AN INTERPRETATION OF NEWTON'S WORK IN CALCULUS

TRISTRAM DE PIRO

Abstract. This short paper provides an account of Newton's original work in calculus, to be found in his published and unpublished papers. In particular, we provide an interpretation of his notion of tangency and differentiation, which reverses the conventional limit model and is more compatible with the use of infinitesimals. We also give a different proof of the Fundamental Theorem of Calculus, along the lines which Newton gave in an unpublished paper of 1665. We argue that Newton's work in calculus was entirely independent of his contemporary Liebniz, and support the view that Newton's work was motivated by deep geometric and aesthetic considerations.

The science of calculus is well known to any college student of mathematics. Its modern rigorous formulation is primarily due to the work of the 19th century mathematicians Cauchy, Riemann and Weierstrass. However, the geometrical ideas underlying the theory are, essentially, due to the work of the English mathematician Isaac Newton. I hope to make this point of view clearer in the course of the paper. The foundations of the calculus are the ideas of differentiation and integration. The idea of differentiation originated in computing the tangent line to a curve at a fixed point O, as a sequence of approximations of lines of the form OP, where P approaches infinitely close to O. The modern formulation, due to the German mathematician Karl Weierstrass, can be expressed as follows;

Definition 0.1. *Differentiation*

Let $f(x)$ be a real-valued continuous function on the open interval (a,b), then we say that $f(x)$ is differentiable on (a,b) if;

$lim_{h\rightarrow 0}\frac{f(x+h)-f(x)}{h}$

exists for every $x\in(a,b)$. In this case, we define the derivative of f to be $\frac{df}{dx}$.

The reader should consult an elementary textbook for the definitions of limit and continuity, again first formulated rigorously by Weierstrass. The idea of integration originated in computing the area under the arc of a curve, as a sequence of approximations of areas under a series of rectangles. The modern formulation is due to the work of the German mathematician Bernhard Riemann;

Definition 0.2. *Integration*

Let $f(x)$ be a real-valued continuous function on the closed interval $[a, b]$, then, if $\epsilon_n = \frac{b-a}{n}$, and;

$s_n = \epsilon_n \sum_{j=0}^{n-1} f(a + j\epsilon_n)$

we define the integral;

$\int_a^b f(x)dx = lim_{n\rightarrow\infty} s_n$

The fundamental theorem of calculus relates the notions of differentiation and integration. Its modern formulation and proof is credited to the French mathematician Augustin Louis Cauchy;

Theorem 0.3. *Fundamental Theorem of Calculus*

Let $f(x)$ be a real-valued continuous function on the closed interval $[a, b]$, then, if;

$F(x) = \int_a^x f(y)dy$

$F(x)$ is differentiable on the open interval (a, b) and continuous on the closed interval $[a, b]$. Moreover;

$\frac{dF}{dx}(x) = f(x)$ *for* $x \in (a, b)$

Proof. For $x \in (a, b)$, we have that;

$F(x + h) - F(x) = \int_a^{x+h} f(y)dy - \int_a^x f(y)dy = \int_x^{x+h} f(y)dy$

Here, an elementary and intuitive property of integration is used. Then, we have that;

$$|\frac{F(x+h)-F(x)}{h} - f(x)| = |\frac{\int_x^{x+h} f(y)dy - \int_x^{x+h} f(x)dy}{h}|$$

$$\leq \int_x^{x+h} \frac{|f(y)-f(x)|}{h} dy$$

$$\leq max_{y \in [x,x+h]} |f(y) - f(x)| \ (*)$$

As $h \to 0$, by continuity of f, the expression given in $(*)$ vanishes. Hence;

$$\frac{dF}{dx}(x) = lim_{h \to 0} \frac{F(x+h)-F(x)}{h} = f(x)$$

The continuity of F on the open interval (a,b) follows immediately from differentiability. It is an elementary exercise, using straightforward properties of integration, to show that F is also continuous on the closed interval $[a,b]$.

□

As a result of the theorem, one obtains a simple method of computing integrals;

Theorem 0.4. *Suppose that $f(x)$ is a continuous function on a closed interval $[a,b]$, and $G(x)$ is an antiderivative of $f(x)$, that is a continuous function on $[a,b]$, with the property that $\frac{dG}{dx}(x) = f(x)$, on (a,b). Then;*

$$\int_a^b f(x)dx = G(b) - G(a)$$

Proof. If F is the function given by the previous theorem, then the function $F - G$ is continuous and;

$$\frac{d(F-G)}{dx} = \frac{dF}{dx} - \frac{dG}{dx} = f(x) - f(x) = 0 \text{ on } (a,b)$$

It follows easily that $F - G = c$, where c is a constant, and;

$$\int_a^b f(x)dx = F(b) - F(a) = (G(b) - c) - (G(a) - c) = G(b) - G(a)$$

□

Newton's approach to calculus is, in many ways, radically different from the modern approach. His method can be unravelled from the unpublished papers that he wrote between 1665 and 1666, collected in [9];

(i). The Calculus Becomes an Algorithm. (Middle 1665?)

(ii). The General Problems of Tangents, Curvature and Limit-Motion Analysed by the Method of Fluxions. (October 1665-May 1666)

(iii). The October 1666 Tract on Fluxions (October 1666)

and his published papers;

(iv). Analysis of Equations of an Infinite Number of Terms (written in 1669, a published version (in Latin) appearing in 1711, and a translated published version appearing in 1745)

(v). The Method of Fluxions and Infinite Series (written in 1671, a translated published version appearing in 1736)

(vi). The Quadrature of Curves (written in 1676, first published with his "Optics" in 1704)

(vii). Enumeration of Lines of the Third Order (first published with his "Optics" in 1704)

(viii). Principia Matematica (written and published in Latin, 1687, a translated published version appearing in 1729)

Newton's major work in the field of calculus occurred in close proximity to that of his contemporary, the German mathematician, Gottfried Liebniz, who published his results in 1684. This led to an acrimonious priority dispute between the two figures. Newton didn't give a full printed account of his version of the calculus until 1704, but Newton's circle of friends claimed that he had obtained the results earlier, around 1667, and before Liebniz. In 1711, members of the Royal Society accused Liebniz of plagiarism, based on previous allegations of Facio and Kiel, and the fact that he had obtained a copy of Newton's manuscript (iv) from a colleague of Newton, Tschirnhaus, in 1675. It is not known whether Liebniz made use of this manuscript in his work on calculus, or, whether he had, in fact, already invented the calculus previously. However, those who question Liebniz's good faith allege that, to a man of his ability, the manuscript sufficed to give him a clue to the methods of calculus.

The first point of departure, in the methods of both Newton and Liebniz, with the modern approach is the replacement of the notion of a limit with that of an infinitesimal quantity. Roughly speaking, an infinitesimal quantity, which Newton usually denoted by o or δ, is a quantity which is non-vanishing, and yet smaller than any other finite quantity. Newton essentially justified such quantities on a geometric level, by using them to find the tangent line to a curve at a given point. In the paper (vi), sections 5 and 6 of the Introduction, we find his explanation of this method, see accompanying figure (calculus1);

"5. Let the Ordinate BC advance from its Place into any new Place bc. Complete the Parallelogram BCEb, and draw the right Line VTH touching the curve in C, and meeting the two lines bc and BA produced in T and V: and Bb, Ec and Cc will be the Augments now generated of the Abciss AB, the Ordinate BC and the Curve Line ACc; and the Sides of the Triangle CET are in the *first Ratio* of these Augments considered as nascent, therefore the fluxions of AB, BC and AC are as the Sides CE, ET and CT of the triangle CET, and may be expounded by these same Sides, or, which is the same thing, by the sides of the Triangle VBC, which is similar to the triangle CET.

6. It comes to the same purpose to take the Fluxions in the *ultimate Ratio* of the evanescent Parts. Draw the right line Cc, and produce it to K. Let the Ordinate bc return into it's former place BC, and when the points C and c coalesce, the right line CK will coincide with the tangent CH, and the evanescent triangle CEc in its ultimate Form will become similar to the Triangle CET, and its evanescent Sides CE, Ec and Cc will be *ultimately* among themselves as the Sides CE, ET and CT of the other triangle CET, are, and therefore the Fluxions of the lines AB, BC and AC are in the same Ratio. If the points C and c are distant from one another by any small Distance, the right line CK will likewise be distant from the Tangent CH by a small Distance. That the right Line CK may coincide with the Tangent CH, and the ultimate Ratios of the lines CE, Ec and Cc may be found, the Points C and c ought to coalesce and exactly coincide. The very smallest Errors in mathematical Matters are not to be neglected".

Newton argues that, in order to find the slope of the tangent line to the curve at C, given by the ratio $\frac{ET}{EC}$, it is necessary to find the "ultimate" ratio $\frac{Ec}{EC}$, as c "coalesces and exactly coincides" with C. Clearly, if c were identified with C, there would be no "evanescent triangle" CEc for which to compute such a ratio. Whereas, if c and C are "distant from one another by any small Distance", one still only obtains a line CK, "distant from the tangent CH by a small Distance". Newton is suggesting at an infinitesimal quantity to solve the problem, and, indeed, in Section 11 of the same Introduction, he shows how to compute the gradient (or the derivative of Definition 0.1) of the function $f(x) = x^n$, which he refers to as its Fluxion;

"11. *Let the Quantity x flow uniformly, and let it be proposed to find the Fluxion of* x^n.

In the same Time that the Quantity x, by flowing, becomes $x + o$, the Quantity x^n will become $(x + o)^n$, that is, by the Method of infinite Series's, $x^n + nox^{n-1} + \frac{n^2-n}{2}oox^{n-2}+$ etc. And the Augments o and $nox^{n-1} + \frac{n^2-n}{2}oox^{n-2}+$ etc are to one another as 1 and $nx^{n-1} + \frac{n^2-2}{2}ox^{n-2} + etc$. Now let these Augments vanish, and their ultimate Ratio will be 1 to nx^{n-1}."

In other words, if the curve in the previous figure is given by the graph of the function $y = x^n$, Newton computes the gradient of the line Cc as the ratio;

$$\frac{y(x+o)-y(x)}{o} = \frac{(x+o)^n-x^n}{o}$$

He then expands the expression in the numerator, and cancels the quantities involving o. In order for this calculation to make sense, it is clearly necessary that o should represent a non-zero quantity. After arriving at the final expression;

$$nx^{n-1} + \frac{n^2-n}{2}ox^{n-2} + \ldots$$

Newton then supposes that o may be taken to be so small, that it can be set to 0 in the above expression, leaving the final fluxion to be(correctly) nx^{n-1}.

On a purely logical level, there is a problem with Newton's argument in 11. Newton concerns himself with a single infinitesimal quantity o, which he needs to be both zero and non-zero at different stages of the calculation in 11. This logical paradox was heavily criticized by the philosopher George Berkeley, in his tract, "The Analyst; Or, A Discourse Addressed to an Infidel Mathematician" (1734);

"XIV...Hitherto I have supposed that x flows, that x hath a real increment, that o is something. And I have proceeded all along on that

Supposition, without which I should not have been able to have made so much as one single Step. From that Supposition it is that I get at the increment of x^n, that I am able to compare it with the Increment of x, and that I find the Proportion between the two Increments. I now beg leave to make a new Supposition contrary to the first, i.e I will suppose that there is no Increment of x, or that o is nothing; which second Supposition destroys my first, and is inconsistent with it, and therefore with every thing that supposeth it. I do nevertheless beg leave to retain nx^{n-1}, which is an Expression obtained in virtue of my first Supposition, which necessarily presupposeth such Supposition, and which could not be obtained without it: All which seems a most inconsistent way of arguing, and such as would not be allowed of in Divinity"

There is also a similar latent logical paradox in Newton's argument $(5,6)$ on tangent lines. On the one hand, Newton needs to find a non-vanishing triangle CEc, in order to compute the ratio $\frac{Ec}{EC}$, while, on the other hand, he needs this triangle to collapse to the point C, in order for this ratio to coincide with the slope of the tangent line given by $\frac{ET}{EC}$. In the same tract, Berkeley again observes this paradox;

"XXXIV...It is supposed that the Ordinate bc moves into the place BC, so that the Point c is coincident with the Point C; and the right Line CK, and consequently the Curve Cc, is coincident with the Tangent CH. In which case the mixtilinear evanescent Triangle CEc will, in its last form, be similar to the triangle CET: And its evanescent Sides CE,Ec and Cc will be proportional to CE, ET and CT the Sides of the Triangle CET. And therefore it is concluded, that the Fluxions of the lines AB,BC, and AC, being in the last Ratio of their evanescent Increments, are proportional to the Sides of the triangle CET, or, which is all one, of the triangle VBC similar thereunto. [NOTE: Introd. ad Quad. Curv.] It is particularly remarked and insisted on by the great Author, that the points C and c must not be distant from one another, by any the least interval whatsoever: But that, in order to find the ultimate Proportions of the Lines CE, Ec, and Cc (i.e. the Proportions of the Fluxions or Velocities) expressed by the finite sides of the triangle VBC, the points C and c must be accurately coincident, i.e one and the same. A Point therefore is considered as a Triangle, or a Triangle is supposed to be formed in a Point. Which to conceive seems quite impossible. Yet some there are, who, though they shrink at all other Mysteries, make no difficulty of their own, who strain at a

Gnat and swallow a Camel."

Leaving aside the additional vitriol that Berkeley pours into his argument, the logical inconsistencies that he observes are valid. However, there is clearly a sense that Newton's geometrical intuition is correct. The rigorous formulation of the calculus in the 19th century was able to resolve the logical problems by replacing the notion of an infinitesimal quantity with that of a limit, arriving at Definition 0.1 to replace Newton's arguments on tangents that we have considered ([1]). However, the notion of a limit draws on certain topological properties of real numbers, that they are in a sense infinitely divisible, which, perhaps, Newton was keen to avoid. Indeed, he gives the following description of Time in (vi);

" 3. Fluxions are very nearly as the Augments of the Fluents generated in equal but very small Particles of Time,..."

In the last fifty years, the theory of infinitesimals has enjoyed a modern renaissance, in the field of what is now referred to as non-standard analysis. The first major pioneer in this area was Abraham Robinson, whose book "Non-Standard Analysis" is still a definitive account of the subject. More recently, the mathematicians Zilber and Hrushovski have further developed the theory of infinitesimals in the context of Zariski structures, finding new approaches to current problems in algebraic geometry.

Robinson was able to resolve the paradox, observed by Berkeley, by finding a logically consistent structure $\mathcal{R}^*$, extending the real numbers

[1]There is no possibility that Newton could have derived his arguments on tangents from Liebniz's work, even though his paper (vi) was published later than Liebniz's work on calculus. One can find Newton's first use of the infinitesimal o notation in his unpublished paper (ii), where he uses the method outlined above, to compute the tangent line of the curve defined by $rx + x^2 - y^2 = 0$;

"Now if the Equation expressing the relation of the lines x and y be $rx+xx-yy = 0$. I may substitute $x + o$ and $y + \frac{qo}{p}$ into the place of x and y because (by the lemma) they as well as x and y doe signify the lines described by the bodys A and B. By doeing so there results $rx + ro + xx + 2ox + oo - yy - \frac{2qoy}{p} - \frac{qqoo}{pp} = 0$. But $rx+xx-yy = 0$ by supposition: there remaines therefore $ro+2ox+oo-\frac{2qoy}{p}-\frac{qqoo}{pp} = 0$. Or divideing it by o tis $r + 2x + o - \frac{2qy}{p} - \frac{oqq}{pp} = 0$. Also those terms in which o is are infinitely less than those in which o is not therefore blotting them out there rests $r + 2x - 2\frac{qy}{p} = 0$. Or $pr + 2px = 2qy$."

Here, the quantities p and q denote the fluxions $\frac{dx}{dt}$ and $\frac{dy}{dt}$. Hence, Newton derives the correct formula for the derivative $\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{q}{p} = \frac{r+2x}{2y}$.

$\mathcal{R}$, which contains infinitesimal elements. ([2]). In such a structure, every bounded element r^* has a uniquely defined standard part, which he denoted $st(r^*)$, with the property that the difference $r - st(r^*)$ is an infinitesimal. Real analytic functions $f(x)$, defined on $\mathcal{R}$, extend to well-defined functions on the non-standard structure $\mathcal{R}^*$. Each element $r \in \mathcal{R}$ has a set of elements which are "infinitely close" to it, in $\mathcal{R}^*$, which is now called an infinitesimal neighborhood $\mathcal{V}_r$;

Definition 0.5. *Infinitesimal Neighborhood*

If $r \in \mathcal{R}$, we define its infinitesimal neighborhood $\mathcal{V}_r$, to be;

$\{r^* \in \mathcal{R}^* : st(r^*) = r\}$

One may then give a consistent definition of differentiation, using infinitesimals, in the following way;

Definition 0.6. *Non-Standard Differentiation*

Let $f(x)$ be a real-valued analytic function on the open interval (a,b), then we say that $f(x)$ is non-standard differentiable on (a,b), if, for every $x \in (a,b)$, there exists $c_x \in \mathcal{R}$, such that;

$\frac{f(x+o)-f(x)}{o} \in \mathcal{V}_{c_x}$, *for every infinitesimal $o \in \mathcal{V}_0$.*

We then define the non-standard derivative of f at x to be c_x.

It is not hard to show that the non-standard definition of differentiation is equivalent to the modern Definition 0.1. The interested reader should look at Robinson's book [7] for a detailed account of his construction or my notes [1]. With this definition, it is easy to see that Newton's calculation, in (5,6) of (vi), is no longer paradoxical. If the ordinate b is taken an infinitesimal distance away from B, the gradient of the corresponding line Cc lies infinitely close to the gradient of the tangent line CT. By Robinson's construction, the gradient of the tangent CT is then determined uniquely from this information.

Although the logical problems with Newton's original method are now resolved, there is still something geometrically unsatisfactory about the resulting use of infinitesimals. This originates in Newton's comment

[2]The reader should look at the paper [1], for the construction of $\mathcal{R}^*$. For technical reasons, it is convenient to work with what I refer to as analytic non-standard extensions, see [1], this assumption will be in force throughout the paper.

from (6) of (vi), that "the points C and c ought to coalesce and exactly coincide". In the Definition 0.1 of differentiation using limits, given any prescribed neighborhood of B, it is possible to take the ordinate b to any point within this neighborhood. In this sense the line Cc genuinely approaches and converges to the tangent line CT. This is not the case with infinitesimals. If I was to choose two distinct infinitesimal quantities, say $\{o, o'\}$, consider the ordinates $B+o$ and $B+o'$, and the corresponding values $\{c, c'\}$, then, although the gradients of the lines Cc and Cc' both lie infinitely close to the gradient of the line CT, it is not important that one gradient lies closer to the gradient of CT than the other. In this picture, there is no *motion* of the line Cc towards the tangent line CT.

That Newton intended the geometrical picture of a sequence of lines converging towards the true tangent line, is supported by his description of mathematical quantities in 1. of (vi);

"I consider mathematical Quantities in this Place not as consisting of very small Parts; but as decrib'd by a continued Motion. Lines are describ'd, and thereby generated not by the Apposition of Parts, but by the continued Motion of Points..."

his description of fluxions as velocities in 2. of (vi);

"Therefore considering that Quantities, which increase in equal Times, and by increasing are generated, become greater or less according to the greater or less Velocity with which they increase and are generated; I sought a Method of determining Quantities from the Velocities of the Motions or Increments, with which they are generated; and calling the Velocities of the Motions or Increments, with which they are generated; and calling these Velocities of the Motions or Increments Fluxions, and the generated Quantities Fluents, I fell by degrees upon the Method of Fluxions..."

and his description of the Method of Fluxions in (v);

"In Finite Quantities so to frame a calculus, and thus to investigate the Prime and Ultimate Ratios of Nascent or Evanescent Finite Quantities, is agreeable to the Ancients; and I was willing to shew, that in the Method of Fluxions there's no need of introducing Figures infinitely small into Geometry. For this Analysis may be performed in any Figures whatsoever, whether finite or infinitely small, so that they are imagined to be similar to the Evanescent Figures..."

Professor Goldblatt makes the important observation in [6], that this last passage supports the view that Newton was prepared to dispense with the use of infinitesimals, if he had a coherent notion of a limit available. In his later paper (viii), Newton developed what he referred to as "The method of first and last ratios of quantities", in which he comes close to formulating a reasonable definition of a limit, and, therefore, avoiding the logical paradoxes of infinitesimals;

From (viii), Lemma 1 of Section 1, Book 1; "Quantities, and the ratios of quantities, which in any finite time converge continually to equality, and before the end of time approach nearer to each other than by any given difference, become ultimately equal.

If you deny it, suppose them to be ultimately unequal, and let D be their ultimate difference. Therefore they cannot approach nearer to equality than by that given difference D; which is against the supposition."

In the case of ratios of quantities, if one takes the definition of "ultimately equal" to be the one provided by his definition of fluxions in Section 11 of (vi), that we considered above, and the definition of a limit as that provided by the remaining statement of the lemma, then Newton's proof attempts to show that these definitions coincide. Again, both the definition and proof are unrigorous by modern standards, but further support the view that Newton favoured the use of the geometrical model using limits, introducing infinitesimals as a practical technical solution.

However, I wish to consider an alternative, more geometrically satisfying picture, which preserves the use of infinitesimals, see attached figure, (calculus2). As will become clearer below, Newton also considered this model, and, hence, had two geometric approaches to the theory of tangency. In this sense, the remark that he made on Time from (vi), which we considered above, reflects a more profound understanding of the geometric significance of the use of infinitesimals. In this example, I consider two complex ([3]) plane algebraic curves C and D, that is curves defined by the polynomials;

[3]This is a different situation to real plane curves, which Newton studied extensively in the cited papers (i)-(vi). We will consider this case later in the paper

$p(x,y)=\sum_{(i+j)\leq n}a_{ij}x^{i}y^{j}$, with $a_{ij}\in\mathcal{C}$

$q(x,y)=\sum_{(i+j)\leq m}b_{ij}x^{i}y^{j}$, with $b_{ij}\in\mathcal{C}$

in the complex plane $\mathcal{C}^2$, intersecting in a point O, and sharing a common tangent line l. ([4]). I now consider what happens if I vary the coefficients of the polynomials by infinitesimal amounts([5]), that is I choose infinitesimal quantities $\{\epsilon_{ij}:i+j\leq n\}$, $\{\delta_{ij}:i+j\leq m\}$ and consider the curves C_ϵ and D_δ, defined by the polynomials;

$p_\epsilon(x,y)=\sum_{(i+j)\leq n}(a_{ij}+\epsilon_{ij})x^{i}y^{j}$

$q_\delta(x,y)=\sum_{(i+j)\leq m}(b_{ij}+\delta_{ij})x^{i}y^{j}$

In general, ([6]), one would expect to obtain 2 points of intersection, marked by the points $\{u,v\}$ in the diagram, which are at an infinitesimal distance from the point O. This is, in fact, an intuition that Newton observes in his discussion of curvature from (v);

"54. And now I have finish'd the Problem; but having made use of a Method which is pretty different from the common ways of operation, and as the Problem itself is of the number of those which are not very frequent among Geometricians: For the illustration and confirmation of the Solution here given, I shall not think much to give a hint of another, which is more obvious, and has a nearer relation to the usual Methods of drawing Tangents. Thus if from any Center, and with-any Radius, a Circle be conceived to be describd, which may cut any Curve in several points; if that Circle be suppos'd to be contracted, or enlarged, till two of the Points of intersection coincide, it will there

[4]If a plane algebraic curve C, (real or complex), is defined by a polynomial $p(x,y)$, with $p(x_0,y_0)=0$, then we say that C is singular at (x_0,y_0), if $\frac{\partial p}{\partial x}(x_0,y_0)=\frac{\partial p}{\partial y}(x_0,y_0)=0$, otherwise we say that C is non-singular. If C is non-singular at (x_0,y_0), we define its tangent line l, to be the line defined by the equation $\frac{\partial p}{\partial x}(x_0,y_0)x+\frac{\partial p}{\partial y}(x_0,y_0)y=0$. In the particular case of a real function $y=f(x)$, which we have considered, this rule gives the tangent line l to be $\frac{df}{dx}(x_0)x-y=0$, as we saw in the explanation of differentiation above.

[5]The construction of non-standard extensions of the complex numbers $\mathcal{C}$ may be done algebraically. The reader should look at my paper [2] for more details.

[6]By which I mean, the infinitesimal quantities are chosen generically, that is there are no algebraic relations between the elements of the tuple $\{\bar{a}+\bar{\epsilon},\bar{b}+\bar{\delta}\}$, and the tuple $\{\bar{a},\bar{b}\}$ defining the original curves C and D is generic in the space of curves, tangent to the line l at O

touch the Curve. And besides, if its Center be suppos'd to approach towards, or recede from, the Point of Contact, till the third Point of intersection shall meet with the former in the Point of Contact; then will that Circle be aequicurved with the Curve in that Point of Contact..." ([7])

A more precise formulation of this geometric intuition is the following;

Theorem 0.7. *Let C and D be plane complex algebraic curves, intersecting at a point O, which is non-singular for both curves. Then C and D share a common tangent line at O iff;*

For a generic choice of infinitesimal quantities $\{\bar{\epsilon}, \bar{\delta}\}$;

$$Card(C_{\bar{\epsilon}} \cap D_{\bar{\delta}} \cap \mathcal{V}_O) \geq 2$$

The proof of this result may be found in my paper ([3]), ([8]) One can reformulate this result in the particular case of a curve intersecting a line;

Theorem 0.8. *New Geometric Formulation of Tangency*

If C is an irreducible complex algebraic curve, in particular, if C is the graph of a polynomial function $f(x)$, that is an algebraic curve of the form $y - f(x) = 0$, passing through the point $O = (0,0)$, then a line of the form $y = c_O x$ is tangent([9]) to the curve at O, iff, for any non-zero choice of infinitesimal $\epsilon \in \mathcal{V}_0$, there exists a non-empty collection of points $\{O_1(\epsilon), \ldots, O_n(\epsilon)\}$, distinct from O, such that;

$$C \cap (y - (c_0 + \epsilon)x = 0) \cap \mathcal{V}_O = \{O_1(\epsilon), \ldots, O_n(\epsilon)\}$$

[7]Newton defines the centre of curvature in (ii), as the limit meet of normals to a curve C. The proof that this is well defined is intrinsically connected to his proof of "The Fundamental Theorem of Calculus", that we consider below.

[8]In this paper, I show an even more general result; that the notion of algebraic intersection multiplicity between curves C and D, coincides with a non-standard notion of intersection multiplicity, defined, as in Theorem 0.7.

[9]One may still formulate a coherent definition of tangency for singular points on an irreducible curve C, the reader should look at the paper [4] for such a definition.

The proof of this result may be found in the paper [4]. The reader should consider the attached figure (calculus3), in which, by moving the tangent line l to the curve C at O, to the new position l_ϵ, one obtains two new intersection points $\{A(\epsilon), B(\epsilon)\}$ infinitely close to the original point O. In this geometric picture, it is not important that these points converge to the fixed point O, but only that they are *released* from the original position O. The distinction is a subtle one, but important for more advanced geometrical constructions. Moreover, it is the most natural geometric picture to use, in conjunction with infinitesimal quantities. The method is explored, in greater detail, in the paper [4].

In order to extend these considerations to the case of real plane curves, that is curves defined by a polynomial $p(x,y)$ in the plane $\mathcal{R}^2$, one needs to overcome certain technical difficulties resulting from the existence of real solutions to such polynomials. A simple example is given by the polynomial $x^2+1=0$, which has *no* solutions in $\mathcal{R}^2$. One would, therefore, hesitate to call this a curve. Another problematic example is given by a polynomial such as $x^2+y^8+x^4+y^2=0$, which only has one real solution at $(0,0)$, due to the fact that this is a singular point for the polynomial. In such cases, it is impossible to formulate a coherent notion of tangency. However, the following lemma clarifies this problem;

Lemma 0.9. *Let C be a real plane curve, defined by a polynomial $G(x,y)$, passing through the origin O, such that O is a non-singular point for C. Then, possibly after a linear change of variables, one can find a power series of the form;*

$s(x)=\sum_{j=1}^{\infty}a_jx^j$, *with* $a_j\in\mathcal{R}$

such that $G(x,s(x))\equiv 0$, *in particular, such a curve has an infinite number of solutions in an infinitesimal neighborhood* $\mathcal{V}_O$.

Proof. After making a linear change of variables, one can assume that the tangent line to C at $(0,0)$ (defined algebraically as in Footnote 4.) does not coincide with the axis $x=0$. Equivalently, one may assume that $G(0,0)=0$ and $\frac{\partial G}{\partial y}(0,0)\neq 0$, $(\dagger)$. The first part of the proof is a consequence of Newton's method of constructing power series solutions by successive approximations, see Paragraph 36 of his paper (v), "The Praxis of Resolution" ([10]). The assumption (†) allows us to write;

$$G(x,y) = p_m(x)y^m + \ldots + p_1(x)y + p_0(x) \ (*)$$

with $p_0(0) = 0$ and $p_1(0) \neq 0$. For (x,y) "small", Newton observes that it, therefore, makes sense to take as a first approximate solution to this equation;

$$y_0 = \frac{-\lambda_0 x^{i_0}}{p_1(0)}$$

where $\lambda_0 x^{i_0}$, $(i_0 \geq 1)$, is the first term in the expression for $p_0(x)$. Now, Newton makes the substitution $y = (y' + y_0)$ in $(*)$, this results in a further polynomial equation of the same form;

$$q_m(x)y'^m + \ldots + q_1(x)y' + q_0(x) = 0 \ (**)$$

By a straightforward algebraic calculation, using the fact that $i_0 \geq 1$, one checks that $q_1(0) \neq 0$ and $ord(q_0(x)) > ord(p_0(x))$. Hence, one can take as the second quote;

$$y_1 = \frac{-\lambda_1 x^{i_1}}{q_1(0)}$$

where $\lambda_1 x^{i_1}$, $(i_1 > i_0)$, is the first term in the expression for $q_0(x)$. Continuing in this way, one obtains a sequence of approximate solutions;

$$s_n(x) = y_0(x) + y_1(x) + \ldots + y_n(x), \text{ for } n \geq 0$$

Either this process terminates after a finite number of approximations, giving a polynomial solution to $(*)$, or one obtains an infinite power series;

[10]Newton also gives a general method for finding power series solutions to polynomial equations of the form $G(x,y) = 0$, without the simplifying assumption that $\frac{\partial G}{\partial y}(0,0) \neq 0$. This is done by the introduction of "Newton's parallelogram", in Paragraph 29 of his paper (v), and fractional exponents $x^{\frac{1}{t}}$, for $t \geq 1$. The parallelogram method is further explained in his paper (iv). This leads to a general method of finding n solutions $y_j(x) \in \mathcal{C}((x^{\frac{1}{t}}))$, for $1 \leq j \leq n$, of a polynomial equation $G(x,y) = 0$ of degree n. The result, now known as Newton's theorem, provides a general method of fragmenting a complex(real) plane curve of degree n into n distinct branches. Newton combines this method with the use of asymptotes extremely effectively in his paper (vii)

$s(x) = \sum_{i\geq 0} y_i(x)$

A straightforward algebraic calculation shows that;

$ord(G(x, s_{n+1}(x))) \geq ord(G(x, s_n(x))) + 1$

Hence, by elementary completeness arguments for the power series ring $\mathcal{R}[[x]]$, we are guaranteed that $G(x, s(x)) \equiv 0$, ([11]).

In order to finish the proof, one can show that the constructed power series $s(x)$ defines an analytic function on an open interval U containing 0, see the proof of the following theorem for a more detailed explanation. As such a function is analytic, it extends to the non-standard model $\mathcal{R}^*$. It follows, from the non-standard definition of continuity, see [1], that, for any infinitesimal ϵ in $\mathcal{V}_0$, the value $s(\epsilon)$, also lies in $\mathcal{V}_0$. Hence, the tuple $(\epsilon, s(\epsilon))$ lies in $C \cap \mathcal{V}_0$. As there exist infinitely many infinitesimals in $\mathcal{V}_0$, the result follows. □

[11]Newton's method of constructing a solution to the polynomial equation $G(x, y) = 0$, given in Lemma 0.9, is closely related to the Newton-Raphson method. Namely, one considers the function;

$G : \mathcal{R}[x] \rightarrow \mathcal{R}[x]$, $G(y) = p_m(x)y^m + \ldots + p_1(x)y + p_0(x)$ (†)

Having obtained a first approximation $y_0 = s_0$ to the equation $G(y) = 0$, the Newton-Raphson method gives the further approximation;

$s_1 = y_0 - \frac{G(y_0)}{G'(y_0)} = y_0(x) - \frac{q_0(x)}{q_1(x)}$

where $q_0(x)$ and $q_1(x)$ are obtained from the transformed polynomial $(**)$ in Lemma 0.9. By a similar argument to the above, replacing the successive approximations $-\frac{\lambda_1 x^{i_1}}{q_1(0)}$ by $-\frac{q_0(x)}{q_1(x)}$ and noting that;

$\frac{q_0(x)}{q_1(x)} = \frac{\lambda_1 x^{i_1}}{q_1(0)} u(x)$ for a unit $u(x) \in \mathcal{R}[[x]]$

one is similarly guaranteed that this method also yields the same power series solution $s(x)$ to the equation $G(y) = 0$ in (†). The Newton-Raphson method is usually applied to functions of a real variable, rather than the rings $\mathcal{R}[x]$ or $\mathcal{R}[[x]]$. However, that Newton intended his method of finding power series solutions to polynomial equations of the form $G(x, y) = 0$, ("species"), to be a partial generalisation of this method is borne out by his consideration of the case of "affected equations", at the beginning of (v). An affected equation is just the case where the indeterminate x is replaced by an explicit numerical value. Newton's work in (v), and similar calculations in (iv), should, therefore, be considered as a principal origination of this idea.

This result shows that any real curve contains a "real branch" in a neighborhood of a real non-singular point, hence, one can formulate a version of tangency, similar to the one given in Theorem 0.8;

Theorem 0.10. *Let C be a real plane curve, such that $O = (0, 0)$ is a non-singular point for C, then a line of the form $y = c_O x$ is tangent to the curve at O, iff, in a non-standard extension of $\mathcal{R}$;*

Either*, for any positive infinitesimal ϵ, there exists at least one solution $O(\epsilon)$, distinct from O in;*

$C \cap (y - (c_O + \epsilon)x = 0) \cap \mathcal{V}_O$

Or*, for any positive infinitesimal ϵ, there exists at least one solution $O(-\epsilon)$, distinct from O in;*

$C \cap (y - (c_O - \epsilon)x = 0) \cap \mathcal{V}_O$

Proof. We consider, first, the case when C is defined by the equation $y - f(x) = 0$. By the algebraic definition of tangency, or the versions of differentiation, given in Definitions 0.6 or 0.1, the tangent line to C at O is given by $y = c_O x$, where $c_O = \frac{dy}{dx}|_O$. Suppose that $m \in \mathcal{R}$ is distinct from c_O, then, we claim that, for any infinitesimal ϵ, there are *no* solutions, distinct from O, to the equation;

$C \cap (y - (m + \epsilon)x = 0) \cap \mathcal{V}_O$ (†)

Making the substitutions, $y = f(t)$ and $x = t$, and using the fact that $f(t) = c_O t + t^r u(t)$, where $u(0) \neq 0$ and $r \geq 2$, this reduces to finding solutions to;

$t^{r-1}u(t) = d + \epsilon$, where $d \in \mathcal{R}_{\neq 0}$, for $t \in \mathcal{V}_0 \setminus \{0\}$

If δ is a non-zero infinitesimal solution to such an equation, then by elementary properties of the standard part mapping st, we would have that $0 = st(\delta^{r-1}u(\delta)) = st(d + \epsilon) = d$, $(*)$ which is a contradiction, hence (†) is shown. If $m = c_O$, then, for an infinitesimal ϵ, we need to find solutions to;

$t^{r-1}u(t) = \epsilon$, for $t \in \mathcal{V}_0 \setminus \{0\}$ (††)

We consider the function $F(t)$ defined by $t^{r-1}u(t)$. As we may assume that $F(t)\neq 0$, we can find an open interval of the form $(0,a)$, such that $F(t)|_{(0,a)}>0$ or $F(t)|_{(0,a)}<0$, $F(t)$ is continuous on $[0,a]$ and $F(a)=max_{t\in[0,a]}F(t)$ or $F(a)=min_{t\in[0,a]}F(t)$, respectively. By the Intermediate Value Theorem, it follows that, for any $y_0\in(0,F(a))$, or $y_0\in(F(a),0)$ respectively, there exists $t_0\in(0,a)$, such that $F(t_0)=y_0$. We can formulate these propositions in the language $(\mathcal{R},+,.,<)$, by either of the statements;

$$\sigma_1\equiv\forall y((0<y<F(a))\rightarrow\exists x((0<x<a)\wedge F(x)=y))$$

$$\sigma_2\equiv\forall y((F(a)<y<0)\rightarrow\exists x((0<x<a)\wedge F(x)=y))$$

By definition of a non-standard extension of $\mathcal{R}$, see [1], one of these statements must also be true in $(\mathcal{R}^*,+,.,<)$. In the former case that σ_1 holds, given a positive infinitesimal ϵ, we can find $t_0\in(0,a)$ such that $F(t_0)=\epsilon$. We claim that $t_0\in\mathcal{V}_0$. If not, then applying the standard part map to the relation $F(t_0)=\epsilon$, we would obtain $F(st(t_0))=0$, $(**)$, which implies that $st(t_0)=0$, by the assumption that $F|_{(0,a)}\neq 0$. Hence, we obtain a solution $O(\epsilon)=(t_0,F(t_0))$, of the form required in the first part of the Theorem. A similar argument, in the case that σ_2 holds, gives a solution of the form $O(-\epsilon)$, for any positive infinitesimal ϵ, as required by the second part of the Theorem.

In order to handle the general case, that C defines a real plane curve defined by a polynomial $G(x,y)$, we use Lemma 0.9, to find a power series $s(x)\in\mathcal{R}[[x]]$ such that $G(x,s(x))\equiv 0$. By a reasonably straightforward algebraic calculation (left to the reader), one can show that there exists a fixed exponent N, dependent on the degree of $G(x,y)$, such that $|c_{n+1}|\leq|c_n^N|$, where c_n denotes the n'th coefficient of the power series $s(x)$. It follows that $limsup_{n\rightarrow\infty}|c_n|^{\frac{1}{n}}\leq|c_0|^N$, and, by standard convergence tests, the power series defines $s(x)$ defines an analytic function on the open interval $(-\frac{1}{d},\frac{1}{d})$, where $d\leq|c_0|^N$, see [8]. The reader should now follow through the rest of the above argument, replacing the polynomial $f(x)$ by the power series $s(x)$. In the course of the proof, one should use the following facts;

(i). Any analytic function on an open interval extends to a universal non-standard extension of $\mathcal{R}$, see previous footnote 2.

(ii). The non-standard definition of continuity of analytic functions, given in [1], to obtain the contradiction given in $(*)$ above and the result $(**)$.

(iii). Standard facts on analytic functions, in order to apply the Intermediate Value Theorem.

□

Remarks 0.11. *One should think of a line passing through the curve at O as rotating, either "clockwise" or "anti-clockwise" about the fixed point O. Tangency is then characterised by an intersection $O(\epsilon)$, moving away continuously from the fixed point O, along the curve, in one of these cases. This form of definition is much closer to the non-standard definition of continuity and seems to improve on the slightly clumsy use of infinitesimals appearing in Newton's original definition of tangency. One can give a more algebraic proof of Lemma 0.9 and Theorem 0.10, using the method of power series (now referred to as analytic functions, see Definition 0.13 below) that Newton introduced successfully in his papers (iv) and (v).*

One should, first, construct a non-standard extension (not necessarily an analytic extension, in the sense of [1]*) of $\mathcal{R}$, containing the power series ring $\mathcal{R}[[x]]$. This can be done by considering the field;*

$$\mathcal{R}^* = \bigcup_{1\leq n<\omega} \mathcal{R}((x^{\frac{1}{n}}))$$

This admits a linear ordering of the following form. For a power series $s(x) = \sum_{j=m_1}^{\infty} c_j x^{\frac{j}{N}}$, in $\mathcal{R}((x^{\frac{1}{N}}))$, we declare that $s(x) < 0$ iff $c_{m_1} < 0$, and, for power series $\{s(x), t(x)\}$ in $\mathcal{R}((x^{\frac{1}{N}}))$, we declare that $s(x) < t(x)$ iff $s(x) - t(x) < 0$. With this requirement, the structure $(\mathcal{R}^*, +, ., <)$ is a totally ordered real closed field. By standard facts about the theory of real closed fields, due to Tarski, it is an elementary extension of the structure $(\mathcal{R}, +, ., <)$, hence, setting any power series of the form $s(x) = \sum_{j=1}^{\infty} c_j x^j$ to be a positive infinitesimal, it is a non-standard extension by the definition given in [1].

The proof of Lemma 0.9 follows immediately from this construction of $\mathcal{R}^*$. The obstacle in showing Theorem 0.10 is to find real power series solutions to the equation $(\dagger\dagger)$. In the particular case that $r = 2$ and $u(t)$ is a polynomial, this follows immediately from Newton's power series method, explained in Lemma 0.9. Otherwise, one should proceed *as follows;*

Let $a_0 = u(0)$*, and let* $u(t) = a_0 v(t)$*, with* $v(0) = 1$ *and* $a_0 \neq 0$*. If* $a_0 > 0$*, one should take the infinitesimal* ϵ *to be positive and seek a solution to the equation;*

$t^{r-1}v(t) = \frac{\epsilon}{a_0}$ $(*)$

As $v(0) = 1$*, we can find a power series* $w(t) \in \mathcal{R}[[t]]$ *such that* $w(t)^{r-1} = v(t)$*, with* $w(0) = 1$*. Similarly, as* $\frac{\epsilon}{a_0}$ *is positive, we can find a positive infinitesimal* δ *in* $\mathcal{R}^*$ *such that* $\delta^{r-1} = \frac{\epsilon}{a_0}$*. We are then reduced to solving the equation;*

$tw(t) = \delta$

Using a straightforward extension of Newton's method, given in Lemma 0.9, this admits a power series solution $h(\delta) \in \mathcal{R}[[\delta]]$*, with* $h(0) = 0$*, which is also, therefore, a solution to* $(*)$ *and to* $(\dagger\dagger)$ *of Theorem 0.10. Given an arbitrary choice of positive infinitesimal* ϵ *in* $\mathcal{R}^*$*, it is easily checked that* $h(\delta)$ *must correspond to a fractional power series in* $\mathcal{R}[[x^{\frac{1}{N}}]]$*, for some fixed exponent* N*. In particularly, it defines an infinitesimal solution to* $(\dagger\dagger)$ *of Theorem 0.10 in* $\mathcal{R}^*$*. In the case that* $a_0 < 0$*, one should take the infinitesimal* ϵ *to be negative and apply a similar argument.*

The methods of Newton and Liebniz also differ considerably from the modern approach to theory of integration. Both formulated a non-standard definition of integration, based on the idea of finding the area under a curve by summation over a series of rectangles of infinitely small width, see attached figure, calculus4. The formal non-standard definition is the following;

Definition 0.12. *Non-Standard Integration*

Let $f(x)$ *be a continuous function on the closed interval* $[a, b]$*, and let* R_f *be the Riemann sum, defined for a real number with* $0 < c < (b-a)$*, by;*

$R_f(c) = \sum_{j=0}^{N(c)} f(a + jc)c$

where $N(c)$ is the greatest positive integer n such that $(a+nc)<b$. Then, we define;

$\int_a^b f(x)dx = st(R_f(\epsilon))$

where ϵ is a positive infinitesimal.

The proof that this is a good definition, that is doesn't depend on the choice of infinitesimal ϵ, and gives the same value as Definition 0.2, can be found in my notes [1]. Arguably, Liebniz was the first to formulate this definition in what would be called a rigorous way. However, that the geometric idea was known to Newton, before Liebniz's publications on calculus, is clear from his from his proof of the Fundamental Theorem of Calculus, given in his unpublished documents (i), see p304 of [9], which we will consider shortly.

Newton's first published explanation of integration can be found in (vi). Newton uses the term quadrature to mean integration, the term quadrature referring obliquely to approximating the area under a curve by a series of quadrangles. However, the paper is essentially a tabulation of integrals for various curves, and there is no recognisable proof of the Fundamental Theorem of Calculus (Theorem 0.3) and its corollary (Theorem 0.4), which allows one to compute explicit integrals. Newton gives a much clearer explanation of quadrature in his earlier paper (iv), see attached figure (calculus5);

" The Demonstration of the Quadrature of Simple Curves belonging to Rule the first.

Preparation for demonstrating the first Rule.

54. Let then $AD\delta$ be any curve whose Base $AB=x$, the perpendicular Ordinate $BD=y$, and the area $ABD=z$, as at the Beginning. Likewise put $B\beta=o$, $BK=v$; and the Rectangle $B\beta HK(ov)$ equal to the Space $B\beta\delta D$.

Therefore it is $A\beta=x+o$, and $A\delta\beta=z+ov$: Which Things being premised, assume any Relation betwixt x and z that you please, and seek for y in the following Manner.

Take at Pleasure $\frac{2}{3}x^{\frac{3}{2}}=z$; or $\frac{4}{9}x^3=z^2$. Then $x+o$ $(A\beta)$ being substituted for x, and $z+ov$ $(A\delta\beta)$ for z, there arises $\frac{4}{9}$ into $x^3+3xo^2+3xo^2+o^3=$ (from the Nature of the Curve) $z^2+2zov+o^2v^2$. And taking away Equals ($\frac{4}{9}x^3$ *and* z^2) and dividing the Remainders by o, there arises $\frac{4}{9}$ into $3x^2+3xo+oo=2zv+ovv$. Now if we suppose $B\beta$

to be diminished infinitely and to vanish, or o to be nothing, v and y, in that Case will be equal, and the Terms which are multiplied by o will vanish: So that there will remain $\frac{4}{9}\times 3x^2 = 2zv$, or $\frac{2}{3}x^2(=zy) = \frac{2}{3}x^{\frac{3}{2}}y$; or $x^{\frac{1}{2}} = \frac{x^2}{x^{\frac{3}{2}}} = y$. Wherefore conversely if it be $x^{\frac{1}{2}} = y$, it shall be $\frac{2}{3}x^{\frac{3}{2}} = z$.

55. Or universally, if $\frac{n}{m+n}\times ax^{\frac{m+n}{n}} = z$; or, putting $\frac{na}{m+n} = c$, and $m+n = p$, if $cx^{\frac{p}{n}} = z$; or $c^nx^p = z^n$: Then, by substituting $x+o$ for x, and $z+ov$ (or which is the same $z+oy$) for z, there arises c^n into $x^p + pox^{p-1}$, etc $= z^n + noyz^{n-1}etc$, the other Terms, which would at length vanish being neglected. Now taking away c^nx^p and z^n which are equal, and dividing the Remainders by o, there remains $c^npx^{p-1} = nyz^{n-1}(=\frac{nyz^n}{z}) = \frac{nyc^nx^p}{cx^{\frac{p}{n}}}$, or, by dividing by c^nx^p, it shall be $px^{-1} = \frac{ny}{cx^{\frac{p}{n}}}$; or $pcx^{\frac{p-n}{n}} = ny$; or by restoring $\frac{na}{m+n}$ for c, and $m+n$ for p, that is m for $p-n$, and na for pc, it becomes $ax^{\frac{m}{n}} = y$. Wherefore conversely, if $ax^{\frac{m}{n}} = y$, it shall be $\frac{n}{m+n}ax^{\frac{m+n}{n}} = z$. Q.E.D"

Here, Newton demonstrates how to find the quadrature of simple curves. If a curve is given by the equation $y = x^{\frac{m}{n}}$, in Article 55, he deduces correctly the formula for the integral of the curve as $\frac{n}{m+n}x^{\frac{m+n}{n}}$. The argument he gives for this formula, however, would not be called rigorous by modern standards. In the diagram pertaining to the problem which Newton gives, (calculus5), Newton sets the area $B\beta\delta D$ to be $o.v$, where $o = length(B\beta)$ and $v = length(BK) = y$ (†). Assuming the quadrature(area) of the curve (area(ABD)) can be expressed as a polynomial expression $z(x) = x^q$, q rational, of the base $x = length(AB)$, Newton then forms the equation;

$$z + o.v = z(x+o) = (x+o)^q$$

By expanding this expression and setting the term o to be nothing, (††), he then derives the expression $y(x) = qx^{q-1}$. Finally, he argues that these steps may be reversed to give the formula for quadrature, given only the equation of the curve $y(x)$, (†††). The first argument (†) relies on the intuition that, as $length(B\beta)$ becomes sufficiently small, the quadrangle $B\beta HK$ is a good approximation to the area under the curve $BD\delta\beta$. In order to make the idea rigorous, one needs to formulate a definition of integration, using infinitesimals, similar to that given in Definition 0.12. However, the intuition is still one of two main geometric components behind such a definition, and, as we will find the other geometric component in an earlier unpublished paper, it is reasonable to say that Newton had, at this stage, formulated a clear geometric idea of integration, although, he fails, here, to formulate the idea in its entirety. The second argument (††) contains the germ of the idea of fluxions, but the exposition of it, here, is unclear. Finally, the third argument (†††), suffers from the same deficiencies as the first (†), that of a clear definition of integration, which is required, here, to make the argument rigorous.

The contents of this paper, in particular the argument given here, was the centre of the later dispute, as to whether Liebniz had obtained the idea of the differential calculus from Newton, previous to his publication of 1684. However, it is interesting to note that in Liebniz's "Excerpta from Newton's De Analisi",([12]), October 1676, he makes no comment whatsoever on this fundamental passage. Given the lack of justification in certain steps of Newton's argument, it seems reasonable that Liebniz's formulation of a rigorous definition of integration and differentiation, using infinitesimals, was a genuine independent achievement. This last conclusion is similar to that which Charles Bossut gives in "A General History of Mathematics from the Earliest Times to the Middle of the Eighteenth Century" (1802);

"All these considerations appear to me to evince that, if the piece De Analysi per Aequationes and the letter of 1672 contain the method of fluxions, it was at least enveloped in great darkness." ([13])

I wish to, finally, consider an argument that Newton gives in his paper (i), see attached figure (calculus6), which I will, first, quote in full;

Prop:

Haveing an equation of 2 dimensions to find w^t crooke line it is whose area it dothe expresse. suppose y^e equation is $\frac{x^3}{a}$. naming y^e quantitys, $a = dh = kl$. $bg = y$. $db = mk = x = gp$. y^e superficies $dbg = \frac{x^3}{a}$. supose y^e square $dhkl$ is equall to y^e superficies gbd; y^n $dk = z = bm = lh = \frac{x^3}{aa}$, and $aaz = x^3$. w^{ch} is an equation expressing

[12] De Analisi was the latin title of Newton's paper (iv)

[13] The letter referred to is one that Newton wrote to Collins, claiming to have found a general method of finding the tangent to a curve, but without giving any demonstration. As we have also observed, the method of quadrature(integration), is also obscure. Given the argument which we will establish below, that Newton had, by this stage, a full understanding of the methods of calculus, it seems likely that this obscurity was deliberate.

y^e nature of y^e line fmd.

Next makeing $nm = s$ a line w^{ch} cutteth dmf at right angles. $nd = v$.

$$ss - vv + 2vx - xx = \frac{x^6}{a^4} = mb \text{ squared.}$$

0. 0. 1. 2. 6.

(w^{ch} is an equation haveing 2 equall rootes and therefore multiplyed according to Huddenius his method, produceth another.)

$$2vx = 2xx + \frac{6x^6}{a^4}.$$

$$v = x + \frac{3x^5}{a^4}. \text{ and } nb = v - x = \frac{3x^5}{a^4}.$$

Now supposeing $mb : bn :: dh : bg$. that is, $\frac{3x^5}{a^4} : \frac{x^3}{aa} :: y : a$. $3xx = ay$ and $3xxa = a^2y$. Which is y^e nature of y^e line dgw. and y^e area $dbg = dklh = \frac{x^3}{aa}$, makeing $db = x$. $dh = a$. Or. $diw = deoh = \frac{x^3}{a}$, determining ($di$) to be ($x$), etc.

The Demonstration whereof is as followeth.

Suppose $w\Pi\Omega$, Ωmz, zfv etc are tangents of y^e line dmf. and from theire intersections z, Ω, v, w draw $va, zq, \Omega s, wx$ and from theire touch points draw $fw, mg, \Pi\xi$ all parallell to kp. also from y^e same points[s] of intersection draw $v\sigma, z\lambda, \Omega v, \omega\zeta$. And $mb : nb :: bt : bm :: \Omega\beta : \beta m :: kl : bg$. wherefore $\Omega\beta \times bg = \beta m \times kl$. that is y^e rectangle $kl\nu\mu = b\rho sg$. And $\pi\rho s\breve{o} = \theta\lambda\nu\mu$. in like manner it may be demonstrated y^t $aq\pi n = \theta\lambda\sigma\rho$, and $\rho\omega xy = \mu d\nu h$ etc so y^t y^e rectangle $\rho\sigma hd$ is equall to any number of such like squares inscribed twixt y^e line $n\psi$ and y^e point d, w^{ch} squares if they bee infinite in number, they will bee equall to y^e superficies $dn\psi\omega g\xi$, (∗). ([14])

[14]Newton uses the shorthand notation w^t for which, y^e for the, y^t for that, and y^n for then. I have made the occasional modifications to Newton's original text, as done in footnotes 48-50 of Whiteside's commentary on the paper. Finally, one should observe that the letter ζ in Newton's argument coincides with the letter h of the attached diagram. In an original, cancelled version of the figure, Newton drew the line ωd distinct from rdh, and $\omega\zeta$ a little above the horizontal axis rdh, see footnote 47 of Whiteside. Also the letter n denotes both the intersection of the vertical line $va\psi$ with the horizontal axis rdh, *and* the intersction of the normal to the curve at m with the base rdh, see Whiteside's footnote 45 in [9]

The argument is essentially a proof of The Fundamental Theorem of Calculus. However, before considering it in detail, we will make a preliminary observation. This is the argument $(*)$, at the end, that the area $dn\psi\omega g\xi$ (that is the area above the lower curve between d and n) is equal to the sum of an infinite number of rectangles inscribed between the lower curve and the axis rdh. This observation, together with the argument we considered above, from Newton's paper (iv), suggest that Newton had, at least, an intuitive idea of a formal definition of integration, using infinitesimals, by the time he wrote (iv), ([15]). Given his use of infinitesimal arguments in the context of differentiation, see footnote 1, and his proofs, here and in (iv), of The Fundamental Theorem of Calculus, it seems clear that Newton had formulated his own version of the calculus, by 1669, independently from Liebniz, and, even at the same level of precision. (The logical paradoxes observed by Berkeley are problematic for both the work of Liebniz and Newton).

Now, considering the argument in more detail, I will show that not only does Newton give a reasonable proof of The Fundamental Theorem of Calculus, but also one which is geometrically superior to the modern proofs of the result. Newton begins his argument, by setting the fixed length dh to be a, the area dbg to be $\frac{x^3}{a}$, the length of the ordinate db to be x, the length of the coordinate bg to be y and the length of the coordinate bm to be $\frac{x^3}{a^2}$. In modern terminology, Newton takes the bottom curve to be described by the equation $y(x)$ and the function;

$$z(x) = \frac{1}{a}\int_0^x y(x)dx \ (\dagger)$$

which describes the top curve. He assumes that $z(x) = \frac{x^3}{a^2}$ $(**)$. Newton now makes the unusual step of drawing the *normal* to the top curve at m. Setting $length(nm) = s$ and $length(nd) = v$, he calculates;

$$(v-x)^2 + length(mb)^2 = s^2 \text{ (Pythagoras' Theorem)}$$

deriving the formula;

$$length(mb)^2 = s^2 - (v-x)^2 = s^2 - v^2 + 2vx - x^2 = (\frac{x^3}{a^2})^2 = \frac{x^6}{a^4} \ (\dagger\dagger)$$

[15]In his paper (viii), Lemmas 2,3 of Section 1, Book 1, Newton gives a much clearer account of a formal definition of integration. Although he previously introduces the notion of a limit in this paper, his definition is closer to the previous Definition 0.12 using infinitesimals, than Definition 0.2, using limits

He then differentiates the expression (††), to obtain;

$(v-x)=\frac{3x^5}{a^4}$([16])

In modern terminology, $v-x$ gives the length of the subnormal to the top curve defined by $z(x)$. The general formula for the length of the subnormal is $z\frac{dz}{dx}$, (***),([17]) as, indeed, Newton deduces correctly in this particular case.

Newton continues by assuming that $\frac{length(bn)}{length(bm)}=\frac{length(bg)}{length(dh)}$. By $(***)$, this is equivalent to;

$\frac{dz}{dx}=\frac{z\frac{dz}{dx}}{z}=\frac{y}{a}$ $(****)$

By Newton's previous argument, which we reformulated in (†), the assumption is exactly a statement of The Fundamental Theorem of Calculus, Theorem 0.3. From this assumption, and the explicit equation of the function $z(x)$, given in $(**)$, Newton is then able to derive the formula for the original "crooked line" $y(x)$, namely $y(x)=\frac{3x^2}{a}$, which was the purpose of the original proposition.

The crux of Newton's argument is, then, contained in his proof of $(****)$, "The Demonstration whereof is as followeth". Newton begins by drawing a series of tangent lines to the top curve, centred at the points $\{f,m,\Pi\}$. From the intersections of these tangent lines and the points of tangency themselves, he draws a series of lines, parallel to the lines defined by kp and dh. He then claims that the ratio $\Omega\beta:\beta m$ is equal to the ratio $kl:bg$, from which he deduces that $\Omega\beta\times bg=\beta m\times kl$,

[16]The reference to Huddenius' method is unclear. The sequence of numbers 0.0.1.2.6 refer to the weights of x in the expression (††)

[17]The proof is an elementary exercise in differentiation and trigonometry, which we give for the convenience of the reader. Suppose that $z(x)$ defines a differentiable function and let (x_0,z_0) be a fixed coordinate on the curve C defined by $z-z(x)=0$. The vector defining the tangent to C at (x_0,z_0) is given by $(1,\frac{dz}{dx}|_{(x_0,z_0)})$. Hence, if (α,β) is the vector defining the normal to C at (x_0,z_0), we obtain, by the use of the dot product, $\alpha+\beta\frac{dz}{dx}|_{(x_0,z_0)}=0$. Hence, $\frac{\beta}{\alpha}=\frac{-1}{\frac{dz}{dx}|_{(x_0,z_0)}}$. The equation of the normal through (x_0,z_0) is given by $z-z_0=-\frac{(x-x_0)}{\frac{dz}{dx}|_{(x_0,z_0)}}$. The point of intersection x_1 of this line, with the axis $z=0$, is, therefore, $x_1=x_0+z_0\frac{dz}{dx}|_{(x_0,z_0)}$, hence, the length of the subnormal is $x_1-x_0=z_0\frac{dz}{dx}|_{(x_0,z_0)}$, as required.

(†††). ([18]) The rectangles formed by the series of lines form a partition, which approximates the area under the bottom curve. Using the equality (†††), he shows that the area defined by this partition is equal to the area defined by the rectangle $\rho\sigma hd$. By refining the partition, that is taking an infinite number of tangent lines and intersections, he deduces the formula (†).

The circularity in Newton's argument, see footnote 18, is easily remedied by, instead, defining the function $z(x)$ by the formula $(*\,*\,**)$ and deducing the formula (†). We now formulate and prove a more precise version of Newton's geometric idea. We first require some straightforward definitions and lemmas;

We, first, recall the notion of an analytic function, that we used previously in the paper, the reader should look at [8] for more details;

Definition 0.13. *Let $f(x)$ be a real valued function, defined on an open interval $U \subset \mathcal{R}$, then we say that f is analytic, if, for every $x_0 \in U$, there exists an open interval $U_{x_0} \subset U$, such that f is defined by a convergent power series on U_{x_0}. If $f(x)$ is defined on a closed interval $V \subset \mathcal{R}$, then we say that f is analytic, if there exists an open interval $U \supset V$, and an analytic function g, defined on U, whose restriction $g|_V = f$.*

Remarks 0.14. *It is easily shown that an analytic function f, defined on an open interval U, has the property that its higher-order derivatives $f^n = \frac{df^n}{dx^n}$ all exist and define analytic functions, and that an analytic function f, defined on a closed (bounded) interval V, has finitely many zeroes.*

We then make the following definition;

Definition 0.15. *If $f(x)$ is analytic on a closed interval V, then we define its derivative $\frac{df}{dx}$ to be the restriction of $\frac{dg}{dx}$ to U, where g is given in the previous definition.*

Remarks 0.16. *It is a simple exercise to check that this is a good definition, and that, again, all the higher order derivatives $f^n = \frac{df^n}{dx^n}$ exist and define analytic functions on V. We will restrict ourselves, from now on, to the class of analytic functions. It is clear, from what*

[18] Unfortunately, this claim is equivalent to the assumption $(*\,*\,**)$ that he is trying to demonstrate. However, we will show presently how to remedy Newton's argument.

we have said above, that this is the class of functions that Newton was primarily concerned with, in his work on calculus. However, the reader is invited to generalise the following results to a larger class of functions, if desired.

The following lemma is straightforward;

Lemma 0.17. *Let $f(x)$ be a non-constant, analytic function, defined on the closed interval $[a,b]$, then there exists a finite collection of points $\{c_1, \ldots c_n\}$, with $a < c_1 < \ldots < c_n < b$, such that on each closed interval $[a,c_1], \ldots, [c_i, c_{i+1}], \ldots [c_n, b]$, with $1 \leq i \leq n-1$, $f(x)$ is either strictly increasing or strictly decreasing.*

Proof. As $f(x)$ is non-constant, its derivative $f'(x)$ is not identically zero. As $f'(x)$ is analytic, by Remarks 0.14 and 0.16, there exist finitely many points $d_1 < \ldots < d_m$, for which $f'(d_j) = 0$, $1 \leq j \leq m$. We consider $f(x)$, restricted to the closed interval $[d_j, d_{j+1}]$. If $x_1 < x_2$ belong to this interval, then, applying the intermediate value theorem, we have that;

$f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$, for some $c \in (x_1, x_2)$

In the case that $f'(x)|_{(d_j,d_{j+1})} > 0$, we obtain $f(x_2) - f(x_1) > 0$, and, in the case that $f'(x)|_{(d_j,d_{j+1})} < 0$, we obtain $f(x_2) - f(x_1) < 0$. Hence, the result follows, after possibly relabelling the points $\{d_1, \ldots d_m\}$, if either a or b belongs to this set.

□

We make the following definition;

Definition 0.18. *Let $f(x)$ be an analytic function, defined on the closed interval $[a,b]$, then we call $f(x)$ a simple function, if $f'(x)$ is either strictly increasing or strictly decreasing on $[a,b]$.*

Then;

Lemma 0.19. *Let $f(x)$ be a non-constant analytic function, defined on the closed interval $[a,b]$, then there exists a finite collection of points, $\{e_1, \ldots e_r\}$, with $\{a < e_1 < \ldots < e_r < b\}$, such that on each closed interval $\{[a,e_1], \ldots, [e_i, e_{i+1}], \ldots, [e_r, b]\}$, with $1 \leq i \leq r-1$, $f(x)$ is simple.*

Proof. The result follows immediately, by applying Lemma 0.17 to the analytic function $f'(x)$.

□

We now show;

Lemma 0.20. *Let $z(x)$ be an analytic function, defined on the closed interval $[a,b]$, with the property that $z'(x)$ is either strictly increasing or strictly decreasing on this interval. Then, for any $c \in [a,b)$, the interval $I_c = (c,b)$ has the following property;*

For any $d \in I_c$, if l_c and l_d denote the tangent lines to the curve C defined by $z(x)$, and p_{cd} denotes their point of intersection, then $c < x(p_{cd}) < d$. $(*)$

Proof. We assume, first, that $z(c) = 0$ and $\frac{dz}{dx}(c) = 0$, $(\dagger)$. Working in coordinates $\{x, w\}$, the tangent line l_c to the curve C at c is given by $w = 0$, while the tangent line l_{x_0} to a general point $(x_0, z(x_0))$ on the curve is given by the equation;

$\frac{w - z(x_0)}{x - x_0} = z'(x_0)$ $(*)$

The intersection point $x(p_{cx_0})$, in this case, is then obtained from $(*)$, by setting $w = 0$, which gives;

$x(p_{cx_0}) = \frac{x_0 z'(x_0) - z(x_0)}{z'(x_0)} = x_0 - \frac{z(x_0)}{z'(x_0)}$

Setting $x_0 = c + \epsilon$, we need to show that, for $\epsilon \in (0, b - c)$;

$c < (c + \epsilon) - \frac{z(c+\epsilon)}{z'(c+\epsilon)} < c + \epsilon$

or equivalently;

$0 < \frac{z(c+\epsilon)}{z'(c+\epsilon)} < \epsilon$ $(\dagger\dagger)$

By hypothesis, on the open interval (c,b), $z'(x)$ is either strictly increasing(i) or strictly decreasing(ii), in particular, either strictly positive or strictly negative. If $0 < \epsilon < (b - c)$, then, using the fact that $z(c) = 0$, and, applying the Intermediate Value Theorem, we have;

$\frac{z(c+\epsilon)}{\epsilon}=z'(c+\epsilon')$ for some $0<\epsilon'<\epsilon$

Supposing, then, for contradiction, that $\frac{z(c+\epsilon)}{z'(c+\epsilon)}\geq\epsilon$, then in case (i), we would have $\frac{z(c+\epsilon)}{\epsilon}\geq z'(c+\epsilon)$ and, therefore, $z'(c+\epsilon')\geq z'(c+\epsilon)$, contradicting the assumption (i). and in case (ii), one obtains the same contradiction, apply the above argument to the function $-z(x)$, instead of $z(x)$. This shows that $\frac{z(c+\epsilon)}{z'(c+\epsilon)}<\epsilon$, for $0<\epsilon<(b-c)$. A similar argument shows that, in case (i), we must have that $z(c+\epsilon)$ is strictly positive, and, in case (ii), we must have that $z(c+\epsilon)$ is strictly negative, for $0<\epsilon<(b-c)$. In both cases, we obtain that $0<\frac{z(c+\epsilon)}{z'(c+\epsilon)}$, for $0<\epsilon<(b-c)$. Hence, (††) is shown, and, the lemma is proved, with the assumption (†). In order to handle the general case, suppose that $z'(c)=m$, and make the invertible change of coordinates, defined by;

$$\theta : x\mapsto(x+c), w\mapsto(mx+w+z(c))$$

If the graph of the function $z(x)$ is defined by $Z(x,w)=w-z(x)=0$, then, one obtains a corresponding function, under this coordinate change, defined by;

$$\overline{Z}(x,w)=(Z\circ\theta)(x,w)=w+(mx-z(x+c))+z(c)$$

which is the graph of the function $\overline{z}(x)=z(x+c)-mx-z(c)$.

The function $\overline{z}(x)$ has the property that $\overline{z}(0)=0$ and $\overline{z}'(0)=0$. Moreover, $\overline{z}'(x)=z'(x+c)-m$, which is, again, either strictly increasing, or strictly decreasing, on the interval $[a-c,b-c]$.

We can, therefore, apply the previous result, in order to obtain an interval $(0,b-c)$, for which the result of the lemma holds for this function $\overline{z}$ at 0. As the invertible change of coordinates, defined by θ, has the property that it translates the x coordinate by c, and maps the tangent lines of $\overline{z}$ to the corresponding tangent lines of z, it is a straightforward exercise to see that the interval (c,b) has the property required for the function $z(x)$ at c. Hence, the lemma is shown.

□

Lemma 0.21. *Let $z(x)$ be an analytic function, defined on the closed interval $[a,b]$, satisfying the conditions of the previous lemma. Let*

$a = c_0 < d_0 < \ldots < c_j < d_j < \ldots < d_{n-1} < b = c_n$ *define a partition, where;*

$c_j = a + \frac{j(b-a)}{n}$, *for* $0 \leq j \leq n$.

$d_j = x(p_{c_j c_{j+1}})$, *for* $0 \leq j \leq n-1$, *as defined in the previous lemma.*

Then;

$z(b) - z(a) = \sum_{0 \leq j \leq (n-1)} (d_j - c_j) z'(c_j) + \sum_{0 \leq j \leq (n-1)} (c_{j+1} - d_j) z'(c_{j+1})$ (†)

Proof. Consider the closed interval $[c_j, c_{j+1}]$, for some $0 \leq j \leq n-1$, so that $c_j < d_j < c_{j+1}$. Let Ω denote the point of intersection of the tangent lines l_{c_j} and $l_{c_{j+1}}$. Let $d_j = x(\Omega)$ and $w(\Omega)$ denote the coordinates of Ω. Then we have;

a. $z(c_{j+1}) - z(c_j) = (z(c_{j+1}) - w(\Omega)) + (w(\Omega) - z(c_j))$

b. $z(c_{j+1}) - w(\Omega) = (c_{j+1} - d_j) z'(c_{j+1})$

c. $w(\Omega) - z(c_j) = (d_j - c_j) z'(c_j)$

where *b.* and *c.* follow immediately from the definition of the derivative $z'(x)$, and *a.* is immediate. Summing *b.* and *c.*, we obtain;

$z(c_{j+1}) - z(c_j) = (d_j - c_j) z'(c_j) + (c_{j+1} - d_j) z'(c_{j+1})$

The result of the lemma then follows by summing over all the closed intervals $[c_j, c_{j+1}]$, for $0 \leq j \leq n-1$.

□

Theorem 0.22. *Newton's Version of The Fundamental Theorem of Calculus*

Let $y(x)$ *be an analytic function, defined on the closed interval* $[a, b]$, *and let* $z(x)$ *be an analytic function, defined on* $[a, b]$, *with the additional property that* $\frac{dz}{dx} = y$. *Then;*

$\int_a^b y(x)dx = z(b) - z(a)$

Proof. We first prove the theorem with this assumption that $z(x)$ is simple, (†). Let $\{y_n : n < \infty\}$ be the sequence of functions, defined by;

$y_n(x) = z'(c_j)$ if $x \in [c_j, d_j]$, for $(0 \leq j \leq n-1)$

$y_n(x) = z'(c_{j+1})$ if $x \in [d_j, c_{j+1}]$, for $(0 \leq j \leq n-1)$

where $\{c_0, \ldots, c_n\}$ and $\{d_0, \ldots, d_{n-1}\}$ are provided by the previous Lemma 0.21. By the assumption that $z'(x) = y(x)$, it is a straightforward exercise to check that, for $x \in [a, b]$, $lim_{n\rightarrow\infty} y_n(x) = y(x)$. Moreover, by the fact that $y(x)$ is either strictly increasing or strictly decreasing on the interval $[a, b]$, each function $|y_n(x)|$ is dominated by the value $max\{|y(a)|, |y(b)|\}$. It follows, using the dominated convergence theorem, ([19]) that;

$\int_a^b y(x)dx = lim_{n\rightarrow\infty} \int_a^b y_n(x)dx$ $(*)$

However, by the previous Lemma 0.21, we have that;

$z(b) - z(a) = \int_a^b y_n(x)dx$

Hence, (†) follows. The general result then follows by applying Lemma 0.19.

□

Remarks 0.23. *Newton's Theorem 0.22, although slightly different in structure to Theorem 0.3, is essentially equivalent for analytic functions. For suppose that $f(x)$ is an analytic function on the closed interval $[a, b]$, and we have shown Theorem 0.22. Let $F(x)$ be the function defined in Theorem 0.3, and let $G(x)$ be an antiderivative of $f(x)$, that is $G'(x) = f(x)$ on $[a, b]$, ([20]). By Theorem 0.22, we have that $F(x) = G(x) - G(a)$. In particular, we obtain $F'(x) = G'(x) = f(x)$,*

[19] One could avoid the technicalities of this theorem to show $(*)$, by taking $(*)$ as a definition of integration. This was clearly Newton's intuition in his proof, here, of The Fundamental Theorem of Calculus. The major technical difficulty in such a definition is that the quadrangles defined by each function $y_n(x)$ have variable width. Newton, in fact, considers this problem in Lemmas 3 and 4, Section 1, Book 1 of his paper $(viii)$. The interested reader might try to formulate such a definition, either using infinitesimal partitions, (some work in this direction was done in [7], Theorem 3.5.2), or more conventional limit arguments. See also footnote 15.

[20] Such a function is easily constructed for an analytic function f, by integrating each term of its power series expansion. This technique was also developed by Newton in (v).

for $x \in (a,b)$. Hence, Theorem 0.3 holds for such an analytic function f. Conversely, suppose that Theorem 0.3 is shown for f, let $G(x)$ and $F(x)$ be as defined above, then we obtain that both $G'(x) = F'(x) = f(x)$ on (a,b). Applying elementary results, not depending on the theory of integration, we have that $G(x) = F(x) + c$, where c is a constant. Then $G(b) - G(a) = F(b) - F(a) = \int_a^b f(x)dx$, by definition of F. Therefore, Theorem 0.22 holds for the analytic function f as well.

Remarks 0.24. *The attentive reader may, at this stage, be wondering why Newton makes the step of introducing the normal to the curve defined by $z(x)$ in his Proposition which we considered above. Although unnecessary for his calculation to go through, Newton makes a connection between his definition of curvature, see footnote 7, and his proof of The Fundamental Theorem of Calculus. The relationship is the following;*

(i). His calculus proof depends on the fact that;

For a given analytic function $f(x)$, if l_{x_0} denotes the tangent line to the curve C at $O = (x_0, f(x_0))$, defined by f, then, for an infinitesimal ϵ, if $l_{x_0+\epsilon}$ denotes the tangent line to the curve at $(x_0 + \epsilon, f(x_0 + \epsilon))$, the intersection $O_\epsilon = l_{x_0} \cap l_{x_0+\epsilon}$ lies in the infinitesimal neighborhood $\mathcal{V}_O$.

(ii). In his definition of curvature;

With the same conditions on f, if n_{x_0} denotes the normal to the curve at $O = (x_0, f(x_0))$, defined by f, then, for an infinitesimal ϵ, if $n_{x_0+\epsilon}$ denotes the normal to the curve at $(x_0 + \epsilon, f(x_0 + \epsilon))$, the intersection $N_\epsilon = n_{x_0} \cap n_{x_0+\epsilon}$ lies in the infinitesimal neighborhood $\mathcal{V}_K$, where K is the centre of curvature of the curve C at O.

The interested reader can find out more about Newton's work on curvature in, for example, the paper (ii), his "Optics" and the paper $(viii)$.

Although the mechanism of Newton's proof of The Fundamental Theorem of Calculus is longer than the modern approach, Theorem 0.3, the geometric idea behind it is more elegant, for the following reasons. First, his argument relies on a "global" geometric relationship between the function $y(x)$ and its antiderivative $z(x)$, thus the global definition of integration is incorporated directly into the proof. In Theorem 0.3, a local property of the integrated function $y(x)$ is used, which, as we will see presently, is, on a geometric level, slightly unsatisfactory. Secondly, Newton explores, in his argument, a connection between curvature and integration. It is not my intention, here, to discuss this connection in great depth, but will make the following remark. The notion of locus of curvature may be defined for any algebraic curve C, using Newton's method, ([21]). Understanding the geometry of the locus of curvature was, for Newton, see final footnote, intrinsically related to understanding the geometry of the curve C itself, and Newton could be said to establish some geometric link with his theory of integration, here. This is not only interesting in itself, but, further confirms the primary role that aesthetic and geometric connections play in Newton's work. A deeper discussion of these connections can be found in my book [5] or on my website, http://www.curveline.net

References

[1] Tristram de Piro, Some Notes on Non-Standard Differentiation and Integration, available at http://www.magneticstrix.net

[2] Tristram de Piro, Infinitesimals in a Recursively Enumerable Prime Model, available at http://www.magneticstrix.net

[3] Tristram de Piro, A Non-Standard Bezout Theorem for Curves, available at http://www.magneticstrix.net

[4] Tristram de Piro, A Theory of Branches for Algebraic Curves, available at http://www.magneticstrix.net

[5] Tristram de Piro, Christian Geometry, in preparation.

[6] Rob Goldblatt, Lectures on the Hyperreals: An Introduction to Nonstandard Analysis Graduate Texts in Mathematics 188, Springer-Verlag, New York, 1998.

[7] Abraham Robinson, Non-Standard Analysis, Princeton University Press, (1996)

[8] Walter Rudin, Real and Complex Analysis, McGraw-Hill, (1987)

[9] D.T Whiteside, The Mathematical Paper of Isaac Newton, Volume 1, C.U.P

Flat 1, 98 Prestbury Road, Cheltenham, GL52 2DJ
E-mail address: `tristam.depiro@unicam.it`

[21] Newton's interest in curvature is related to his idea that, if the curve C could be considered as a perfect refractive surface, light shining on the curve would be focused along the locus of curvature