III : 8 Rectangular coordinate systems

Ever since the invention of rectangular coordinate systems in the seventeenth century, rectangular coordinates have proven to be a powerful tool for understanding and proving geometric facts; some of this was implicit in Greek studies of ellipses, hyperbolas and parabolas by Apollonius and others, but the decisive steps in formulating the modern use of coordinates really date back to the work of R. Descartes (1596–1650), P. de Fermat (1601–1669) and their numerous successors during the next hundred years. The purpose of this section is to formulate the main results on finding convenient systems of rectangular coordinates. Usually there are many choices, and using a good coordinate system is often crucial to coming up with a proof that is both concise and convincing. As noted in the *Wikipedia* article on coordinate systems, "Knowledge of how to erect a coordinate system where there is none [or no *a priori* choice] is essential to applying" coordinate geometry to a specific problem.

The main result

Here is the theorem for coordinatizing planes.

Theorem 1. *Let* $\mathbf{P}$ *be a system satisfying the axioms for Euclidean geometry, and let* a, b, c *be three noncollinear points in* $\mathbf{P}$*. Then there is a* $1-1$ *correspondence* $\varphi : \mathbf{P} \to \mathbb{R}^2$ *such that the following hold:*

(1) *Under the mapping* φ*, the Cartesian lines, distances and angle measures correspond to the abstract lines, distances and angle measures in* $\mathbf{P}$*.*

(2) $\varphi(a) = (0,0)$, $\varphi(b) = (x,0)$ *where* $x > 0$, *and* $\varphi(c) = (x_2, y)$ *where* $y > 0$.■

A complete proof of this result is implicit in Chapter 17 of Moise, *Elementary Geometry from an Advanced Standpoint*, Third Edition (Addison-Wesley, Reading MA *etc.*, 1990). We shall verify a special case later in this document.

Typical examples

In the applications of this result one frequently can read off some worthwhile information about the coordinate system fairly quickly.

Example 1. Suppose that we have two parallel lines L and M in $\mathbf{P}$, and let a, b, c be three noncollinear points in $\mathbf{P}$ such that the first two lie on L and the third point lies on M. If we apply the theorem to these three points, then the line L, which joins the first two points, will correspond to the x-axis, and the line M, which contains c, will be defined by an equation of the form $y = h$, where $h > 0$.■

Example 2. Suppose that we have points a, b in $\mathbf{P}$. If δ is the distance between the two points and m is the midpoint of $[ab]$, then we can find a coordinate system such that a corresponds to $(\delta/2, 0)$ and m corresponds to $(0,0)$. The conditions on the coordinate system then implies that b corresponds to $(-\delta/2, 0)$.■

The file `locus-problems.pdf` proves a standard result in coordinate geometry, first using classical methods and then using coordinate gteometry with a judiciously chosen rectangular coordinate system; the second proof is much simpler than the first, but the discussion in `locus-problems.pdf` explains why we are not really getting something for nothing. Much of the hard work is being done by the implicit properties of the rectangular coordinate system. Additional examples showing the power of rectangular coordinates appear in `locus-problems2.pdf` and `locus-problems3.pdf`.

Three-dimensional analog

There is a corresponding result in solid geometry, with the following changes: In the first part, planes must be included as part of the structural data. In the second part, one starts with a quadruple of noncoplanar points a, b, c, d and the mapping φ satisfies $\varphi(a) = (0,0,0)$, $\varphi(b) = (x,0,0)$ where $x > 0$, $\varphi(c) = (x_2, y, 0)$ where $y > 0$, and $\varphi(d) = (x_3, y_3, z)$ where $z > 0$.■

Proof of the main result

We shall do this when $\mathbf{P}$ is a coordinate plane $\mathbb{R}^n$ with distance and angle measurement defined as in linear algebra. Most of the reasoning is contained in the following elaboration of the Gram-Schmidt orhonormalization process:

Proposition 2. *Let V be an n-dimensional real inner product space, and let $y_1, \cdots, y_n$ be a basis for V. Then there is an orthonormal basis $u_1, \cdots, u_n$ for V such that the following hold:*

(*i*) *For each ksuch that $1 \leq k \leq n$ the spans of $y_1, \cdots, y_k$ and $u_1, \cdots, u_k$ are equal.*

(*ii*) *For each ksuch that $1 \leq k \leq n$ the inner product $\langle y_k, u_k \rangle$ is positive.*

Proof that Proposition 2 implies Theorem 1. Recall that we are restricting ourselves to the case where $\mathbf{P}$ is the standard coordinate plane $\mathbb{R}^2$; the general case follows from the uniqueness result cited in Theorem V.5.4 of the file `geometrynotes05b.f13.pdf`.

Since a, b and c are noncollinear points, it follows that $y_1 = b - a$ and $y_2 = c - a$ form a basis for $\mathbb{R}^2$. Let $\{u_1, u_2\}$ be the orthonormal basis given by Proposition 2, and let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the unique orthogonal linear transformation which maps the standard unit vectors $\{e_1, e_2\}$ to $\{u_1, u_2\}$. Then the isometry $\varphi(x) = L^{-1}(x - a)$ has the required properties. ■

Note that one can extend this argument to coordinate n-spaces $\mathbb{R}^n$ for all $n \geq 3$.

Proof of Proposition 2. The first step is to apply the Gram-Schmidt orthonormalization process to the basis $y_1, \cdots, y_n$. This yields an orthonormal basis $v_1, \cdots, v_n$ which satisfies Property (i) in the proposition. Hence it is only necessary to modify the construction so that (ii) also holds.

To see that this is possible, let k be an arbitrary integer between 1 and n, and consider the inner product $\langle y_k, v_k \rangle$. We claim that this inner product is nonzero. If $k = 1$ this is true because v_k is a nonzero multiple of the nonzero vector y_k, and if $k \geq 2$ this is true because $\langle y_k, v_k \rangle = 0$ and the standard formula

$$y_k \quad = \quad \sum_{j=1}^{k} \langle y_k, v_j \rangle$$

would imply that y_k lies in the span of $v_1, \ \cdots \ , v_{k-1}$, which is equal to the span of $y_1, \ \cdots \ , y_{k-1}$. The latter would imply that the vectors $y_1, \ \cdots \ , y_n$ are linearly dependent, which contradicts the assumption that the latter form a basis. Hence the inner product is nonzero as claimed.

Finally, modify the orthonormal basis $v_1, \ \cdots \ , v_n$ as follows: If $\langle y_k, v_k \rangle$ is positive let $u_k = v_k$, and if $\langle y_k, v_k \rangle$ is negative let $u_k = -v_k$. Clearly the spans of $v_1, \ \cdots \ , v_k$ and $u_1, \ \cdots \ , u_k$ are identical since the vectors in one are nonzero multiples of the vectors in the other, so Property (i) is still satisfied. Furthermore, our modification guarantees that $\langle y_k, u_k \rangle$ is always positive, so that Property (ii) is also satisfied.■