Graphic Insight into Mathematical Concepts[1]

David Tall
Mathematics Education Research Centre
Warwick University, COVENTRY CV4 7AL, U.K.

Beverly West
Mathematics Department,
Cornell University,
Ithaca, NY 14583, U.S.A.

The human brain is powerfully equipped to process visual information. By using computer graphics it is possible to tap this power to help students gain a greater understanding of many mathematical concepts. Furthermore, *dynamic* representations of mathematical processes furnish a degree of psychological reality that enables the mind to manipulate them in a far more fruitful way than could ever be achieved starting from a static text and pictures in a book. Add to this the possibility of student exploration using prepared software and the sum total is a potent new force in the mathematics curriculum.

In this paper we report on the development of interactive high resolution graphics approaches to various areas in mathematics. The first author has concentrated initially on the calculus in the U.K. (Tall 1986, Tall *et al* 1990) and the second is working in the USA on differential equations with John H. Hubbard (Hubbard and West 1990).

An interactive visual approach is proving successful in other areas, for example, in geometry (*The Geometric Supposer*, *Cabri Géomètre*), in data manipulation (e.g. *Macspin*, *Mouse Plotter*), in probability and statistics (e.g. Robinson & Bowman 1987) and, more generally, in a wide variety of topics (such as the publications in the *Computer Illustrated Text* series, which use computer programs to provide dynamic illustrations of mathematical concepts).

New approaches to mathematics

The existence of interactive visual software leads to the possibility of exploratory approaches to mathematics which enables the user to gain intuitive insight into concepts, providing a cognitive foundation on which meaningful mathematical theories can be built. For example, the notion of a limit has traditionally caused students problems (e.g. Cornu 1981, Tall & Vinner 1981). The computer brings new possibilities to the fore; we may begin by considering the gradient not of the tangent, or of a chord as it approaches a tangential position, but simply *the gradient of the graph itself*. Although a graph may be curved, under high magnification a small part may well look almost straight. In such a case we may speak of the gradient of the graph as being the gradient of this magnified (approximately straight) portion. For instance, a tiny part of the graph $y=x^2$ near $x=1$ magnifies to a line segment of gradient 2 (figure 1).

To represent the changing gradient of a graph, it is a simple matter to calculate the expression $(f(x+c)-f(x))/c$ for a small fixed value of c as x varies. As the chord clicks along the graph for increasing values of x, the numerical value of the gradient for each successive chord can be plotted as a point and the points outline the graph of the gradient function (figure 2). In this case the chord gradient function of $\sin x$ for small c

[1]The authors are grateful to Professor John H. Hubbard for his assistance in the preparation of an earlier version of this article.

Published in *The Influence of Computers and Informatics on Mathematics and its Teaching* (ed. Howson G. & Kahane J-P), C.U.P (1986).

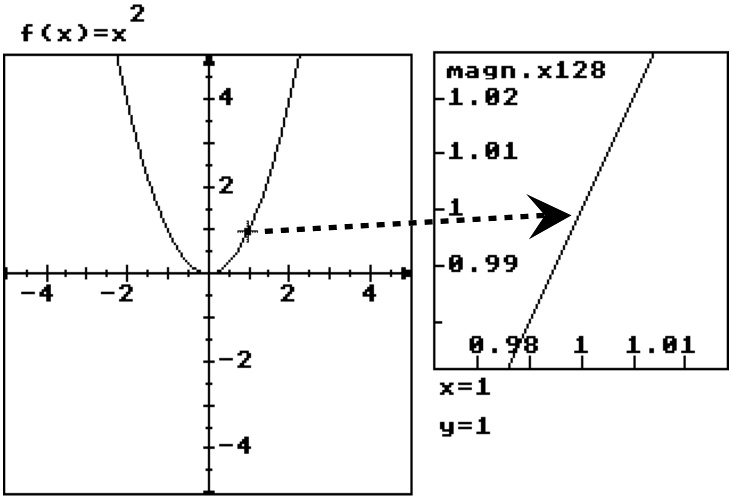


figure 1

approximates to cos$x$, which may be checked by superimposing the graph of the latter for comparison. Thus the gradient of the graph may be investigated experimentally before any of the traditional formalities of limiting processes are introduced.

The moving graphics also enable the student to get a dynamic idea of the changing gradient. Students following this approach can see the gradient as a global *function*, not simply something calculated at each individual point.

The symbols dx, dy can also be given a meaning as the increments in x,y to the tangent. Better still, (dx,dy) may be viewed as the *tangent vector*, a valuable idea when we come to the meaning of differential equations.

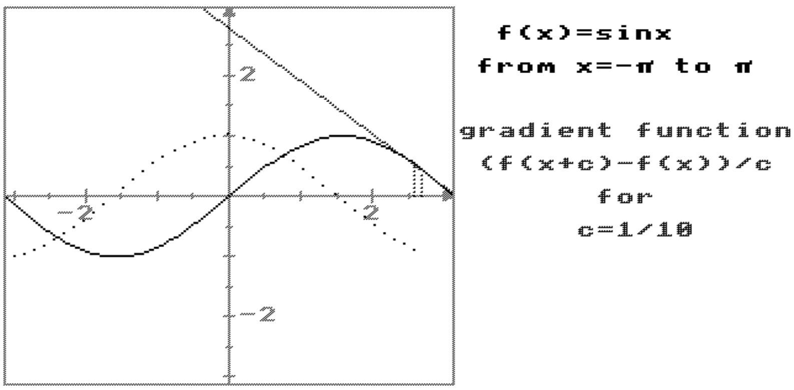


figure 2

Conceptualizing non-differentiable functions

In a traditional course, non-differentiable functions would not be considered until much later on, if at all. However, if one views a differentiable function as one which is “locally straight”, then a non-differentiable function is simply one which is *not* locally straight. For instance, the graph of $|x\text{-}1|$ at x=1, or $|\sin x|$ at x=π, each have a “corner” at the point concerned with different gradients to the left and right. More generally, it is possible to draw a function that is so wrinkled that it never looks straight *anywhere* under high magnification.

An example is the *blancmange function* $bl(x)$, first constructed by Takagi in 1903. First a saw-tooth $s(x)$ is constructed by taking the decimal part $d=x$-INTx of x and defining

$$s(x)=d \text{ if } d<\tfrac{1}{2}, \text{ otherwise } s(x)=1-d.$$

The sequence of functions

$$b_1(x)=s(x),$$
$$b_2(x)=s(x)+s(2x)/2,$$
$$\ldots$$
$$b_n(x)=s(x)+\ldots+s(2^{n-1}x)/2^{n-1}, \ldots$$

tends to the blancmange function (figure 3).

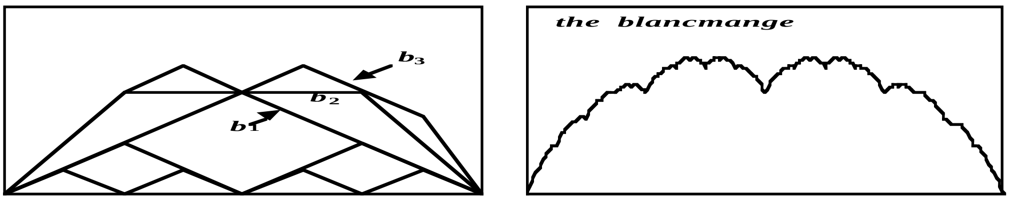


figure 3

The process may be drawn dynamically on a VDU; we regret that it cannot be pictured satisfactorily in a book, not even this one. But higher magnification of the blancmange function using prepared software shows it nowhere magnifies to look straight, so it is nowhere differentiable. This intuitive approach can easily be transformed into a formal proof of disarming simplicity (Tall 1982).

Visualizing solutions of first order differential equations

In graphical terms, a first order differential equation dy/dx=f(x,y) simply states the gradient of a solution curve at any point (x,y) and a solution is simply a curve which has the required gradient everywhere. The *Solution Sketcher* (Tall 1990) allows the user to point at any position in the plane and draws a small line segment of the appropriate direction. This line-segment may be marked on-screen and successive line segments fitted together to build up an approximate solution curve. More broadly, it is possible to draw a direction diagram with an array of such segments and to trace a solution by following the given directions (figure 4).

The differential equation

$$y\frac{dy}{dx} = -x$$

has implicit solutions in the form $x^2+y^2=k$, rather than a global explicit solution in the form $y=f(x)$. At points where the flow-lines meet the x-axis the tangents are vertical and the interpretation of dy/dx as a function fails, but the vector direction (dx,dy) is valid with $dx=0$ and $dy\neq 0$. Thus a first-order differential equation is sometimes better viewed in terms of the direction of the tangent to a solution curve rather than specifying the derivative.

Existence of solutions

There comes a time in every university course on differential equations when honesty should compel the teacher to admit that the cookbook methods for solving differential equations are inadequate. Such innocent looking equations as

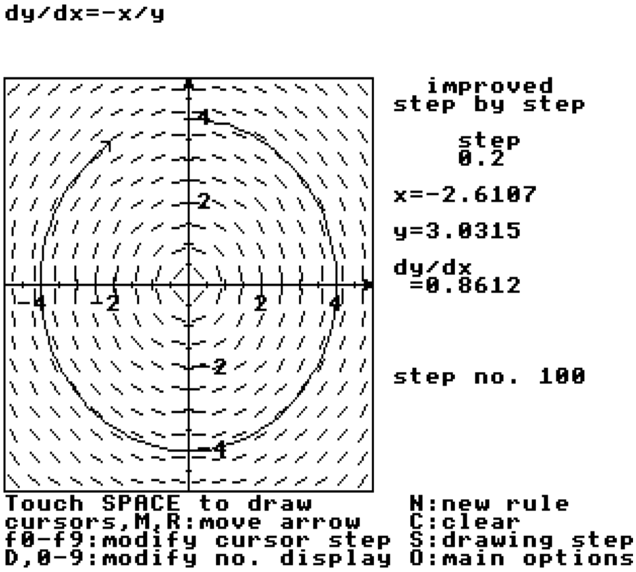


figure 4

$$dy/dx=y^2-x,\ dy/dx=\sin(xy),\ dy/dx=e^{xy}$$

do not have solutions that can be written in terms of elementary functions. Students often mistakenly confuse this with the idea that the equations have no solutions at all. However, if they are able to interact with a computer program that plots a direction field and then draws solutions numerically following the direction lines, the phenomenon takes a genuine meaning: “Of course the equations have solutions: we can *see* them!” From this cognitive base it is possible to use the computer to analyse solutions in an entirely new way.

Qualitative analysis of differential equations

New forms of analysis emerge now we can see as many solutions as we wish all at the same time. In figure 5, notice how the solutions tend to “funnel” together moving to the lower right-hand side; in the upper right they spray apart (an “antifunnel”). Qualitatively descriptive terms such as “funnel” and “antifunnel” can be defined precisely to give powerful theorems with accurate quantitative results (Hubbard & West 1990). For example, the equation $dy/dt=y^2-t$ in figure 5 has two overall behaviours: solutions either approach vertical asymptotes for finite t or fall into the funnel and approach $y=-\sqrt{t}$ as $t\to+\infty$. In the antifunnel there is a unique solution approaching $y=+\sqrt{t}$ which separates the two usual behaviours. Furthermore, the qualitative techniques enable us to estimate the vertical asymptote for a solution through any given point with good precision.

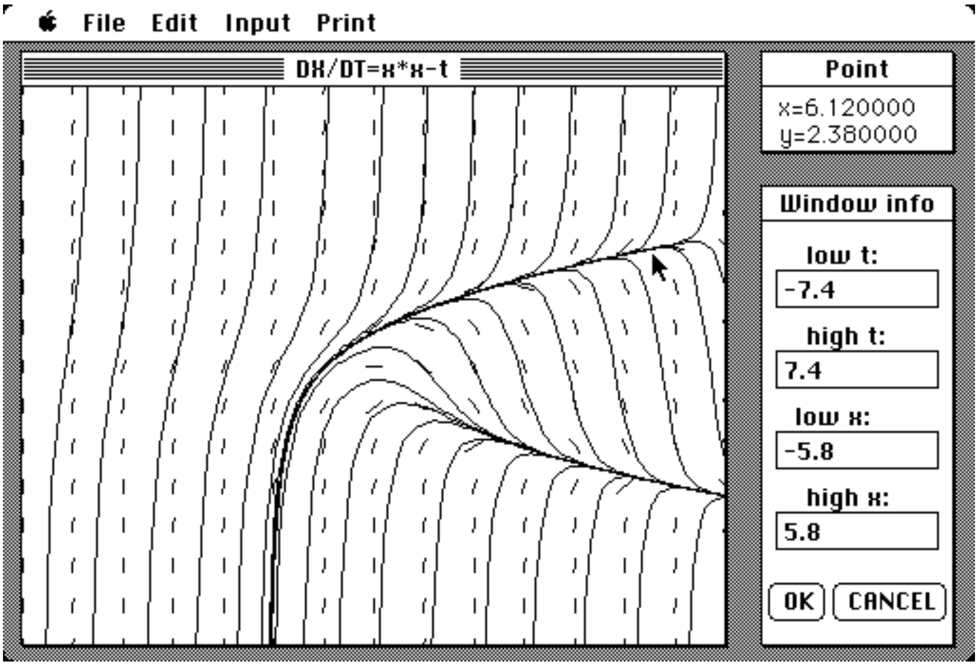


figure 5

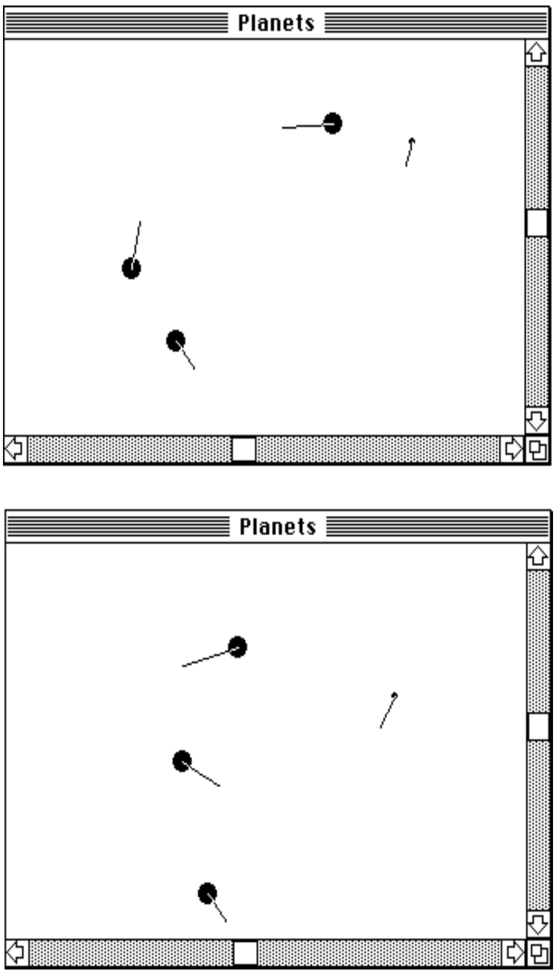
Newton's Laws

The classical three body problem defies elementary analysis, yet a computer program can cope with relative ease. The program *Planets* (Hubbard & West 1990) takes a configuration of up to ten bodies with specified mass, initial position and velocity and displays the movement under Newton's laws (figure 6). The data can be input either graphically with the cursor, or numerically in a table. The program allows exploration of possible planetary configurations and it soon becomes plain that stability is the exception rather than the rule. One may wonder under what circumstances stability occurs. Other questions arise, such as the reason for the braided rings of Saturn that were a great surprise when first observed by the Voyager space flight. Nobody had imagined such a behaviour beforehand, yet braided behaviour showed up in the very first experiments with the Planets program.

Figure 7 shows a model of a possible orbit of a tiny satellite round two larger bodies, alternately oscillating between revolving round one then moving into a position of superior gravitational pull of the other and moving, for a time, to revolve round the other (Koçak 1986). Once again, computer exploration shows vividly how three bodies move in a complex pattern.

The theory of dynamical systems and chaos is a paradigmatic example of a new branch of mathematics in which the complementary roles of computer generated experiments to suggest theorems and formal mathematical proof to establish them with logical precision go hand in hand.

> Chaos has become not just a theory but also a method, not just a canon of beliefs but also a way of doing science. ... To chaos researchers, mathematics has become an experimental science, with the computer replacing laboratories full of test tubes and microscopes. Graphic images are the key. “It’s masochism for a mathematician to do without pictures,” one chaos specialist would say. “How can they see the relationship between that motion and this, how can they develop intuition?”. (Gleick 1987, pp. 38-39)



masses in initial position
with velocity vectors

a little later
under the action
of Newton’s Laws

figure 6

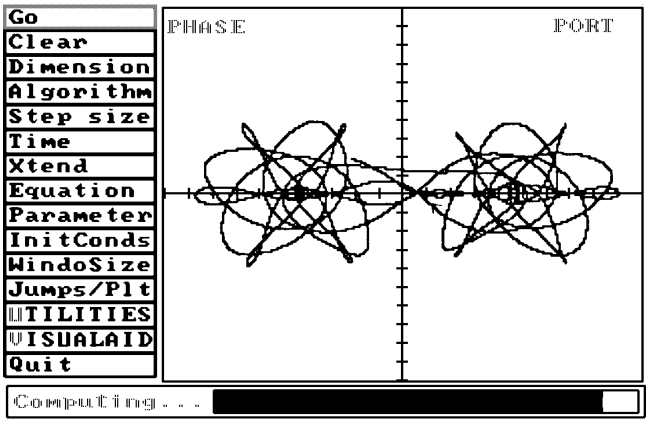


figure 7

Systems of differential equations

The software of Hubbard & West (1990) draws solutions of systems of differential equations $dx/dt=f(x,y)$, $dy/dt=g(x,y)$ in the x,y plane and also locates singular points using Newton's method, drawing separatrices for saddle points (figure 8).

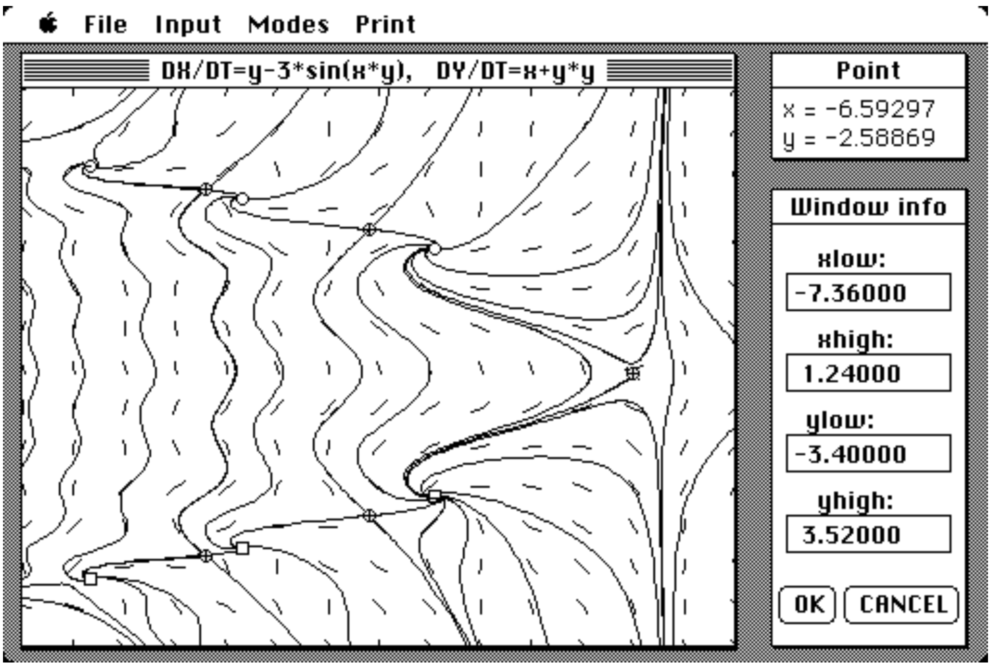


figure 8

In this way the computer may be used to draw solutions of systems of differential equations that are far too complicated to draw by hand. As a further example, Artigue and Gautheron (1983) draw the solutions of the polar differential equations

$$dr/dt=\sin r,\ d\theta/dt=\cos r$$

which exhibit limit cycles for $r=k\pi$ (figure 9).

figure 9

Generalizing the concept of visual solutions

A second order differential equation such as

$$d^2x/dt^2=-t$$

no longer has a simple direction field in (t,x) space, because through each point (t,x) there is a different solution for each starting direction $v=dx/dt$. However, this differential equation is equivalent to the simultaneous linear equations:

$$dx/dt=v$$

$$dv/dt=-t,$$

and in three dimensional (t,x,v) these equations determine a unique tangent vector

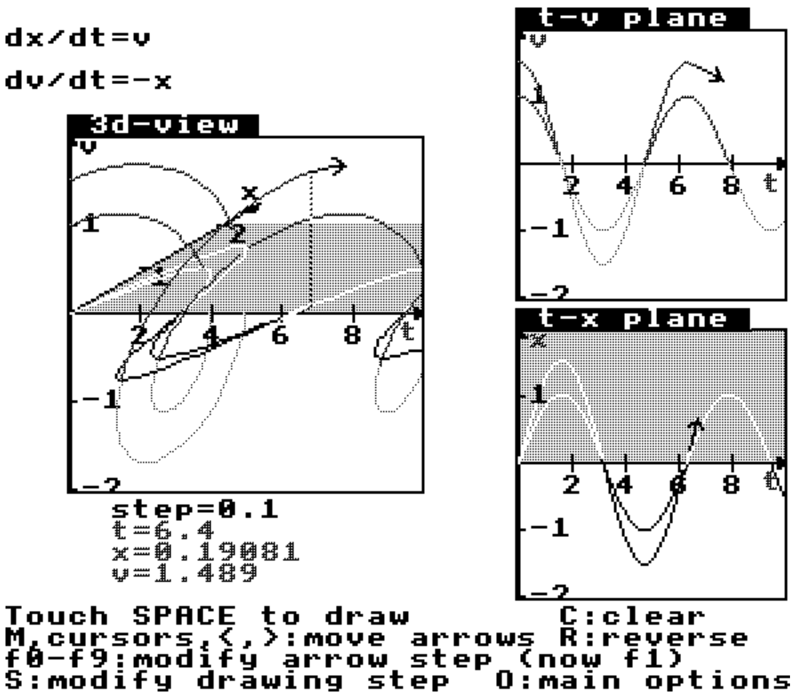


figure 10

(dt,dx,dv) in the direction (1,v,–t). Hence the idea of a direction field *does* generalize, but it must be visualized in three dimensional (t,x,v) space. Figure 10 shows two solutions of the differential equation spiralling through (t,x,v) space and their projections onto the t–x and t–v planes.

Visual exploration in geometry

Euclidean geometry traditionally served to introduce students to a deductive system. In many countries (such as the United Kingdom) it has all but disappeared from the mathematics curriculum. Computers now give the opportunity to manipulate geometrical figures to build up intuitions for possible theorems (*Geometric Supposer* Schwartz & Yerushalmy, 1985, *Cabri Géomètre* 1987). The initial phase of study of geometry can now be an experimental science, in which the student can use the computer to construct a figure and experiment with it.

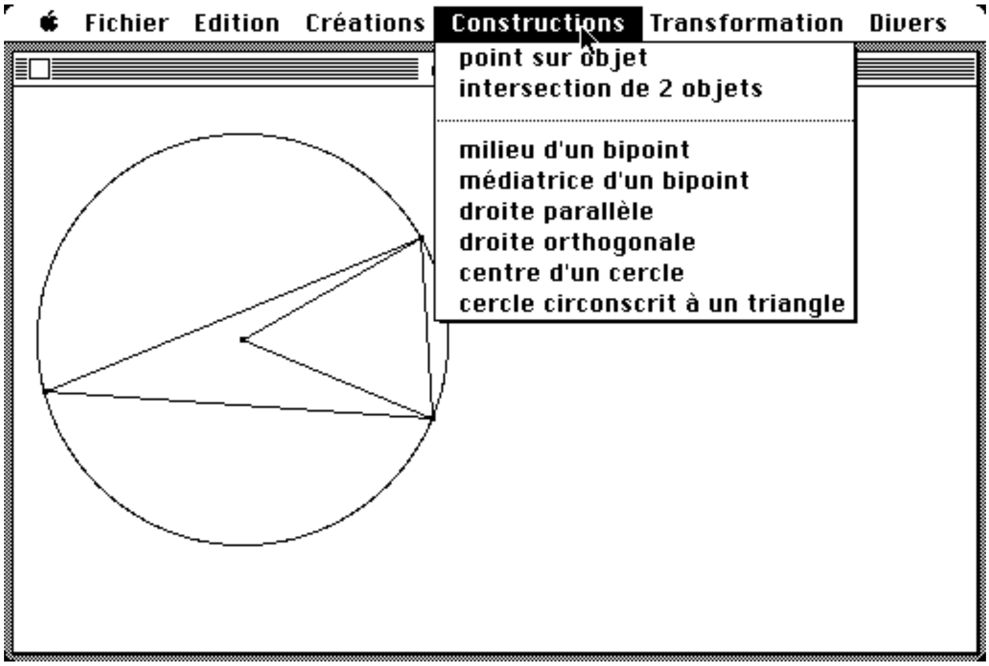


figure 11

Visual Data Processing

It is now possible to explore data visually, for example, to see a line of best fit for data in two or three dimensions. *MacSpin* allows up to ten categories of data, from which any three can be selected and displayed. Though only represented as a projection of three dimensions onto the two-dimensional screen, the data may be rotated and viewed dynamically from any angle to give a sense of depth that is not visible in a static picture (figure 12). Individual points may be selected and inspected to see where the data originates to identify interesting information, such as outlying values. Rotating the data in the figure suggests that it clusters together in a way which intimates that the three components are correlated.

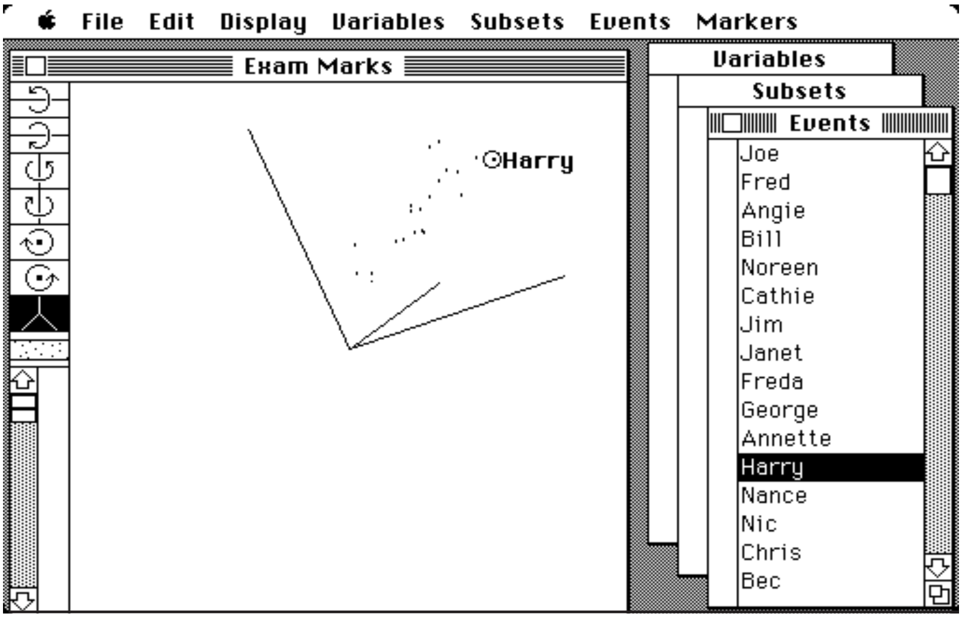


figure 12

Modern spreadsheets, statistical packages and data handling packages now include visual representation of data which encourages the user to explore and communicate complex information in visual ways.

The ability to present and manipulate information visually is becoming widely available in a many different areas in mathematics. For example, Robinson & Bowman (1986) introduce probability and statistics using computer graphics with the intention of giving a 'feel' for probability distributions rather than go into details of the mathematics. More generally, the *Computer Illustrated Texts* (starting with Harding 1985) are designed to use simple computer programs to provide interactive illustrations of mathematical ideas which can be explored by the student in place of static pictures in a book.

Is programming essential?

We have not explicitly mentioned programming so far. In the U.K. a growing body of expertise is growing in which students are expected to handle short programs (usually in structured BASIC or Logo) to carry out mathematical algorithms. From here it is intended that they move on to prepared software that uses the underlying algorithms in a more interactive manner. The early computer illustrated texts assumed that the programming would be sufficiently simple that it would allow the student to modify the programs, but this became an impossible ideal in later texts. Programming requires a serious investment in time and effort. However, it can pay vast dividends if the investment is sufficiently generous.

Dubinsky has evidence that having students make certain programming constructions (in the computer language ISETL) can lead to their making parallel mathematical constructions in their minds and thereby come to understand various mathematical concepts (see, for example, Dubinsky & Schwingendorf 1990). Clearly a spectrum of approaches may be possible with varying amounts of programming, depending on the time and commitment available.

New Styles of Learning

Software is becoming widely available to give graphical representations in calculus, differential equations, geometry, data handling, numerical analysis, and many other areas in mathematics. This is usually predicated on a new kind of learning experience — one in which the student may explore and manipulate the ideas, to *investigate* patterns, to *conjecture* theorems and to *test* their theories experimentally before going on to *prove* them in a more formal context.

For instance, in the calculus students may investigate the gradients of functions such as sine, cosine, tangent, exponential and logarithm, and conjecture their formulae before they are derived formally (Tall 1986, 1987). In differential equations they may explore problems at the boundaries of research (such as the rings of Saturn) and make the mental link between the friendly world of (mostly linear) equations that can be solved by formulae and the strange world of those (usually non-linear) that can not (Hubbard and West 1990).

This form of learning is not a *replacement* for formal deduction, but a *precursor* and a *complement* to it. It enables the less able student to grasp essential ideas that would previously be too difficult when framed in a purely formal theory and for the more able student to build a cognitive base for the formal theory to follow. It enables a wide range of students to integrate their knowledge structure through their powers of visualization.

References

Artigue M. & Gautheron V. 1983: *Systèmes Différentielles: Étude Graphique*, CEDEC, Paris.

Cabri Géomètre 1987: IMAG, BP 53X, Université de Grenoble (for the Macintosh computer).

Cornu B. 1981 : 'Apprentissage de la notion de limite : modèles spontanés et modèles propres', *Actes due Cinquième Colloque du Groupe International P.M.E.*, Grenoble 322-326.

Dubinsky E. & Schwingendorf K. E. 1990: "Constructing calculus concepts: cooperation in a computer laboratory", *MAA Notes Series*, (ed. Leinbach C.), Math. Assoc. Amer.

Gleick J. 1987: *Chaos: Making a New Science*, Penguin.

Harding R. 1985: *Fourier Series and Transforms, A Computer Illustrated Text* Adam Hilger, Bristol & Boston (for the BBC, IBM and Apple Computers).

Hubbard J. H. & West B. 1990 : *Differential Equations* (with software for the Macintosh computer).

Koçak H. 1986: *Differential and Difference Equations through Computer Experiments*,Springer-Verlag (for the I.B.M. Computer).

MacSpin 1985: D^2 Software Inc., Austin, Texas (for the Macintosh Computer).

Phillips R. 1988: *Mouse Plotter*, Shell Centre, Nottingham (for the Archimedes computer).

Robinson D.A. & Bowman A.W. 1986: *Introduction to Probability*, Adam Hilger, Bristol & Boston (for the BBC and I.B.M computers).

Schwartz J. & Yerushalmy M. 1985: *The Geometric Supposer,* Sunburst Communications, Pleasantville, N.Y. (for the Apple Computer).

Takagi T. 1903: A simple example of a continuous function without derivative, *Proc. Phys.-Math.* Japan, 1, 176-177.

Tall D. O. 1982: The blancmange function, continuous everywhere but differentiable nowhere, *Mathematical Gazette*, 66, 11-22.

Tall D. O. 1986:*Graphic Calculus I -III*, Glentop Press, London (for BBC compatible computers).

Tall D.O. 1987:*Readings in Mathematical Education: Understanding the calculus*, Association of Teachers of Mathematics (collected articles from ‘Mathematics Teaching’, 1985-7)

Tall D.O. 1990: *Real Functions & Graphs: SMP 16-19*, Rivendell Software (for BBC compatible computers), prior to publication by Cambridge University Press.

Tall D.O., Blokland P. & Kok D. 1990: *A Graphic Approach to the Calculus*, Sunburst Inc, USA (for I.B.M. compatible computers).

Tall D.O. & Vinner S. 1981: ‘Concept image and concept definition in mathematics, with special reference to limits and continuity’, *Educational Studies in Mathematics*, 12 151-169.