ECE3340

Numerical Fitting, Regression, Interpolation and Approximation

PROF. HAN Q. LE

Note: PPT file is the main outline of the chapter topic – associated Mathematica file(s) contain details and assignments

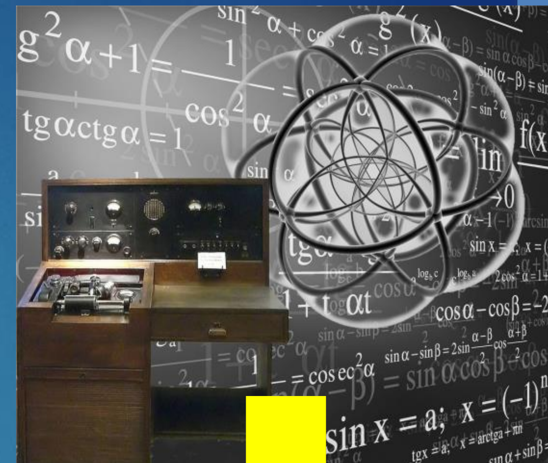
Overview

from data

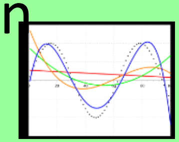
simple model

$f(x_1, x_2 \ldots, t)$

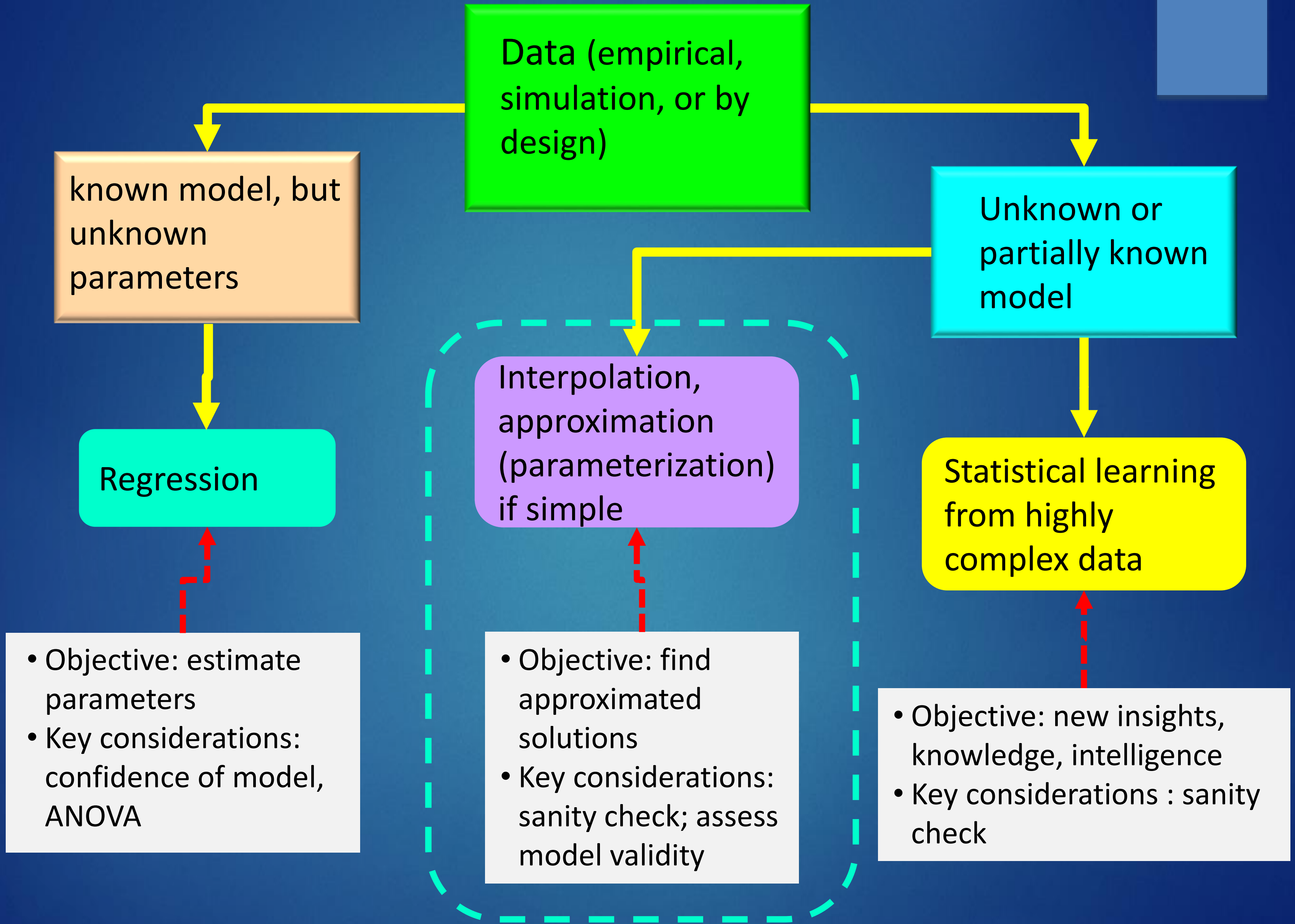
from complex system



simplified, fitting, approximation



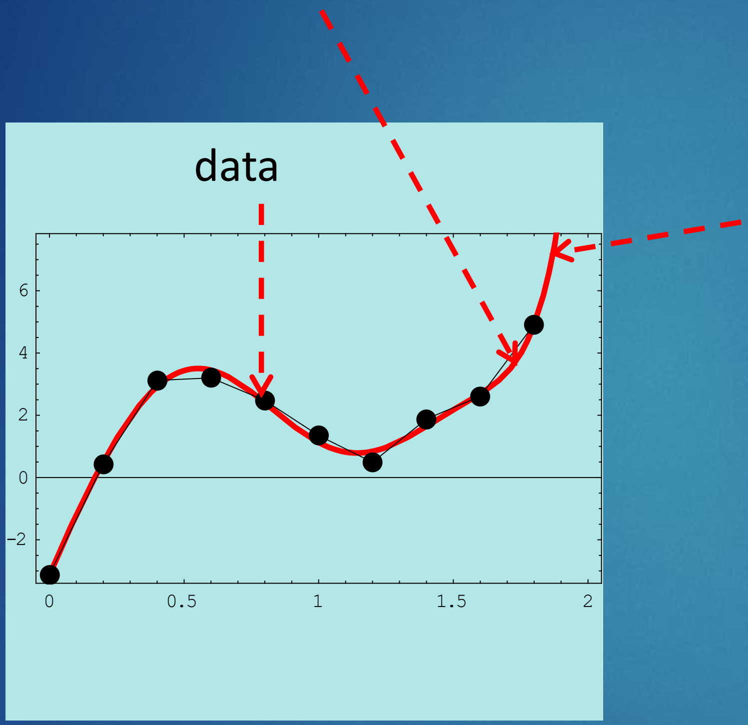
The underlying motivation is to simplify - to make things easy to understand or to implement


Data (empirical, simulation, or by design)
known model, but unknown parameters
Unknown or partially known model
Regression
Interpolation, approximation (parameterization) if simple
Statistical learning from highly complex data
• Objective: estimate parameters
• Key considerations: confidence of model, ANOVA
• Objective: find approximated solutions
• Key considerations: sanity check; assess model validity
• Objective: new insights, knowledge, intelligence
• Key considerations : sanity check

Part B – Fitting, Interpolation, Spline (smoothing)

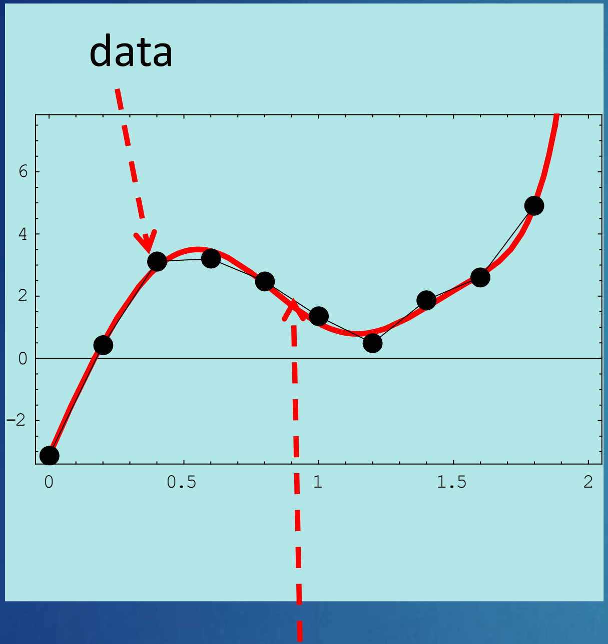
Interpolation/Extrapolation

Estimate the function value at a point not in the data set



- Extrapolation if the point is outside the data range.
- It is never a good idea to extrapolate if there is no information or guiding model what the function should behave outside the known range.
- Use only there is partial knowledge or some confidence level of what can be expected in the extrapolated range. Best when the function vanishes over the extrapolated range

Common methods for interpolation



Interpolation:
- globally (over all points)
- piecewise (by segments)

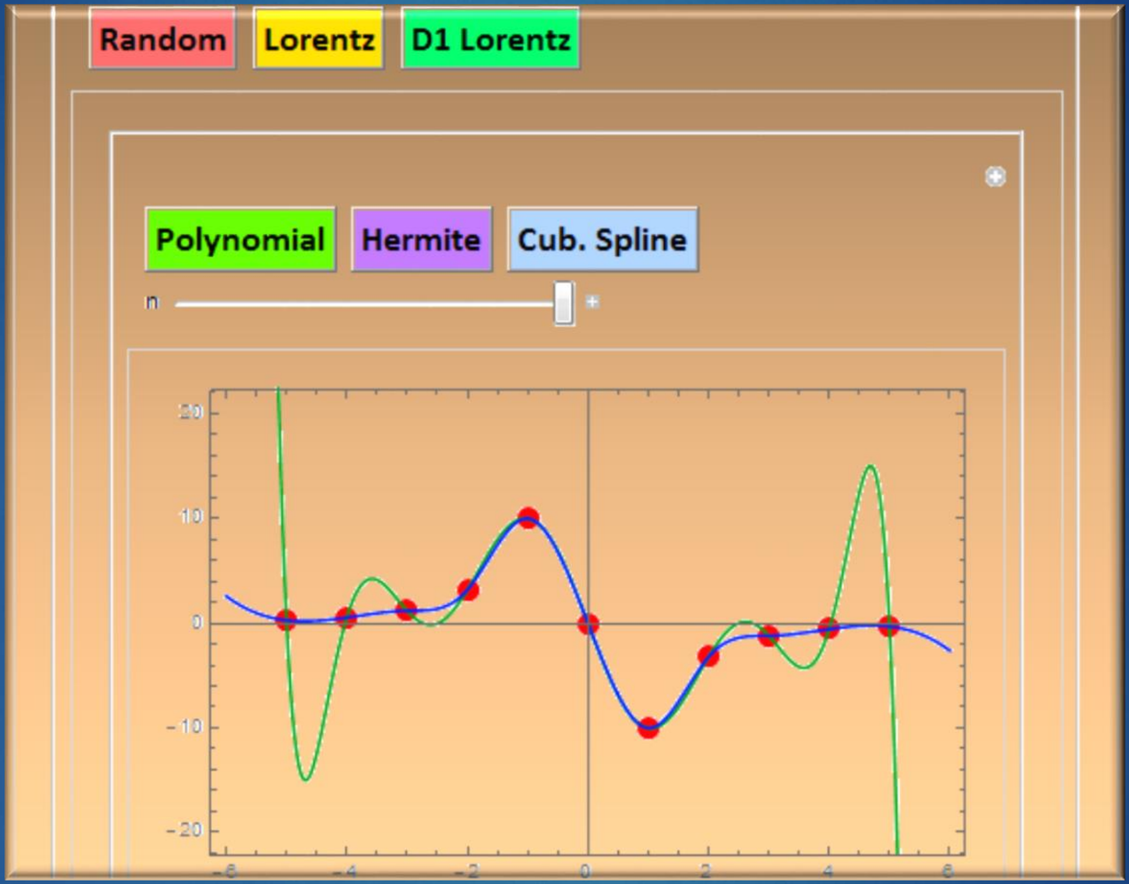
$$y = a_0 + a_1\,x + a_2\,x^2 + \ldots + a_{n-1}\,x^{n-1} + a_n\,x^n$$

Polynomial:
- Newton
- Lagrange
- Chebyshev
- Hermite

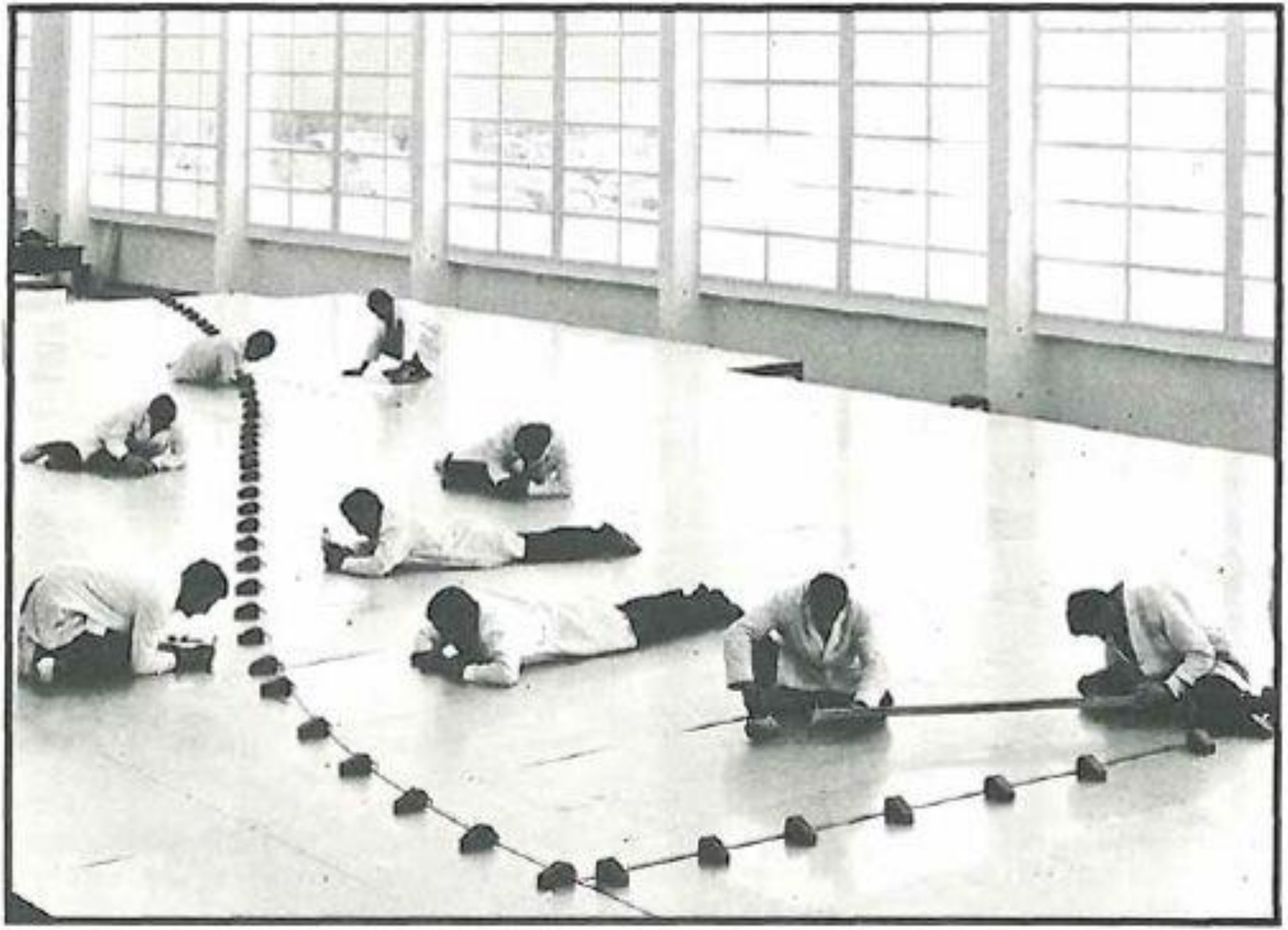
These have limited use because of potential large errors.

Piece-wise adaptive: Spline

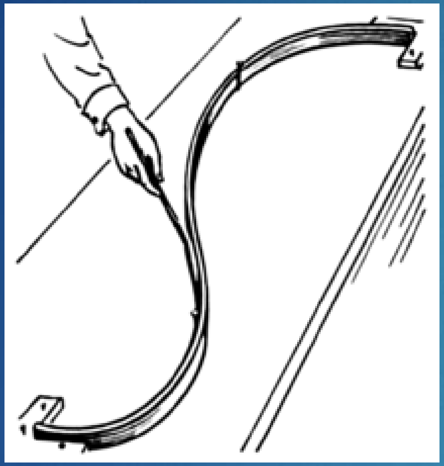
These avoid errors of the polynomial method; grow rapidly and widespread in numerous fields, especially with the emergence of computer graphics

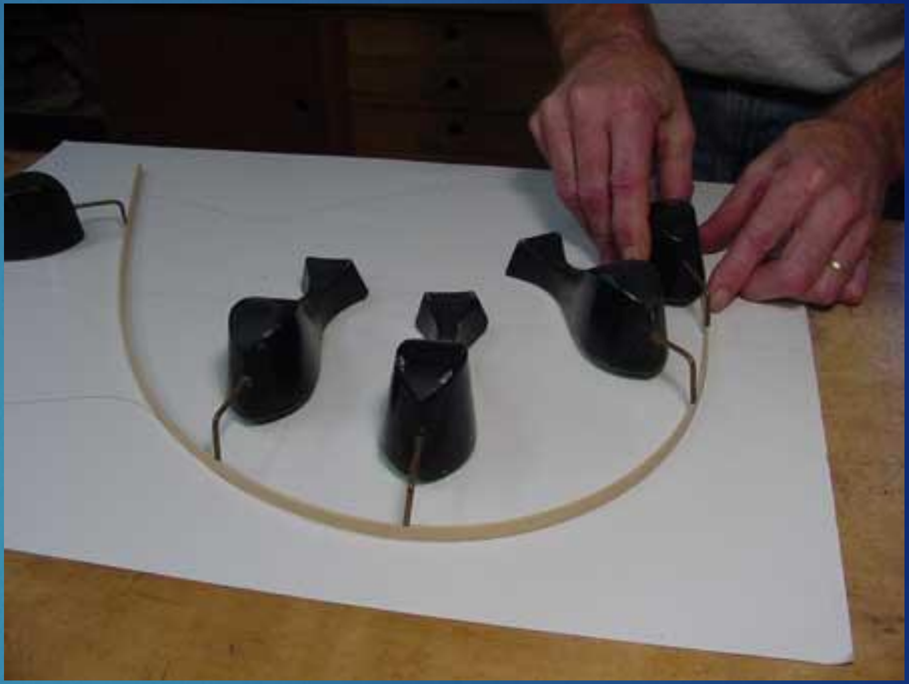

Random
Lorentz
D1 Lorentz
Polynomial
Hermite
Cub. Spline
n
20
10
0
−10
−20
−6
−4
−2
0
2
4
6

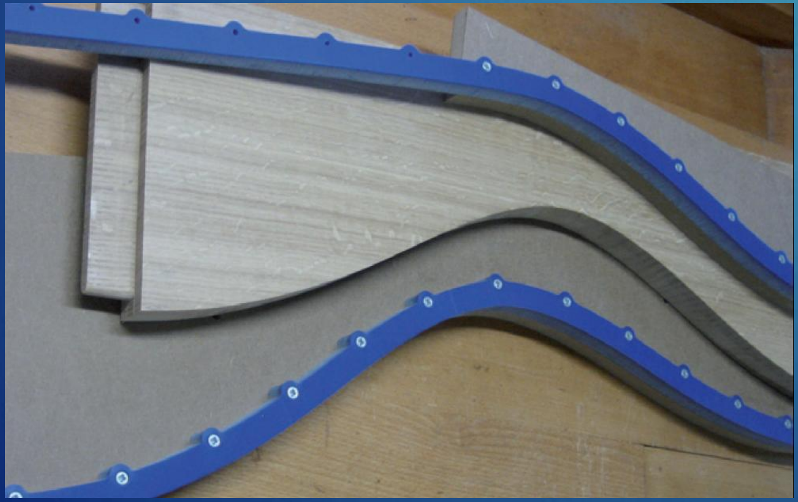
Spline introduction



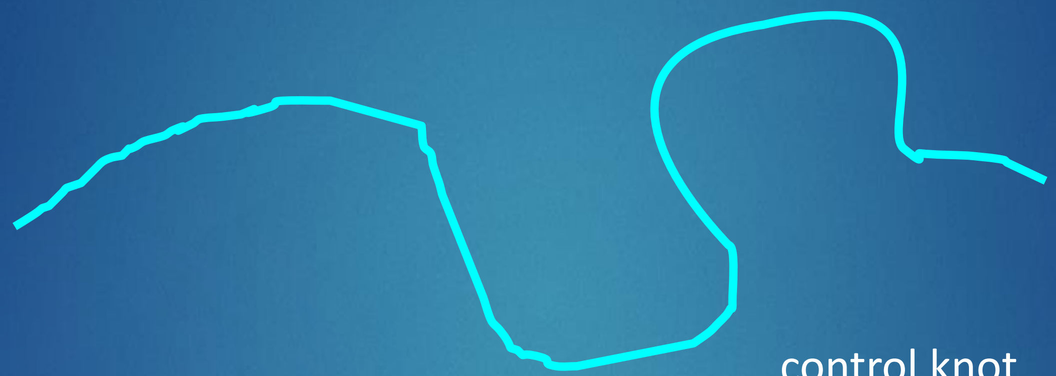
Model building - Spline and smooth curve design



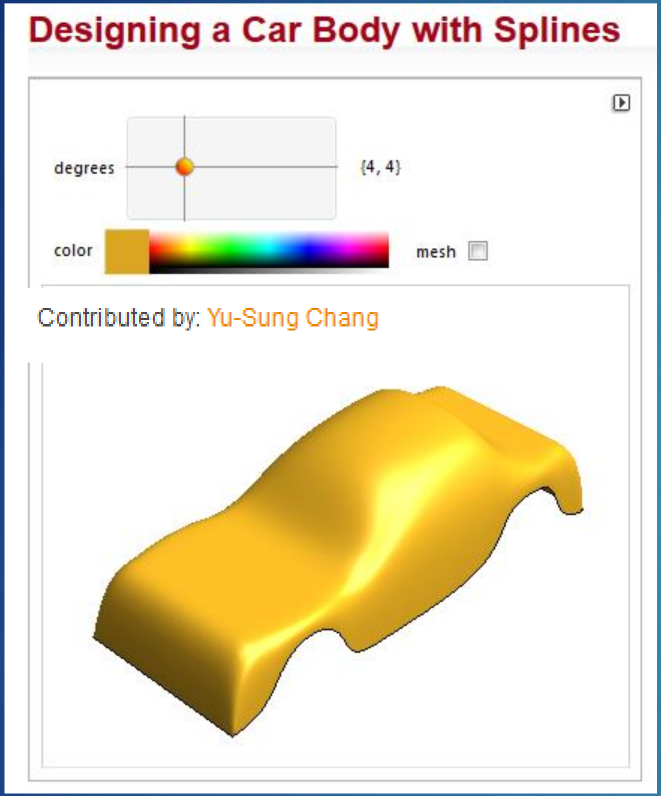


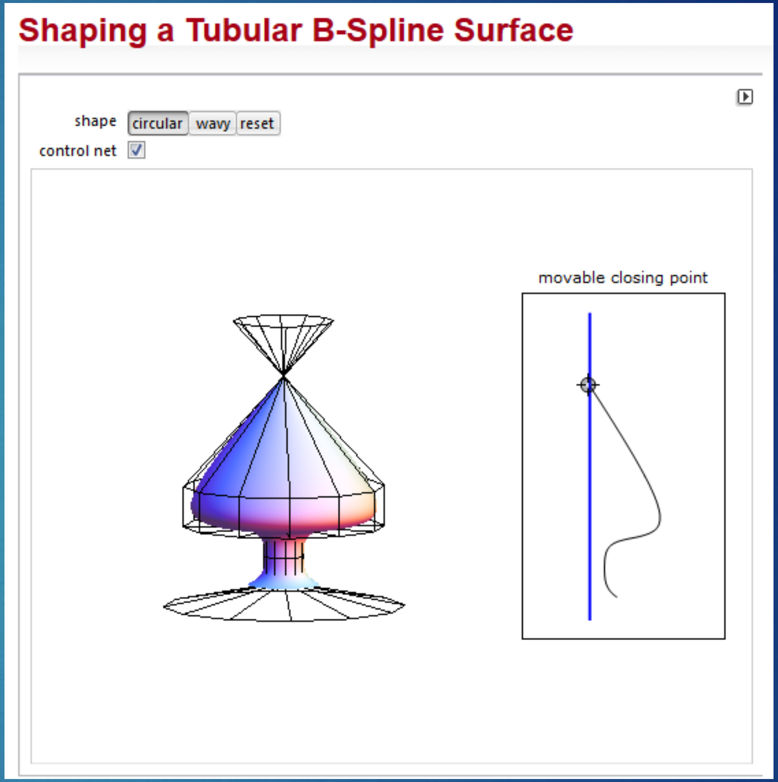


How do we draw with common computer software like ppt?



Spline applications





General concepts

An interpolating approximation that:

- efficient and avoid the error and deficiency of the polynomial methods (especially to avoid high power-order oscillatory behavior)
- inspired by the old tried-and-true draftsman technique of spline that is known to be useful. (controllable derivatives)

Key ideas:

- piecewise interpolation: each segment between “knots” (data points) is approximated independently from knots far away
- a constraint to limit the oscillatory behavior: roughness penalty forces the fit to minimize the highly oscillatory or rough behavior of the approximation:

$$S[f] = \sum_{i=0}^{n} (y_i - f[x_i])^2 + \alpha \int_{x_0}^{x_n} f''[x]^2 \, dx$$

equivalent with minimum spline strain energy

for least square data fitting

for least “roughness”

Applications of Spline

$$S[f] = \sum_{i=0}^{n} (y_i - f[x_i])^2 + \alpha \int_{x_0}^{x_n} f''[x]^2 \, dx$$

for least square data fitting

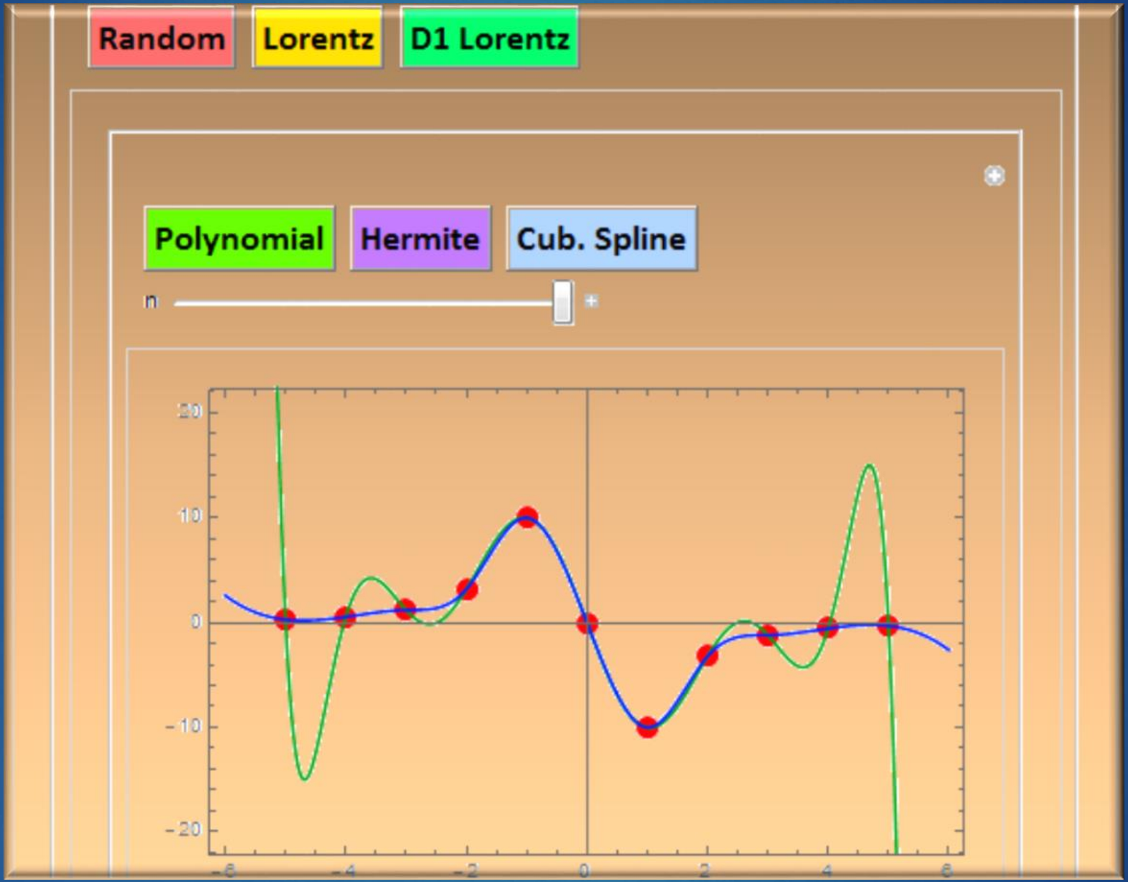
for least "roughness"

Computer graphics, including curve drawing

For data fitting and unknown model building: select a criterion for balancing of the two terms (can be subjective)

For pure interpolation and design: the first term = 0

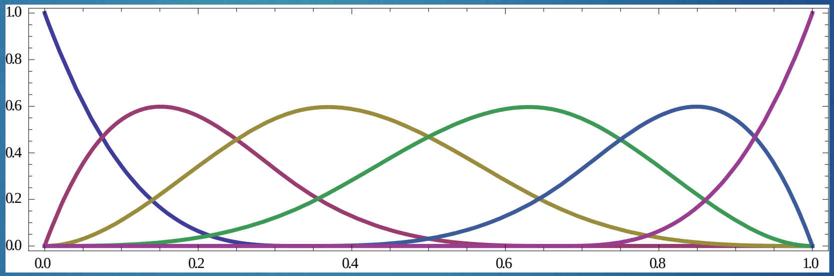
- criterion on the power order of each segment: cubic (3rd order) is the most tried-and-true spline function
- smoothness criterion: order of derivative continuity at each knot.
- basis function for each segment and the total function is a linear combination of basis functions: B-spline
- many types of spline has been developed for different requirements: especially for design (Bezier spline functions)



Random
Lorentz
D1 Lorentz
Polynomial
Hermite
Cub. Spline
n
20
10
0
−10
−20

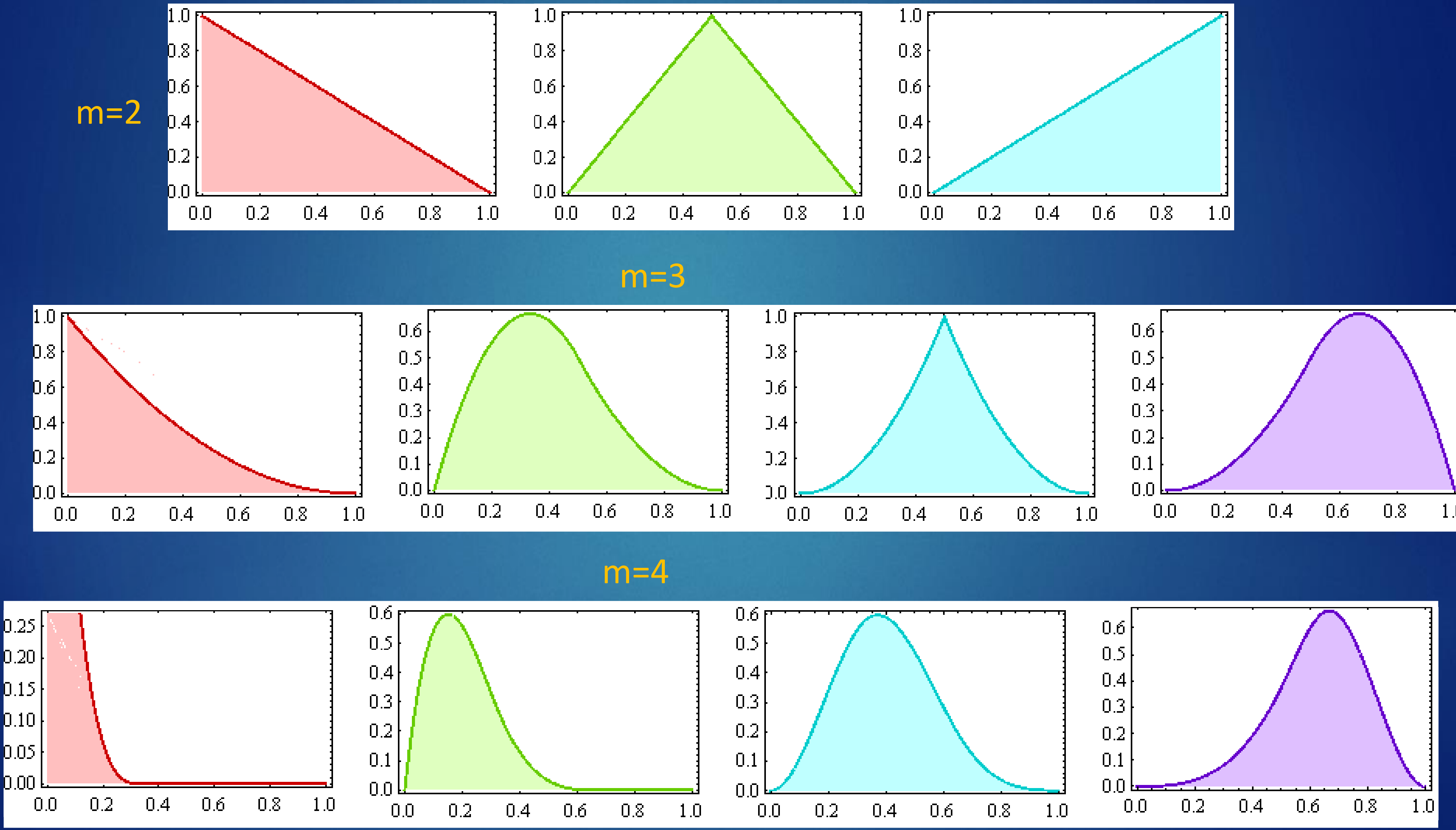
B-spline basis functions

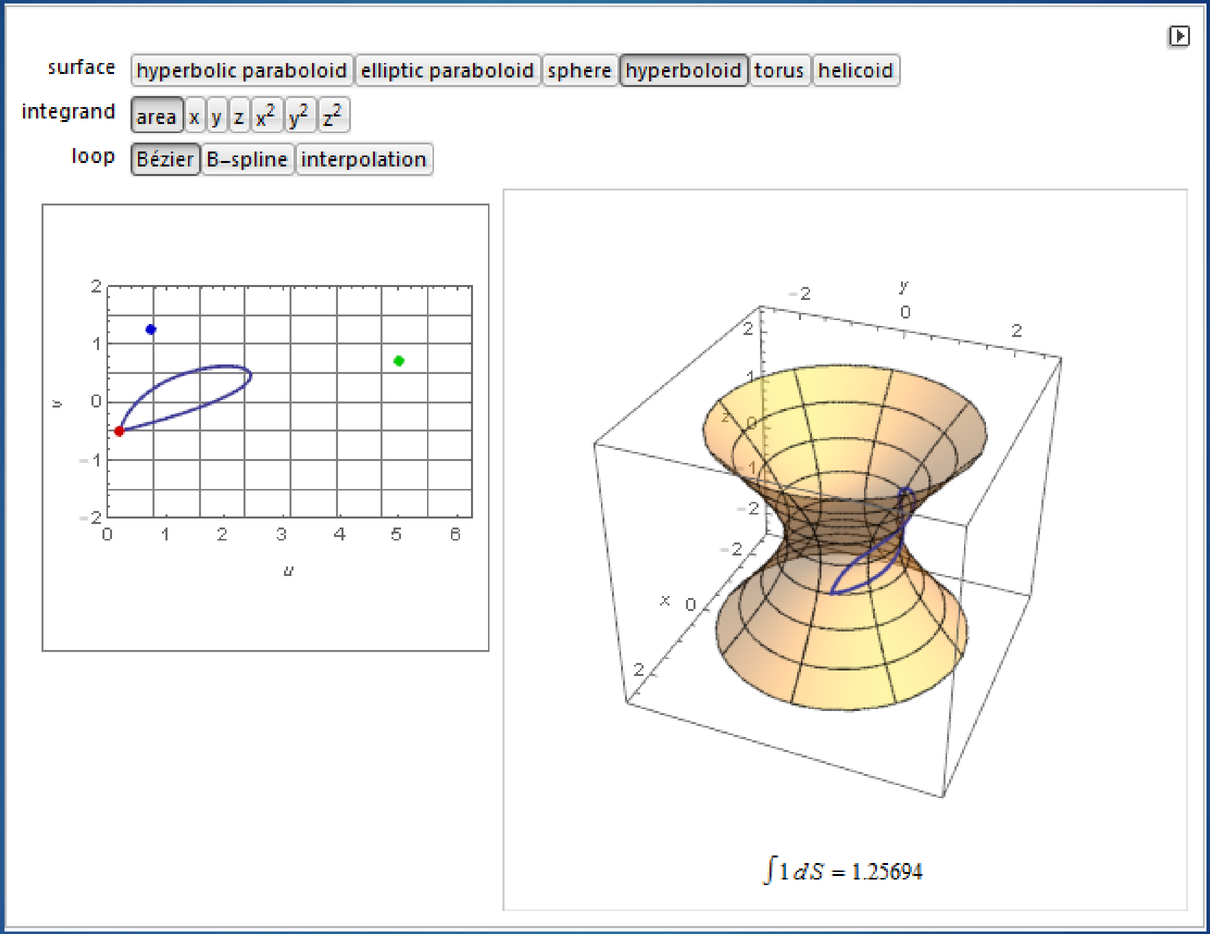
- Basic concept: To approximate a function defined by a sequence of knots {x0,x1,x2,….,xn} by treating it as a linear combination of a set of basis functions.
- Each basis function is localized on a sub-sequence of knots of certain number (which defines a partition) within the global knot sequence.



- The basis functions are designed for smoothness with increasing order to join the partitions

B-spline basis - illustration




surface
hyperbolic paraboloid
elliptic paraboloid
sphere
hyperboloid
torus
helicoid
integrand
area
x
y
z
x2
y2
z2
loop
Bézier
B-spline
interpolation
u
v
∫1 dS = 1.25694

Summary (takeaways)

- Least square regression is used to fit experimental data with models
  - Key metrics: confidence of model parameter estimates (fit coefficients statistics)
- When data is not available or the solutions are known only at a limited number of points, interpolation/extrapolation can be used to fill in the gaps
  - Interpolation: polynomial and cubic spline are two main methods. Spline allows constraint on derivatives.
  - Extrapolation is intrinsically risky – but can be applied if there is sufficient understanding of the problem
- For designing a solution, the spline method can be used with considerations for smoothness and minimizing derivatives to mimic physically reasonable solutions.