Original Lecture #5: 14th October 1993
Topics: Polynomials
Scribe: Mark P. Kust*

# 1 Roots of polynomials

A polynomial function $f(x)$ of degree $n$ can be expressed in terms of $n+1$ coefficients $a_i$ as

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n. \tag{1}$$

The *roots* of the polynomial $f(x)$ are those values of $x$ for which

$$f(x) = 0.$$

Consider the cubic polynomial $g(x)$ shown in Figure 1 and given by

$$g(x) = \frac{16}{3}x^3 - \frac{10}{3}x + C.$$

Depending on the value of $C$, $g(x)$ can have a variety of flavors of roots. For example

1. $C = 0$: The curve crosses the $X$ axis three times.

2. $C = 1$: The curve crosses the $X$ axis once and is tangent to the $X$ axis at one point.

3. $C = 2$: The curve crosses the $X$ axis once.

In the first case, $g(x)$ is said to have 3 simple real roots. In the second case, $g(x)$ is said to have 1 simple real root and 1 double real root. In the third case, $g(x)$ is said to have 1 simple real root and a pair of complex conjugate roots. For the case of a repeated root $r$, the multiplicity of the root is the number of times $(x-r)$ is a factor of $g(x)$. Equivalently, that is 1 plus the order of the highest derivative of $g(x)$ for which $(x-r)$ is a root.

If we allow multiple roots and complex roots, then we can state the fundamental theorem of algebra.

**Theorem 1 (The Fundamental Theorem of Algebra).**
*Every polynomial of degree $n$ has $n$ roots.*

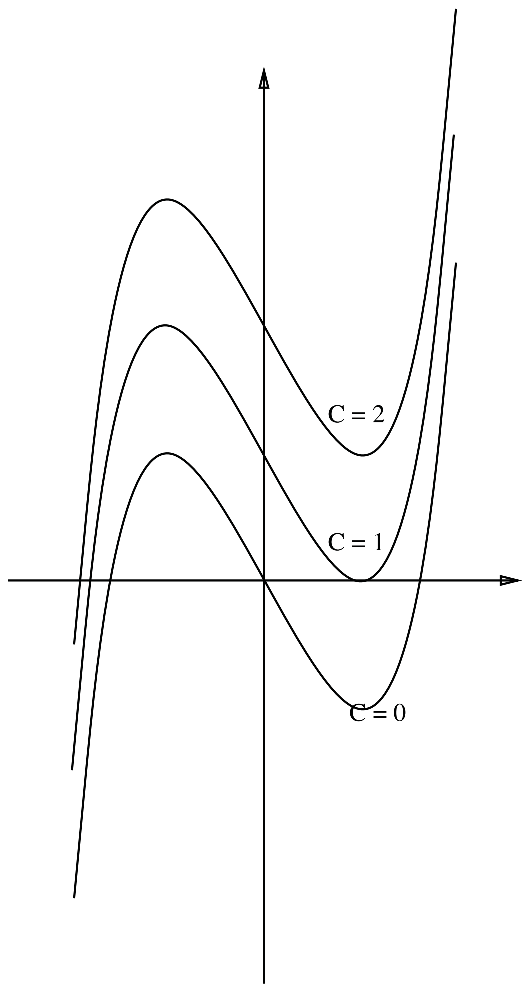


Figure 1: $g(x) = \frac{16}{3}x^3 - \frac{10}{3}x + C$.

# 2 Intersections of two polynomials

Consider intersections between linear functions and the parabola $Y = X^2$. First let the linear functions be horizontal lines, $Y = C$. Then, there are various types of intersections, as seen in Figure 2.

1. $C > 0$: Two distinct intersections.
2. $C = 0$: One repeated intersection of multiplicity two.
3. $C < 0$: Two complex conjugate intersections.

If we now choose a point on the parabola and pass a line through the point, then gradually rotate this line, we see a new family of intersections. As in Figure 3, these seem to be first

*Revised from the 1992 notes by Nilay Banker

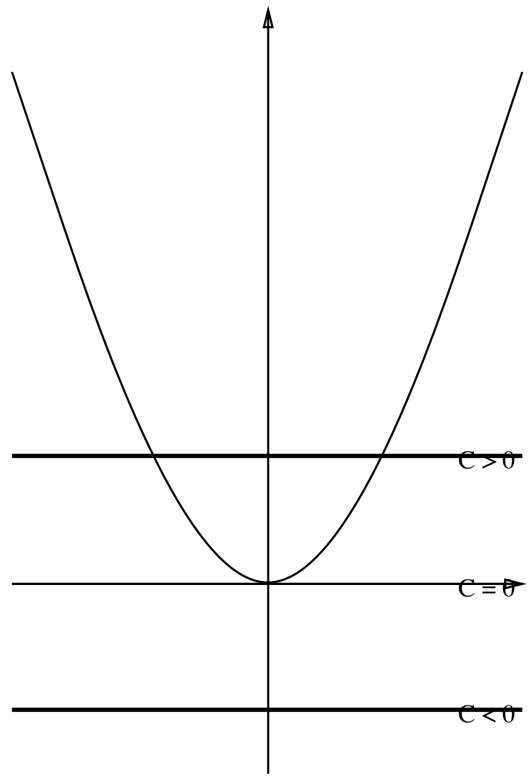


Figure 2: Intersections between horizontal lines and the curve $Y = X^2$.

two distinct real intersections, then a repeated intersection, then one real intersection. In the latter case if we move to the projective plane, we obtain an additional intersection at a point at infinity.

To see this, we homogenize the equations and get $wy = x^2$ and $x = 2w$. Then

$$wy = (2w)^2 = 4w^2$$

$$wy - 4w^2 = 0$$

$$w(y - 4w) = 0$$

From which we obtain the solutions $y = 4w$ and $w = 0$. These correspond to a finite solution and a solution at infinity. The finite solution is $(1;2,4)$ and the solution at infinity is $(0;0,1)$.

In general, for intersections, if we count multiple intersections, complex intersections and intersections at infinity, we have *Bézout's Theorem.*

**Theorem 2 (Bézout's Theorem).**

*A curve of degree $n$ and a curve of degree $m$ intersect at exactly $mn$ points.*

**Example 1 (Intersections between a hyperbola and straight lines).**

*In Figure 4 we see the possible types of intersections between a hyperbola and straight*

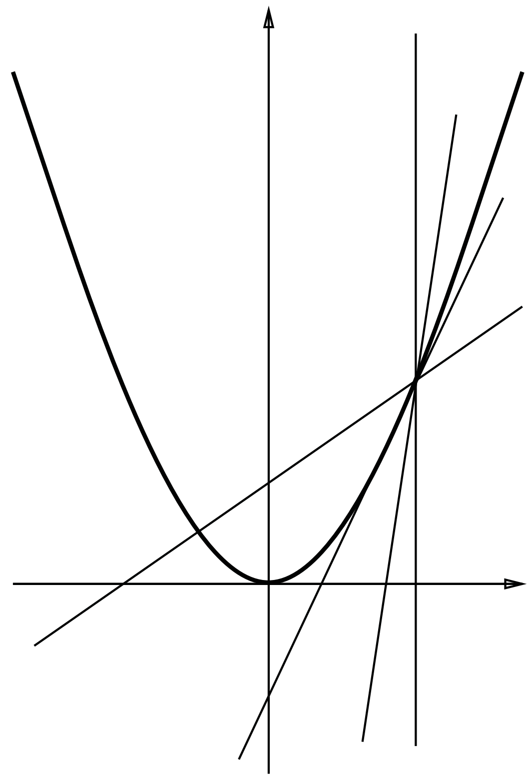

Figure 3: Intersections at a point on $Y = X^2$.

*lines. By Bézout's Theorem we expect two intersections since the hyperbola has degree $m = 2$ and the straight lines have degree $n = 1$. Here are the cases that arise:*

1. *Two distinct real roots.*
2. *A repeated real root.*
3. *One real root, one root at vertical infinity.*
4. *One real root, one root at horizontal infinity.*
5. *A repeated root at vertical infinity.*
6. *A repeated root at horizontal infinity.*

# 3 Interpolation of polynomials

Quite often the polynomial $f(x)$ may be given as a table of values instead of as a closed form expression. In such a case the value of the function $f(x)$ is tabulated at a discrete set of values

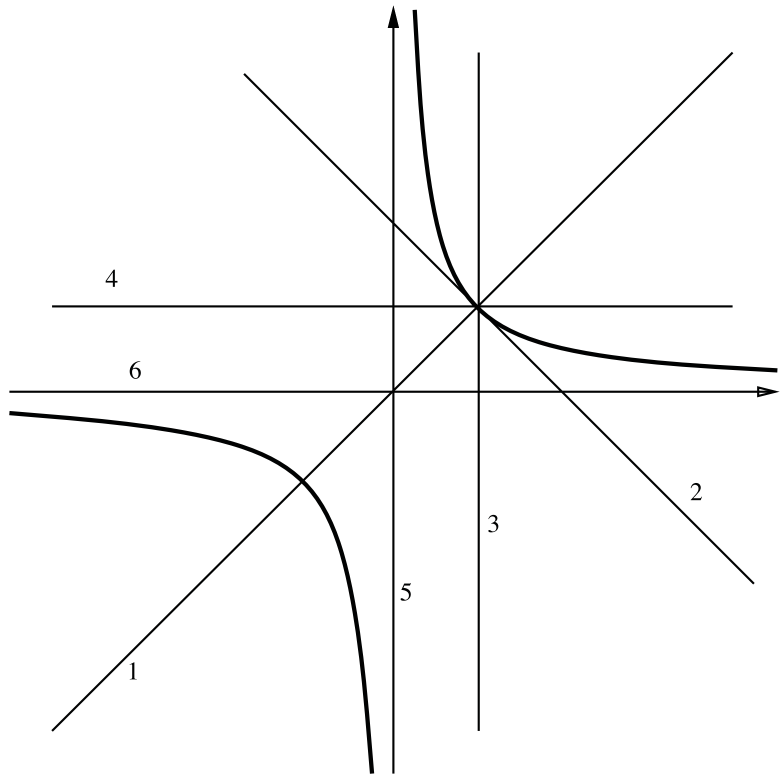


Figure 4: Intersections between a hyperbola and straight lines

of the argument $x$ as shown below

$$\begin{array}{cccc} x_0 & x_1 & \dots & x_n \\ \vdots & \vdots & & \vdots \\ f(x_0) & f(x_1) & \dots & f(x_n) \end{array}$$

If we wish to find the value of $f(x)$ at some point $y$ which is not one of the the tabulated points, then the value is computed by a process called *interpolation*. To interpolate a polynomial at all points we need to determine its coefficients, given the above table of values.

## 3.1 Interpolation and the Vandermonde determinant

The coefficients $a_i$ of the $n$-th degree polynomial which passes through $n+1$ specified points $(x_i, y_i)$ is given by the solution of the following system of $n+1$ linear equations:

$$\begin{aligned} y_0 &= a_0x_0^n + a_1x_0^{n-1} + \dots + a_{n-1}x_0 + a_n \\ y_1 &= a_0x_1^n + a_1x_1^{n-1} + \dots + a_{n-1}x_1 + a_n \\ &\vdots \\ y_n &= a_0x_n^n + a_1x_n^{n-1} + \dots + a_{n-1}x_n + a_n. \end{aligned}$$

Note than in this system, the $a_i$ are the unknowns and the $x_i, y_i$ are the given data.

To recover the polynomial $f(x)$ we compute the determinant of the system's coefficients, namely:

$$V(x_0, x_1, \ldots, x_n) = \begin{vmatrix} x_n^n & x_n^{n-1} & \ldots & x_n & 1 \\ x_{n-1}^n & x_{n-1}^{n-1} & \ldots & x_{n-1} & 1 \\ & & \vdots & & \\ x_0^n & x_0^{n-1} & \ldots & x_0 & 1 \end{vmatrix}.$$

The determinant of this system is known as the *Vandermonde determinant*. Our system will have a solution if and only if this determinant is nonzero.

Consider evaluating $V$ by substituting $\zeta$ for $x_n$, as follows:

$$V(x_0, x_1, \ldots, x_{n-1}, \zeta) = \begin{vmatrix} \zeta^n & \zeta^{n-1} & \ldots & \zeta & 1 \\ x_{n-1}^n & x_{n-1}^{n-1} & \ldots & x_{n-1} & 1 \\ & & \vdots & & \\ x_0^n & x_0^{n-1} & \ldots & x_0 & 1 \end{vmatrix}.$$

This is a polynomial in $\zeta$ of degree $n$. Clearly if $\zeta$ is set equal to $x_i$ for $i \neq n$, two rows of $V$ will become equal, and hence the determinant will vanish. Since $x_0, \ldots, x_{n-1}$ are all the roots of $V(\zeta)$, we conclude that we can express $V$ as:

$$V(x_0, x_1, \ldots, x_{n-1}, \zeta) = \alpha \prod_{j=0}^{n-1} (\zeta - x_j).$$

We observe that $\alpha$ is the coefficient of the $\zeta^n$ term which is given by the cofactor of the $\zeta^n$ term in $V$. This gives the following recursion, after replacing $\zeta$ with $x_n$:

$$V(x_0, x_1, \ldots, x_{n-1}, x_n) = V(x_0, x_1, \ldots, x_{n-1}) \prod_{j=0}^{n-1} (x_n - x_j),$$

which simplifies to

$$V(x_0, x_1, \ldots, x_{n-1}, x_n) = \prod_{i=0}^{n} \prod_{j=0}^{i-1} (x_i - x_j).$$

This shows that if the the $x_i$'s are distinct, the determinant is non-zero—and hence a unique solution vector of $a_i$'s can be found.

In fact, the corresponding polynomial expression $y - f(x)$ can itself be written as the determinant

$$\begin{vmatrix} x^n & x^{n-1} & \ldots & x & 1 & y \\ x_n^n & x_n^{n-1} & \ldots & x_n & 1 & y_n \\ x_{n-1}^n & x_{n-1}^{n-1} & \ldots & x_{n-1} & 1 & y_{n-1} \\ & & \vdots & & & \\ x_0^n & x_0^{n-1} & \ldots & x_0 & 1 & y_1 \end{vmatrix}.$$

as it is easy to check. However, this is is quite a complex method as the determinant is large and so does not represent a useful way to compute numerically a solution.

## 3.2 Lagrange interpolation

Rather than generate the interpolating polynomial in one step, Lagrange interpolation fits each point $(x_i, y_i)$ with an degree $n$ polynomial which is constructed in such a way as to be "orthogonal" to the other interpolating basis polynomials of the other points. We construct such a polynomial according to the following rule:

$$L_i(x_j) = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{otherwise.} \end{cases}$$

To see how we construct $L_0(x)$, we note $x_1, x_2, \ldots, x_n$ are roots and hence we have

$$L_0(x) = \frac{(x - x_1)(x - x_2)\ldots(x - x_n)}{(x_0 - x_1)(x_0 - x_2)\ldots(x_0 - x_n)},$$

where the denominator normalizes the function for $x = x_0$, that is, makes $L_0(x_0) = 1$.

The resulting Lagrange interpolating polynomial is a superposition of the $L_i(x)$ polynomials, weighted by the $y_i$ ordinates:

$$P(x) = \prod_{i=0}^{n} y_i L_i(x).$$

## 3.3 Hermite interpolation

Hermite interpolation is an extension of the polynomial interpolation problem to include constraints on the derivatives of the function at each point. In general we are given the following constraints

$$\begin{array}{llll} P(x_0) = y_0^{(0)}, & P^{(1)}(x_0) = y_0^{(1)}, & \ldots, & P^{(m_0)}(x_0) = y_0^{(m_0)} \\ \vdots & & & \\ P(x_k) = y_k^{(0)}, & P^{(1)}(x_k) = y_k^{(1)}, & \ldots, & P^{(m_k)}(x_k) = y_k^{(m_k)} \end{array}$$

These equations represent $\sum_{i=0}^{k}(m_i + 1) = n + 1$ constraints.

Note that a contiguous sequence of derivatives must be specified at each $x_i$, starting with the 0th derivative and going up. Trying to generalize the interpolation problem further by specifying arbitrary derivatives at arbitrary points quickly leads to problems. Beach (pp. 37) shows, for example, that specifying a quadratic $P$ by giving $P(x_1)$, $P(x_2)$, and the first derivative $P'(\frac{x_1+x_2}{2})$ leads to an underdetermined or inconsistent specification, since we have

$$P'\left(\frac{x_1 + x_2}{2}\right) = \frac{P(x_2) - P(x_1)}{x_2 - x_1}$$

for any quadratic $P$.

# 4 Elimination

We are given a set of polynomials. Under what conditions do they have a common root? This is the fundamental question that elimination theory tries to answer.

Given polynomials $f(x)$ and $g(x)$ in an arbitrary field $K$

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$
$$g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_m,$$

we wish to find the necessary and sufficient conditions for these equations to have a nonconstant common factor $\phi(x)$. We shall not exclude the possibility that $a_0 = 0$ or $b_0 = 0$, i.e. the degree of $f(x)$ or $g(x)$ is less than $n$ or $m$ respectively. If the polynomial $f(x)$ is written as above then $a_0$ is called the leading coefficient and $n$ the formal degree of the polynomial.

Initially we assume that $a_0 \neq 0$ and $b_0 \neq 0$. Under this assumption we shall first show that $f(x)$ and $g(x)$ have a nonconstant common divisor $\phi(x)$ if and only if an equation of the form $h(x)f(x) = k(x)g(x)$ holds, where $h(x)$ is at most of degree $m-1$ and $k(x)$ is at most of degree $n-1$, and where both polynomials $h, k$ do not vanish identically. All prime factors of $f(x)$ must divide the right side of the above equation just as often as $f(x)$. Yet they cannot divide $k(x)$ as often as they do $f(x)$, for $k(x)$ is at most of degree $n-1$. Hence at least one prime factor of $f(x)$ occurs also in $g(x)$.

If conversely $\phi(x)$ is a nonconstant common factor $f(x)$ and $g(x)$, it is merely necessary to put

$$\begin{aligned} f(x) &= \phi(x)k(x), \\ g(x) &= \phi(x)h(x), \end{aligned}$$

and the equation above will be satisfied.

In order to investigate the above equation further, we set

$$h(x) = c_0x^{m-1} + c_1x^{m-2} + \ldots + c_{m-1},$$
$$k(x) = d_0x^{n-1} + d_1x^{n-2} + \ldots + d_{n-1}.$$

Substituting in $h(x)f(x) = k(x)g(x)$ and equating coefficients gives the following set of equations

$$\begin{aligned} c_0a_0 &= d_0b_0 \\ c_0a_1 + c_1a_0 &= d_0b_1 + d_1b_0 \\ \vdots \qquad & \qquad \vdots \\ c_{m-1}a_n &= d_{n-1}b_m \end{aligned}$$

This is a homogeneous set of $n+m$ linear equations in the $c_i$ and $d_i$. A nontrivial solution exists only if the following determinant is zero.

$$R(f,g) = \begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_n & & & \\ & a_0 & a_1 & \dots & a_{n-1} & a_n & & \\ & & & \vdots & & & & \\ & & & & a_0 & a_1 & \dots & a_n \\ b_0 & b_1 & b_2 & \dots & b_m & & & \\ & b_0 & b_1 & \dots & b_{m-1} & b_m & & \\ & & & \vdots & & & & \\ & & & & b_0 & b_1 & \dots & b_m \end{vmatrix}$$

where there are $m$ rows of $a_i$ coefficients and $n$ rows of $b_i$ coefficients. The determinant $R(f,g)$ is called the *resultant* and can be expressed as

$$R(f,g) = a_0^m b_0^n \prod_i \prod_j (x_i - y_j)$$

where the $x_i$ are the roots of $f(x)$ and the $y_j$ are the roots of $g(x)$. This shows that the resultant will be zero precisely when a common root exists.

# 5 Issues in Modeling a Shape

The main goal in *Computer-Aided Geometric Design* (CAGD) is to generate smooth shapes. Our goal is to use systems of polynomial equations to model smooth curves in the plane and smooth surfaces in 3-space. Hence we address a number of questions concerning the choices which need to be made before proceeding with the modeling of a curve or surface.

## 5.1 One Piece or Many Pieces ?

Single polynomial or rational functions can be used to represent smooth shapes; however in applications where *local control* over the curves is desired it is not possible to represent the shapes using a single polynomial or rational. This is so because if two two such functions agree on a small arc, they will have to be the same everywhere. In CAGD we don't have single polynomial equations but we use a collection of arcs from different polynomial or rational curves. The pieces are then joined together smoothly by the imposition of certain constraints on the systems of equations. The shapes constructed using this method are in general called **splines**, i.e. piecewise models of shapes in which the pieces have been constrained to be joined with a certain degree of continuity.

## 5.2 Parametric or Implicit Representations

Consider our curved shape $S$ and let $Q$ be some larger, flat space in which $S$ lies. We will call $Q$ the *object space*. Let $s$ and $q$ denote the dimensions of $S$ and $Q$, respectively. Our system of equations mathematically specifies a function which describes $S$. Two standard ways in which to do this are through the use of *parametric models*, or *implicit models*. The former are characterized by the fact that the modeling function maps from some other space into $Q$, while in the later the function maps from $Q$ into some other space.

In *parametric models*, the shape $S$ is defined as the range $F(P)$ of a function $F : P \to Q$, where $P$ is an auxiliary, $s$-dimensional space called the *parameter space*. Thus, for a parametric model, the input to the modeling function is a point in parameter space—that is, values for the $s$ parameters—and the output of the modeling function is a point on the shape $S$.

In *implicit models*, on the other hand, the shape $S$ is defined by the formula $S = F^{-1}(\langle 0,\ldots,0\rangle)$, where $F : Q \to R$ is a function from the object space $Q$ to an auxiliary space $R$ of dimension $q-s$. The auxiliary space $R$ is sometimes called the *gauge space*. For an implicit model, the input to the modeling function is a point in the object space, while the output of the modeling function is a point in gauge space—given by the $q-s$ "coordinates" of the gauge value.

To make these ideas more concrete, consider as an example the standard parabola $y = x^2$. This shape is a smooth curve in the plane $Q = R^2$, so $s = 1$ and $q = 2$.

The function $G : R \to R^2$ given by $G(t) = \langle t, t^2\rangle$ is a parametric model for that parabola. The variable $t$ here denotes a parameter value. The parameter space $P = R$ of a curve is one-dimensional, so it is convenient to think of the parameter as time. The modeling function $G : P \to Q$ maps times to points on the curve $S$.

The function $H : R^2 \to R$ given by $H(z) = H(\langle x,y\rangle) = y - x^2$, where $z$ denotes a point in the object plane, is an implicit model for the same parabola. The corresponding gauge value $H(z)$ is positive, zero, or negative according as the point $z$ lies above, on, or below the curve $S$.

Note that both parametric and implicit models of shapes encode some extra information, over and above the shape $S$ itself. A parametric model, in addition to determining the shape $S$, also provides a 'roadmap' of $S$—a way to 'name' each point of $S$ easily. It is natural to use the parametric form to look at select pieces of a shape, because one then only needs to specify the corresponding pieces of the parameter space $P$ and look at the corresponding images under $G$. On the other hand, an implicit model, in addition to determining $S$, also associates gauge values with all of the other points in the object space.

## 5.3 What functions do we use for basic shapes ?

Whether we are modeling our shape $S$ parametrically or implicitly, we need to decide which class of functions we will use.

The simplest and most well understood class of functions is the class of *polynomial functions*. These are functions $F$ such that each (Cartesian) coordinate of the output point $F(p)$ is a polynomial function of the (Cartesian) coordinates of the input point $p$. Notice that both the parametric model $G(t) = \langle t, t^2\rangle$ and the implicit model $H(\langle x,y\rangle) = y - x^2$ used above are examples of polynomial functions. Polynomial functions have the computational advantage that they can be built up using just the operations of addition, subtraction, and multiplication.

If we also allow division to be a legal operation, we get a larger class of functions called the *rational functions*. These are functions for which each coordinate of an output point may be written as the ratio of two polynomials in the coordinates of the input point. As an example, note that the unit circle cannot be parameterized by polynomials. However, the rational function

$$t \mapsto \left\langle \frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right\rangle$$

from $R$ to $R^2$ does the job.

The polynomial and rational functions are what we will use in this course.

By allowing the operation of solving polynomial equations of any degree we arrive at the class of functions called the *algebraic functions*. If we then also allow the operation of summing an absolutely convergent, infinite series, we get the *real analytic functions*. This class of functions allows, for instance, $\sin(t), \cos(t)$, etc.

## 5.4 What degree of Polynomials/Rational functions?

Having decided to remain within the classes of polynomial and rational functions, we need to consider the effect choosing a degree. (When we say that a polynomial $f(t) = a_0 t^n + \ldots + a_{n-1}t + a_n$ is of degree $n$ (where $a_0$ through $a_n$ are real coefficients) we do *not* require that $a_0 \neq 0$. In other words, the function $f$ is of degree *at most n*.)

Using a function of a high degree will mean gaining more flexibility over functions of lower degrees, but it will also mean an increase in ‘wiggles’ and instabilities. Splines allow a tradeoff between number of pieces and the complexity of each piece. A designer may use a large number of simple pieces or a few pieces of relatively high degree. Whether using polynomial or rational functions, placing a bound on the degrees of the polynomials involved will serve as a way of controlling the complexity of the shape being modeled.

For most applications of CAGD, polynomials of fairly low degree seem to be sufficient. In this class we will work mainly with degrees 1, 2, and 3.