1$, then, for any two vectors $\mathbf{U}$ and $\mathbf{V}$, we have

$$T(\lambda\mathbf{U} + \mu\mathbf{V}) = \lambda\mathbf{U} + \mu\mathbf{V} + \mathbf{D} = \lambda(\mathbf{U} + \mathbf{D}) + \mu(\mathbf{V} + \mathbf{D}) = \lambda T(\mathbf{U}) + \mu T(\mathbf{V})$$

and so T is affine. We can define a related mapping $\tilde{T} : \mathcal{P}_3 \to \mathcal{P}_3$ by $\tilde{T}(\mathbf{P}) = \mathbf{P} + \mathbf{D}$. And the calculations we did above show that $\tilde{T}$ is again affine. Obviously $\tilde{T}$ just translates (displaces) its input point $\mathbf{P}$ using the vector $\mathbf{D}$.

1.5 Scaling

Let α be any real number, and define a mapping $T : \mathcal{V}_3 \to \mathcal{V}_3$ by $T(\mathbf{V}) = \alpha\mathbf{V}$. In other words, the mapping T simply scales its input vector, changing its length by a factor of α. If $\mathbf{U}$ and $\mathbf{V}$ are any two vectors, and λ and μ are any two numbers, then we have

$$T(\lambda\mathbf{U} + \mu\mathbf{V}) = \alpha(\lambda\mathbf{U} + \mu\mathbf{V}) = \lambda(\alpha\mathbf{U}) + \mu(\alpha\mathbf{V}) = \lambda T(\mathbf{U}) + \mu T(\mathbf{V})$$

so T is a linear mapping.

If $\mathbf{Q}$ is some fixed point, then we can define a related mapping $\tilde{T} : \mathcal{P}_3 \to \mathcal{P}_3$ by $\tilde{T}(\mathbf{P}) = \alpha(\mathbf{P} - \mathbf{Q})$. This is a scaling transformation with center at $\mathbf{Q}$. Routine computations show that $\tilde{T}$ is affine.

1.6 Projection onto a Vector

Let $\mathbf{N}$ be a unit vector, and let L be a line in the direction of $\mathbf{N}$. We saw in section (**??**) that the projection of any given vector $\mathbf{V}$ onto the line L is given by the mapping

$$T(\mathbf{V}) = (\mathbf{V}\cdot\mathbf{N})\mathbf{N}$$

It is easy to verify that T is a linear mapping.

If we choose an origin $\mathbf{O}$, then we can define a related mapping $\tilde{T}$ whose effect on any given point $\mathbf{P}$ is

$$\tilde{T}(\mathbf{P}) = T(\mathbf{P} - \mathbf{O}) = [(\mathbf{P} - \mathbf{O})\cdot\mathbf{N}]\mathbf{N}$$

Again, showing that $\tilde{T}$ is affine is straightforward.

1.7 Projection onto a Plane

Let π be the plane with unit normal $\mathbf{N}$ passing through the point $\mathbf{Q}$. The projection of a given vector $\mathbf{V}$ onto $\mathbf{N}$ is $(\mathbf{V}\cdot\mathbf{N})\mathbf{N}$, so its projection onto π is given by

$$T(\mathbf{V}) = \mathbf{V} - (\mathbf{V}\cdot\mathbf{N})\mathbf{N}$$

Similarly, the projection of a given point $\mathbf{P}$ onto π is

$$\tilde{T}(\mathbf{P}) = \mathbf{P} - [(\mathbf{P} - \mathbf{Q})\cdot\mathbf{N}]\mathbf{N}$$

1.8 Reflection in a Plane

Again, let π be the plane with unit normal $\mathbf{N}$ passing through the point $\mathbf{Q}$. The reflection (mirror image) of a given vector $\mathbf{V}$ in the plane π is given by

$$T(\mathbf{V}) = \mathbf{V} - 2(\mathbf{V}\cdot\mathbf{N})\mathbf{N}$$

Similarly, the reflection of a given point $\mathbf{P}$ in π is

$$\tilde{T}(\mathbf{P}) = \mathbf{P} - 2[(\mathbf{P} - \mathbf{Q})\cdot\mathbf{N}]\mathbf{N}$$

1.9 Reflection in a Line

1.10 Reflection in a Point

1.11 Scaling About a Line

Let α be a real number (positive?), let $\mathbf{N}$ be a unit vector, and let L be a line through the origin in the direction of $\mathbf{N}$. Consider the mapping that performs axial scaling around L, in other words, we scale the distance of a point from the line by a factor α. It is easy to show that the transform is

$$T(\mathbf{P}) = \alpha\mathbf{P} + (1-\alpha)[(\mathbf{P} - \mathbf{O})\cdot\mathbf{N}]\mathbf{N}$$

1.12 Scaling About a Plane

1.13 Rotation Around a Vector

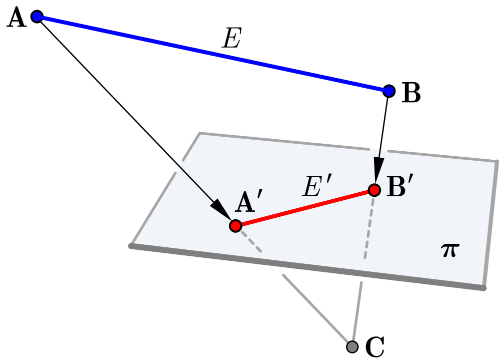
Let $\mathbf{N}$ be a unit vector. We consider rotations about a line though the origin in the direction of $\mathbf{N}$. The rotation of a given vector $\mathbf{V}$ though an angle θ is given by

$$T(\mathbf{V}) = (\cos\theta)\mathbf{V} + (1-\cos\theta)(\mathbf{N}\cdot\mathbf{V})\mathbf{N} - (\sin\theta)(\mathbf{N}\times\mathbf{V})$$

The term $1-\cos\theta$ is subject to subtractive cancellation error when θ is near zero, so it's better to write it in the form $\frac{1}{2}\sin^2\frac{1}{2}\theta$.

1.14 Central Projections

Central projections are used in perspective viewing and in the definition of rational curves and surfaces. The projection is based on a plane, π, called the **projection plane**, and a point $\mathbf{C}$, called the **center of projection**. Objects are projected onto the plane π along lines passing through the point $\mathbf{C}$. So, in the pictures, the point $\mathbf{A}$ is projected to the point $\mathbf{A}'$, the point $\mathbf{B}$ is projected to the point $\mathbf{B}'$, and the line E is projected to the line E', and so on. This isn't really how our eyes work, so it's not an accurate representation of perspective viewing, but it's close enough for most situations.



The line E can be either above the projection plane or below it. This doesn't matter – the projection idea still applies. Actually, it would still work even if the line E crossed over the projection plane.

In perspective viewing, we can think of the eye or camera being positioned at the point $\mathbf{C}$. In this situation, we sometimes call the point $\mathbf{C}$ the "eye point".

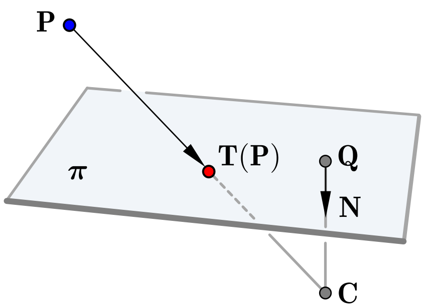
The central projection mapping has some interesting properties:

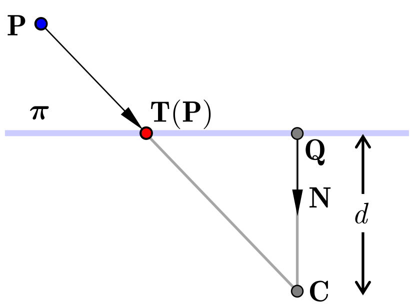
- Lines are mapped to lines (fortunately)
- Circles are mapped to conic section curves (not necessarily circles)
- Conic section curves are mapped to conic section curves
- Quadric surfaces are mapped to quadric surfaces (in some sense)
- Polynomial splines are mapped to rational splines
- The mapping is not linear — the mid-point of line E is not mapped to the mid-point of E'
- The mapping of a line is known once you know its effect on three points of the line
- The mapping preserves cross ratios (see any book on projective geometry)

Note that the mapping is defined only on *points* in 3D space. Defining its effect on vectors is tricky. Suppose we have a vector $\mathbf{V}$. The obvious approach is to find two points $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{V} = \mathbf{A} - \mathbf{B}$, and define $\mathbf{T}(\mathbf{V})$ to be $\mathbf{T}(\mathbf{A}) - \mathbf{T}(\mathbf{B})$. But this won't work, because different choices of $\mathbf{A}$ and $\mathbf{B}$ will give different values for $\mathbf{T}(\mathbf{A}) - \mathbf{T}(\mathbf{B})$.

1.15 The Projection Formula

Given a point $\mathbf{P}$, we want to find a formula that gives us its projection $\mathbf{T}(\mathbf{P})$ onto the plane π, as shown in the figures below:





First, we need an equation of the plane π. If $\mathbf{N}$ is a unit normal vector for the plane, and $\mathbf{R}$ is some fixed given point on the plane, then the plane can be described by the equation of the form $(\mathbf{X} - \mathbf{R}) \cdot \mathbf{N} = 0$. But, many different equations of this form are possible, and we can make our formulas simpler by some judicious choices. First, we choose $\mathbf{R} = \mathbf{Q}$, where $\mathbf{Q}$ is the foot of the perpendicular from $\mathbf{C}$ to the plane π. Then we reverse the direction of $\mathbf{N}$, if necessary, so that $\mathbf{C} = \mathbf{Q} + d\mathbf{N}$, where $d \geq 0$. In other words, we choose to make $\mathbf{N}$ point in the direction of

$\mathbf{C} - \mathbf{Q}$ (from the plane towards the center of projection). Since $\mathbf{T}(\mathbf{P})$ lies on the line through $\mathbf{P}$ and $\mathbf{C}$, there is some value t such that

$$\mathbf{T}(\mathbf{P}) = \mathbf{C} + t(\mathbf{P} - \mathbf{C}) \tag{1.2}$$

Also, since $\mathbf{T}(\mathbf{P})$ lies on the plane π, we have

$$(\mathbf{T}(\mathbf{P}) - \mathbf{Q})\cdot\mathbf{N} = 0 \tag{1.3}$$

We take the expression for $\mathbf{T}(\mathbf{P})$ from equation (1.2), substitute it into equation (1.3), and solve for t, giving

$$t = \frac{(\mathbf{C} - \mathbf{Q})\cdot\mathbf{N}}{(\mathbf{C} - \mathbf{P})\cdot\mathbf{N}} = \frac{d}{(\mathbf{C} - \mathbf{P})\cdot\mathbf{N}}$$

Substituting this value of t back into equation (1.2), we get

$$\mathbf{T}(\mathbf{P}) = \mathbf{C} - \frac{d(\mathbf{P} - \mathbf{C})}{(\mathbf{P} - \mathbf{C})\cdot\mathbf{N}} \tag{1.4}$$

Of course, there will be trouble if $(\mathbf{P} - \mathbf{C})\cdot\mathbf{N} = 0$, which will happen if the point $\mathbf{P}$ lies on a plane through $\mathbf{C}$ that is parallel to the projection plane π. The projection of these kinds of points lies “at infinity” in some sense.

For later purposes, it is useful to write our projection formula in terms of $\mathbf{Q}$, rather than $\mathbf{C}$. If we substitute $\mathbf{C} = \mathbf{Q} + d\mathbf{N}$ in equation (1.4), and do some algebra, we get:

$$\mathbf{T}(\mathbf{P}) = \mathbf{Q} + \frac{d\{(\mathbf{P} - \mathbf{Q}) - [(\mathbf{P} - \mathbf{Q})\cdot\mathbf{N}]\mathbf{N}\}}{d - (\mathbf{P} - \mathbf{Q})\cdot\mathbf{N}} \tag{1.5}$$

We can simplify this a little if we let $r = 1/d$ (the reciprocal of the distance), and write $\mathbf{D} = \mathbf{P} - \mathbf{Q}$. Then we get:

$$\mathbf{T}(\mathbf{P}) = \mathbf{Q} + \frac{\mathbf{D} - (\mathbf{D}\cdot\mathbf{N})\mathbf{N}}{1 - r(\mathbf{D}\cdot\mathbf{N})} \tag{1.6}$$

Given a point $\mathbf{P}$, we can calculate the projected point $\mathbf{T}(\mathbf{P})$ using any one of the equations (1.4), (1.5), or (1.6).

Note that these equations are all entirely coordinate-free — they do not make any reference to any coordinate system whatsoever. So, we can do the calculations in any coordinate we choose. As long as $\mathbf{P}$, $\mathbf{C}$, $\mathbf{Q}$ are expressed in the same coordinate system, the formulas will give us $\mathbf{T}(\mathbf{P})$, in this same coordinate system.