19 Interpolation and splines

How are smooth surfaces drawn (when polygonal approximation fails)?

Aim: understand principles of interpolation and extrapolation.

Reading:

- Foley Sections 9.2 Parametric bubic curves, 9.2.1 Basic characteristics, 9.2.3 Bézier curves, 9.3 Parametric bicubic surfaces, 9.3.1 Hermite surfaces, 9.3.2 Bézier surfaces and 9.3.4 Normals to surfaces, 9.4 Quadric surfaces.

Further reading:

- Heath's lecture 7 on interpolation, (see http://www.cse.uiuc.edu/heath/scicomp/notes/) from Scientific Computing: An Introductory Survey, Second Edition by Michael T. Heath, McGraw-Hill 2002; ISBN 0-07-239910-4.

Introduction to interpolation

Say know value of a function $f(x)$ at a set of points $x_1, x_2, \ldots, x_n$ $(x_1 < x_2 < \ldots < x_n)$ but $f(x)$ not known in analytic form.

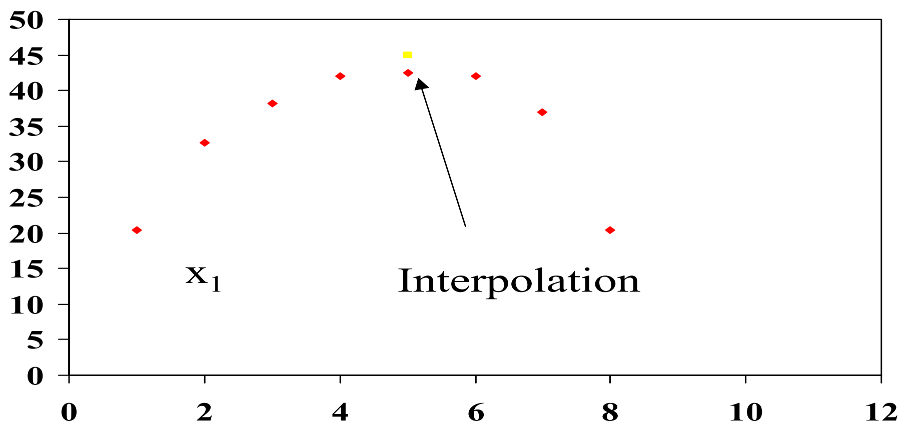
i.e. don't know equation for $f(x)$.

e.g. may get $f(x_i)$ values from a physical measurement in an experiment, or from a long, complicated calculation.

How do we estimate $f(x)$ for arbitrary x?

If desired x within range of x_i then use **interpolation**

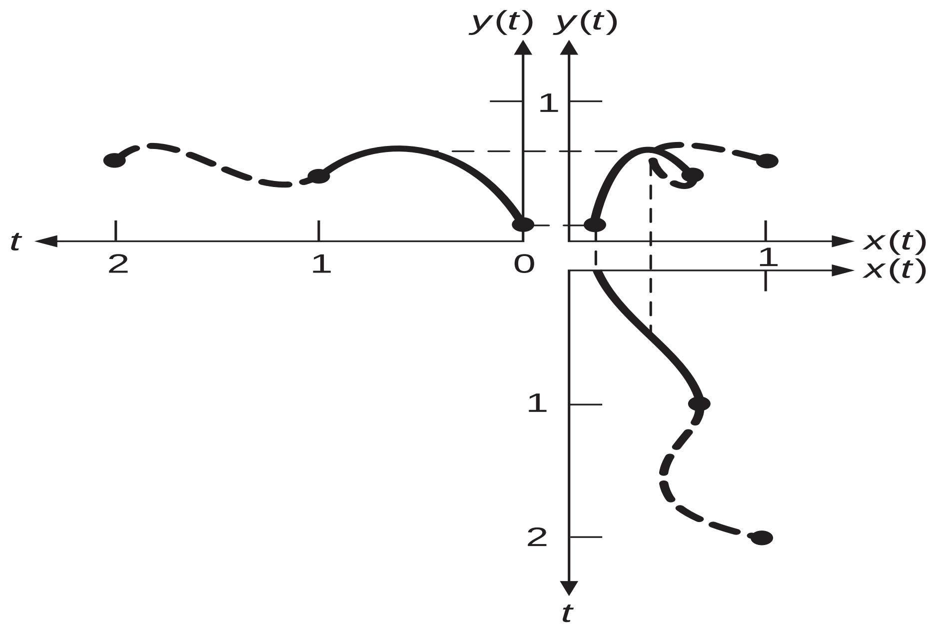
If desired x outside range of x_i then use **extrapolation**



Most common functional form used are **polynomials**.

Also use

- **rational functions** (quotients of polynomials) and
- **trigonometric functions** (sines, cosines etc.)
- as well as others, e.g. $\frac{ax^2+bx+c}{dx+c}$



Two joined 2D parametric curve segments (Foley Figure 9.7).

Parametric curve

$$x(t) = a_x t^3 + b_x t^2 + c_x t + d_x$$
$$y(t) = a_y t^3 + b_y t^2 + c_y t + d_y$$
$$z(t) = a_z t^3 + b_z t^2 + c_x t + d_z$$

where $0 \leq t \leq 1$

$$C = \begin{bmatrix} a_x & b_x & c_x & d_x \\ a_y & b_y & c_y & d_y \\ a_z & b_z & c_z & d_z \end{bmatrix}$$

Can rewrite parametric curve as

$$Q(t) = \begin{bmatrix} x(t) & y(t) & z(t) \end{bmatrix}^T = C \cdot T$$

Tangent vectors

The derivative of $Q(t)$ is the parametric tangent vector of the curve.

$$x(t) = a_x t^3 + b_x t^2 + c_x t + d_x$$
$$y(t) = a_y t^3 + b_y t^2 + c_y t + d_y$$
$$z(t) = a_z t^3 + b_z t^2 + c_x t + d_z$$

where $0 \leq t \leq 1$

$$\frac{d}{dt} Q(t) = Q'(t) = \begin{bmatrix} \frac{d}{dt}x(t) & \frac{d}{dt}y(t) & \frac{d}{dt}z(t) \end{bmatrix}^T$$

$$= \frac{d}{dt} C \cdot T = C \cdot \begin{bmatrix} 3t^2 & 2t & 1 & 0 \end{bmatrix}^T$$

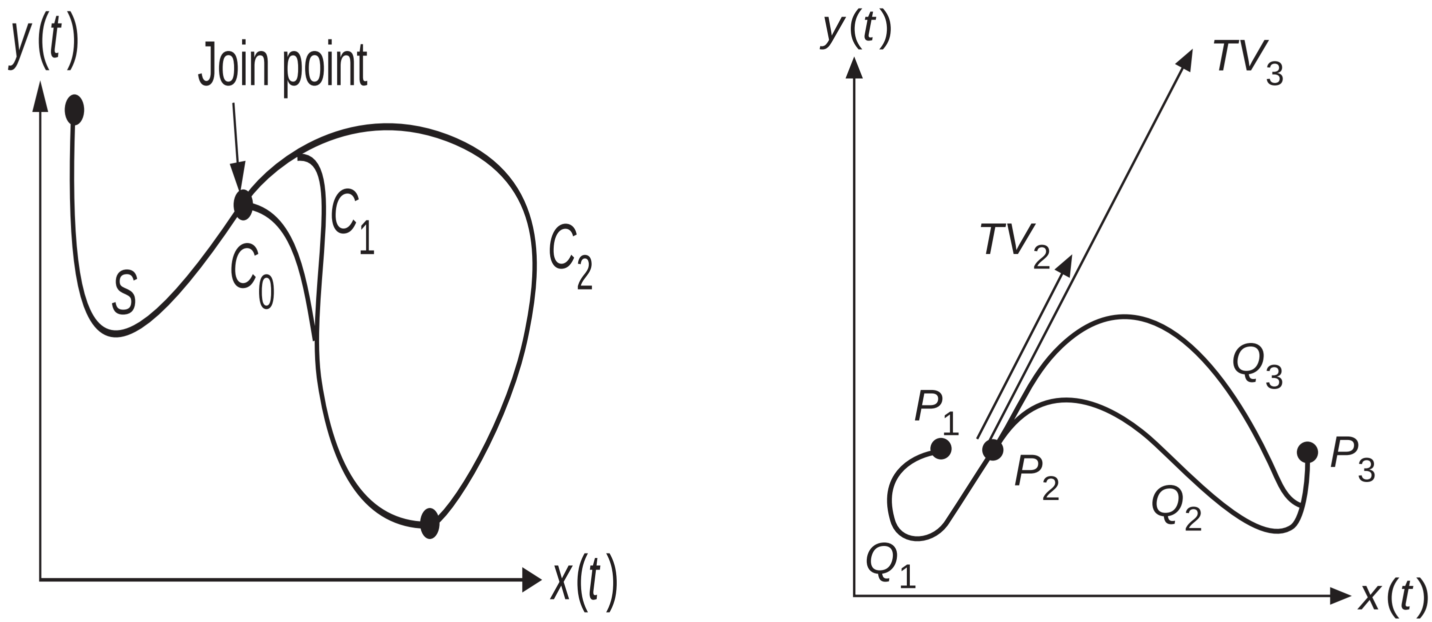
$$= \begin{bmatrix} 3a_x t^2 + 2b_x t + c_x & 3a_y t^2 + 2b_y t + c_y & 3a_z t^2 + 2b_z t + c_z \end{bmatrix}^T$$

Parametric continuity

If the direction and magnitutde of the

$$\frac{d^n}{dt^n}\left[Q(t)\right]$$

through the nth derivative are equal at the join point, the curve is called c^n continuous.



Left: Curve segment S joined to segments C_0, C_1 and C_2.

Right: Curve segments Q_1, Q_2 and Q_3 join at the point P_2 and are identical except for their tangent vectors at P_2 (Foley Figures 9.8 and 9.9).

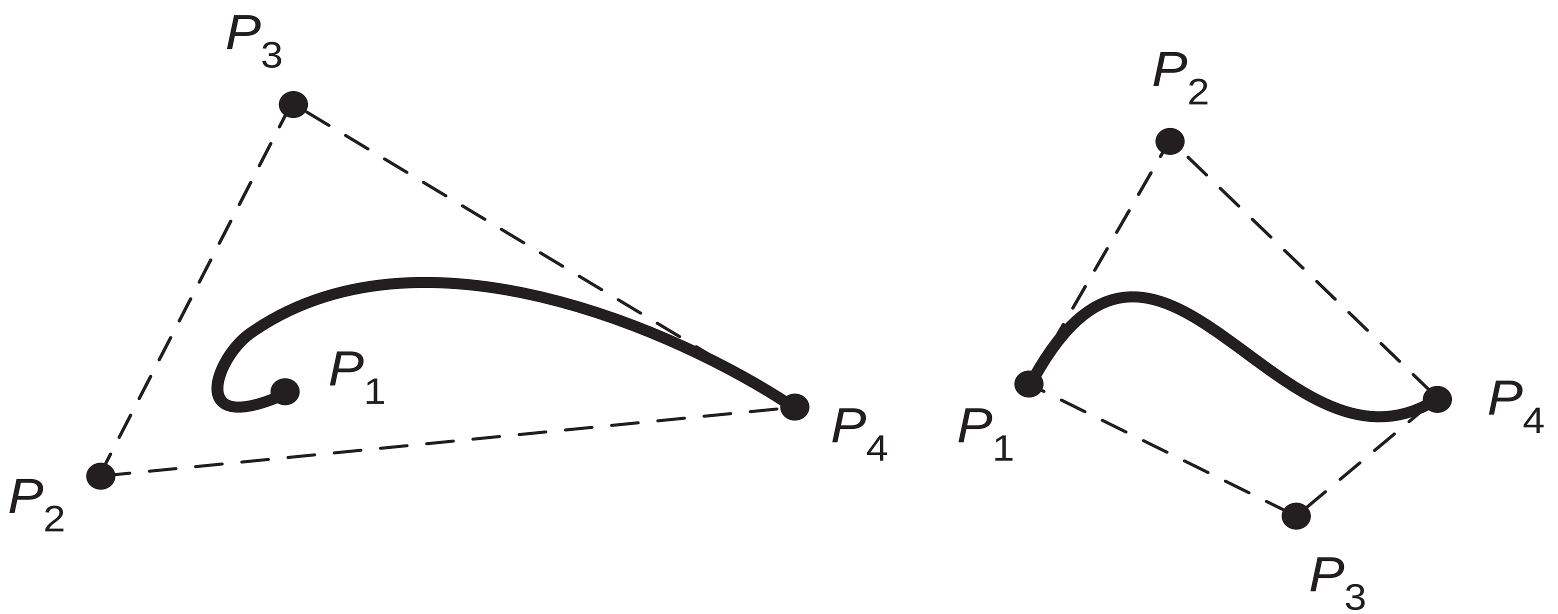
Bézier curves

Bézier curves are defined by the starting and ending vectors of a curve,

P_1P_2 and P_3P_4

and are therefore determined by four *control* points.

19.1 Bézier curves



Two Bézier curves and their control points (Foley Figure 9.15).

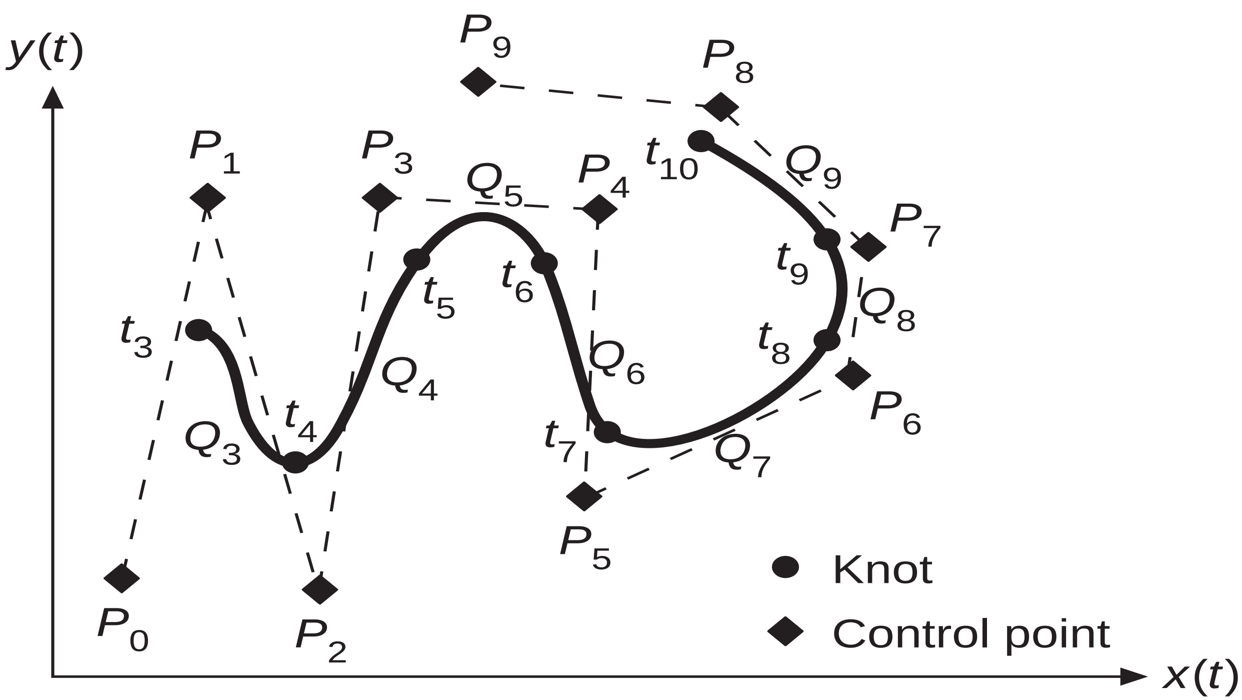
19.2 B-Splines

The term *spline* relates to strips of metal used by draftsmen to lay out surfaces in aeroplanes, cars and ships. These metal splines had second-order continuity.

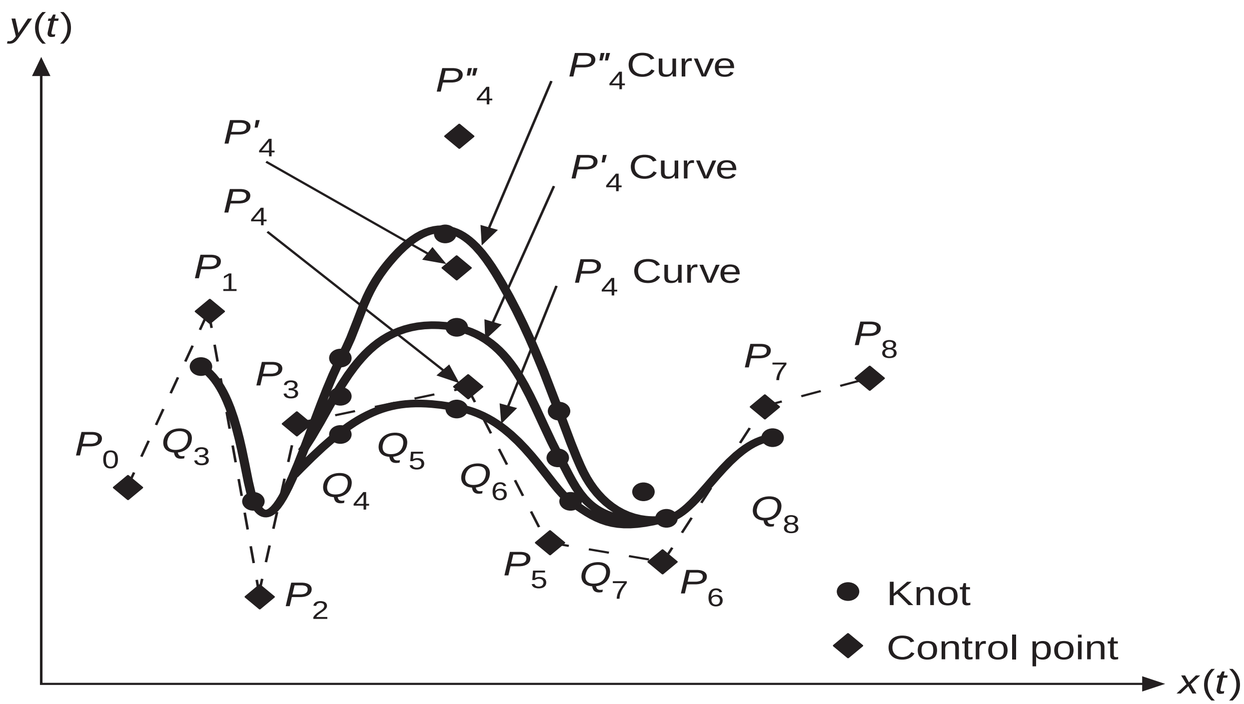
The mathematical equivalent of these are C^0, C^1, and C^2 continuous cubic polynomial that interpolates (passes through) the control points.

Splines have one more degree of continuity that Bézier forms, thus splines are smoother.

19.3 Uniform nonrational B-splines

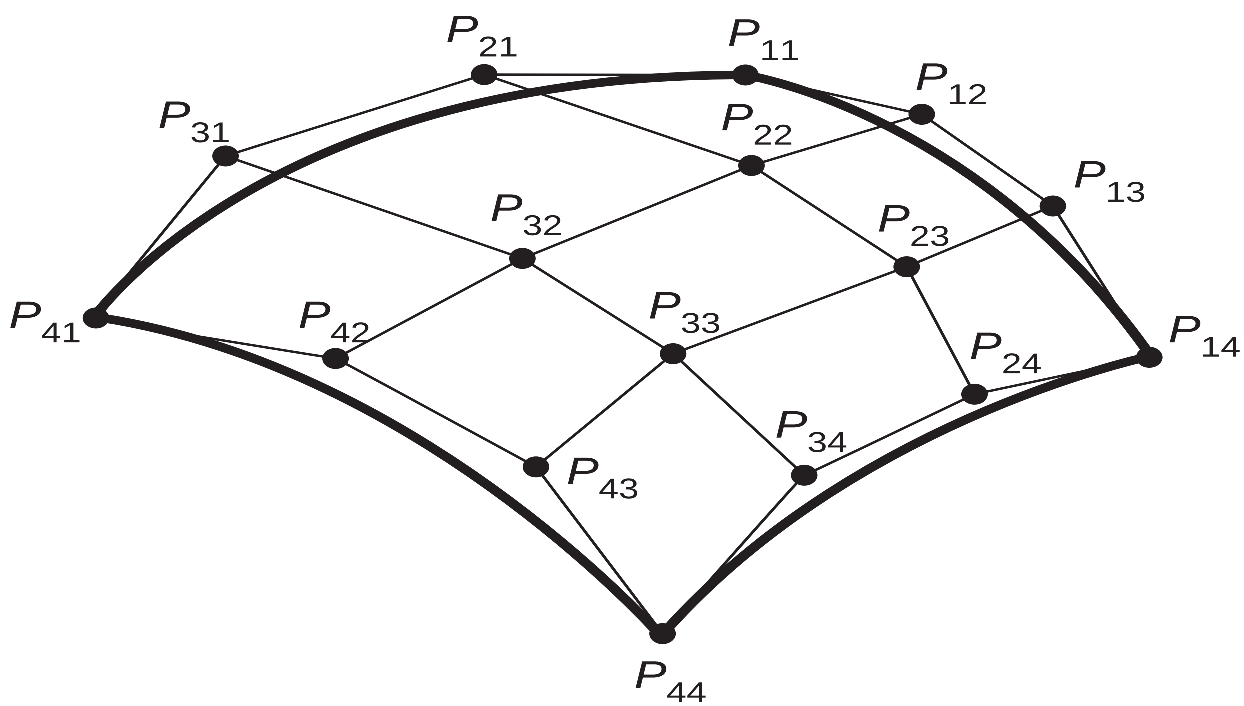


(Foley Figure 9.18)



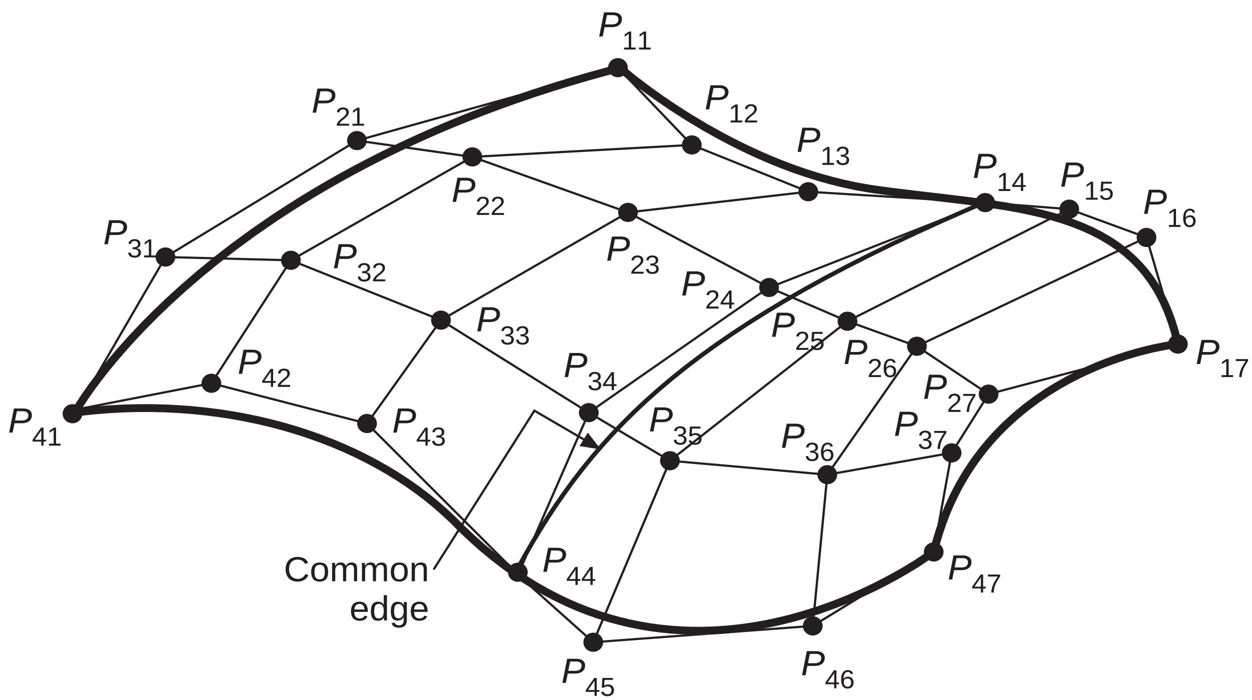
Moving control point P_4 to different positions (Foley Figure 9.19).

19.4 Parametric bicubic surfaces



Sixteen control points of a Bézier bicubic surface (Foley Figure 9.24).

19.5 Parametric bicubic surfaces



(Foley Figure 9.25).

19.6 Computing parametric curves and surfaces

Despite the computational complexity involved in rendering Parametric curves and surfaces (see Foley Section 9.3.5 Displaying bicubic surfaces) they clearly have benefits, including

- the production of very high resolution, photo-realistic images for visualising designs (in the automotive industry bicubic splines have been extensively used for visualising new body shapes).
- the production of high-tolerance surfaces with very few errors or artifacts (useful for the design of industrial components such as engine parts).

Although many of these applications are presently limited either expensive hardware accelerators and/or batch processing before the final results are ready to visualise.

19.7 Parametric in real-time graphics?

What kinds of real-time rendering and shading applications can you think of that might benefit from parametric surfaces?

What kinds of parametric surfaces do you think might be good candidates for real-time computer graphics why?

What advantages might a parametric approach have over a polygonal approach to modelling surfaces in animation?