Note: this document is excerpted from a 1995 draft paper. Quite a lot has changed has changed in the CURVE lisp library since then, but a lot remains the same as well. An update seems due.

Appendix A
REFERENCE TO ANALYTICAL SOFTWARE

The computer tools described below are organized as a collection given the name of *Curve*, which includes all those used for the time-series data analysis presented along with some drawing and graphic utilities and templates. They were developed in the AutoLISP[i] programming language for use with AutoCAD[i] running on DOS or Windows computers. AutoCAD is a general graphical database product used extensively in architecture and engineering design, mapping and control applications. The Curve analysis collection operates only in an AutoCAD environment and is not directly transferable to other platforms. Any graphing or database software with a reasonable customization capability could be used to develop similar tools following the basic principles outlined here.

Correspondence:
Philip F. Henshaw, ph@idt.net, tel (212)579-2914

Queries: Questions from researchers and developers seeking help with programming and analysis problems are welcome. Copies of the software are available for research use.

Data: Interesting data for experimental analysis would be very welcome. Any well documented time series suspected of having underlying structure would be of interest. Ideal data sets would cover the entire history of some subject event from before its beginning to after its end. Data in comma separated ASCII text field format with title and comment heading lines, column headings and optional end field notes for individual data points would be preferred.

'CURVE' Programming Notes/Command List

1) Command Operation

- Once the functions are loaded, the basic user operation sequence is to enter an analytic function name, select data graphs on the screen to operate on and then choose from various options. The programs then inspect the data and the attached history of prior operations and then scan and interpret the point values. Finally, when the analysis is complete the results are plotted and the command name and options used are added to the operations history. Some functions scan forward only and others both forward and back, using 2, 3, 5 or more points in the analytical point bracket, and may perform single or multiple iterations.

2) Presentation of Results

- Every analytic function creates a new graph, as every operation on a mathematical equation creates a new formula. When several steps are taken to produce a desired result the intermediates are typically hidden or discarded unless they present useful visual information.

---

[i]AutoCAD and AutoLISP are registered trade names of AutoDesk Inc.

- Because the scale of derivatives and integrals is often quite different from that of the original data, derivatives are sometimes automatically rescaled with a peak value of 4/5 of the peak value of the original and integrals at 5/4 of the original.   When a comparison of different data sets or methods of analysis is desired, a fixed scaling factor is selected so that the results have matched presentation scales.

3)   **Major Functions**

- Derivative Interpolation - DIN locates a new point in the middle of a four point bracket that makes the third derivatives on either side equal, using sub-function (F_3SYM). This creates a curve, including the original data points, which has greater continuity than the original and approximates a curve with the minimum scale and reversals of underlying acceleration necessary for a continuous path between the given points.

- Trend Line Bridging - TLIN- draws a graph between local inflection points as defined by reversals in the sign of the second derivative. This corresponds to the principle that fluctuations about a trend will cross the trend line with a maximum slope and have a slope equal to that of the trend when at a maximum distance from it.   This is commonly the case for homeostatic processes.   .

  Options are provided to select every, alternating inflection points or to prompt inflection points for individual user selection, and to set a maximum length of fluctuation to recognize.

  The best large scale trend results took several steps.   First interpolation and derivative smoothing were done and then the short period fluctuations filtered out. After additional derivative smoothing the long period fluctuations could be identified and bridged.

  If the last inflection point was close to the end of the data the end point was placed according to a weighted projection from the preceding trend third derivative using (FF_TENDS).   This substantially reduced endpoint distortions.

  Each line segment of a trend bridge graph has a separate point on the bridge line for each original data point during the period.   When bridge lengths were large the number of points was reduced before derivative smoothing using (GPAR).   This is required to maintain a regular frequency of points and then to eliminate long periods of false constant slope which derivative smoothing would only reinforce. (See Figure A1)

- **Double Derivative Smoothing -** The objective of derivative smoothing is to reduce the number and scale of underlying derivatives while leaving the integral of the curve unchanged.   **DDSM** combines simultaneous forward and backward scans adjusting the middle point of a 5 point bracket using third derivative smoothing (F_3SYM), or a 3 point bracket using 1st derivative smoothing, to equalize the rates of change on either side of the center point.   Sometimes referred to as curve fitting, rather than smoothing, the routine does slightly effect local maxima and minima but has rapidly declining effect on repeated iteration.   It provides strong local smoothing but has little effect on overall curve shape or timing.(see Figure A.2)

- **Graphic Function Calculator -**The **GCAL** function will perform calculations using the names of graphs and user selected points and distances as variables in equations written as text with a general set of mathematical functions.   One curve is used for the time value set and the implied scale of the other curves at those points is used in the calculation. The function used can be read from and to selected lines of text and is recorded with the resultant curve as part of its operation history record . Graph variable names are identified by starting with a double letter (i.e. GG1, etc.) and points variable names by starting with 'pt' (i.e. PT1, etc.). Other preset AutoLISP variables may also be used by name.   This utility can be used to produce function graphs such as exponentials or log plots of data.   Its primary intended use is for applying theoretical relations between dynamic measures, such as to examine energy flow and temperature relations, etc.

4)   Subfunctions

- **3rd Derivative Smoothing (F_3SYM)** - adjusts the scale of the midpoint of a five point sequence..   The calculation shown here assumes equal time periods. In the program the result is then adjusted for unequal periods.   The equation yields the Δy (**ΔYi**) for the middle point that will make the 3rd derivative in the first three periods equal to that in the last three periods :

$$\Delta y_i = \frac{1}{2}\left(\Delta y_1 + \Delta y_2 - \frac{1}{6}\left(\frac{\Delta y_3}{\Delta x_3} - \frac{\Delta y_0}{\Delta x_0} \right)\!\left(\Delta x_1 - \Delta x_2 \right) \right) \qquad \text{A.1}$$

Algorithm for 3rd Derivative Smoothing
(refer to point and ratio naming convention in figure A3)

Point variables in the sequence are labeled from 0 to 4 and differences from 0 to 3.   The routine also makes corrections at points where new second derivative reversals would be introduced and at the mid-points of double reversal periods which produce inconsistent sign and magnitude errors.

- (**F_DOGRAF**) The Simpler graphing functions were written to record results as each successive point is read and interpreted according to a rule fed to this graph drawing sub-function.

- (**F_DXYVAR**) and (**F_LISTVARS**) The more complex functions maintain numbered variable names for each place in the bracket being considered with data values successively moved from one place to the next as the curve is scanned.

- (**FF_TENDS**) Start and end points of a series are sometimes retained unchanged from the original curve and sometimes projected according to a damped first, second, or third derivative implied by preceding or following points..

5)   Basic Operations

- **Importing Data Files - GRFIL** imports comma separated two column text data along with title, headings, scaling factors and notes.

- **Derivative - DIF** plots the ratio of $\Delta y/\Delta t$ with the starting point assigned a value equal to the average of the second and third.   The proportional difference, or growth rate, option provided is $\Delta y/y$. The auto-scaling option sets the peak value of the derivative to 4/5 that of source.

- **Integral - INT** plots $\Sigma \Delta y * \Delta t$ the area under the curve of each period with the starting point set at a user picked or entered value. The auto-scaling option sets the peak at 5/4 of that of source.

- **Graph Scaling - GSC** scales a graph, in proportion to the ratio of one distance to another

- **Increasing Point Frequency - GSEG** (linear interpolation) inserts a variable number of points equally spaced on a straight line between existing points. Used to increase the point density of a data set..

- **Reducing Point Frequency - GPAR** Creates a subset partition of the data set by skipping points in a given range or replacing a range of points with an average. Used to reduce the point density of a data set. This was primarily applied to a TLIN curve to create an equally spaced point set that could be effectively smoothed by DDSM.

- **Average - ASM** averages a variable bracket of points, with each point given equal weight or a weight in proportion to its position in the point bracket.

- **Symmetric Double Average - DASM** is the same as **ASM,** but scans both forward and back and averages the results to reduce directional bias.

- **Derivative Smoothing - DSM** with 1st or 3rd derivative symmetry with a variable number of iterations. Same as DDSM but without combining simultaneous forward and backward scans.

6) **Other Functions**

- **Trend Separation - TSEP** is a special application tool that separates trend periods in a data sequence and inserts constant periods at maxima, minima and axis crossing points. This is done either by simply adding new points between the developmental trend periods (increasing the range), or with linear compression of the trend period so the range of the data is the same as the original.

- **Step Curves - GSTEP** creates a step curve passing through the points of any graph. It is primarily a graphic tool.

- **Graph Recording - GREC** records the accelerations of the screen pointer for one movement of the digitizer and scales the point set to fit a pre-defined window.