

COMPUTER AIDED GEOMETRIC DESIGN

Thomas W. Sederberg

August 19, 1997

Contents

- 1 INTRODUCTION** **1**
- 1.1 Points, Vectors and Coordinate Systems 1
 - 1.1.1 Vector Algebra 2
 - 1.1.2 Points vs. Vectors 6
 - 1.1.3 Homogeneous Coordinates 7
- 1.2 Lines 7
 - 1.2.1 Parametric equations of lines 7
 - 1.2.2 Implicit equations of lines 9
- 1.3 Conic Sections 11
 - 1.3.1 Parametric equations of conics 12

- 2 BÉZIER CURVES** **13**
- 2.1 Bézier Curves over Arbitrary Parameter Intervals 16
- 2.2 Subdivision of Bézier Curves 16
- 2.3 Degree Elevation 18
- 2.4 Distance between Two Bézier Curves 20
- 2.5 Derivatives 20
- 2.6 Continuity 21
- 2.7 Three Dimensional Bézier Curves 22
- 2.8 Rational Bézier Curves 22
- 2.9 Tangency and Curvature 24
- 2.10 Reparametrization 26
- 2.11 Circular Arcs 27
- 2.12 Explicit Bézier Curves 28
- 2.13 Integrating Bernstein polynomials 28
- 2.14 Forward Differencing 30

- 3 POLYNOMIAL EVALUATION** **33**
- 3.1 Power Basis 33
- 3.2 Bernstein Basis 33
- 3.3 Basis Conversion 34
 - 3.3.1 Example 35
 - 3.3.2 Closed Form Expression 36

4	BLENDING FUNCTIONS	37
4.1	Timmer's Parametric Cubic	39
4.2	Ball's Rational Cubic	40
4.3	Overhauser Curves	40
5	B-SPLINE CURVES	43
5.1	Polar Form	44
5.1.1	Subdivision of Bézier Curves	45
5.2	Knot Vectors	46
5.3	Extracting Bézier Curves from B-splines	47
5.4	Multiple knots	49
5.5	Periodic B-splines	49
5.6	Bézier end conditions	49
5.7	Knot insertion	51
5.8	The de Boor algorithm	51
5.9	Explicit B-splines	51
5.10	B-spline hodographs	53
5.11	Cubic B-Splines	53
5.11.1	Splitting a B-spline into Bézier Curves	53
5.11.2	Knot Insertion	54
5.11.3	Combining Bézier curves into a B-spline	55
5.12	B-spline blending functions	56
6	POLYNOMIAL INTERPOLATION	57
6.1	Undetermined Coefficients	57
6.2	Lagrange Interpolation	59
6.3	Error Bounds	60
6.3.1	Chebyshev Polynomials	63
6.4	Interpolating Points and Normals	63
7	ALGEBRAIC GEOMETRY FOR CAGD	65
7.1	Preliminaries	66
7.2	Implicitization	68
7.3	Brute Force Implicitization	68
7.4	Resultants	69
7.5	Determining the Common Root	71
7.6	Implicitization and Inversion	74
7.7	Curve-Curve Intersections	76
7.8	Surfaces	78
7.9	Base Points	79
7.10	Implicitization in Bézier Form	80
8	CURVE INTERSECTION	81
8.1	Timing Comparisons	81
8.2	Bézier subdivision	82
8.3	Interval subdivision	82
8.4	Implicitization	83
8.5	Bézier Clipping method	83
8.5.1	Fat Lines	83

8.5.2	Bézier Clipping	87
8.5.3	Iterating	88
8.5.4	Clipping to other fat lines	89
8.5.5	Multiple Intersections	89
8.5.6	Rational Curves	89
8.6	Tangent Intersections	90
9	FREE-FORM DEFORMATION	93
10	PLANAR ALGEBRAIC CURVES	97
10.1	Genus and Parametrization	97
10.2	Detecting Double Points	97
10.3	Implicit Curve Intersections	100
10.4	Discriminants	101
10.5	Parametrizing Unicursal Curves	102
10.6	Undetermined Coefficients	104
11	POLYNOMIAL ROOT FINDING IN BERNSTEIN FORM	107
11.1	Convex Hull Marching	108
11.2	Bernstein Combined Subdivide & Derivative Algorithm	108
12	INTERVAL BÉZIER CURVES	113
12.1	Interval arithmetic and interval polynomials	113
12.2	Interval Bézier curves	115
12.2.1	Affine maps	116
12.2.2	Centered form	118
12.2.3	Error monotonicity	121
12.2.4	Envelopes of interval Bézier curves	121
12.2.5	Interval hodographs	122
12.3	Approximation by interval polynomials	123
12.3.1	Remainder formulae and interval approximants	123
12.3.2	Hermite interpolation	124
12.3.3	Estimating bounds on derivatives	127
12.4	Approximation by interval Bézier curves	127

List of Figures

1.1	Equivalent Vectors	2
1.2	Unit Vectors	3
1.3	Vector in Component Form	3
1.4	Relative Position Vectors	4
1.5	Vector Addition	4
1.6	Vector Subtraction	4
1.7	Line given by $\mathbf{A}_0 + \mathbf{A}_1 t$	7
1.8	Line given by $\mathbf{P}(t) = \frac{(t_1-t)\mathbf{P}_0+(t-t_0)\mathbf{P}_1}{t_1-t_0}$	8
1.9	Affine example	8
1.10	Line defined by point and normal.	10
1.11	Normalized line equation.	11
2.1	Center of mass of four points	13
2.2	Cubic Bézier blending functions	14
2.3	Cubic Bézier curve	14
2.4	Examples of cubic Bézier curves	15
2.5	Bézier curves of various degree	15
2.6	Subdividing a cubic Bézier curve.	17
2.7	Recursively subdividing a quadratic Bézier curve.	17
2.8	Subdividing a quadratic Bézier curve.	18
2.9	Degree elevation.	19
2.10	Repeated degree elevation.	19
2.11	Difference curve.	20
2.12	Hodograph.	21
2.13	C^2 Bézier curves.	22
2.14	Endpoint curvature.	23
2.15	Rational Bézier curve.	24
2.16	Rational curve as the projection of a 3-D curve	25
2.17	Circular arcs	27
2.18	Circle as Degree 5 Rational Bézier Curve	27
2.19	Explicit Bézier curve	28
4.1	Convex Hull Property	38
4.2	Variation Diminishing Property	39
4.3	Timmer's PC	40
4.4	Ball's Cubic	41
4.5	Overhauser curves	42

5.1	Spline and ducks	43
5.2	Bézier curves labeled using polar form.	44
5.3	B-spline curve labeled using polar form.	45
5.4	Affine map property of polar values.	46
5.5	Subdividing a cubic Bézier curve.	47
5.6	First step in Böhm algorithm.	48
5.7	Second step in Böhm algorithm.	48
5.8	Double knot.	49
5.9	Periodic B-spline	50
5.10	Bézier end condition.	50
5.11	Before and after.	51
5.12	Knot insertion.	52
5.13	De Boor algorithm	52
5.14	Extracting a Bézier curve from a B-spline	54
5.15	Knot insertion	54
6.1	Interpolating Four Points	57
6.2	Interpolating Four Points	59
6.3	Error Bounds	61
6.4	Piecewise linear approximation of a Bézier curve	62
6.5	Two cases of $(x - x_0)(x - x_1) \cdots (x - x_9)$ for $0 \leq x \leq 1$	63
7.1	Two cubic curves intersecting nine times	77
8.1	Convex Hulls	82
8.2	Three iterations of Bézier subdivision	82
8.3	Interval preprocess and subdivision	83
8.4	Fat line bounding a quartic curve	84
8.5	Fat line for a polynomial quadratic curve	85
8.6	Fat lines for polynomial cubic curves	86
8.7	Bézier curve/fat line intersection	86
8.8	Explicit Bézier curve	87
8.9	After first Bézier clip	88
8.10	Two intersections	89
8.11	Two intersections, after a split	89
8.12	Collinear normal line	91
8.13	Tangent intersection	91
9.1	FFD example	93
9.2	FFD example	94
9.3	Continuity control	94
9.4	FFD local coordinates	95
9.5	FFD undisplaced control points	95
10.1	Irreducible Cubic Curve	98
10.2	Crunode: $x^3 + 9x^2 - 12y^2 = 0$	99
10.3	Cusp: $x^3 - 3y^2 = 0$	99
10.4	Acnode: $x^3 - 3x^2 - 3y^2 = 0$	100
10.5	Circle and Hyperbola	101

10.6	Parametrizing a Circle	103
10.7	Parametrizing a Cubic Curve	103
11.1	Bernstein root finding	109
11.2	Root isolation heuristic (a-d).	110
11.3	Root isolation heuristic (e-h).	111
12.1	A cubic interval Bézier curve.	116
12.2	The affine map of two scalar points.	117
12.3	The affine map of two scalar intervals.	117
12.4	The affine map of two vector intervals.	118
12.5	Interval de Casteljau algorithm.	119
12.6	The envelope of an interval Bézier curve.	122
12.7	Approximate arc length parametrization of circle.	129

Chapter 1

INTRODUCTION

Computer aided geometric design (CAGD) concerns itself with selected aspects of geometry, computer graphics, numerical analysis, approximation theory, data structures and computer algebra. It considers all problems in which geometry is involved in a computer algorithm. Practical examples include defining the shape of a car or airplane to facilitate the design and manufacturing process.

CAGD is a young field. The first work in this field began in the mid 1960s. The term *computer aided geometric design* was coined in 1974 by R.E. Barnhill and R.F. Riesenfeld in connection with a conference at the University of Utah.

1.1 Points, Vectors and Coordinate Systems

Consider the simple problem of writing a computer program which finds the area of any triangle. We must first decide how to uniquely describe the triangle. One way might be to provide the lengths l_1, l_2, l_3 of the three sides, from which Heron's formula yields

$$Area = \sqrt{s(s-l_1)(s-l_2)(s-l_3)}, \quad s = \frac{l_1 + l_2 + l_3}{2}.$$

An alternate way to describe the triangle is in terms of its vertices. But while the lengths of the sides of a triangle are independent of its position, we can specify the vertices to our computer program only with reference to some *coordinate system* — which can be defined simply as any method for representing points with numbers.

Note that a coordinate system is an artificial device which we arbitrarily impose for the purposes at hand. Imagine a triangle cut out of paper and lying on a flat table in the middle of a room. We could define a *Cartesian coordinate system* whose origin lies in a corner of the room, and whose coordinate axes lie along the three room edges which meet at the corner. We would further specify the unit of measurement, say centimeters. Then, each vertex of our triangle could be described in terms of its respective distance from the two walls containing the origin and from the floor. These distances are the Cartesian coordinates (x, y, z) of the vertex with respect to the coordinate system we defined.

Vectors A *vector* can be pictured as a line segment of definite length with an arrow on one end. We will call the end with the arrow the *tip* or *head* and the other end the *tail*.

Two vectors are equivalent if they have the same length, are parallel to each other, and point in the same direction (have the same *sense*) as shown in Figure 1.1. For a given coordinate system, we

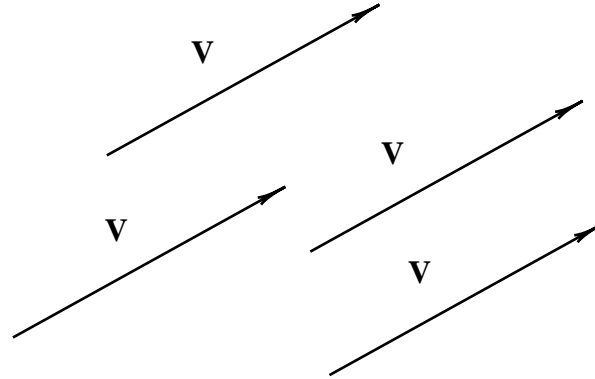


Figure 1.1: Equivalent Vectors

can describe a three-dimensional vector in the form (a, b, c) where a (or b or c) is the distance in the x (or y or z) direction from the tail to the tip of the vector.

Unit Vectors The symbols \mathbf{i} , \mathbf{j} , and \mathbf{k} denote vectors of “unit length” (based on the unit of measurement of the coordinate system) which point in the positive x , y , and z directions respectively (see Figure 1.2). Unit vectors allow us to express a vector in component form (see Figure 1.3):

$$\mathbf{P} = (a, b, c) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}.$$

An expression such as (x, y, z) can be called a *triple* of numbers. In general, an expression (x_1, x_2, \dots, x_n) is an n -tuple, or simply a tuple. As we have seen, a triple can signify either a point or a vector.

Relative Position Vectors Given two points \mathbf{P}_1 and \mathbf{P}_2 , we can define

$$\mathbf{P}_{2/1} = \mathbf{P}_2 - \mathbf{P}_1$$

as the vector pointing from \mathbf{P}_1 to \mathbf{P}_2 . This notation $\mathbf{P}_{2/1}$ is widely used in engineering mechanics, and can be read “the position of point \mathbf{P}_2 relative to \mathbf{P}_1 ” (see Figure 1.4).

In our diagrams, points will be drawn simply as dots or small circles, and vectors as line segments with single arrows. Vectors and points will both be denoted by bold faced type.

1.1.1 Vector Algebra

Given two vectors $\mathbf{P}_1 = (x_1, y_1, z_1)$ and $\mathbf{P}_2 = (x_2, y_2, z_2)$, the following operations are defined:

Addition:

$$\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{P}_2 + \mathbf{P}_1 = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

Subtraction:

$$\mathbf{P}_1 - \mathbf{P}_2 = (x_1 - x_2, y_1 - y_2, z_1 - z_2)$$

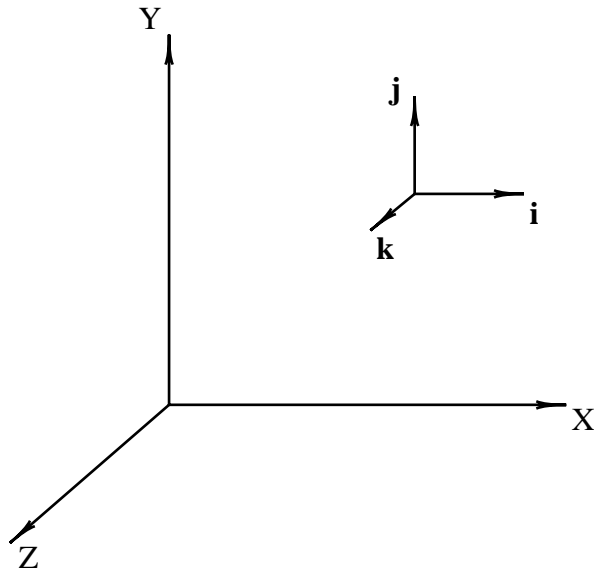


Figure 1.2: Unit Vectors

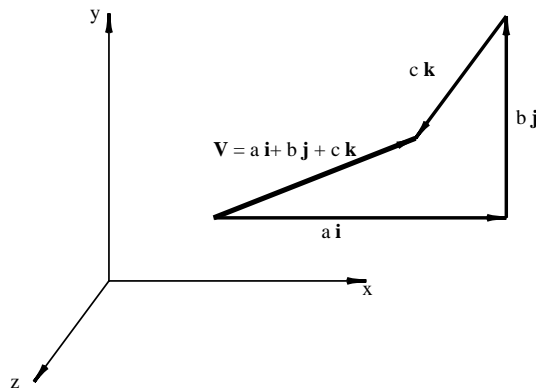


Figure 1.3: Vector in Component Form

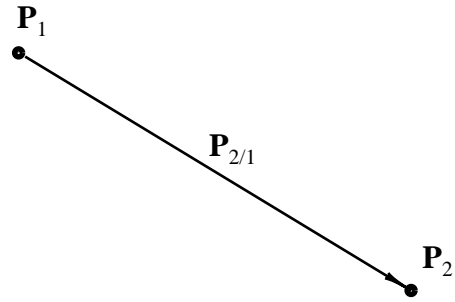


Figure 1.4: Relative Position Vectors

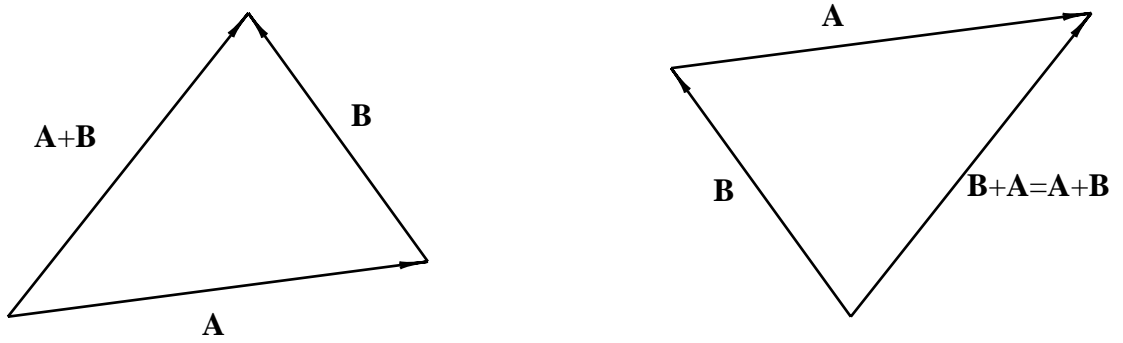


Figure 1.5: Vector Addition

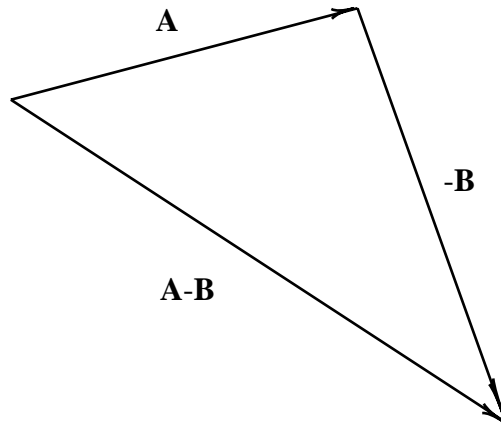


Figure 1.6: Vector Subtraction

Scalar multiplication:

$$c\mathbf{P}_1 = (cx_1, cy_1, cz_1)$$

Length of a Vector

$$|\mathbf{P}_1| = \sqrt{x_1^2 + y_1^2 + z_1^2}$$

Dot Product The dot product of two vectors is defined

$$(1.1) \quad \mathbf{P}_1 \cdot \mathbf{P}_2 = |\mathbf{P}_1||\mathbf{P}_2| \cos \theta$$

where θ is the angle between the two vectors. Since the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are mutually perpendicular,

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1$$

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k} = \mathbf{j} \cdot \mathbf{k} = 0.$$

Since the dot product obeys the distributive law

$$\mathbf{P}_1 \cdot (\mathbf{P}_2 + \mathbf{P}_3) = \mathbf{P}_1 \cdot \mathbf{P}_2 + \mathbf{P}_1 \cdot \mathbf{P}_3,$$

we can easily derive the very useful equation

$$(1.2) \quad \begin{aligned} \mathbf{P}_1 \cdot \mathbf{P}_2 &= (x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \cdot (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) \\ &= (x_1 * x_2 + y_1 * y_2 + z_1 * z_2) \end{aligned}$$

The dot product allows us to easily compute the angle between any two vectors. From equation 1.1,

$$\theta = \cos^{-1} \left(\frac{\mathbf{P}_1 \cdot \mathbf{P}_2}{|\mathbf{P}_1||\mathbf{P}_2|} \right).$$

Example. Find the angle between vectors $(1, 2, 4)$ and $(3, -4, 2)$.

Answer.

$$\begin{aligned} \theta &= \cos^{-1} \left(\frac{\mathbf{P}_1 \cdot \mathbf{P}_2}{|\mathbf{P}_1||\mathbf{P}_2|} \right) \\ &= \cos^{-1} \left(\frac{(1, 2, 4) \cdot (3, -4, 2)}{|(1, 2, 4)|| (3, -4, 2)|} \right) \\ &= \cos^{-1} \left(\frac{3}{\sqrt{21}\sqrt{29}} \right) \\ &\approx 83.02^\circ \end{aligned}$$

Cross Product: The cross product $\mathbf{P}_1 \times \mathbf{P}_2$ is a vector whose magnitude is

$$|\mathbf{P}_1 \times \mathbf{P}_2| = |\mathbf{P}_1||\mathbf{P}_2| \sin \theta$$

(where again θ is the angle between \mathbf{P}_1 and \mathbf{P}_2), and whose direction is mutually perpendicular to \mathbf{P}_1 and \mathbf{P}_2 with a sense defined by the right hand rule as follows. Point your fingers in the direction

of \mathbf{P}_1 and orient your hand such that when you close your fist your fingers pass through the direction of \mathbf{P}_2 . Then your right thumb points in the sense of $\mathbf{P}_1 \times \mathbf{P}_2$.

From this basic definition, one can verify that

$$\mathbf{P}_1 \times \mathbf{P}_2 = -\mathbf{P}_2 \times \mathbf{P}_1,$$

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}$$

$$\mathbf{j} \times \mathbf{i} = -\mathbf{k}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}.$$

Since the cross product obeys the distributive law

$$\mathbf{P}_1 \times (\mathbf{P}_2 + \mathbf{P}_3) = \mathbf{P}_1 \times \mathbf{P}_2 + \mathbf{P}_1 \times \mathbf{P}_3,$$

we can derive the important relation

$$\begin{aligned} \mathbf{P}_1 \times \mathbf{P}_2 &= (x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \times (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) \\ &= (y_1z_2 - y_2z_1, x_2z_1 - x_1z_2, x_1y_2 - x_2y_1) \\ (1.3) \quad &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} \end{aligned}$$

Area of a Triangle. Cross products have many important uses. For example, finding a vector which is mutually perpendicular to two other vectors. Also, the cross product provides a straightforward method for finding the area of a triangle which is defined by three points $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ in space.

$$(1.4) \quad Area = \frac{1}{2}|\mathbf{P}_{1/2}||\mathbf{P}_{1/3}|\sin\theta_1 = \frac{1}{2}|\mathbf{P}_{1/2} \times \mathbf{P}_{1/3}|$$

For example, the area of a triangle with vertices $\mathbf{P}_1 = (1, 1, 1)$, $\mathbf{P}_2 = (2, 4, 5)$, $\mathbf{P}_3 = (3, 2, 6)$ is

$$\begin{aligned} Area &= \frac{1}{2}|\mathbf{P}_{1/2} \times \mathbf{P}_{1/3}| \\ &= \frac{1}{2}|(1, 3, 4) \times (2, 1, 5)| \\ &= \frac{1}{2}|(11, 3, -5)| = \frac{1}{2}\sqrt{11^2 + 3^2 + (-5)^2} \\ &\approx 6.225 \end{aligned}$$

1.1.2 Points vs. Vectors

A point is a geometric entity which connotes position, whereas a vector connotes direction and magnitude. From a purely mathematical viewpoint, there are good arguments for carefully distinguishing between triples which refer to points, and triples that signify vectors [Goldman '85]. However, no problem arises if we recognize that a triple connoting a point can be interpreted as a vector from the origin to the point. Thus, we could call a point an *absolute position vector*, and the difference between two points a *relative position vector*. These phrases are often used in engineering mechanics, where vectors are used to express quantities other than position, such as velocity or acceleration.

1.1.3 Homogeneous Coordinates

The *homogeneous* Cartesian coordinates (X, Y, W) of a point are defined

$$x = \frac{X}{W}; \quad y = \frac{Y}{W}.$$

Homogeneous coordinates are useful, among other things, for expressing points at infinity: The point $(X, Y, 0)$ is an infinite distance from the origin (or, from any finite point, for that matter) in the direction $X\mathbf{i} + Y\mathbf{j}$. Obviously, the homogeneous coordinates of a point are only unique to within a scale factor. For example, the point $(x, y) = (2, 3)$ has homogeneous coordinates $(X, Y, W) = (2, 3, 1)$, or $(4, 6, 2)$, or in general, $(2W, 3W, W)$. The point $(X, Y, W) = (0, 0, 0)$ is undefined.

1.2 Lines

A line can be defined using either a parametric equation, or an implicit equation.

1.2.1 Parametric equations of lines

Linear parametric equation. A line can be written in parametric form as follows:

$$x = a_0 + a_1t; \quad y = b_0 + b_1t$$

In vector form,

$$(1.5) \quad \mathbf{P}(t) = \left\{ \begin{array}{c} x(t) \\ y(t) \end{array} \right\} = \left\{ \begin{array}{c} a_0 + a_1t \\ b_0 + b_1t \end{array} \right\} = \mathbf{A}_0 + \mathbf{A}_1t.$$

In this equation, \mathbf{A}_0 is a point on the line, and \mathbf{A}_1 is the direction of the line (see Figure 1.7)

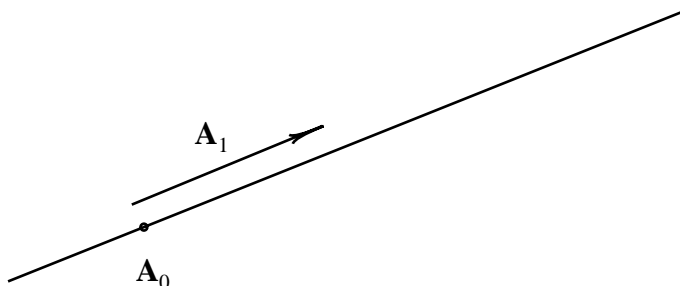


Figure 1.7: Line given by $\mathbf{A}_0 + \mathbf{A}_1t$.

Affine parametric equation of a line. A straight line can also be expressed

$$(1.6) \quad \mathbf{P}(t) = \frac{(t_1 - t)\mathbf{P}_0 + (t - t_0)\mathbf{P}_1}{t_1 - t_0}$$

where \mathbf{P}_0 and \mathbf{P}_1 are two points on the line and t_0 and t_1 are any parameter values. Note that $\mathbf{P}(t_0) = \mathbf{P}_0$ and $\mathbf{P}(t_1) = \mathbf{P}_1$. Note in Figure 1.8 that the line *segment* \mathbf{P}_0 - \mathbf{P}_1 is defined by restricting the parameter:

$$t_0 \leq t \leq t_1.$$

Sometimes this is expressed by saying that the line segment is the portion of the line in the *parameter interval* or *domain* $[t_0, t_1]$. We will soon see that the line in Figure 1.8 is actually a degree one Bézier

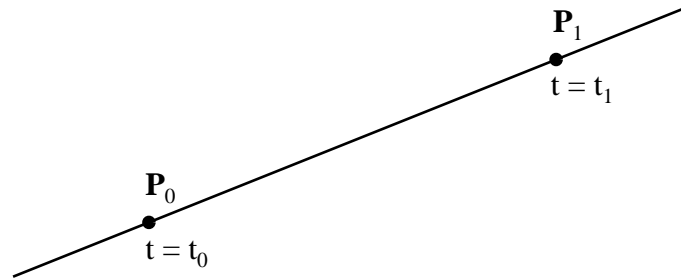


Figure 1.8: Line given by $\mathbf{P}(t) = \frac{(t_1-t)\mathbf{P}_0+(t-t_0)\mathbf{P}_1}{t_1-t_0}$

curve. Most commonly, we have $t_0 = 0$ and $t_1 = 1$ in which case

$$(1.7) \quad \mathbf{P}(t) = (1-t)\mathbf{P}_0 + t\mathbf{P}_1.$$

Equation 1.7 is called an *affine* equation, whereas equation 1.5 is called a *linear* equation. An affine equation is coordinate system independent, and is mainly concerned with ratios and proportions. An affine equation can be thought of as answering the question: “If a line is defined through two points \mathbf{P}_0 and \mathbf{P}_1 , and if point \mathbf{P}_0 corresponds to parameter value t_0 and point \mathbf{P}_1 corresponds to parameter value t_1 , what point corresponds to an arbitrary parameter value t ?” Figure 1.9 shows a line on which \mathbf{P}_0 corresponds to parameter $t = t_0 = 1$ and \mathbf{P}_1 is assigned parameter value $t = t_1 = 4$. For example, the point corresponding to $t = 2$ is one third of the way from \mathbf{P}_0 to \mathbf{P}_1 .

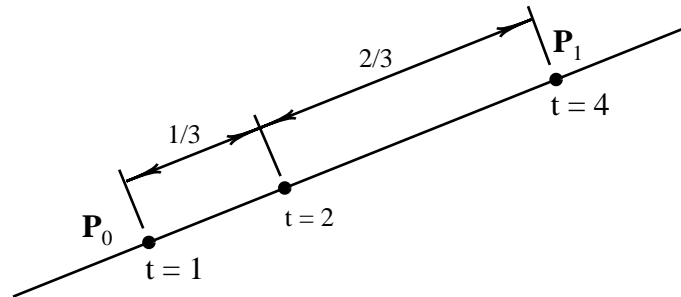


Figure 1.9: Affine example

Note that an affine equation can be derived from any two points on a line, given the parameter values for those points. If $\mathbf{P}(\alpha)$ is the point corresponding to parameter value $t = \alpha$ and if $\mathbf{P}(\beta)$ is the point corresponding to parameter value $t = \beta$ ($\alpha \neq \beta$), then the point corresponding to parameter value γ is

$$\mathbf{P}(\gamma) = \mathbf{P}(\alpha) + \frac{\gamma - \alpha}{\beta - \alpha} [\mathbf{P}(\beta) - \mathbf{P}(\alpha)] = \frac{(\beta - \gamma)\mathbf{P}(\alpha) + (\gamma - \alpha)\mathbf{P}(\beta)}{\beta - \alpha}$$

Rational parametric equations and homogeneous form. A line can also be defined using the following parametric equations:

$$(1.8) \quad x = \frac{a_0 + a_1 t}{d_0 + d_1 t}; \quad y = \frac{b_0 + b_1 t}{d_0 + d_1 t}.$$

This is normally called a *rational* parametric equation (because of the *ratio* of linear equations involved).

It is often convenient to write such equations in homogeneous form. Recall that the homogeneous Cartesian coordinates (X, Y, W) of a point are related to its Cartesian coordinates by

$$(x, y) = \left(\frac{X}{W}, \frac{Y}{W} \right).$$

Thus, we can rewrite equation 1.8 as

$$X = a_0 + a_1 t; \quad Y = b_0 + b_1 t; \quad W = d_0 + d_1 t.$$

Furthermore, rational parametric equations are sometimes written in terms of *homogeneous parameters* (T, U) where $t = \frac{T}{U}$. Thus,

$$X = a_0 + a_1 \frac{T}{U}; \quad Y = b_0 + b_1 \frac{T}{U}; \quad W = d_0 + d_1 \frac{T}{U}.$$

But since we can scale (X, Y, W) without changing the point (x, y) which it denotes, we can scale by U to give

$$X = a_0 U + a_1 T; \quad Y = b_0 U + b_1 T; \quad W = d_0 U + d_1 T.$$

Homogeneous parameters (T, U) facilitate working with infinite parameter values. For example, the point on the line corresponding to $t = \infty$ is easily handled using homogeneous parameters $(T, U) = (\text{anything}, 0)$. This has direct practical value in CAGD: a floating point algorithm attempting to evaluate a point of infinite parameter value would face a floating point overflow unless it made use of homogeneous parameters.

1.2.2 Implicit equations of lines

A line can also be expressed in what is known as an *implicit* equation:

$$f(x, y) = ax + by + c = 0; \quad \text{or} \quad F(X, Y, W) = aX + bY + cW = 0.$$

The line defined by an implicit equation is the set of all points which satisfy the equation $f(x, y) = 0$.

An implicit equation for a line can be derived given a point $\mathbf{P}_0 = (x_0, y_0)$ on the line and the normal vector $\mathbf{n} = a\mathbf{i} + b\mathbf{j}$. As shown in Figure 1.10, a point $\mathbf{P} = (x, y)$ is on this line if

$$(\mathbf{P} - \mathbf{P}_0) \cdot \mathbf{n} = 0$$

from which

$$(1.9) \quad f(x, y) = (x - x_0, y - y_0) \cdot (a, b) = ax + by - (ax_0 + by_0) = 0.$$

From equation 1.9, a line whose implicit equation is $ax + by + c = 0$ has the normal vector $\mathbf{n} = a\mathbf{i} + b\mathbf{j}$.

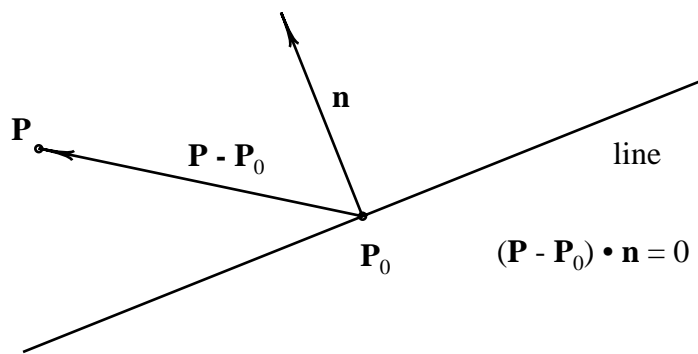


Figure 1.10: Line defined by point and normal.

Distance from point to line. If \mathbf{n} is a unit vector (that is, if $a^2 + b^2 = 1$), then the value $f(x, y)$ in equation 1.9 indicates the perpendicular distance of a point (x, y) to the line. This can be seen from equation 1.9 and Figure 1.10. The dot product $(\mathbf{P} - \mathbf{P}_0) \cdot \mathbf{n}$ is the projection of vector $(\mathbf{P} - \mathbf{P}_0)$ onto the unit normal \mathbf{n} , which is the perpendicular distance from \mathbf{P} to the line.

Since the coefficients of an implicit equation can be uniformly scaled without changing the curve (because if $f(x, y) = 0$, then $c \times f(x, y) = 0$ also), the implicit equation of a line can always be *normalized*:

$$f(x, y) = a'x + b'y + c' = \frac{a}{\sqrt{a^2 + b^2}}x + \frac{b}{\sqrt{a^2 + b^2}}y + \frac{c}{\sqrt{a^2 + b^2}} = 0.$$

Then, $f(x, y)$ is the signed distance from the point (x, y) to the line, with all points on one side of the line having $f(x, y) > 0$ and the other side having $f(x, y) < 0$. Note that $|c'| = |f(0, 0)|$ is the distance from the origin to the line. Thus, if $c = 0$, the line passes through the origin. The coefficients a' and b' relate to the slope of the line. Referring to Figure 1.11, $a' = \cos(\theta)$, $b' = \sin(\theta)$, and $c' = -p$.

Implicit equation of line through two points.

Three points (X_1, Y_1, W_1) , (X_2, Y_2, W_2) and (X_3, Y_3, W_3) are collinear if

$$\begin{vmatrix} X_1 & Y_1 & W_1 \\ X_2 & Y_2 & W_2 \\ X_3 & Y_3 & W_3 \end{vmatrix} = 0.$$

Thus, the equation of the line through two points is

$$\begin{vmatrix} X & Y & W \\ X_1 & Y_1 & W_1 \\ X_2 & Y_2 & W_2 \end{vmatrix} = 0.$$

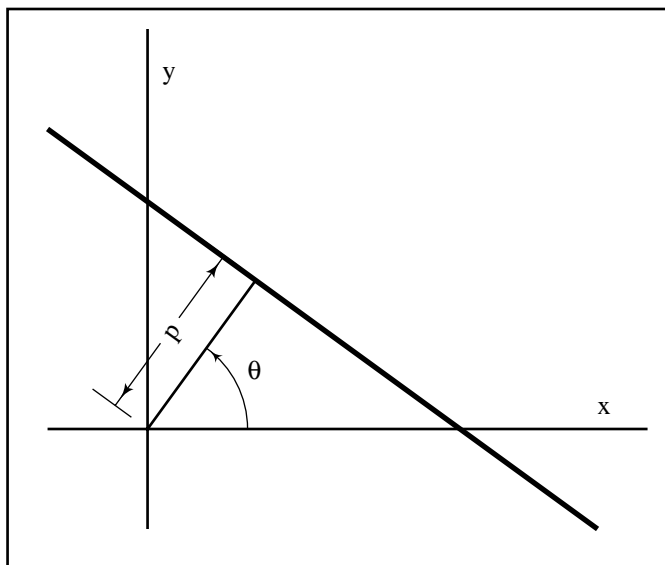


Figure 1.11: Normalized line equation.

1.3 Conic Sections

A conic section (or, simply *conic*) is any degree two curve. Any conic can be expressed using a degree two implicit equation:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

or, in homogeneous form:

$$(1.10) \quad aX^2 + bXY + cY^2 + dXW + eYW + fW^2 = 0.$$

Conics can be classified as hyperbolas, parabolas and ellipses (of which the circle is a special case). What distinguishes these cases is the number of real points at which the curve intersects the line at infinity $W = 0$. A hyperbola intersects $W = 0$ in two real points. Those points are located an infinite distance along the asymptotic directions. A parabola is tangent to the line at infinity, and thus has two coincident real intersection points. This point is located an infinite distance along the parabola's axis of symmetry. Ellipses do not intersect the line at infinity at any real point – all real points on an ellipse are finite.

To determine the number of real points at which a conic intersects the line at infinity, simply intersect equation 1.10 with the line $W = 0$ to get:

$$aX^2 + bXY + cY^2 = 0$$

from which

$$\frac{Y}{X} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2c}.$$

The two values Y/X are the slopes of the lines pointing to the intersections of the conic with the line at infinity. Thus, if $b^2 - 4ac > 0$, there are two distinct real intersections and the conic is a

hyperbola. If $b^2 - 4ac = 0$, there are two coincident real intersections and the conic is a parabola, and if $b^2 - 4ac < 0$, there are no real intersections and the conic is an ellipse. The value $b^2 - 4ac$ is known as the *discriminant* of the conic.

1.3.1 Parametric equations of conics

The parametric equation of any conic can be expressed:

$$x = \frac{a_2t^2 + a_1t + a_0}{d_2t^2 + d_1t + d_0}; \quad y = \frac{b_2t^2 + b_1t + b_0}{d_2t^2 + d_1t + d_0}.$$

or, in homogeneous form,

$$\begin{aligned} X &= a_2T^2 + a_1TU + a_0U^2; \\ Y &= b_2T^2 + b_1TU + b_0U^2; \\ W &= d_2T^2 + d_1TU + d_0U^2. \end{aligned}$$

It is also possible to classify a conic from its parametric equation. We again identify the points at which the conic intersects the line at infinity. In the parametric form, the only places at which (x, y) can be infinitely large is at parameter values of t for which

$$d_2t^2 + d_1t + d_0 = 0.$$

Thus, we note that $d_1^2 - 4d_0d_2$ serves the same function as the discriminant of the implicit equation. If $d_1^2 - 4d_0d_2 > 0$, there are two real, distinct values of t at which the conic goes to infinity and the curve is a hyperbola. If $d_1^2 - 4d_0d_2 < 0$, there are no real values of t at which the conic goes to infinity and the curve is an ellipse. If $d_1^2 - 4d_0d_2 = 0$, there are two real, identical values of t at which the conic goes to infinity and the curve is a parabola.