A Long-Term Learning Schema for Calculus and Analysis[1]

David Tall

Mathematics Institute, University of Warwick

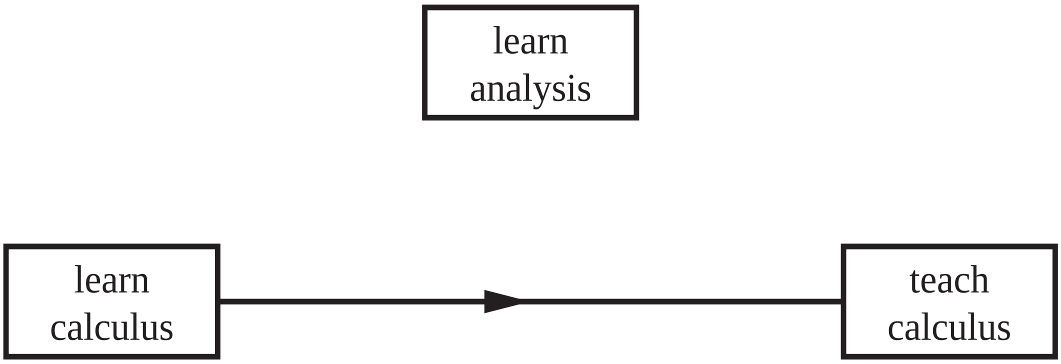
1. The basic problem

The understanding and teaching of calculus in schools and analysis in universities have long been considered subjects of great difficulty. The Assistant Masters Handbook [6] says “its ideas are so novel that there must be no attempt to rush the early stages”. Vast numbers of textbooks are available, seemingly covering every conceivable approach, but many problems remain.

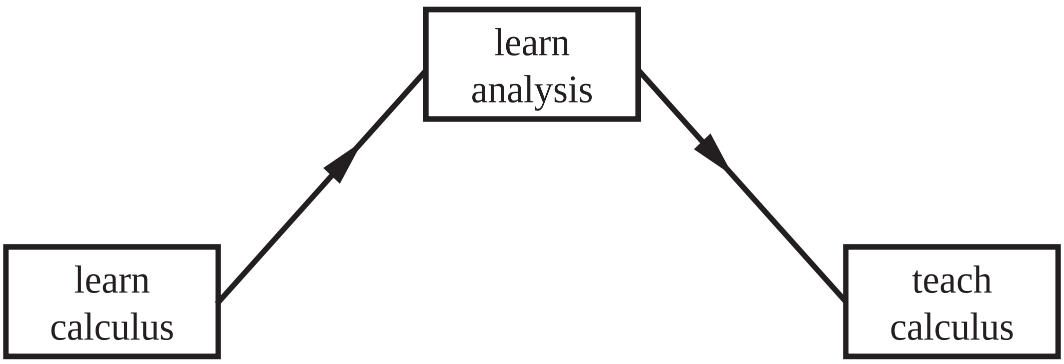
Most solutions restrict themselves to one particular area of development, a school text, a university lecture course and so on. In this article we look at the possibility of establishing a long term *schema* which unites an intuitive approach to the calculus with a rigorous approach to analysis arising naturally from the earlier intuition: It cannot be a universal panacea, because the complexity of the subject remains, but it is hoped that the approach will be more coherent and realistic at each stage of development. Its origin lies in the realisation that analysis as taught in most colleges and universities is totally inappropriate for the teaching of calculus in schools. Analysis done badly can involve a logical, analytic approach which destroys the student’s faith in geometric intuition and drawing pictures. A student who has done an initial course in calculus can differentiate and integrate many complicated functions; he adds very few more to his repertoire in analysis. Instead he spends endless time on long and arduous proofs of results which seemed intuitively obvious, often ending up confused about ideas which he thought he understood before the course started. If he becomes a teacher he finds that the sophistications of analysis are inappropriate for teaching calculus, but he is now worse off than when he was at school because he has developed inner phobias about the subject. So he totally rejects analytic ideas and returns to teaching calculus in the manner he learnt in school, now coloured with apprehension at the back of his mind. He can easily pass his fear on to his pupils and the cycle of difficulty is set up

[1] First published in *Mathematical Education for Teaching* 2, 2, 1975.
This was my first published article on the teaching of calculus and analysis. At the time I was a mathematics lecturer steeped in formal mathematical ideas, anxious to communicate them to undergraduates. Since that time my views have changed to a more realistic consideration of what students can achieve. However, this article contains the seeds of important ideas, so it is reproduced here with all its strengths and manifold weaknesses to reveal the starting point of various theories. *David Tall*

again in the next generation. It is this cycle which lecturers training future teachers must break. It can be represented by the diagram:



Standard solutions to this syndrome usually involve attempts at simplifying the sophisticated analytical notions to a level that college students might understand. This is not only demeaning to the students, it is also no answer to the problem because it rarely relates analysis fully to the ideas of calculus used in schools. Such a solution is simply putting the cart before the horse because calculus as a technique precedes analysis both historically and within the development of the subject in schools. Instead of forcing analytic ideas down into calculus, surely it might be more appropriate to look at calculus and see how intuitive geometric ideas can be refined to give rigorous notions in analysis. Instead of trying to *replace* the calculus schema by a formal analysis schema, leading to conflict between the two, the whole system should be realigned so that the concepts of the calculus are *accommodated* within a broader schema of analysis. Given such an approach, one might hope that the study of analysis could put the future teacher in a position to revitalise the teaching of calculus as represented in the following diagram:



2. Extrapolating Piaget

In teaching mathematics we are increasingly giving weight to the developmental level of the learner. For example we realise that in the concrete operational period (approximate age range 8–11 years) that logical thought of a sort is possible, provided that it is rooted in concrete experience. After this period children enter a stage of development of formal thought, where abstract logical deduction becomes increasingly possible.

Formal methods of thought are essential to the study of analysis. For example there are many instances of proof by contradiction; this involves supposing that the impossible could happen and showing that such an assumption leads to a contradiction. A method of proof like this is beyond the realms of anyone who cannot detach his thinking from actual concrete experience. It follows that much of analysis is beyond his understanding.

Let us extrapolate Piaget's theories a little with specific reference to mathematics, according to the experience of teachers in colleges and universities. We will postulate the existence of a stage of development beginning when isolated examples of formal deductions become possible. As an observation of the problems of college and university students I suggest that there is a stage where isolated deductions are possible, even short chains of such deductions, but there comes a point where enough formalism becomes too much and the mind withdraws from the contest. This corresponds to the plea "I can follow the proof step by step but I don't understand it."

In a sense this is a formal analogue of what happens in the earlier period of intuitive thought and I hope the expert reader will bear with me as I explain the analogy. If a child at this stage sees a glass of lemonade poured into two smaller glasses he says there is still the same amount, but poured into a much larger number of glasses (four or five) he is seduced by the visual impact into saying there is *more* lemonade. This is a "catastrophe" situation in which there are conflicting forces acting on the child, his external visual experience and his growing internal conviction that the amount must be conserved. Before the period of intuitive thought only the former experience operates, but during this period the latter comes into play as a conflicting factor. With only two glasses his internal feeling is uppermost and his visual experience is not overstretched; but increase the number of glasses and the effect of the visual experience becomes more significant. His internal feeling is tested until he snaps and jumps to his old belief of non-conservation. This seemingly irrational jump from one belief to the other is the natural consequence of a mind under two conflicting pressures where first one pressure is adhered to and then a steady increase in the other can cause a sudden jump to it (called a catastrophe).

In the later stages of formal operational thought we may be in a similar position. At an earlier stage the learner is content to follow his intuitive understanding, but with the coming of greater maturity the necessity for formal proof becomes increasingly apparent. Just as there is a transitional period of intuitive thought in earlier life where conflict occurs, all the ingredients are present once more in this formal stage. Faced with a mathematical fact that is intuitively evident, the growing need for formal proof can be acknowledged but if the chain of reasoning in the proof is too long for the learner, he rejects it mentally and is content to rely on his intuition. At this stage a completely formal development of analysis is not viable because the mind is not capable of taking it all in. From experience a large number of first year college and university students seem to be in such a position. Even if the diagnosis is not confirmed by future research, we still have the observation that some students can follow a proof line by line, yet be uneasy about the proof as a whole (especially where contradiction is involved). This may be due to other factors (for instance the sheer cognitive strain of handling too many concepts) but the remedy for the situation may still be the same as that about to be suggested. Just as in the concrete operational period the child's logical thought must be related to concrete experience, so in this formal transition period, logic must be attached to intuitive abstract ideas with chains of formal deductions between such attachments being as short as possible.

As a long-term schema for calculus and analysis therefore the beginnings of calculus should be intuitive, related directly to concrete experience, the transition to analysis in first year degree courses should not be completely formal, but attached at frequent intervals to intuitive experience, and complete formalism should come with increasing mathematical maturity, if and when this capability is developed. In line with the first section of this article we should look at the intuitive ideas of calculus carefully to see if they are capable of progressive refinement to give analytic ideas rather than operate in the opposite direction.

The best place to begin developing a long-term schema is not at the beginning, but at the transition from calculus to analysis with the teachers of the future. Until these teachers are convinced that a new approach is both possible and necessary we cannot hope to change the current situation. It is to this area that the rest of this article is addressed. The ideas should be suitable for first year university, honours B. Ed, and for mathematically inclined sixth-formers. Pass degree students could follow the same course but concentrate more on the intuitive links and the ideas of calculus. Whoever understands this approach is likely to be much better equipped to teach elementary calculus and see its position in a long-term learning schema.

3. Practical drawing

The whole schema grows out of a consideration of the factors involved in drawing graphs. In the first place, no matter how accurately the values of a function are known, a graphical representation has very limited accuracy. A fine drawing pen marks ink lines approximately 0.1 millimetres thick, so marking a "point" really means a blob about 0.1 millimetres across. The theoretical point is really somewhere in the middle of such a blob, so if two theoretical points are much less than 0.1 millimetres apart, the blobs representing them on a drawing are liable to overlap. Even using a large piece of paper with 1 metre representing a unit length, two such points 0.1 millimetres apart on a number line represent numbers which differ by $0.1/1000 = 1/10^4$. To such a scale it is difficult to distinguish points which only differ beyond the fourth decimal place. Accuracy to four decimal places is as much as we can hope for and even this may not be achieved. Using a much larger scale leads to surprisingly little increase in accuracy. If the distance from the equator to the north pole (10,000 kilometres) is taken as a unit length then using a fine drawing pen gives the ratio $0.1/10{,}000{,}000{,}000 = 1/10^{11}$ and a plausible accuracy of at most 11 places of decimals. This means that we have little chance of distinguishing between any real number and its approximation to, say, 12 decimal places, in particular we cannot distinguish between an irrational number and a rational approximation to it. In a drawing we cannot distinguish between the real line of modern analysis and the rational line of the ancient Greeks.

Suppose that we think of a continuous function as one whose graph can be drawn freely without removing the pen from the paper. (If the domain or range is unbounded, we will only be able to draw a finite part of the graph.) Such a function defined on a closed interval $[a,b]$ with $f(a)$ negative and $f(b)$ positive must cross the axis somewhere in between, it is not physically possible to draw it otherwise. For example the function $f(x)=x^2-2$ has this property on the interval [1,2]. If we considered this function to only be defined for rational values of x, then the result would be false. There is no rational number x in [1,2] such that $f(x)=0$, although no drawing of the graph would actually exhibit this property, because the inaccuracy of drawing prevents us from distinguishing rationals from irrationals. Spatial intuition in this case relates to the *real* line, not the rationals. The moral is that in dealing with continuous functions if we want the theorems to align with our intuition, then we must introduce the completeness property for the reals (in some suitable form) *right at the beginning.*

At the same time it is totally at variance with spatial intuition to do this by saying that there are “gaps” on the rational line because numbers like $\sqrt{2}$ are missing. On a physical drawing we cannot distinguish the real line from the rationals, so in an actual picture there are *no* gaps. The simplest way to introduce completeness is to imagine a real number x on the real line and then successively divide the scale into tenths, hundredths, thousandths and so on to read off the value of x to one, two, three, … decimal places. Of course the inaccurate drawing will only allow us to find a few decimal places but our imagination allows us to think that we can get as many places as we like. having found x to n places, i.e. $a_0{\cdot}a_1\ldots a_n$, where a_0 is an integer and $a_1, \ldots, a_n$ are integers between 0 and 9, then the next place is given by

$$a_0{\cdot}a_1\ldots a_n a_{n+1} \le x < a_0{\cdot}a_1\ldots a_n a_{n+1} + 1/10^{n+1}\,.$$

The completeness property can be interpreted as saying that the real numbers consist precisely of infinite decimals which arise in this way. This property allows us to refine our intuition towards a more formal definition of the real numbers later on.

At this stage it would be convenient to talk about convergent sequences. The formulation "we can make a_n as close as we please to the limit ℓ by making n sufficiently large" has unfortunate hidden connotations. For example "close" usually means "near, but not equal to". When later dealing with the limit $\lim_{x\to a} f(x)$, then x is allowed to get near but not equal to a although $f(x)$ can equal $f(a)$. This fine distinction can cause very deep subliminal confusion to students, even to those who seem on the surface to understand it. They often feel that $f(x)$ gets close to the limit but never quite gets there. To avoid the problem, one only has to look at a drawing. If one draws a sequence converging to a limit, the terms not only get close to the limit, because of the built-in inaccuracy of drawing they actually become *indistinguishable* from it. To conceive of the abstract idea of convergence, one just must imagine that this happens, no matter how accurately one draws, realising that if the drawing is more accurate then it may be necessary to draw many more terms before they become indistinguishable from the limit. This can be refined to the usual definition by saying two real numbers x and y are "ε-indistinguishable" if $|x-y| < \varepsilon$. Then $a_n \to \ell$ means:

> given any accuracy $\varepsilon>0$, there exists N such that if $n>N$, then a_n and ℓ are ε-indistinguishable.

Even if you find these words a bit gimmicky, it is still possible to use this approach to get the usual definition.

Given a real number x, then working out x to one, two, three, and more decimal places gives a sequence of rational approximations which tends to x. Within this scheme we can fit decimals which end in an infinite sequence of 9's. For example the sequence 0·9, 0·99, 0·999, … tends to the limit 1, so the infinite decimal 0·999. . . is another way of representing the real number 1. (Teachers who talk about equivalence relations might be interested in this example of equivalence of decimals which represent the same number. The equivalence classes contain either one, or two infinite decimals. The latter occurs with terminating decimals which can also be written ending in an infinite sequence of 9's.)

With this basic foundation it is perfectly feasible to deduce other forms of the completeness property in ways which appeal strongly to geometric intuition. (See [4] or [5].) These include the property that increasing sequences bounded above have a limit, as do decreasing sequences bounded below. The least upper bound property can be linked to intuitive geometric idea, and also follows from the use of infinite decimals. I personally would forget about Dedekind cuts because they assume that the rationals are there and describe the reals in terms of them. Since our spatial intuition suggests properties like the intermediate value theorem which correspond to the reals, not the rationals, it is foolish at an early stage in the development of the subject to construct something which intuitively the students believe to exist anyway. When dealing with graphs and the calculus, the basic number system that students should be asked to believe in is the real line, with its properties of arithmetic, order and completeness. This has the advantage that the basic number system ***R*** has the natural numbers ***N***, the integers ***Z*** and the rationals ***Q*** as subsets.*

4. Continuous functions and drawing graphs

It is significant that most early approaches to calculus do not mention continuity. The fact that a differentiable function must be continuous contributes to this, but the main reason is surely that formal continuity is a difficult topic to understand at an early stage. The intuitive idea that a real function is continuous if its graph can be drawn freely without taking the pencil off the paper is easy to grasp. Little credence is given to this concept because it is not usually properly linked to the formal definition. In fact the two are interchangeable, provided that one interprets them in the appropriate context. (See [5]). The school teacher who understands this need no longer feel uneasy about using the simple geometric idea of a freely drawn graph to describe continuity because it is now part of a coherent long term schema which develops the formal definition later on.

One of the "pay-offs" of this approach is that it clearly labels continuity as a global property where a function is continuous everywhere in its domain. The usual formal approach describing continuity at a point first by the ε–δ definition (or neighbourhoods) often lays too much emphasis on this restricted concept. As such it does not

---

* The current emphasis that the positive integers are different from the natural numbers because the former are equivalence classes of ordered pairs of the latter is a gross mistake. It is only a *temporary* distinction which occurs should one wish to construct ***Z*** from ***N***. Mathematicians invented the notion of isomorphism precisely to describe the idea of the same basic mathematical system turning up in different guises. The natural numbers can be used both for counting and for naming points at equal intervals on a line; this can be considered as the same basic mathematical system having two isomorphic manifestations. Since counting numbers are visibly different from points on a line, this interpretation is in accord with the students' intuition.

accord with the notion of a freely drawn graph which gives the ideas of continuity over an interval. Functions like

$f(x) = 1/q$ when $x = p/q$ in lowest terms, $f(x)=0$ when x is irrational,

which is continuous only at irrational points, are all right when the students' intuition has been refined, but they may be too disturbing to be introduced at the beginning. If at first we consider functions continuous on *connected*[2] sets then their graphs look like those of our naive intuition and if we then restrict to closed bounded intervals we can actually draw the graph on a piece of paper. (See [5]).

5. How difficult?

Here is the rub. What is formally difficult is often closer to intuition and vice versa.

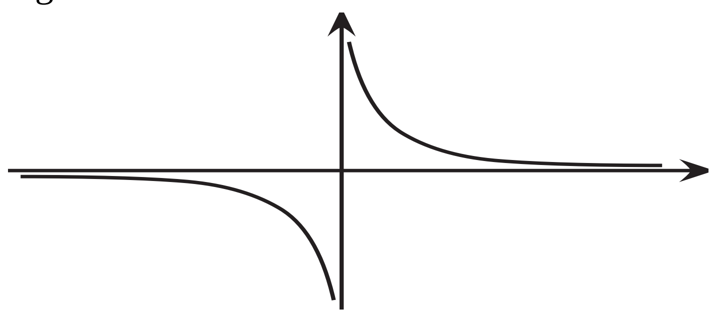
To a mathematician an axiomatic system is often logically simpler if it contains fewer axioms. Because the completeness axiom is a little involved, it may be considered logically simpler to omit it at the beginning of an honours degree course. Formally this means dealing with an ordered field instead of a complete ordered field.

Developmentally using fewer axioms in this manner can be *more* complex because there are many different ordered fields (including the reals, the rationals and other more exotic examples) whereas there is only *one* complete ordered field. If a student is incapable of long chains of deduction and we wish him to retain a link to his intuitive understanding of the real line we must therefore include some form of the completeness axiom.

Likewise the notions of connectedness of a subset and compactness (introduced in a more restricted form using closed bounded intervals) are best introduced right at the beginning. If a set S is not connected, then the graph of a continuous function defined on S can "jump" at missing points e.g.

$f(x) = 1/x \qquad S = \boldsymbol{R}\backslash\{0\}$

jumps at the origin.

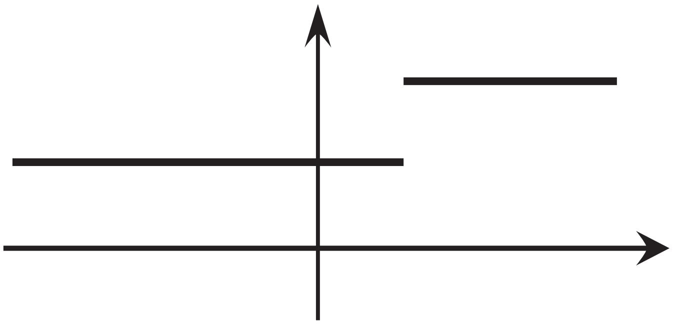


[2] A set $S \subseteq \boldsymbol{R}$ is connected if given $a, b \in S$, $a < x < b$, then $x \in S$. This definition generalises topologically to the notion of a path-connected set.

More insidiously, if we only deal with rational numbers then the function f:$\mathbf{Q}\to\mathbf{Q}$ given by

$$f(x)=\begin{cases}1 & x<\sqrt{2},\ x \text{ rational}\\ 2 & x>\sqrt{2},\ x \text{ rational}\end{cases}$$

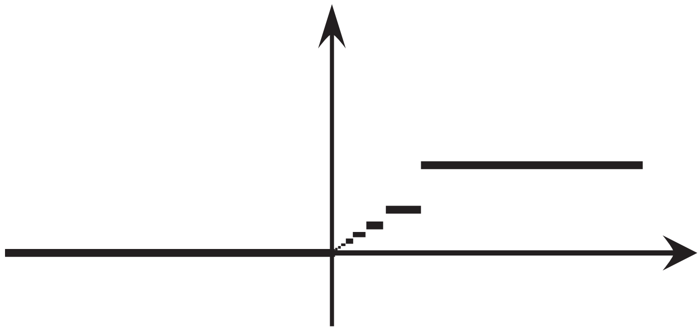
is continuous according to the ε-δ definition:



It jumps over the missing point √2 and there is no reason why we should not concoct continuous functions defined on the rationals which jump at many places.

For example:

$$c(x)=\begin{cases}0 & \text{for } x\le 0\\ 1/(n+1) & \text{for } \dfrac{\sqrt{2}}{n+1}<x<\dfrac{\sqrt{2}}{n},\ n \text{ a positive integer}\\ 1 & \text{for } x>\sqrt{2}\end{cases}$$



Continuous functions on the rationals cannot always be drawn without taking the pencil off the paper, yet in a practical picture we cannot distinguish rational and irrational points! To get the formal idea of continuity to correspond to the practical drawing idea we need to work with the reals, including completeness and connectedness. To be able to draw the graph on a piece of paper we need the domain to be bounded and the range to be bounded. This can be achieved by working over a closed bounded interval.

In practice all the notions of completeness, connectedness, closed and bounded are very simple and directly linked to a student's intuitive idea of the real line. Of course the theorems involved are still as difficult but the concept of continuity now links up with spatial intuition. The student can draw pictures and know how to interpret them. This should lessen cognitive strain and prove psychologically more satisfying.

6. Benefits in Calculus

By throwing in the concepts of completeness, connectedness and closed bounded interva1s early on there are other compensations. Differentiation and integration correspond to the physical concepts of gradient of a graph and area under a graph, and these notions can be made rigorous. It is even possible to draw a graph of an everywhere continuous nowhere differentiable function, remembering that such a drawing is only accurate to two or three decimal places so does not represent the function precisely. Drawing such a graph is in general no harder or easier than drawing the graph of a differentiable function. This is because the pencil used for drawing gives a line of finite thickness and the picture glosses over tiny wrinkles which would occur in a theoretical graph. But it is more intellectually honest. (There are far more continuous, non-differentiable functions than differentiable ones.)

Emphasis on connectedness also refines and corrects a point glossed over in early calculus. We will call a function $F:D\to\boldsymbol{R}$ an *antiderivative* of $f:D\to\boldsymbol{R}$ (where $D\subseteq\boldsymbol{R}$) if $F'=f$. In early calculus it is shown that the derivative of a constant is zero, so $(F + c)' = f$ where c is a constant. It is often wrongly concluded that any other antiderivative of f must be given by adding a constant. This is true over a *connected* domain D, but if D is not connected there may be a different constant over each connected component. This follows trivially from the mean value theorem (see [5] chapter VIII).

7. Continuity, limits and differentiation

Earlier modern approaches to the calculus took the notion of a limit $\lim_{x\to a} f(x) = \ell$ as basic and defined a function to be continuous at a if $\lim_{x\to a} f(x) = f(a)$. Recent developments such as [3] take continuity as being basic and define $\lim_{x\to a} f(x)=\ell$ if the function

$$\varphi(x)=\begin{cases} f(x) & x\neq a \\ \ell & x=a \end{cases}$$

is continuous at a. The excellent book [2] has a special point of view in which the notion of limit is seen as an *extension* of continuity. These are both viable approaches, especially when they are so sensitively handled. But let us close our eyes to analysis and look at the situation in a simple practical drawing. We find a very different picture. A freehand graph suggests global continuity, not continuity at a point and the idea of a limit is a non-event. The latter is because $\lim_{x\to a} f(x)$ requires us to consider values of $f(x)$ for x arbitrarily close to a but *not equal* to it. We *cannot* draw this, we can only imagine it.

For example if $f(x)=\dfrac{x^2-4}{x-2}$ for $x \neq 2$, then we must draw

$$f(x)=\begin{cases}\dfrac{x^2-4}{x-2} & x\neq 2,\\ 4 & x=2.\end{cases}$$

With a pencil marking a finite blob instead of a theoretical point we cannot draw the points $(x,x+2)$ on the graph where x is close to 2, without blobbing over the value for $x=2$ and covering the point (2,4). Try drawing the values for $x = 2{\cdot}001$, $1{\cdot}999$, or if that accuracy can be achieved, repeat for 2 ± 10^{-n} for $n = 1, 2, 3, \ldots$. A physical picture will not get very far (certainly $n \leq 11$ for normal drawing implements) before the missing point (2,4) is marked indistinguishably under the blobs representing $(2\pm10^{-n}, 4\pm10^{-n})$.

Assuming the learner has met the idea of a continuous function being one that can be drawn freely without the pencil leaving the paper, a good approach to the derivative is to draw the graph of the slope function at a:

$$s(x)=\frac{f(x)-f(a)}{x-a} \quad (x\neq a)$$

If f is continuous in D, then s is continuous in D except possibly at $x=a$. If the derivative $f'(a)$ exists we cannot avoid blobbing over the point $(a,f'(a))$ in drawing the graph. Here we have a naturally occurring example of a limit. Many examples put in courses to “motivate” the concept are concocted synthetic cases which have little relevance to the student at the time. Simply by looking at the growing spatial intuition of the learner, a more natural sequence of events grows as follows

(i) Draw graphs.

(ii) Consider intuitive continuity – drawing freely by hand without removing the pencil from the paper.

(iii) Find the gradient $(f(x)–f(a))/(x–a)$ of the graph and draw the graph of this gradient $(x \neq a)$.

By suitably chosen examples the graph of the gradient fills in the "missing value" (a, $f'(a)$) to give the gradient of the tangent.

(iv) This is the time for

(a) Drawing the tangent $y=f'(a)(x-a)+f(a)$ to see that it touches the graph at $x = a$.

(b) A discussion of the intuitive idea of limits and its relevance in formalising the notion of differentiation.

This sequence of events can be treated intuitively in school and later formalised at college by refining the concepts to give formal definitions. More important, by drawing pictures it is possible to face squarely the existence of non-differentiable, continuous functions and give a proof about a page long that such functions exist. (See [5].)

8. Integration

This can be handled easily as a way of calculating the area under the curve provided that we have the concepts:

(a) (uniform) continuity,

(b) completeness (preferably in terms of least upper bounds).

Both of these are covered in the approach recommended. (The policy of stepping in feet first with completeness is again paying dividends.) Almost of the difficulties of formal Riemann integration are avoided if we just integrate (uniformly) continuous functions. In a more formal course, this could be generalised to Riemann integration of bounded functions.

It should be remarked that differentiation and integration are quite separate processes and it is possible to cover integration before differentiation. This is not recommended in a first calculus course because without the process of antidifferentiation (and hence its predecessor, differentiation) it is very arduous to actually calculate the simplest of areas, (try $\int_0^1 x^4 dx$ by straight summation!) In an analysis course however this is quite a good approach to adopt because it emphasises the independence of the two operations and underlines the remarkable nature of the fundamental theorem of calculus which links them together.

The fundamental theorem says that if f is continuous on $[a,b]$ and $F(x)=\int_a^x f$ then F is differentiable and $F'=f$. This allows us to take a continuous, non-differentiable function and integrate it to get a function which is once differentiable but nowhere twice differentiable. By repetition we find that there are functions n times differentiable everywhere and n+1 times differentiable nowhere. a salutary warning against indiscriminate application of Taylor's series .

As an application of the fundamental theorem the natural logarithm can be introduced as $\log x = \int_1^x 1/t\,dt$, and the exponential unction as its inverse. It is possible then to go on to series and power series (including complex values to allow for $e^{ix} = \cos x + i \sin x$) and Taylor's series for infinitely differentiable functions. Here the "error term" in the finite Taylor's series must tend to zero. We have the equipment available (if we so desire) to exhibit a function all of whose derivatives are zero, but which is not equal to its Taylor's series. This is

$$\varphi(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

where

$$\varphi'(0) = \varphi''(0) = \ldots = \varphi^{(n)}(0) = \ldots = 0$$

but

$$\varphi(x) \neq \sum_{n=0}^{\infty} \frac{\varphi^{(n)}(x)}{n!} \quad \text{for } x \neq 0.$$

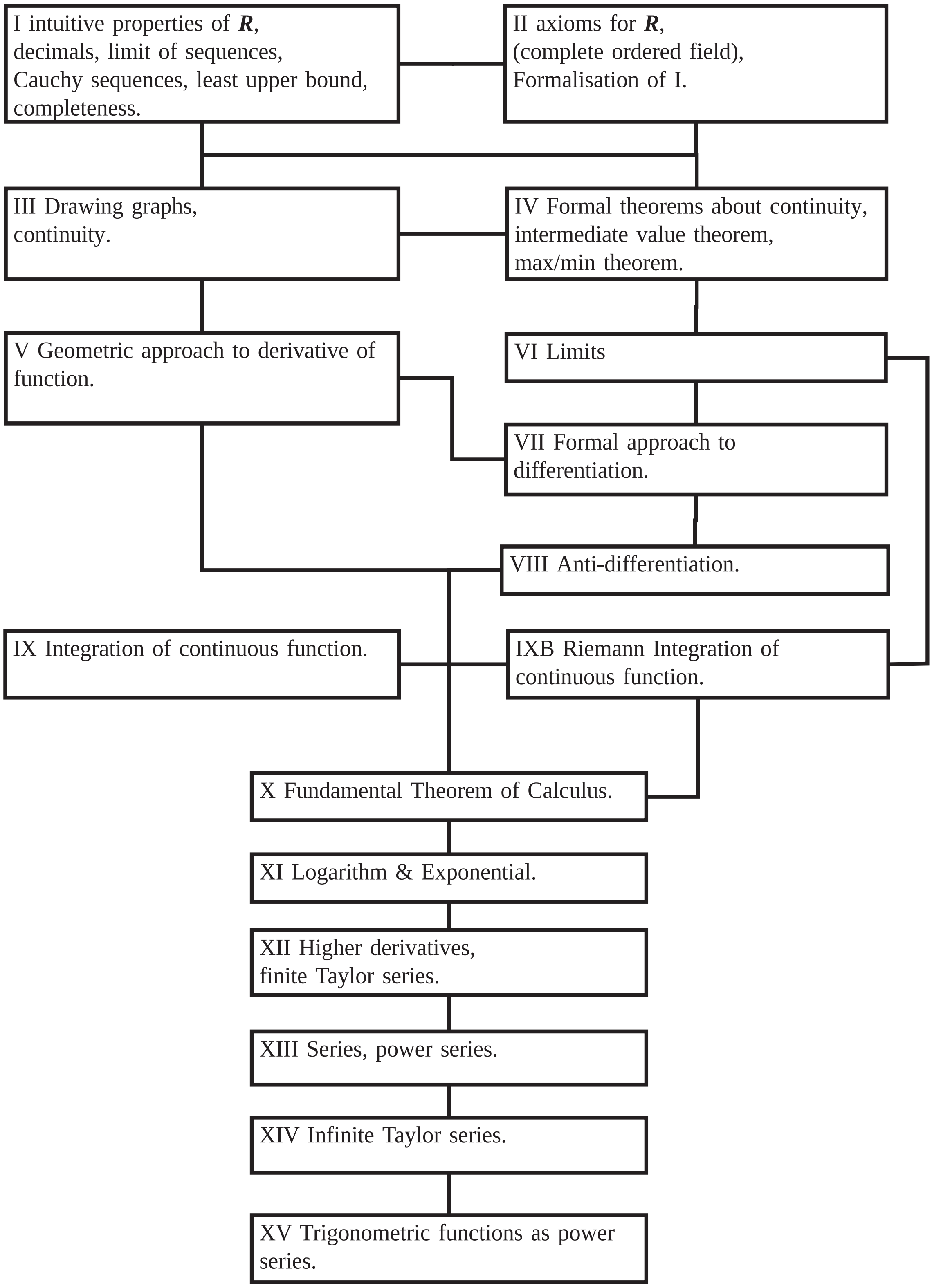
For any function f which has an infinite Taylor's series, the function $f^*=f+\varphi$ has all derivatives at the origin the same as f, but f^* is not equal to its own Taylor's series!

Even at this stage it is possible to return to earth linking up earlier spatial concepts by dealing with trigonometric functions. The basic approach of the course has emphasised inaccuracy of practical measurement, and so it is no surprise to have to look again at the definition of trigonometric functions as ratios of the sides of a right angled triangle. An alternative approach is possible now. In terms of power series, which will give any desired accuracy. A vital part of the course is then to relate this back to the standard definition to make the total development of the calculus/analysis schema a coherent whole.

9. Conclusions

If we try to develop a long term schema for calculus/analysis by refining the intuitive spatial concepts of the calculus into rigorous analysis, then the following sequence of events may occur.

We can begin in school with the idea of a continuous real function as one whose graph can be drawn freely by hand without the pencil leaving the paper. Differentiation can be introduced by drawing the graph of the slope function $s(x) = (f(x)-f(a))/(x-a)$. This is a natural place to discuss the idea of limit. We should continue teaching calculus at school from a geometric point of view, positively developing spatial intuition. At college or university, (especially for future teachers) we can refine the early geometric intuition to develop formal ideas and relate these ideas back again. The method most in accord with spatial intuition would involve looking at the real numbers carefully, contrasting inaccurate drawing with abstract precision and using this physical limitation to motivate the ideas of completeness, limit of a sequence and continuity of a function. The notions of connectedness. and closed, bounded intervals should be introduced from the start so that the formal idea of continuity relates to the intuitive one. Differentiation of continuous functions can be discussed before limits. At this stage it is possible to show how the formal ideas of limit and continuity are really equivalent (and even limit of a sequence is a special case). Differentiation and integration can then be treated formally and related by the fundamental theorem. At regular intervals these formal ideas should be related to earlier intuition. A flow diagram for such an analysis eourse can look like the following. (Pass degree students could omit II, IXB and soft pedal on long formal chains of reasoning such as are found in IV, Vl):


I intuitive properties of *R*, decimals, limit of sequences, Cauchy sequences, least upper bound, completeness.
II axioms for *R*, (complete ordered field), Formalisation of I.
III Drawing graphs, continuity.
IV Formal theorems about continuity, intermediate value theorem, max/min theorem.
V Geometric approach to derivative of function.
VI Limits
VII Formal approach to differentiation.
VIII Anti-differentiation.
IX Integration of continuous function.
IXB Riemann Integration of continuous function.
X Fundamental Theorem of Calculus.
XI Logarithm & Exponential.
XII Higher derivatives, finite Taylor series.
XIII Series, power series.
XIV Infinite Taylor series.
XV Trigonometric functions as power series.

References

1. Beard, R. M., *An Outline of Piaget's Developmental Psychology*, Routledge & Kegan Paul.

2. Flett, T. M., *Mathematical Analysis*, McGraw-Hill, 1966.

3. Moss & Roberts, *A Preliminary Course in Analysis*, Chapman & Hall, 1968.

4. Stewart, I. N., and Tall, D. O., *Foundations of Mathematics*, Oxford University Press: to appear.

5. Tall, D. O., *From Intuition to Rigour in Real Analysis*, available from Warvwick University

6. Assistant Masters Association, *The Teaching of Mathematics in Secondarv Schools*, Cambridge University Press, 1957.

7. Piaget, Jean, *The Child's Conception of Number*, Routledge & Kegan Paul, 1952.