CHAPTER 11

LIMITS AND CONTINUITY

Differential calculus was developed without any explicit definition of either limits or continuity, but with an intuitive assumption that both could in some sense be taken for granted. Widespread use of the calculus during the eighteenth century led to more careful consideration of such matters, but it was not until the early nineteenth century that Bolzano and Cauchy arrived at what are more or less the modern definitions. In this chapter we trace the history of both ideas up to the early 1820s.

11.1 LIMITS

11.1.1 Wallis's 'less than any assignable', 1656

The first writer to work with the concept of a limit in something like the modern sense was Wallis, who in his *Arithmetica infinitorum* in 1656 repeatedly claimed that two quantities whose difference could be made less than any assignable quantity could ultimately be considered equal (see, for example, 3.2.3). In 1656 Wallis stated this as a self-evident fact, but thirty years later, in his *Treatise of algebra*, he attempted to justify it by appealing to Euclidean ratio theory. In the *Elements* Book V (Definition V) Euclid had stated a special property of homogeneous magnitudes (that is, magnitudes of the same kind): given any pair of such quantities, the smaller of them, however tiny, can always be multiplied to exceed the greater. Wallis argued the converse, namely, that if a quantity is (or becomes) so small that it *cannot* be made to exceed a larger quantity, no matter many times it is multiplied, it must be regarded as 'no quantity' or nothing:[1]

> And whatever is so little or nothing in any kind, as that it cannot by Multiplication, become so great or greater than any proposed Quantity of that kind, is (as to that kind of Quantity,) *None at all.*

Wallis then went on to claim something rather stronger: if a difference between two quantities is less than any assignable quantity, then by definition it cannot be multiplied to exceed some given quantity, and therefore by the previous argument it is nothing, and the two original quantities are equal. Again, Wallis claimed Euclid as his authority:[2]

> ...he [Euclid] takes this for a Foundation of his Process in such Cases: That *those Magnitudes* (or Quantities,) *whose Difference may be proved to be Less than any Assignable are equal.* For if unequal, their Difference, how small soever, may be so Multiplied, as to become Greater than either of them: And if not so, then it is nothing.

Though he attributed his arguments to Euclid, Wallis was stretching them considerably further than Euclid or any other Greek author had ever done. The first proposition of Book X of the *Elements* makes the following claim: if from a given quantity there is repeatedly subtracted a half (or more), then what remains will eventually be less than any preassigned quantity. This was crucial to the method of exhaustion; it enables one to prove, for instance, that the space between a circle and an inscribed polygon can be made as small as one pleases by repeatedly doubling the number of sides of the polygon. Nowhere, however, did Euclid or any other Greek mathematician claim that this steadily diminishing quantity could be considered non-existent, or zero. Instead, Proposition X.1 was used in proofs by double contradiction to show, for example, that the space inside a circle was neither greater nor less than some predetermined quantity (see 1.2.3).

Wallis's insight may not have had the classical authority he claimed for it but, like several of his ideas in the *Arithmetica infinitorum*, it was put to particularly good use by Newton.

11.1.2 Newton's first and last ratios, 1687

In the *Principia* in 1687 Newton gave Wallis's idea of 'ultimate equality' the status of a proposition, indeed he made it the opening Lemma of Book I, Section I (see 5.1.2).

At the very end of Section I, Newton introduced the Latin word *limes*, in the everyday sense of a boundary which may not be crossed, just as Barrow had done in 1660 (see 1.2.1). He used 'limes' in a similar sense again in the final sentence when he spoke of quantities decreasing *sine limite*, that is, without end, or indefinitely. Newton also observed that a quantity may approach such a boundary as closely as one pleases; by Lemma I this was equivalent to 'ultimate equality'.

1. Wallis 1685, 281. 2. Wallis 1685, 282.

Newton's idea of a limit

from Newton, *Principia mathematica*, 1687, I, 35–36

[35]

contenta. Præmiſi vero hæc Lemmata ut effugerem tædium deducendi perplexas demonſtrationes, more veterum Geometrarum, ad abſurdum. Contractiores enim redduntur demonſtrationes per methodum indiviſibilium. Sed quoniam durior eſt indiviſibilium Hypotheſis; & propterea Methodus illa minus Geometrica cenſetur, malui demonſtrationes rerum ſequentium ad ultimas quantitatum evaneſcentium ſummas & rationes, primaſq; naſcentium, id eſt, ad limites ſummarum & rationum deducere, & propterea limitum illorum demonſtrationes qua potui breuitate præmittere. His enim idem præſtatur quod per methodum indiviſibilium; & principiis demonſtratis jam tutius utemur. Proinde in ſequentibus, ſiquando quantitates tanquam ex particulis conſtantes conſideravero, vel ſi pro rectis uſurpavero lineolas curvas, nolim indiviſibilia ſed evaneſcentia diviſibilia, non ſummas & rationes partium determinatarum, ſed ſummarum & rationum limites ſemper intelligi, vimq; talium demonſtrationum ad methodum præcedentium Lemmatum ſemper revocari.

Objectio eſt, quod quantitatum evaneſcentium nulla ſit ultima proportio; quippe quæ, antequam evanuerunt, non eſt ultima, ubi evanuerunt, nulla eſt. Sed & eodem argumento æque contendi poſſet nullam eſſe corporis ad certum locum pergentis velocitatem ultimam. Hanc enim, antequam corpus attingit locum, non eſſe ultimam, ubi attigit, nullam eſſe. Et reſponſio facilis eſt. Per velocitatem ultimam intelligi eam, qua corpus movetur neq; antequam attingit locum ultimum & motus ceſſat, neq; poſtea, ſed tunc cum attingit, id eſt illam ipſam velocitatem quacum corpus attingit locum ultimum & quacum motus ceſſat. Et ſimiliter per ultimam rationem quantitatum evaneſcentium intelligendam eſſe rationem quantitatum non antequam evaneſcunt, non poſtea, ſed quacum evaneſcunt. Pariter & ratio prima naſcentium eſt ratio quacum naſcuntur. Et ſumma prima & ultima eſt quacum eſſe (vel augeri & minui) incipiunt & ceſſant. Extat limes quem velocitas in fine motus attingere poteſt, non autem tranſgredi.

F 2 Hæc

[36]

Hæc eſt velocitas ultima. Et par eſt ratio limitis quantitatum & proportionum omnium incipientium & ceſſantium. Cumq; hic limes ſit certus & definitus, Problema eſt vere Geometricum eundem determinare. Geometrica vero omnia in aliis Geometricis determinandis ac demonſtrandis legitime uſurpantur.

Contendi etiam poteſt, quod ſi dentur ultimæ quantitatum evaneſcentium rationes, dabuntur & ultimæ magnitudines; & ſic quantitas omnis conſtabit ex indiviſibilibus, contra quam *Euclides* de incommenſurabilibus, in libro decimo Elementorum, demonſtravit. Verum hæc Objectio falſæ innititur hypotheſi. Ultimæ rationes illæ quibuſcum quantitates evaneſcunt, revera non ſunt rationes quantitatum ultimarum, ſed limites ad quos quantitatum ſine limite decreſcentium rationes ſemper appropinquant, & quas propius aſſequi poſſunt quam pro data quavis differentia, nunquam vero tranſgredi, neq; prius attingere quam quantitates diminuuntur in infinitum. Res clarius intelligetur in infinite magnis. Si quantitates duæ quarum data eſt differentia augeantur in infinitum, dabitur harum ultima ratio, nimirum ratio æqualitatis, nec tamen ideo dabuntur quantitates ultimæ ſeu maximæ quarum iſta eſt ratio. Igitur in ſequentibus, ſiquando facili rerum imaginationi conſulens, dixero quantitates quam minimas, vel evaneſcentes vel ultimas, cave intelligas quantitates magnitudine determinatas, ſed cogita ſemper diminuendas ſine limite.

TRANSLATION

I have put forward these lemmas at the beginning, in order to avoid the tedium of composing intricate demonstrations by contradiction in the manner of the ancient geometers. For the demonstrations are rendered more concise by the method of indivisibles. But since the hypothesis of indivisibles is cruder, and that method therefore judged less geometrical, I have preferred to deduce the demonstrations of what follows by means of first or last sums and ratios of nascent or vanishing quantities, that is, to limits of sums and ratios, and therefore to put forward demonstrations of those limits as briefly as I could. For the same can be shown by these as by the method of indivisibles, and the principles having been demonstrated, we may now more safely use them. Consequently in what follows, whenever I have considered quantities as if consisting of particles, or if I have used little curved lines for straight lines, I do not mean indivisibles but vanishing divisibles, and there should always be understood not sums and ratios of the known parts but the limits of sums and ratios, and the validity of such demonstrations is always to be based on the method of the preceding lemmas.

The objection is that the ultimate ratio of vanishing quantities might not exist; since before they vanish, it is not ultimate; and where they have vanished, it is non-existent. But by the same argument it could equally be contended that the ultimate velocity of a body arriving at a certain place does not exist. For in this case, before the body reaches the place, the velocity is not ultimate; where it reaches it, it does not exist. And the answer is easy. By the ultimate velocity is to be understood that with which the body moves, not before it reaches the final place and the motion ceases, nor after, but as it reaches it; that is, that same velocity with which the body reaches the final place and with which the motion ceases. And similarly by the ultimate ratio of vanishing quantities there must be understood the ratio of quantities not before they vanish, nor after, but with which they vanish. And equally the first ratio of nascent quantities is the ratio with which they originate. And the first or ultimate sum is that with which they begin or cease to be (according as they are increasing or decreasing). There exists a limit which at the end of the motion the velocity may attain, but not exceed. [36] This is the ultimate velocity. And likewise for the limiting ratio of all quantities and proportions beginning or ceasing to be. And since this limit is fixed and definite, the problem is to determine it correctly geometrically. Indeed anything geometric can legitimately be used to determine or demonstrate other things geometrically.

It may also be contended that if ultimate ratios of vanishing quantities are given, so are the ultimate magnitudes; and thus every quantity will consist of indivisibles, contrary to what *Euclid* proved of incommensurables in the tenth book of the *Elements*. But this objection is based on a false hypothesis. Those ultimate ratios with which quantities vanish, are not actually ratios of ultimate quantities, but limits to which the ratios of quantities decreasing without limit always approach, and which they may attain more closely than by any given difference, but never exceed, nor attain before the quantities are infinitely diminished. This may be more clearly understood for the infinitely large. If two quantities, whose difference is given, are infinitely increased, their ultimate ratio will be given, namely the ratio of equality, but nevertheless there will not thereby be given the ultimate or greatest quantities of which this is the ratio. Therefore whenever in what follows, to make things easier to imagine, I speak of quantities as the smallest, or vanishing, or ultimate, avoid thinking of quantities of finite magnitude, but always consider that they are to be decreased without limit.

11.1.3 Maclaurin's definition of a limit, 1742

Maclaurin, writing some sixty years after Newton, continued to use the word 'limit' in much the same sense, as a bound that may be approached as closely as one wishes. Stung by the criticisms of Berkeley and others (see 10.2.2) he took great pains to show that limits were well defined, but his words 'it is manifest …' did nothing to avoid or disguise the fundamental problem of neglecting *o* after dividing by it.

Maclaurin's definition of a limit

from Maclaurin, *A treatise of fluxions*, 1742, I, §502–§503

502. But however ſafe and convenient this method may be, ſome will always ſcruple to admit infinitely little quantities, and infinite orders of infiniteſimals, into a ſcience that boaſts of the moſt evident and accurate principles as well as of the moſt rigid demonſtrations; and therefore we choſe to eſtabliſh ſo extenſive and uſeful a doctrine in the preceeding chapters on more unexceptionable *poſtulata*. In order to avoid ſuch ſuppoſitions, Sir ISAAC NEWTON conſiders the ſimultaneous increments of the flowing quantities as finite, and then inveſtigates the ratio which is the limit of the various proportions which thoſe increments bear to each other, while he ſuppoſes them to decreaſe together till they vaniſh; which ratio is the ſame with the ratio of the fluxions by what was ſhewn in art. 66, 67 and 68. In order to diſcover this limit, he firſt determines the ratio of the increments in general, and reduces it to the moſt ſimple terms ſo as that (generally ſpeaking) a part at leaſt of each term may be independent of the value

value of the increments themſelves; then by ſuppoſing the increments to decreaſe till they vaniſh, the limit readily appears.

503. For example, let a be an invariable quantity, x a flowing quantity, and o any increment of x; then the ſimultaneous increments of xx and ax will be $2xo + oo$ and ao, which are in the ſame ratio to each other as $2x + o$ is to a. This ratio of $2x + o$ to a continually decreaſes while o decreaſes, and is always greater than the ratio of $2x$ to a while o is any real increment, but it is manifeſt that it continually approaches to the ratio of $2x$ to a as its limit; whence it follows that the fluxion of xx is to the fluxion of ax as $2x$ is to a. If x be ſuppoſed to flow uniformly, ax will likewiſe flow uniformly, but xx with a motion continually accelerated: The motion with which ax flows may be meaſured by ao, but the motion with which xx flows is not to be meaſured by its increment $2xo + oo$, (by ax. 1.) but by the part $2xo$ only, which is generated in conſequence of that motion; and the part oo is to be rejected becauſe it is generated in conſequence only of the acceleration of the motion with which the variable ſquare flows, while o the increment of its ſide is generated: And the ratio of $2xo$ to ao is that of $2x$ to a, which was found to be the limit of the ratio of the increments $2xo + oo$ and ao.

11.1.4 D'Alembert's definition of a limit, 1765

When d'Alembert wrote and edited the mathematical sections of the great *Encyclopédie* of Denis Diderot, published between 1751 and 1765, he provided new and useful definitions of many recent mathematical concepts. His definition of 'limit' in Volume IX was close to Newton's idea of a limit as a bound that could be approached as closely as one chose, and because d'Alembert, like Newton, worked with examples that were primarily geometric, there was still no obvious need to consider quantities that might oscillate from one side of a limit to the other.

D'Alembert's definition of a limit

from Diderot and d'Alembert, *Encyclopédie*, 1751–65, IX, 542

TRANSLATION

LIMIT (*Mathematics*). One says that a magnitude is the *limit* of another magnitude, when the second may approach the first more closely than by a given quantity, as small as one wishes, moreover without the magnitude which approaches being allowed ever to surpass the magnitude that it approaches; so that the difference between such a quantity and its *limit* is absolutely unassignable.

For example, suppose we have two polygons, one inscribed in a circle and the other circumscribed; it is clear that one may increase the number of sides as much as one wishes, and in that case each polygon will approach ever more closely to the circumference of the circle; the perimeter of the inscribed polygon will increase and that of the circumscribed polygon will decrease, but the perimeter or edge of the first will never surpass the length of the circumference, and that of the second will never be smaller than that same circumference; the circumference of the circle is therefore the *limit* of the increase of the first polygon and of the decrease of the second.

1. If two magnitudes are the *limit* of the same quantity, the two magnitudes will be equal to each other.

2. Suppose $A \times B$ is the product of two magnitudes A, B. Let us suppose that C is the *limit* of the magnitude A, and D the *limit* of the quantity B; I say that $C \times D$, the product of the *limits*, will necessarily be the *limit* of $A \times B$, the product of the magnitudes A, B.

These two propositions, which one will find demonstrated exactly in the *Institutions de Géométrie*, serve as principles for demonstrating rigorously that one has the area of a circle from multiplying its semicircumference by its radius. *See* the work cited, *p. 331 and following in the second volume.*

The theory of *limits* is the foundation of the true justification of the differential calculus. *See* DIFFERENTIAL, FLUXION, EXHAUSTION, INFINITE. Strictly speaking, the *limit* never coincides, or never becomes equal to the quantity of which it is the *limit*, but the latter approaches it ever more closely, and may differ from it as little as one wishes. The circle, for example, is the *limit* of the inscribed and circumscribed polygons; for strictly it never coincides with them, although they may approach it indefinitely. This notion may serve to clarify several mathematical propositions. For example, one says that the sum of a decreasing geometric progression in which the first term is a and the second b, is $\frac{aa}{a-b}$; this value is never strictly the sum of the progression, it is the *limit* of that sum, that is to say, the quantity which it may approach as closely as one wishes, without ever arriving at it exactly. For if e is the last term in the progression, the exact value of the sum is $\frac{aa-be}{a-b}$, which is always less than $\frac{aa}{a-b}$ because even in a decreasing geometric progression, the last term e is never 0; but as this term continually approaches zero, without ever arriving at it, it is clear that zero is its *limit*, and that consequently the *limit* of $\frac{aa-be}{a-b}$ is $\frac{aa}{a-b}$, supposing $e = 0$, that is to say, on putting in place of e its *limit*. *See* SEQUENCE OR SERIES, PROGRESSION, etc.

11.1.5 Cauchy's definition of a limit, 1821

Cauchy's definition of a limit, first given in his *Cours d'analyse* in 1821, imitated that of d'Alembert and combined the same basic ideas: the existence of a fixed value, and the possibility of approaching it as closely as one wishes. The same definition was repeated, with further examples, at the beginning of his *Résumé des leçons* in 1823.

Cauchy established the concept of a limit as the starting point of textbook expositions of analysis but in most respects his definition was no clearer than Newton's 150 years earlier, for there was still no precise discussion of what it meant to approach a fixed value 'indefinitely', nor of whether a variable quantity might actually attain or even at times surpass its limit. Cauchy offered the well worn illustration of a circle and polygons, but also produced a new and more interesting example, of an irrational number approached by rationals; he did not yet suggest, however, that a limit could be approached from both sides simultaneously.

Cauchy's definition of a limit, 1821

from Cauchy, *Cours d'analyse*, 1821, 4–5

On nomme quantité *variable* celle que l'on considère comme devant recevoir successivement plusieurs valeurs différentes les unes des autres. On désigne une semblable quantité par une lettre prise ordinairement parmi les dernières de l'alphabet. On appelle au contraire quantité *constante*, et on désigne ordinairement par une des premières lettres de l'alphabet toute quantité qui reçoit une valeur fixe et déterminée. Lorsque les valeurs successivement attribuées à une même variable s'approchent indéfiniment d'une valeur fixe, de manière à finir par en différer aussi peu que l'on voudra, cette dernière est appelée la *limite* de toutes les autres. Ainsi, par exemple, un nombre irrationnel est la limite des diverses fractions qui en fournissent des valeurs de plus en plus approchées. En géométrie, la surface du cercle est la limite vers laquelle convergent les surfaces des polygones inscrits, tandis que le nombre de leurs côtés croît de plus en plus; &c....

Lorsque les valeurs numériques successives d'une même variable décroissent indéfiniment, de manière à s'abaisser au-dessous de tout nombre donné, cette variable devient ce qu'on nomme un *infiniment petit* ou une quantité *infiniment petite*. Une variable de cette espèce a zéro pour limite.

Lorsque les valeurs numériques successives

PRELIMINAIRES. 5

d'une même variable croissent de plus en plus, de maniere à s'élever au-dessus de tout nombre donné, on dit que cette variable a pour limite l'*infini positif*, indiqué par le signe ∞, s'il s'agit d'une variable positive, et l'*infini négatif*, indiqué par la notation — ∞, s'il s'agit d'une variable négative. Les infinis positif et négatif sont désignés conjointement sous le nom de *quantités infinies*.

TRANSLATION

One calls a *variable* quantity one that is considered to take successively several values different from each other. One denotes such a quantity by a letter usually taken from amongst the last in the alphabet. On the other hand one calls a *constant* quantity every quantity that takes a fixed and known value, and one usually denotes it by one of the first letters of the alphabet. When the values successively attributed to the same variable approach indefinitely to a fixed value, in such a way as to end by differing from it as little as one wishes, this last is called the *limit* of all the others. Thus, for example, an irrational number is the limit of various fractions that furnish values more and more closely approaching it. In geometry, the area of a circle is the limit towards which converge the areas of inscribed polygons, when the number of their sides increases more and more; etc....

When the successive numerical values of the same variable decrease indefinitely, in such a way as to fall below every given number, this variable becomes what one calls an *infinitesimal* or an *infinitely small* quantity. A variable of this kind has zero for its limit.

When the successive numerical values [5] of the same variable increase more and more, in such a way as to rise above every given number, one says that this variable has for its limit *positive infinity*, indicated by the sign ∞, if one is dealing with a positive variable, and *negative infinity*, indicated by the notation $-\infty$, if one is dealing with a negative variable. Positive and negative infinities are known jointly under the name of *infinite quantities*.

11.2 CONTINUITY

11.2.1 Wallis and smooth curves, 1656

The concept of continuity arises most naturally from the consideration of curves: intuitively a continuous curve is one that can be drawn without the pen leaving the paper. This idea was first made explicit by Wallis in the *Arithmetica infinitorum* of 1656, where he defined a smooth curve (*aequabilis curva*) as one that does not jump about (*non subsultans*). For Wallis this was a property that allowed him to interpolate intermediate values in a reliable way. Thus he was instinctively using what was later called the 'intermediate value theorem', which states that if a function (or its associated curve) is continuous between two finite values then it must pass through all intermediate values.

Wallis's smooth curve

from Wallis, *Arithmetica infinitorum*, 1656, Proposition 192

PROP. CXCII. *Theorema.*

SI ſit æquabilis Curva (non hinc inde ſubſultans) VC, cujus Axis VX, & Tangens in vertice VT; unde ductis ad curvam rectis axi parallelis, & ab invicem æqualibus diſtantiis remotis, harum Secunda, Quarta, Sexta, Octava, &c. (in locis paribus) ſint ut 1, 6, 30, 140, &c. (qui numeri fiunt ex continuâ multiplicatione horum, $1 \times \frac{6}{1} \times \frac{10}{2} \times \frac{14}{3} \times \frac{18}{4}$ &c.) Erit, ut Secunda ad Tertiam (hoc eſt, ut 1 ad numerum ipſis 1, 6, interponendum,) ſic Semicirculus ad Quadratum Diametri.