Chapter 3

Axis Magic

3.1 Introduction

In a sense, the only function that the human mind really understands well is a straight line. One good strategy for making a graph easy to understand is to make it as linear as possible.

If $u(x)$, the function to be graphed, is thought to resemble a function $\mu(x)$, then graph $\mu^{-1}(u(x))$. The graph of the inverse of μ of u will be approximately linear. This is a very general strategy because it can be applied with very general "model" functions μ.

For example, if u resembles the exponential function, then plotting the inverse is then graphing $\log(u)$ instead of u. For this special case, most libraries offer the option of plotting u but on a *logarithmic* $y - axis$. This is mathematically equivalent to graphing $\log(u)$. However, a log-scale has the important practical advantage that the graphing software will automatically label the axis with numbers and tickmarks appropriate for u. If we plotted $\log(u)$ using a linear scale, then the tickmarks and numerical labels would be those for $\log(u)$ rather than u, and these are usually harder to interpret.

It is usually not possible to make a curve *exactly* linear. Furthermore, if the model μ is a very complicated function, it may not be very insightful to make a plot of $\mu^{-1}(u)$ — the curve may be a straight line, but what does this mean if μ is a confluent hypergeometric function? However, it is very common for a function u to have a simple *asymptotic* shape for some range of x. A plot can then easily confirm and display this asymptotic behavior.

For example, suppose

$$u = \sin(5x) \exp(-x^2) + \exp(-x) \tag{3.1}$$

The graph of this function on $x \in [0, 100]$ is rather wiggly because of the sine functions. However, the sinusoidal factor decays rapidly for large x so that

$$u \sim \exp(-x), \qquad x \gg 1 \tag{3.2}$$

If we plot u using a logarithmic scale for u, then u will asymptote to a *linear* function. It follows that the choice of a nonlinear scale has made at least part of the range in x easy to understand.

3.2 Squeezing the Axis

Another motive for using nonlinear axes is that it may be difficult to fit the desired range into the desired space. Two amusing examples are illustrated as Figs. 3.1 and 3.2. The tip

of the curve in Playfair's graph is repeated at the bottom of the graph as well as extending beyond the frame at the top. This eccentric procedure is distracting and requires some effort to decode. The advantage is that emphasizes the sharp rise in the plotted quantity; it has increased so much that it is difficult to show the peak on the same scale as earlier years.
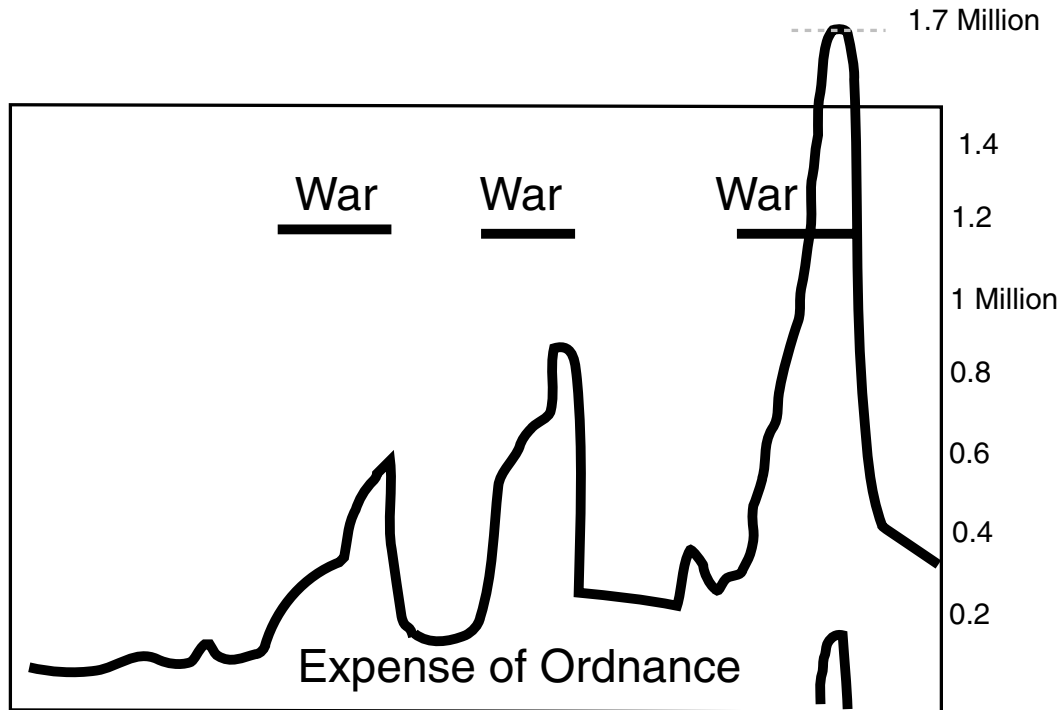

Figure 3.1: A time series of military expenditure versus year in the late eighteenth century. The short horizontal bars mark periods of war when (unsurprisingly!) military costs rose rapidly. Drawn by J. Boyd after an original by William Playfair, published in his *Commercial and Political Atlas*, (London, 1786), plate 14.

Similarly, in the second figure in which the river-length is depicted by horizontal length, the segments for two very long rivers are simply curled around like giant snakes. The disadvantage of this wrap-around is that it is almost impossible to accurately compare the lengths of the wrapped-around rivers with the others. And yet the graph is not a failure; the need to plot two rivers wierdly emphasizes that these rivers are very, very long compared to the rest.

Graphing software is bad at this kind of creativity, which usually must be *drawn* rather than plotted. A logarithmic scale needs no drawing. However, a logarithmic scale, too, makes demands of the reader.
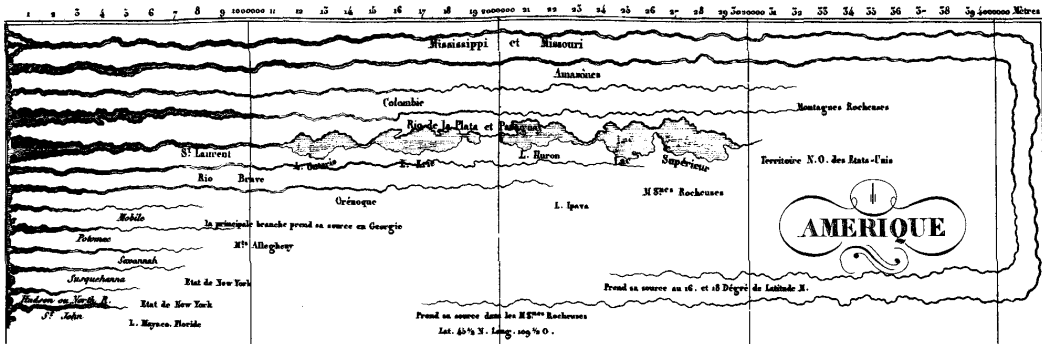

Figure 3.2: Rivers of North America and their lengths in meters. Instead of conventional bars, the wiggly lines attempt to depict the shape each river would have if replaced by a ribbon of rubber which was then straightened out. From Jean Buchon, *Atlas Geographique, Statistique, Historique, et Chronologique des Deux Ameriques* (Paris, 1825), no. LXIII.

3.3   Log, Log-Log and All That

Definition 6 (Linear-Linear Scale) *A plot is said to be a* LINEAR *plot or have a* LINEAR-LINEAR SCALE *if it is a graph of u versus x.*

Definition 7 (Nonlinear Scale/Nonlinear Ordinate) *A plot is said to have a* NONLINEAR *y-scale or* NONLINEAR ORDINATE *if it is a plot of $\mu(u)$ versus x where μ is a nonlinear function of its argument.*

Definition 8 (Log-Linear Plot) *A plot is said to be a* LOG-LINEAR *plot if it is a graph of $\log(u)$ versus x.*

Definition 9 (Log-Log Plot) *A plot is said to be a* LOG-LOG *plot if it is a graph of $\log(u)$ versus $\log(x)$.*

A logarithmic scale has two advantages:

- Much wider range for the plotted variable

- Exponential growth or decay, geometric progression and power-law behavior can all be *linearized*, that is, rendered as a straight line, by using the appropriate type of log or log-log plot.

Example One: Log-in-y/Linear-in-x for Chebyshev Coefficients

Chebyshev spectral methods have the property that if the solution is smooth on the domain of integration, the coefficients of the Chebyshev series for the solution of a differential equation will decrease EXPONENTIALLY FAST with DEGREE. That is, if the coefficients of the spectral series are denoted by a_j, then the coefficients should decrease like

$$a_j \sim \{\} \exp(-q\,j), \qquad j \to \infty \tag{3.3}$$

where the empty braces denotes factors that vary more slowly with j than the exponential, such as a power of j, and where q is a positive constant that depends on the specific problem. In using spectral algorithms to solve differential equations, it is sound practice to periodically check the coefficients to confirm the predicted behavior. (In aviation, the equivalent proverb is: "Even when flying on instruments, look out the window occasionally.")

Fig. 3.3 shows two graphs of the spectral coefficients, identical except that one employs a logarithmic scale and the other does not. The linear-linear plot is a disaster: most of the data is invisible! The log-log plot has two virtues. First, all the coefficients are plotted, even the very tiny ones. This is the advantage of a wide range noted earlier. Second, the coefficients asymptote to a straight line as the degree j increases. This is consistent with the predicted exponential behavior: $\exp(-qj)$ would plot as a straight line with a logarithmic scale.
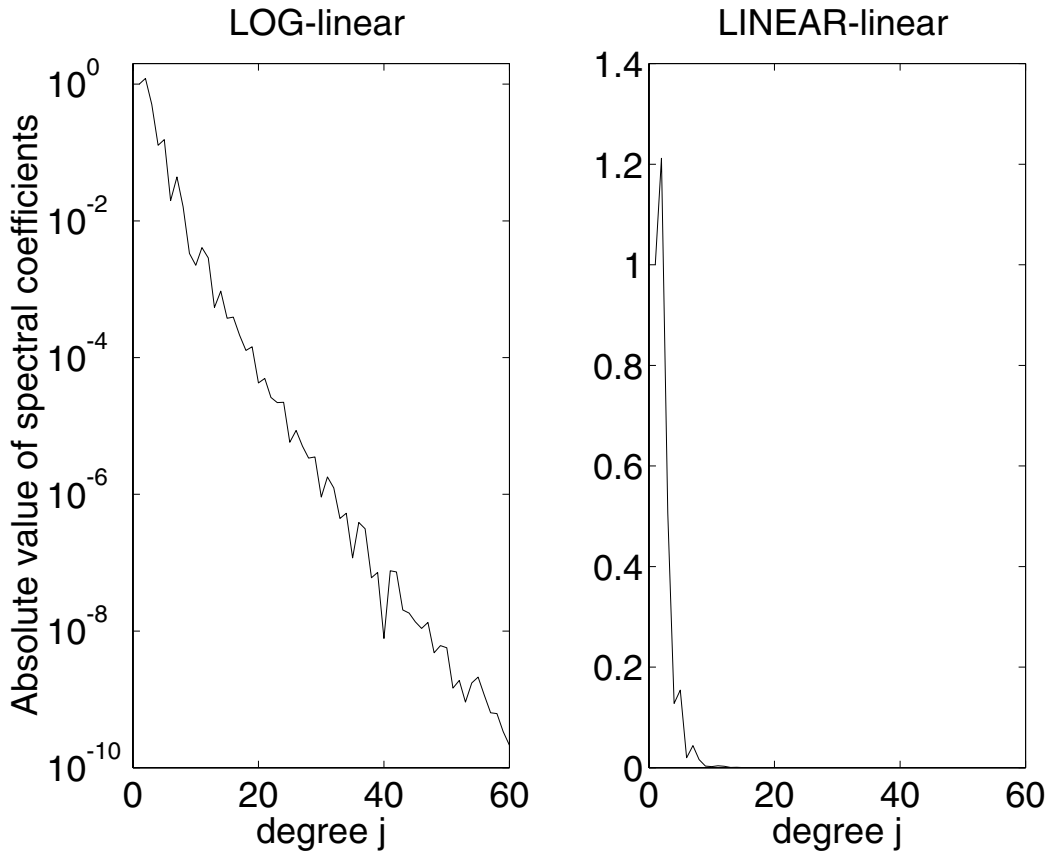
Figure 3.3: Both graphs illustrate a_j for the rational Chebyshev series for the unstable mode of a hydrodynamic instability problem. The linear-linear plot (right) conveys little information; most of the spectral coefficients are so small as to be invisible. When the ordinate scale is logarithmic, (left panel), all the coefficients are visible. Furthermore, the coefficients are APPROXIMATELY linear for LARGE degree which confirms the expected asymptotic exponential decay. [Boussinesq-Charney eigenproblem of meteorology for $r = 1/2$.]

EXAMPLE TWO: Check of Asymptotic Series

Many functions have asymptotic series in negative powers of z which become increasingly accurate for a given number of terms as z increases. The accuracy of such series can be checked against accurate, independent numerical solutions by plotting the errors on a log-log graph.

The absolute error in a k-th order asymptotic approximation should decrease as

$$E_k \sim q \, z^{-k} \tag{3.4}$$

where q is a positive constant. Taking the logarithm of both sides of this asymptotic relation gives

$$log\{E_k\} \sim \log(q) - \log(z) \tag{3.5}$$

which is a straight line with a slope equal to negative k on a log-log plot.

If several different orders of approximation are plotted on a single log-log graph, the errors should not only decrease with increasing z; the errors should ASYMPTOTE TO A STRAIGHT LINE. In addition, the slopes should steadily increase with increasing order.
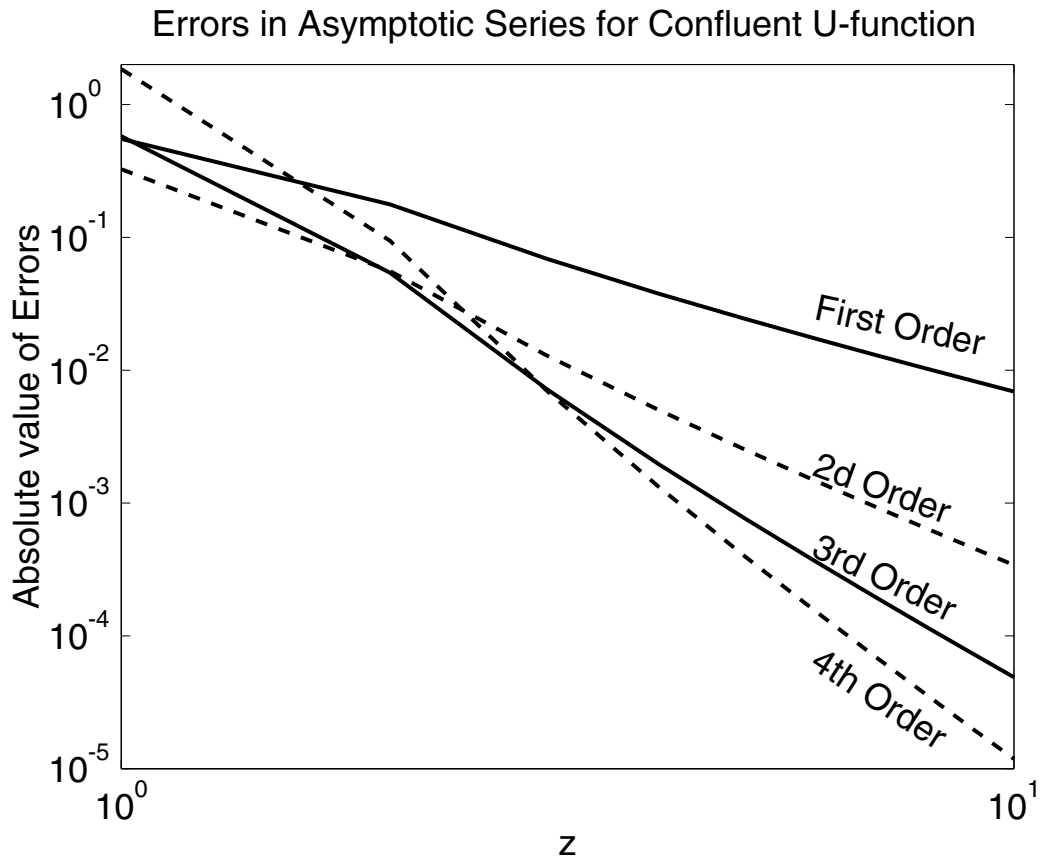

Figure 3.4: The error in asymptotic approximations of first, second, 3rd and 4th orders versus z is plotted. [Confluent hypergeometric function $U(1/2, 2, z)$. the errors are proportional to the first, second, third, and fourth powers of the reciprocal of $|z|$.] On a log-log plot, the errors should therefore asymptote to a straight line whose slope steepens with increasing order. If this pattern is not found, then it screams "ERROR"!]

3.4 Probability, Arctanh and Arcsinh Scales

Fig. 3.5 shows two representations of the same function. The nonlinear scale for the right graph is the "probability function" (more precisely, the "cumulative normal frequency curve"). The probability function itself itself is linearized by this nonlinear scale. The crucial point is that real-word statistical samples will also be *approximately* linearized if their distribution is "normal". The deviation of the curve from from a straight line, when plotted with the "probability" scale, provides a visual test of statistical normality.
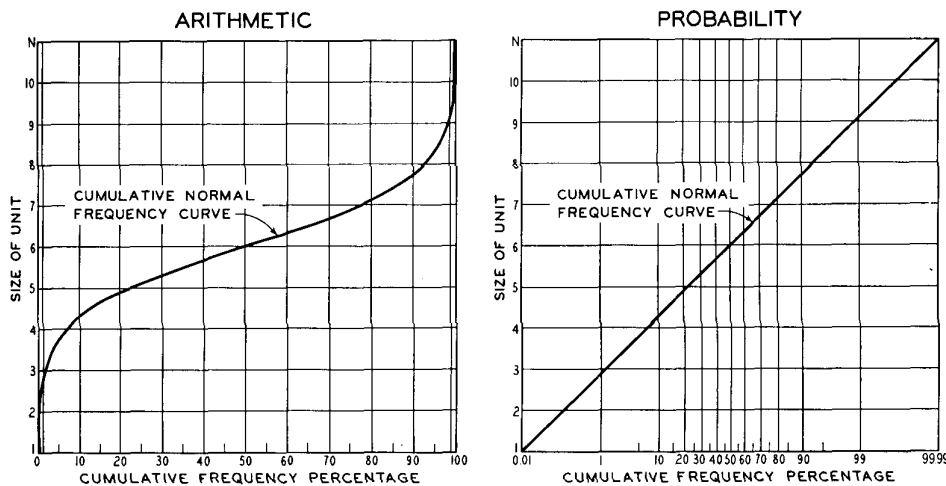

Figure 3.5: Two representations of the same curve (solid), the cumulative normal frequency curve. Left: this function graphed with the usual linear scales (labeled "arithmetic"). Right: same function plotted on "probability paper", which converts this curve into a straight line. From Schmid(1954), *A Handbook of Graphic Presentation.*
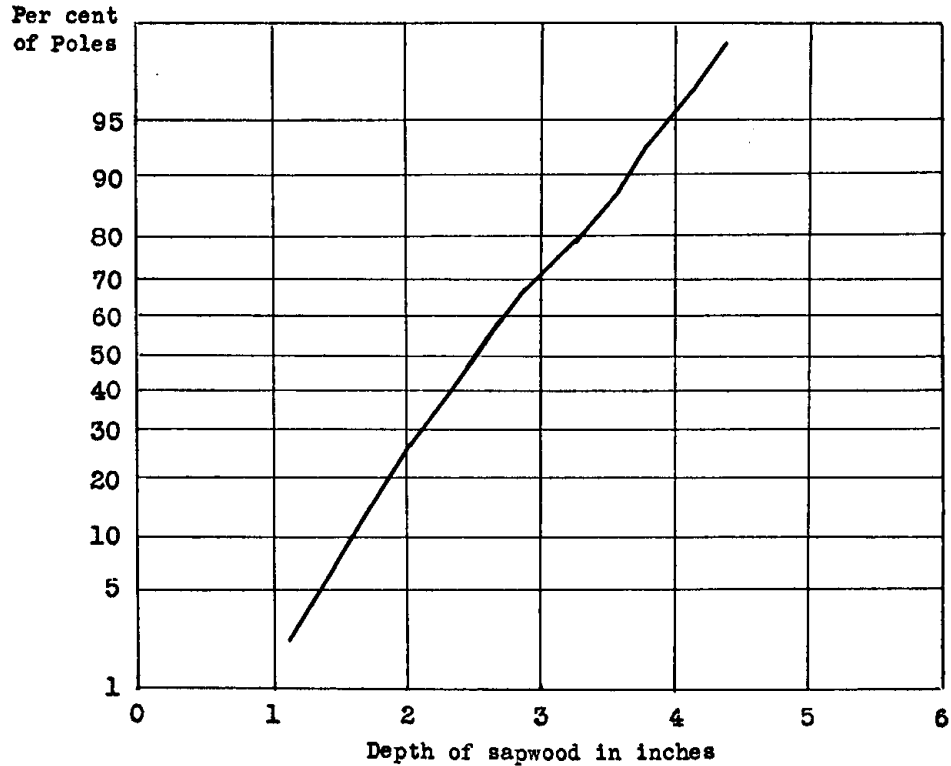
Figure 3.6: The raw data is S-shaped. By using a novel axis (inverse of the "cumulated normal probability curve" or "ogive"), the plotted points are almost linearized where the deviation from linearity illustrates the deviation from the standard probabilistic distribution. From Arkin and Colton(1936), who acknowledge W. A. Shewhart, "Economic Control of Quantity of Production"; the data is the depth of sapwood in a sample of 1370 telephone poles.

Fig. 3.6 is an example. The near-linearity of the data curve shows that the depth of sapwood for this sample of telephone poles is, statistically speaking, close to normal. We plotted a curve, and instead obtained a theory!

It would be much more difficult to see the closeness to normality if the data had been plotted on a linear ordinate scale.

Other nonlinear scales are useful. The error function $\text{erf}(x)$ is a "tanh-like" function in the sense that it is antisymmetric with respect to $x = 0$ and varies monotonically between -1 and 1. One way to see how closely these two functions resemble each other is to plot the error function using an arctanh scale. That is to say, we plot $\text{arctanh}(\text{erf})(x)$ rather than the error function itself. However, the vertical axis is labelled with the values of the error function rather than of its arctanh as illustrated in Fig. 3.7.
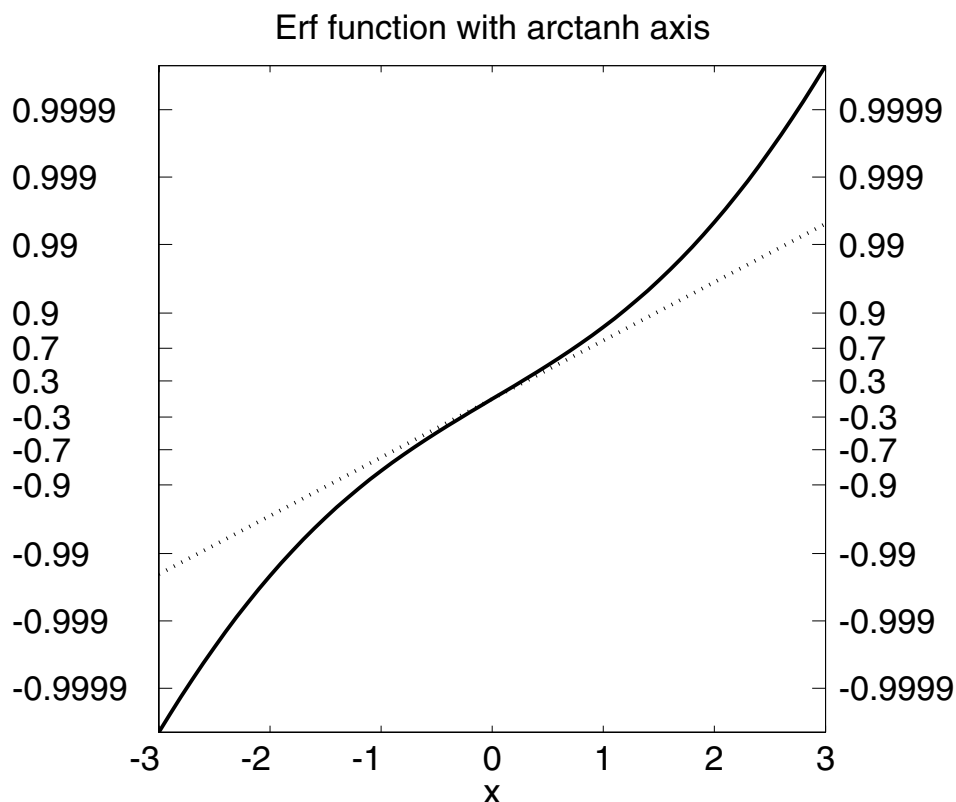
Erf function with arctanh axis

0.9999	0.9999
0.999	0.999
0.99	0.99
0.9	0.9
0.7	0.7
0.3	0.3
-0.3	-0.3
-0.7	-0.7
-0.9	-0.9
-0.99	-0.99
-0.999	-0.999
-0.9999	-0.9999

-3 -2 -1 0 1 2 3

x

Figure 3.7: A graph of the error function using an arctanh scale. That is, the graph is $\text{arctanh}(\text{erf})(x)$ versus x, but the labels give the values of erf, which range between -1 and 1, rather than those of its arctanh. The dotted line shows the graph of $\tanh(x)$ on the same scale for reference.

To compute a function u which is growing exponentially for large $|x|$, the obvious choice is to use a logarithmic scale. However, if $u < 0$, then the logarithm is complex-valued unless we plot the *negative* of u. If u is of both signs, then a logarithmic scale will fail unless we plot the absolute value of u. However, this has two disadvantages. First, there is no simple way to show the sign of u on the plot. (One remedy: split the curve into segments where u is of one-signed and plot each separately,using solid line for segments where $u > 0$ and dashed lines for segments where u is negative.) Second, $\log(|u|)$ runs off to negative infinity at each zero of u.

A more direct alternative is to plot the function using an arcsinh scale. This scale has no difficult with sign changes. In addition, the arcsinh of a very large number is a moderate number so that, like a logarithmic scale, the arcsinh scale can accomodate a huge range. An example is shown in Fig. 3.8.
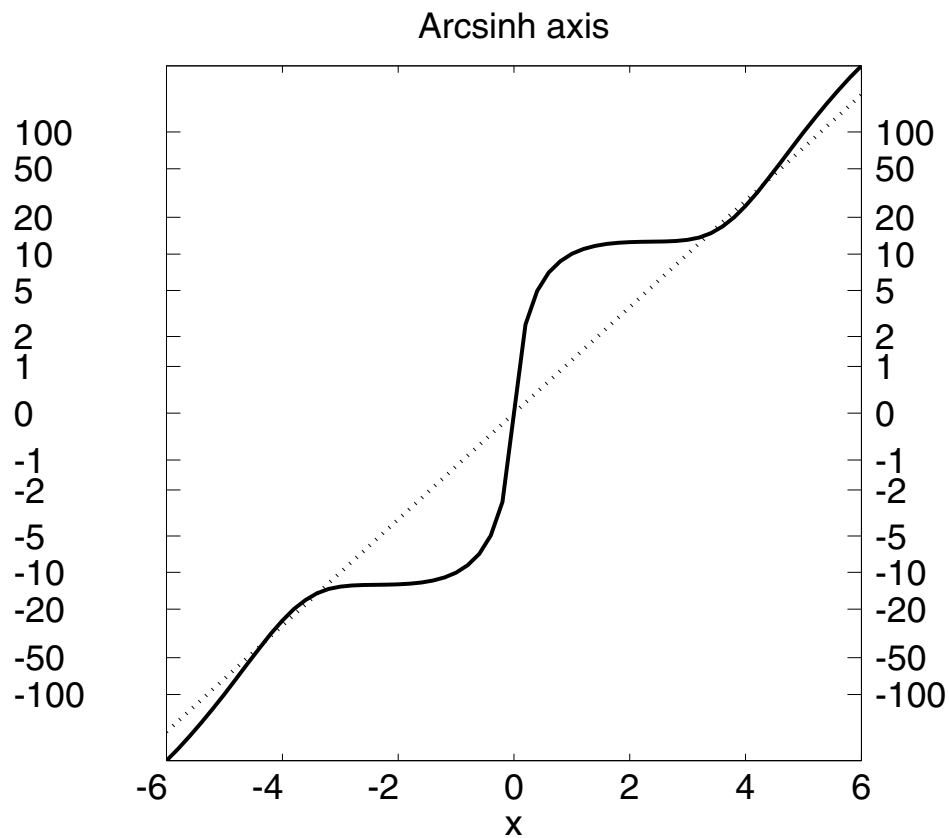

Figure 3.8: A graph of a function $f(x)$ on an arcsinh scale. The vertical location of each point is that of arcsinh(f)(x_j), but the labels are the values of $f(x_j)$ itself. The function f was chosen arbitrarily to be $f = \{1 + 0.7\cos(x)\}\sinh(x) + 10\mathrm{erf}(x)$.

3.5 Monetary Inflation

Comparing costs and salaries for different years is difficult because of inflation, which is defined to be the change in the price of a fixed item, averaged over many categories of items. The Consumer Price Index (C. P. I.), which defined the inflation rates employed in Fig. 3.9, is a number published annually by the U. S. goverment to estimate inflation as felt by the average person. (The inflation rates for individual sectors of the economy, such as real estate or energy, may be considerably different from the C. P. I.)

Thus, a salary that rises at 4% a year may not actually produce any increase in a person's standard of living because the cost of housing, food, automobile, taxes and so on have risen at the same rate. Fig. 3.9, devised by Robert Sherman, employs unusual axes to compensate. The horizontal axes are slanted downward at a rate proportional to the C. P. I. A salary of $30,000 should be plotted on the slanting line beginning at "3" on the left axis; the horizontal location of the point is given by the year. One can then convert this into the dollars that would buy the same amount of goods in 1969 by drawing a horizontal line from the plotted point to the left axis.

Another way to use this graph: a salary which rises just fast enough to keep up with inflation will be a horizontal line on the graph: the absolute amounts rise just fast enough to compensate for the downward slant of the axes.

This artifice of slanted axis lines is different from the nonlinear axis scales used in earlier sections: for logarithmic and arctanh graphs, the axis lines are always horizontal. No standard software package includes a routine to create slanted axis graphs, but it is not hard to modify existing Matlab routines to do this.

A much simpler solution, of course, is to simply plot salary divided by an inflation factor. Indeed, this is the most common tactic because one can simultaneously plot two curves, one showing salary in absolute terms and another showing salary in constant 1969 dollars.

Nevertheless, the slanted axis graph is useful as a reminder. In economics, it is probably much more sensible to compare quantities in different years after adjusting for inflation than without such an adjustment. We can reiterate too often that one of the most important choices in graphics is: What quantity should be plotted? The obvious answer (dollar amounts) is often a misleading answer (because inflation-corrected dollars more accurately reflect the purchasing power of money).
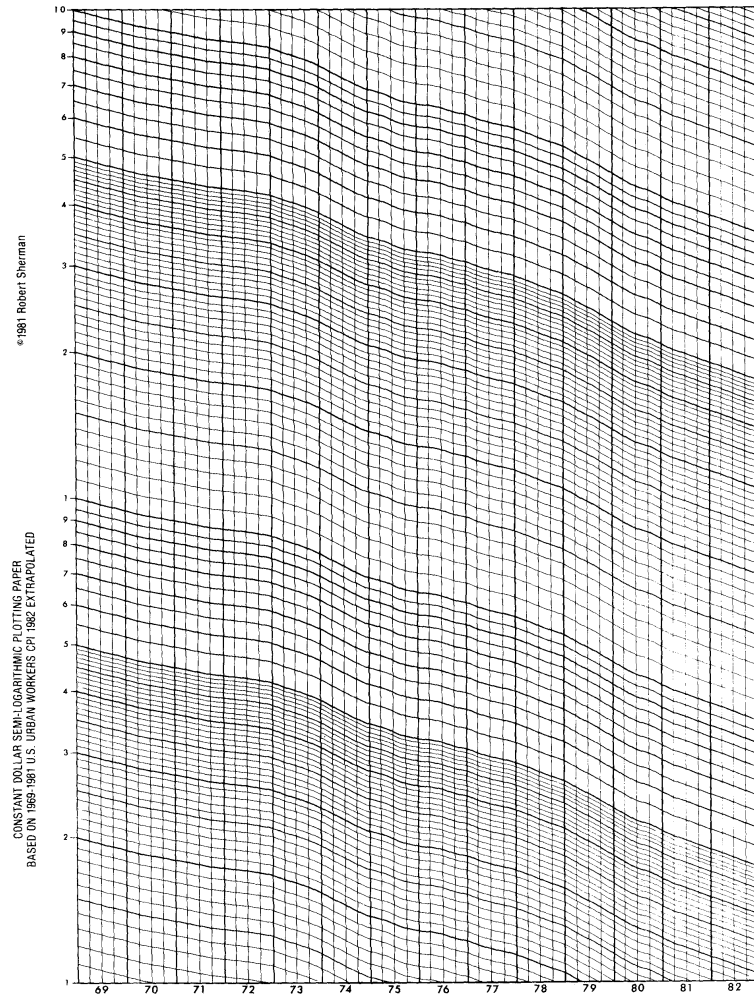
Figure 3.9: Constant-dollar graph paper. The bottom axis specifies the year (from 1969-1982) in a linear scale. The vertical axis is dollars on logarithmic scale. The horizontal axis lines are given a downward slope which is proportional to the yearly inflation rate at that time. The lowest axis line which runs the full width of the graph conveys the information that $1000 in 1969 is equivalent to $3000 in 1982. A salary which rises at a rate to just match inflation will plot as a horizontal line on the graph. Devised by Robert Sherman and reprinted in Wainer(1997), pg. 123.

3.6 Nomography and Graphical Calculators

In the pre-computer days, and continuing into the mainframe period when computers were still huge and unwieldy, "nomography", that is, graphs that could be used to calculate quantities, was a major part of engineering. Often, linear axes failed miserably for such graphs because the answer would be a tiny, flat wiggle, almost impossible to accurately translate into numbers. Therefore, nomographs and graphical calculators made heavy use of very ingenious scales.

Crowhurst (1965), who wrote one of the last books on this topic, has many wildly nonlinear scales in his short book. Two of the more ingenious are displayed in Fig. 3.10.

Nomography has diminished in importance because of the wide availability of computers and hand-held calculators. However, it has not completely disappeared. Crowhurst's short book is a very readable description of the art.

How does one interpret these axes from Crowhurst? The answer can only be obtained by really *studying* the graph. This is a disadvantage of nonlinear scales: they require an investment of study-time by the reader. This investment increases rapidly as the nonlinear scale/s become more exotic.
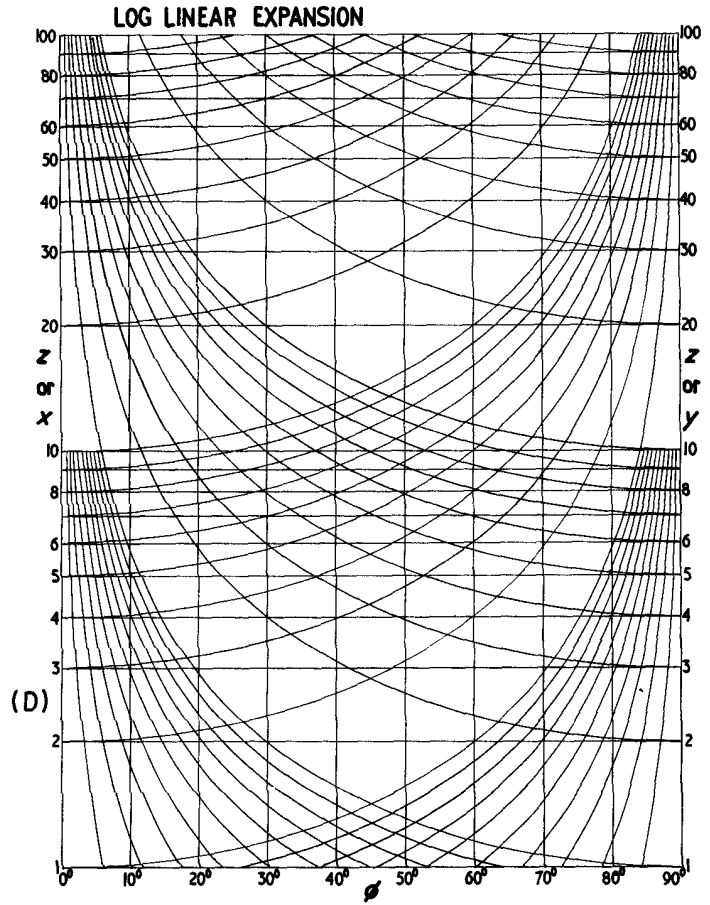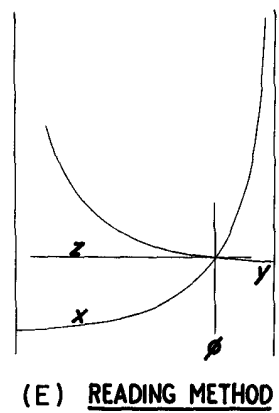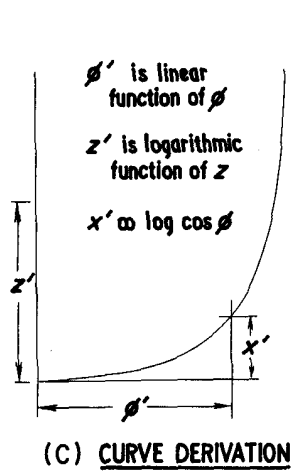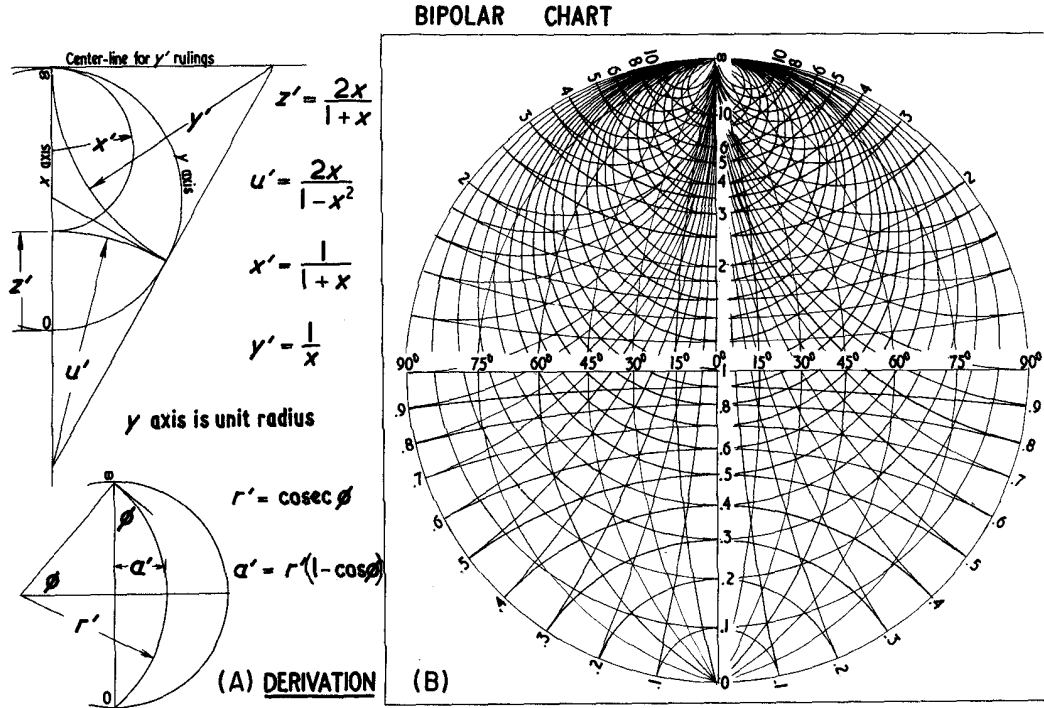
BIPOLAR   CHART

Center-line for y' rulings

$$z' = \frac{2x}{1+x}$$

$$u' = \frac{2x}{1-x^2}$$

$$x' = \frac{1}{1+x}$$

$$y' = \frac{1}{x}$$

y axis is unit radius

$$r' = \operatorname{cosec}\phi$$

$$a' = r'(1-\cos\phi)$$

(A) DERIVATION   (B)

ϕ' is linear
function of ϕ

z' is logarithmic
function of z

x' ∞ log cosϕ

(C) CURVE DERIVATION

(E)   READING METHOD

LOG LINEAR EXPANSION

(D)

Figure 3.10: Two sets of axes for "graphical calculators" from Crowhurst(1965), pg. 56.