Visualizing Differentials in Integration to Picture the Fundamental Theorem of Calculus

David Tall

Mathematics Education Research Centre
University of Warwick
COVENTRY CV4 7AL

Introduction

What is a differential? From my investigations asking sixth-formers arriving at university about concepts in the calculus, the evidence shows a manifest confusion in the meaning of notations such as $\frac{dy}{dx}$ and $\int f(x)\,dx$. Many students repeat the received wisdom that $\frac{dy}{dx}$ means the derivative of a function $y=f(x)$ and should be thought of as a single indivisible symbol, not as a quotient. Those who do give dx a meaning as a separate entity invariably talk of a "very tiny change in x" or an "infinitely small change in x" or even "the limit of δx as it tends to zero". Meanwhile the dx in $\int f(x)\,dx$ means "with respect to x", though students need to be willing to make the substitution $du = \frac{du}{dx}\,dx$, to compute the integral by substitution. A little later they may be faced with the problem to solve the differential equation

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

(where they had been told that $\frac{dy}{dx}$ is an indivisible symbol) by "separating the variables" to get

$$y\,dy = -x\,dx$$

(what does the dx mean here?) then put an integral sign in front to get

$$\int y\,dy = -\int x\,dx$$

(where presumably dx now means "with respect to x"), to obtain the solution(s)

$$\frac{y^2}{2} = -\frac{x^2}{2} + c.$$

Why is it that we seem to teach what should be a logical and clearly defined subject in such a perverse and mystical way? Perhaps it is simply that we belong to a mathematical community and have learned to repeat the litanies of our youth that gave

Published in *Mathematics Teaching*, **137**, 29–32 (1991)

us a passport to become fully fledged members. I confess that, along with most of my colleagues, I learned to cope with the routines and got the right answers until I ceased to ask awkward questions about the meanings of the symbolism.

A perusal of calculus textbooks in current use shows some very neat ways of sidestepping the fundamental issues of meaning, but few seem to have a totally coherent view of the meaning of differential that works throughout the calculus.

In an earlier article[1] I discussed the visual interpretation of the differential in differentiation, but at the time had not yet appreciated a possible analogous use of the differential dx in the integral $\int f(x)\,dx$.

In this article I will show that there is indeed a simple meaning that can be given to dx which works in both differentiation and integration – a meaning that was given by Leibniz in his first publications on the calculus[2,3] – a meaning that takes on a new and vigorous life in our modern computer age. For three hundred years Leibniz has been maligned in the English-speaking world, first as a man who stole Newton's theory of the calculus, and then as someone who invented an incredibly useful but curiously mystical notation. He deserves better treatment. All it requires is to look at the right pictures.

1. The differential in differentiation

This is an easy matter to explain and has been well represented in the literature (see, for example, Quadling[4]). The derivative of a function f at a point x is found by calculating

$$\frac{f(x+h)-f(x)}{h}$$

and considering what happens as h gets small. If it tends to a limiting value, this is denoted by $f'(x)$ and called the *derivative* of f at x. It is, of course, the gradient of the tangent to the graph at x. If dx is any real number, then dy is defined as

$$dy = f'(x)\,dx.$$

This simply says that the tangent vector is in the direction (dx, dy) (figure 1).

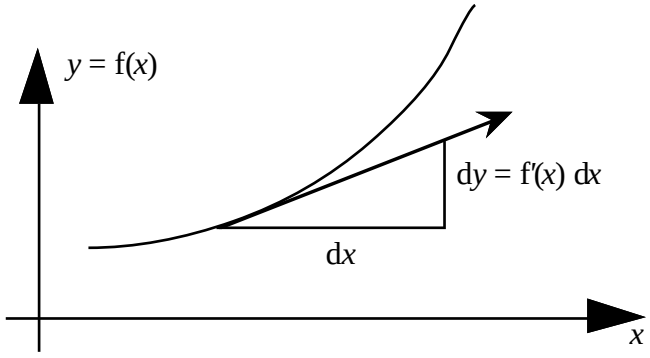

figure 1: dx and dy as components of the tangent vector

Computer graphics now help us visualize that, under high magnification, a suitably small portion of the graph of a differentiable function will look straight. Thus, when dx is suitably small, then dy is a good approximation to f(x+dx)–f(x) (figure 2).

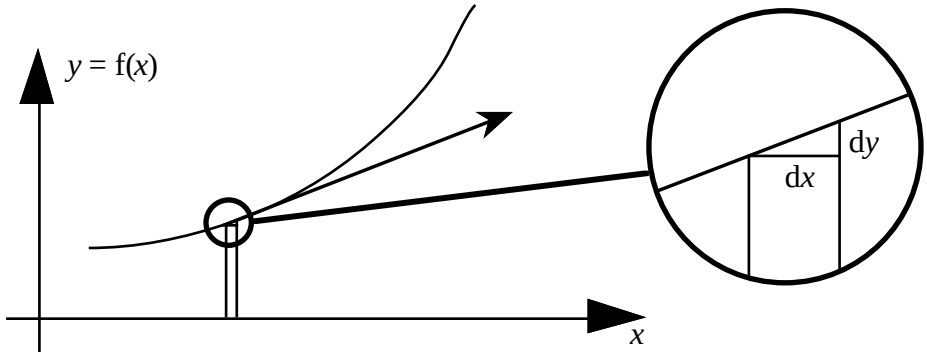

figure 2: magnifying a locally straight portion of the graph of a differentiable function

This insight will prove valuable in linking this use of the differential with that in integration.

2. The differential in integration

Let us begin by attempting to give the differential in integration the same meaning as in differentiation: dx represents a change in x. We will use the notation $\sum_{x=a}^{b} f(x)\,dx$ or $\Sigma_a^b f(x)\,dx$ to denote the sum of the areas of strips width dx, height f(x) between $x=a$ and $x=b$. (In Britain, the more usual symbol is $\sum_{x=a}^{b} f(x)\,\delta x$, where δx denotes a small change in x, but this extra piece of symbolic baggage is not absolutely necessary. It arose in a Cambridge textbook in 1803[5] to distinguish between the value of an increment δx before taking a limit and an infinitesimal change dx "in the limit". Since then it has been used to distinguish between a finite sum and the limit of the sum as the width of the strips diminish in size. We will avoid this notational difficulty shortly by replacing the Σ in the finite sum by the $\int$ in the limit. No other notational change will be necessary.

3. Visualizing the Fundamental Theorem

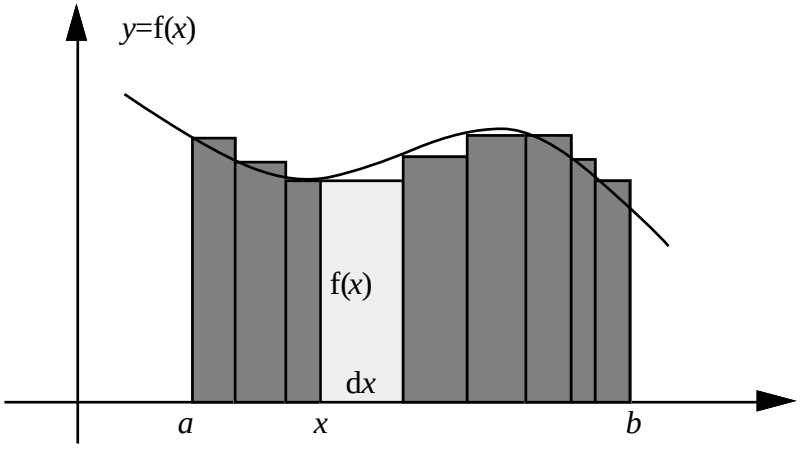

figure 3 : The area sum Σ_a^b f(x) dx

To understand the relationship between the use of d*x* here and that in differentiation, all we have to do is to *look at the right picture*. Figure 3, which for so long has been the standard view is not the right one for the fundamental theorem. We must look not at the graph of $y=f(x)$, but at the graph of $y=I(x)$ where $I'(x)=f(x)$. We will assume for the moment that such a function I can be found, returning to the conditions under which it exists in the next section. Given such an I where $I'(x)=f(x)$ we simply draw the corresponding picture for $y=I(x)$ with the same subdivision of the interval $[a,b]$ into sub-intervals (figure 3).

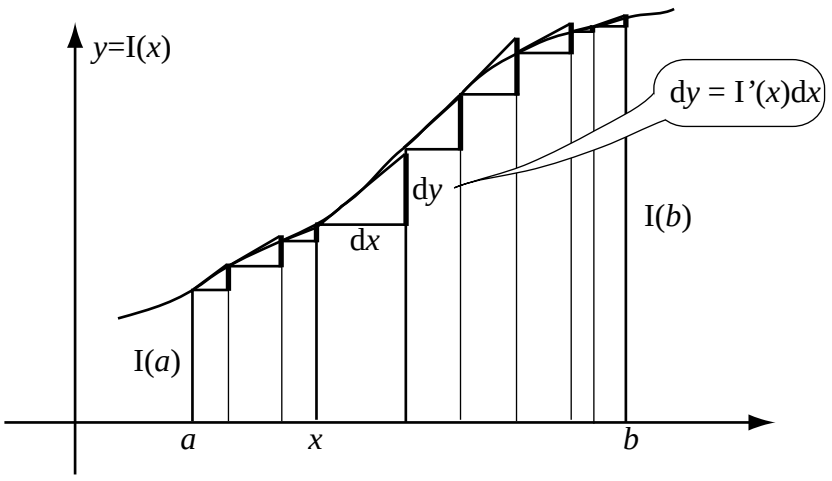

figure 4: The sum Σ_a^b f(*x*) d*x* as a sum of lengths Σ d*y*

At each point *x* of the subdivision we draw the tangent to the curve I(*x*), then the corresponding increment to the tangent is

$$dy = I'(x)\,dx = f(x)\,dx.$$

Thus the sum Σ_a^b f(x) dx is seen as the sum of the lengths Σ dy where each dy is the vertical component of the tangent vector to the graph of y=I(x). The sum Σ dy is the sum of the vertical line segments, and, provided that the dx are taken to be small so that the graph is relatively straight from x to x+dx, then this is approximately equal to the increment to the graph, I(x+dx)–I(x). Adding together the increments *to the graph* from x=a to x=b simply gives I(b)–I(a). Thus adding together the vertical steps dy may in some sense approximate to I(b)–I(a). Figure 4 gives a fair indication of this idea. It fails in part because we had to make the strips fairly wide to see what is going on and this, in turn, means that the graph may be so curved in a strip that dy is clearly different from I(x+dx)–I(x). But imagine it instead as having a large number of strips, and imagine a part of the graph being magnified to see its local straightness with a few strips next to each other (figure 5).

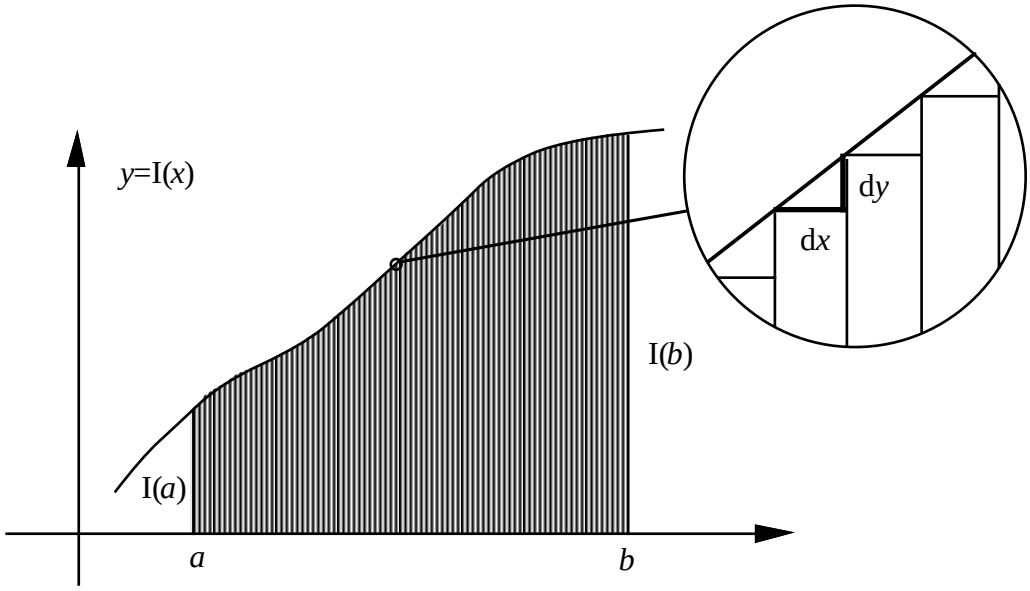

figure 5: looking closely at the summation process

Now it should be possible to imagine that as the lengths dx get smaller, the sum of the lengths Σ dy approximates to I(b)–I(a).

The picture only gives a *sense* of what is going on. But the zooming-in process should hint at something more. As one zooms in, the curved graph gets *less curved*. Students who use a graph-plotting program readily appreciate this phenomenon. It was the first thing that the first students to play with *Graphic Calculus* observed without any prompting.

If we accompany the zooming-in process by taking a smaller value of dx, the corresponding value of dy more closely approximates the step up to the curve. We should be able to see that the *relative error* in the difference between the value of dy

and the actual step to the curve gets less. In this way we get a hint that the errors are now small *in proportion* to the vertical step, so adding them together, the total error is also now small in proportion to the total vertical step. In other words we may begin to see that, as the lengths dx get smaller, the sum of the vertical steps, $\Sigma\, \mathrm{d}y$, gets closer to I(b)–I(a).

The symbol $\int_a^b \mathrm{f}(x)\, \mathrm{d}x$ is used to represent the limit of $\Sigma_a^b \mathrm{f}(x)\, \mathrm{d}x$ as (the maximum size of) dx gets small. The argument just given provides a powerful intuition as to why this limit is likely to be:

$$\int_a^b \mathrm{f}(x)\, \mathrm{d}x = \mathrm{I}(b)-\mathrm{I}(a).$$

4. Proving the Fundamental Theorem

Although we may now have a *sense* of why the fundamental theorem of calculus might be true, this does not yet constitute a *proof*. For instance, what are the conditions on the function f which would *guarantee* that there is a function I such that I′ = f ? Until we can establish this we cannot even draw figures 4 and 5, because we cannot be sure of the existence of the function I. I have earlier shown how one might visualize the likely properties required of f by looking at the graph of f in a different way[6,7]. I suggest that interesting pictures might occur by maintaining a constant y-range whilst taking a much smaller x-range. For instance, figure 6 shows the graph of y=sin$x$ with the same y-range in each (–3 to 3) but x-range being changed from –3 to 3 down to 1 to 1.01. What happens is that the graph in the second case is pulled flat by the stretching of a thin x-range to fill the computer window.

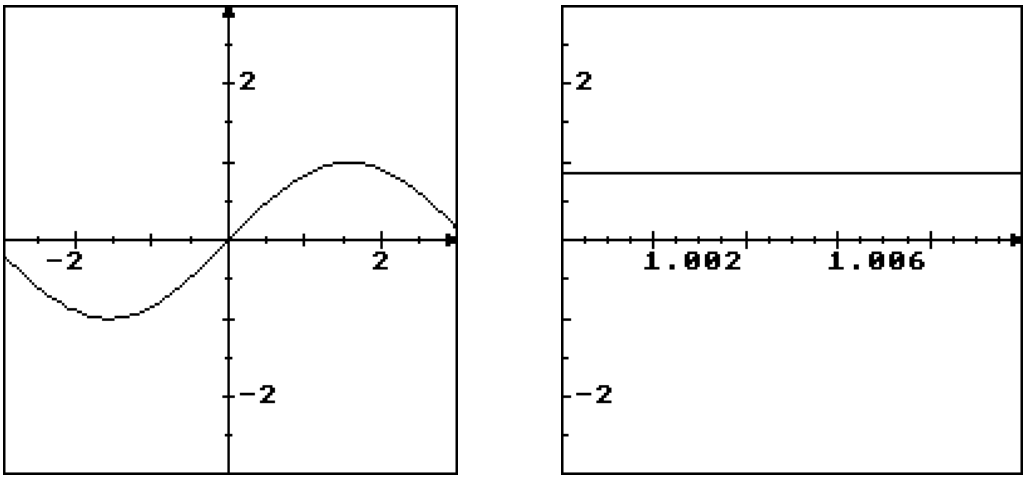

Figure 6 : The graph of y=sin$x$, pulled out flat

If one calculates the area under a flat graph like this, the area from x to x+h is approximately f(x)h. This represents a change in area from A(x) to A(x+h), so

$$A(x+h)-A(x) \approx f(x)h.$$

This suggests that we may have

$$\frac{A(x+h)-A(x)}{h} \approx f(x),$$

and perhaps, as $h \to 0$, we might get

$$A'(x)=f(x).$$

What kind of function, when stretched out horizontally near $x=x_0$, looks flat ? If we suppose this means that the graph lies in a pixel representing a height $f(x)\pm\varepsilon$, then we need to know that, given such an $\varepsilon > 0$, then we can find a small enough x interval, say $x\pm\delta$, so that when t lies between $x-\delta$ and $x+\delta$, then $f(t)$ lies between $f(x)-\varepsilon$ and $f(x)+\varepsilon$. In other words, a natural condition for the function to satisfy the fundamental theorem is that it be *continuous* in the formal sense:

> Given any $\varepsilon>0$, a $\delta>0$ can be found such that whenever $x-\delta < t <x+\delta$ we know that $f(x)-\varepsilon < f(t) < f(x)+\varepsilon$.

For such a function if the width of a strip h is taken positive and less than δ then the value of $f(t)$ will lie between $f(x)-\varepsilon$ and $f(x)+\varepsilon$ throughout the strip and so

$$(f(x)-\varepsilon)h \ < A(x+h)-A(x) < (f(x)+\varepsilon)h.$$

Hence

$$\frac{A(x+h)-A(x)}{h}$$

is sandwiched between $f(x)-\varepsilon$ and $f(x)+\varepsilon$. A similar argument holds for h negative. As ε is arbitrary, this is the formal definition that

$$\lim_{\varepsilon \to 0} \frac{A(x+h)-A(x)}{h} = f(x),$$

i.e. $A'(x) = f(x)$.

Note that this only requires the *continuity* of $f(x)$, not that it be differentiable. The reader who has completed a university mathematics degree will doubtless know this theorem. But can you *see* why it is true with your inner eye?

It is easy to simulate the idea on the computer. Figure 7 shows the blancmange function, which is everywhere continuous and nowhere differentiable. Differentiability

is exhibited by seeing that when magnifying the graph it never looks straight (figure 7) and continuity by stretching the graph horizontally to see it pull out flat (figure 8) [8,9].

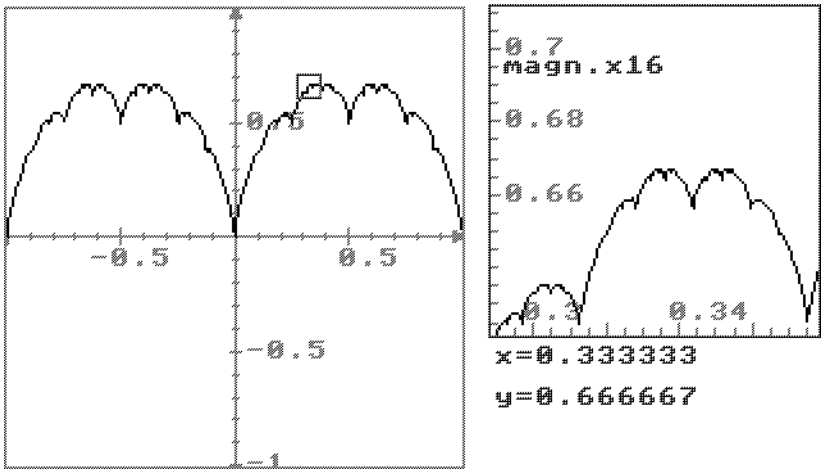

Figure 7 : The blancmange function being magnified near x=1/3

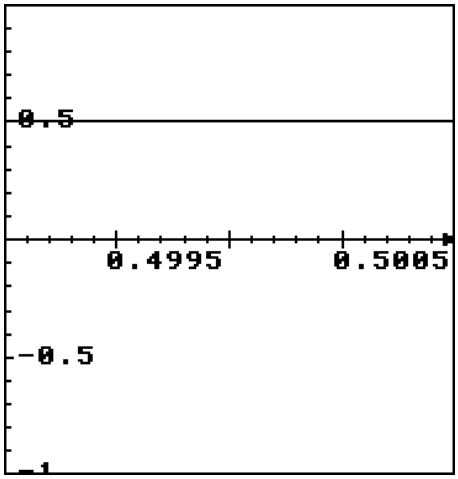

Figure 8 : The blancmange function being stretched horizontally from x=0.499 to x=0.501

Figure 9 shows two graphs[10]: one – the smooth one – is the numerical area a(x) under the blancmange function, calculated using the mid-ordinate rule with strips width 0.1. This is smooth and (within the accuracy of the approximation) locally straight (as can be seen by magnifying it). The other looks like the blancmange function, but it is not. It is the (numerically calculated) gradient function of the numerical area. It is only an approximation, but it clearly intimates that the area function a(x) is differentiable once to get the blancmange function. Thus the area function a(x) is differentiable everywhere once but nowhere twice...

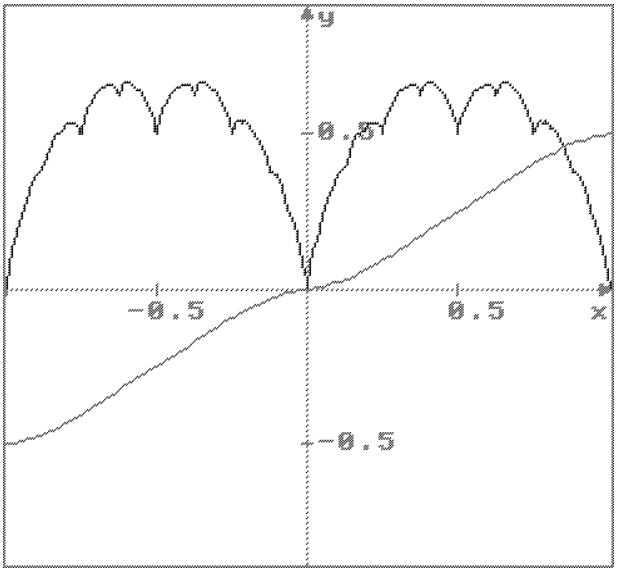

Figure 8 : The area function for the blancmange and its (approximate) derivative

Using *Real Functions and Graphs* [10] on the Archimedes computer, it is possible to draw (a numerical approximation to) the area function for a(x), which is a graph b(x) whose derivative is a(x) and whose second derivative is the blancmange function. This builds an approximation to a graph which is everywhere differentiable twice, but nowhere three times, even though its second derivative is continuous. It is relatively boring to draw, yet illustrates the concept which is otherwise hard to imagine: the possibility of a graph differentiable twice but not three times, or, more generally, n times but not n+1 times.

Leibniz would have never visualized such monsters. In fact, when they appeared in the nineteenth century, they were generally regarded as unintuitive. But, given appropriate computer environments for exploration, their properties become more natural and they become integral parts of the theory. We can now see continuity arising as a natural ingredient of the Fundamental Theorem of the Calculus rather than being a theoretical pre-requisite imposed for some more esoteric reason. And, at last, we see the intuition of Leibniz, which was exhibited in the very first published articles on the calculus, yet vilified for three hundred years in the English press, appearing as a natural conception in a modern computer approach.

References

1. Tall D. O. 1985: “Tangents and the Leibniz notation”, *Mathematics Teaching*, 112 48-52.

2. Leibniz G. W. 1684: “Nova methodus pro maximis et minimis, itemque tangentibus, qua nec fractas, nec irrationales quantitates moratur, & sinulare pro illis calculi genus”, *Acta Eruditorum* 3, 467-473.

3. Leibniz G. W. 1686: “De geometria recondita et analysi indivisibilium atque infinitorum”, *Acta Eruditorum* 5.

4 . Quadling D.A. 1955: *Mathematical Analysis*, O.U.P., Oxford.

5. Woodhouse R. 1803: *The Principles of Analytical Calculation*, Cambridge.

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

7 .Tall D.O. 1986: *Graphic Calculus II* (for BBC, Nimbus, Archimedes computers), Glentop Press, London.

8. Tall D. O. 1986: *Graphic Calculus I* (for BBC, Nimbus, Archimedes computers), Glentop Press, London.

9. Tall D. O., Blokland P. & Kok D. 1990: *A Graphic Approach to the Calculus*, (for IBM compatibles) Sunburst, Pleasantville, NY.

10. Tall D. O. 1991: *Real Functions and Graphs*, (picture drawn using the Archimedes computer), C.U.P., Cambridge.