Available online at www.sciencedirect.com

SCIENCE DIRECT

Historia Mathematica 30 (2003) 432–440

HISTORIA MATHEMATICA

www.elsevier.com/locate/hm

Passage to the limit in Proposition I, Book I of Newton's *Principia*

Herman Erlichson

The College of Staten Island, The City University of New York, Staten Island, NY 10314, USA

Abstract

The purpose of this paper is to analyze the way in which Newton uses his polygon model and passes to the limit in Proposition I, Book I of his *Principia*. It will be evident from his method that the limit of the polygon is indeed the orbital arc of the body and that his approximation of the actual continuous force situation by a series of impulses passes correctly in the limit into the continuous centripetal force situation. The analysis of the polygon model is done in two ways: (1) using the modern concepts of force, linear momentum, linear impulse, and velocity, and (2) using Newton's concepts of motive force and quantity of motion. It should be clearly understood that the term "force" without the adjective "motive," is used in the modern sense, which is that force is a vector which is the time rate of change of the linear momentum. Newton did not use the word "force" in this modern sense. The symbol F denotes modern force. For Newton "force" was "motive force," which is measured by the change in the quantity of motion of a body. Newton's "quantity of motion" is proportional to the magnitude of the modern vector momentum. Motive force is a scalar and the symbol F_{m} is used for motive force.

Résumé

Le but de ce papier est d'analyser la façon dont Newton emploie son modèle polygonale et passe à la limite dans Proposition I, livre I de son *Principia*. Il sera évident de sa méthode que la limite du polygone est l'arc orbitale du corps et que son approximation de l'actuel force continuelle par une série d'impulses passe correctement dans la limite au force continuel centripete. L'analyze du modèle polygone est fait par deux méthodes : (1) utilisant les concepts modernes de force, momentum linéar, impulse linéare, et vélocité, et (2) utilisant les concepts Newtonien de force motive et quantité de motion. Il doit être clairement compris dans le papier entier que dans notre usage du terme « force » sans l'adjectif « motive » nous l'employons dans le sens moderne, que la force est une vecteur qui est le changement avec le temps du momentum linéaire. Newton n'a pas employé le mot « force » dans ce sens moderne. Nous employons la symbole F pour force moderne. Pour Newton « force motive », est mesuré par le changement dans la quantité de motion d'un corps. La quantité de motion de Newton est proportionel au grandeur du momentum vecteur moderne. La force motive est un scaler et nous employons le symbole F_{m} pour la force motive.

E-mail address: erlichson@postbox.csi.cuny.edu.

0315-0860/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0315-0860(02)00008-3

The purpose of this paper is to analyze the way in which Newton uses his polygon model and passes to the limit in Proposition I, Book I of his *Principia* [Newton, I. 1687]. It will be evident from his method that the limit of the polygon is indeed the orbital arc of the body and that his approximation of the actual continuous force situation by a series of impulses passes correctly in the limit into the continuous centripetal force situation. The analysis of the polygon model will be done in two ways: (1) using the modern concepts of force, linear momentum, linear impulse, and velocity, and (2) using Newton's concepts of motive force and quantity of motion.

It should be clearly understood throughout this paper that when we use the term "force" without the adjective "motive," we use it in the modern sense, in which force is a vector that is the time rate of change of the linear momentum. Newton did not use the word "force" in this modern sense. We use the symbol F for modern force. For Newton "force" was "motive force," which is measured by the change in the quantity of motion of a body. Newton's "quantity of motion" is proportional to the magnitude of the modern vector momentum. Motive force is a scalar and we will use the symbol F_m for motive force.

For ready reference we present a brief table of the terms used by Newton and their modern counterparts (see Table 1).

It should also be noted that Newton always talked about motive force in a comparative sense. Thus, for example, in Proposition I, Corollary III, he talks about the ratio of the motive forces at two orbital points. As can be seen from the table above the limiting ratio of two motive forces is the same as the ratio of two modern force magnitudes.

1. Newton's passage to the limit

We begin by considering Newton's approximation of a continuous centripetal force by a series of instantaneous impulses in Proposition I. In modern terminology these instantaneous impulses, which discontinuously change the momentum of the body at a single point, correspond to delta function forces. In terms of Newton "motive force," however, the motive forces associated with Newton's instantaneous impulses are finite motive forces. Newton did not, of course, describe his instantaneous impulses as corresponding to delta function forces; but in terms of the modern force concept this is a correct description, since the instantaneous impulses act to produce finite momentum changes in vanishingly small time intervals. Newton's concept of instantaneous impulses, when described in terms of the modern concept of force, involves the Dirac delta function approximately 250 years before Dirac used it in quantum mechanics [Dirac, 1958, Chap. III, Sect. 15].

The skeptical reader has but to consider Newton's *Principia* diagram illustrating Proposition I of Book I, here shown as our Fig. 1, and to reflect that the abrupt changes of path shown in that diagram have

Table 1

Newton	Modern
Quantity of motion, proportional to the scalar quantity mv	Linear momentum, the vector mv
Motive force, a scalar, measured by the change in the quantity of motion	Force, a vector, measured by the rate of change of the linear momentum
$\lim_{\Delta t\to 0} \frac{\text{Motive force } F_{\mathrm{m}1} \text{ for } \Delta t}{\text{Motive force } F_{\mathrm{m}2} \text{ for } \Delta t}$	$\frac{F_1}{F_2}$

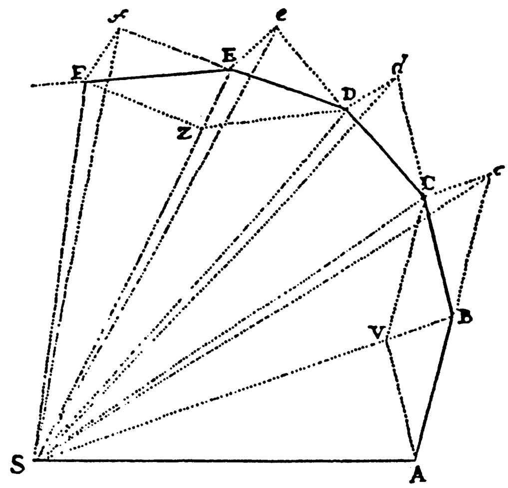

Fig. 1. Newton's diagram for Proposition I (from Newton [1687, Vol. I, p. 40]).

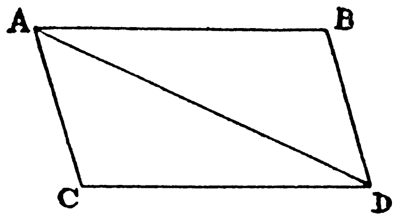

Fig. 2. Newton's diagram for Corollary I (from Newton [1687, p. 14]).

to be associated with velocity discontinuities due to instantaneous impulsive forces. We quote Newton on the nature of the force causing the abrupt velocity change at point B:

> when the body is arrived at B, suppose that a centripetal force acts at once with a great impulse, and, turning aside the body from the right line Bc, compels it afterwards to continue its motion along the right line BC [Newton, 1687, p. 40].

The "great impulse" at B is proportional to the quantity of motion change of the body at B. Moreover, Proposition I is not the first time in the *Principia* that Newton uses instantaneous impulses. In his earlier Corollary I on the composition of forces Newton said (see Fig. 2, which is Newton's diagram for Corollary I):

> If a body in a given time, by the force M impressed apart in the place A, should with uniform motion be carried from A to B, and by the force N impressed apart in the same place, should be carried from A to C, let the parallelogram ABCD be completed, and by both forces acting together, it will in the same time be carried in the diagonal from A to D [Newton, 1687, p. 14].

The force M and the force N in this corollary are instantaneous impulsive forces, just as in Theorem I. They change the velocity of a body instantaneously. If M acts alone the velocity of the body is abruptly changed from zero at point A and the body then moves with constant velocity from A to B; likewise, if N acts alone at point A the velocity of the body is abruptly changed at A and the body then moves with constant velocity from A to C.

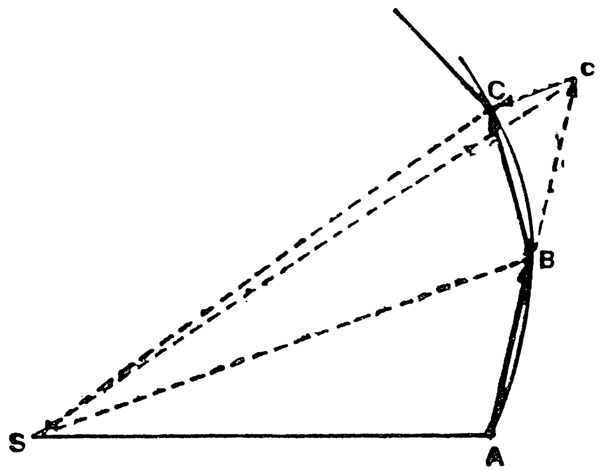

Fig. 3. The time Δt to go from B to C along **BC** equals the time to go from A to B along **AB**.

We return now to Proposition I and Fig. 1. Although Newton did not state it explicitly, it can be assumed that the points A, B, C, ... shown in the figure are actual points on the orbit of a body moving under the action of a centripetal force. This standard method of breaking up a continuous curve into a discrete form was used by Newton in his *Waste Book* analysis of circular motion [Herivel, 1965, pp. 129–130, 146–147] and later in the *Principia*. In Proposition I the approximating polygonal orbit is broken up into finite equal-time segments, each segment being covered in time Δt. It is important to realize that the equal-time points on the actual orbit are slightly different from the equal-time points on the polygonal orbit; they only become identical in the limit $\Delta t \to 0$.

Consider Fig. 3, which shows the actual orbit between A and C and the two straight line sections AB and BC of the polygonal orbit approximation. In this approximation the body moves force-free along AB with constant velocity $v_{\mathrm{B}-}$, which is equal to $\mathbf{AB}/\Delta t$. When it arrives at B with this velocity it is subjected to a delta function centripetal force $\mathbf{F}_{\mathrm{B}\delta}$ (modern force definition) of size $m(\mathbf{v}_{\mathrm{C}-} - \mathbf{v}_{\mathrm{B}-})$, where m is the mass of the body and $\mathbf{v}_{\mathrm{C}-}$ is the constant velocity with which the body travels along **BC**, i.e., the constant velocity of the body just before it receives the next impulse at point C. Thus,

$$\mathbf{F}_{\mathrm{B}\delta} = m(\mathbf{v}_{\mathrm{C}-} - \mathbf{v}_{\mathrm{B}-})\delta(t_{\mathrm{B}}), \tag{1}$$

where $\delta(t_{\mathrm{B}})$ has been written for $\delta(t - t_{\mathrm{B}})$, and $\delta(t)$ is the Dirac delta function, defined by

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1, \quad \delta(x) = 0 \text{ for } x \neq 0.$$

In sharp contrast to the delta function complexities needed if the modern force concept is used, Newton's motive force for the impulse at B is simply measured by the change in the quantity of motion at B,

$$F_{\mathrm{mB}} \alpha m |(\mathbf{v}_{\mathrm{C}-} - \mathbf{v}_{\mathrm{B}-})|.$$

Returning to the modern description, the change in momentum at point B due to the application of the impulse at B is given by

$$\int_{t_{B-}}^{t_{B+}} F_{B\delta}\, dt = m(\mathbf{v}_{C-} - \mathbf{v}_{B-}). \tag{2}$$

The velocity $\mathbf{v}_{C-}$ is controlled by **BC**, since $\mathbf{v}_{C-} = \mathbf{BC}/\Delta t$. The determination of **BC** is done as follows:

(1) **AB** is extended an equal directed distance **Bc** to point c so that **Bc** is equal to the displacement the body would have undergone in time Δt with the constant velocity $\mathbf{v}_{B-}$.
(2) From point c a displacement **cC** parallel to **BS** is drawn. This displacement is that due to the delta function impulse at point B. Since the change in velocity of the body at point B is $\mathbf{v}_{C-} - \mathbf{v}_{B-}$, we have

$$\mathbf{cC} = (\mathbf{v}_{C-} - \mathbf{v}_{B-})\Delta t - (\Delta \mathbf{v})\Delta t.$$

The terminal point of the displacement **cC** is point C, which is on the orbit of the body.
(3) The connecting displacement **BC** is now drawn.

The process described above is repeated at point C and for the subsequent points D, E, F, ... It is important to note that for the polygonal approximating path ABC... the time intervals Δt between points are equal, but that the time intervals between points on the actual curved path are, in general, unequal. In the limit $\Delta t \to 0$ these time intervals on the actual path become equal.

The areas of triangles SAB and SBc are equal because they have equal bases **AB** and **Bc** (each equal to $v_{B-}\Delta t$) and a common height (the perpendicular distance from S to the straight line ABc). The areas of triangles SBc and SBC are equal because they have the same base SB and a common height (because **cC** is parallel to **BS**). Hence,

$$\text{area}\,\Delta\text{SAB} = \text{area}\,\Delta\text{SBC},$$

and Newton has shown thereby that for the polygonal path equal areas are tracked out in equal times.

Now let $\Delta t \to 0$ and the polygonal path will become the actual path of the orbiting body. As the polygon sides become vanishingly small the equal areas in equal times continues to hold. Hence it holds for the actual curved orbit, and Newton has proven Proposition I.

In terms of Newton's motive force concept how did he see the fundamental triangle BcC of Fig. 3? BC is proportional to the orbital motive force (proportional to v_B), cC is proportional to the motive force of the instantaneous impulse at B (proportional to Δv), and BC is proportional to the orbital motive force after the impulse at B has struck (proportional to v_{B+}). If we compare motive forces at two orbital points such as B and E, the motive force at B for the time interval Δt is proportional to cC (or BV) (see Fig. 1) and the motive force at E is likewise proportional to EZ. We have

$$\frac{F_{mB}(\text{for } \Delta t)}{F_{mE}(\text{for } \Delta t)} = \frac{\text{BV}}{\text{EZ}}.$$

When we pass to the limit $\Delta t \to 0$,

$$\lim_{\Delta t \to 0} \frac{F_{mB}(\text{for } \Delta t)}{F_{mE}(\text{for } \Delta t)} = \frac{F_{mB}}{F_{mE}} = \lim_{\Delta t \to 0} \frac{\text{BV}}{\text{EZ}}.$$

This equation is a statement of Newton's Corollary III to Proposition I. We see that passage to the limit is a very simple process in terms of Newton's comparison of motive forces.

The analysis of the passage to the limit in terms of the modern concept of force is considerably more complicated because we have to deal with the delta function. The delta function force given by Eq. (1) becomes, in the limit,

$$\lim_{\Delta t\to 0} \mathbf{F}_{\mathrm{B}\delta} = \mathbf{F}_{\mathrm{B}} = m(d\mathbf{v}/dt)_{\mathrm{B}},$$

where $\mathbf{F}_{\mathrm{B}}$ is the centripetal force at B for the curved orbit. To see that this is the case we start by rewriting Eq. (2). Let $\mathbf{F}_{\mathrm{B}\delta'}$ be a large force which acts only during the very short time interval Δt_{B} centered around t_{B}. Then Eq. (2) can be written

$$\mathbf{F}_{\mathrm{B}\delta'}\Delta t_{\mathrm{B}} = m(\mathbf{v}_{\mathrm{C}-} - \mathbf{v}_{\mathrm{B}-}). \tag{3}$$

The impulse of $\mathbf{F}_{\mathrm{B}\delta'}$ corresponds closely to the great impulse described by Newton in Proposition I when he said "when the body is arrived at B, suppose that a centripetal force acts at once with a great impulse, and, turning aside the body from the right line Bc, compels it afterwards to continue its motion along the right line BC." In the limit as $\Delta t \to 0$ the large force $\mathbf{F}_{\mathrm{B}\delta'}$ passes to the delta function force $\mathbf{F}_{\mathrm{B}\delta}$:

$$\lim_{\Delta t_{\mathrm{B}}\to 0} \mathbf{F}_{\mathrm{B}\delta'} = \mathbf{F}_{\mathrm{B}\delta}. \tag{4}$$

We now consider the limit of Eq. (3) as $\Delta t \to 0$. As $\Delta t \to 0$, $\mathbf{v}_{\mathrm{C}-} \to \mathbf{v}_{\mathrm{B}-} + d\mathbf{v}_{\mathrm{B}}$ and $\mathbf{v}_{\mathrm{C}-} - \mathbf{v}_{\mathrm{B}-} \to d\mathbf{v}_{\mathrm{B}'}$, and we have

$$\lim_{\Delta t\to 0} (\mathbf{F}_{\mathrm{B}\delta'}\Delta t_{\mathrm{B}}) = m\, d\mathbf{v}_{\mathrm{B}}.$$

Now, as $\Delta t \to 0$ the short time interval Δt_{B} must correspondingly go to zero (since Δt_{B} is always small in relation to Δt), so in this limit $\Delta t_{\mathrm{B}} \to dt_{\mathrm{B}}$ and $\mathbf{F}_{\mathrm{B}\delta'} \to \mathbf{F}_{\mathrm{B}\delta}$ by Eq. (4). Hence, we have,

$$\lim_{\Delta t\to 0} (\mathbf{F}_{\mathrm{B}\delta})\, dt_{\mathrm{B}} = m\, d\mathbf{v}_{\mathrm{B}},$$
$$\lim_{\Delta t\to 0} \mathbf{F}_{\mathrm{B}\delta} = m\, d\mathbf{v}_{\mathrm{B}}/dt_{\mathrm{B}} = m(d\mathbf{v}/dt)_{\mathrm{B}}. \tag{5}$$

From Newton's Second Law (modern form) we know that

$$\mathbf{F}_{\mathrm{B}} = m(d\mathbf{v}/dt)_{\mathrm{B}}, \tag{6}$$

where $\mathbf{F}_{\mathrm{B}}$ is the centripetal force exerted on the body at point B in the orbit. Comparing Eq. (5) with Eq. (6), we see that

$$\lim_{\Delta t\to 0} \mathbf{F}_{\mathrm{B}\delta} = \mathbf{F}_{\mathrm{B}}. \tag{7}$$

This result may seem surprising since $\mathbf{F}_{\mathrm{B}\delta}$ is a spike, or delta function force, which is of infinite size for any finite Δt. However, in the limit as $\Delta t \to 0$ this spike becomes the finite force $\mathbf{F}_{\mathrm{B}}$. With the passage to the limit of the polygonal path to the curved orbit the limit of the impulsive delta function force is the actual centripetal force (modern force).

2. Alternate polygonal model for passage to the limit

It is interesting to note that there is an alternate passage to the limit which could have been used by Newton. In Newton's method the impulse applied at B is along the central direction SB and the areas of all the focal triangles SAB, SBC, etc. are all equal for any t. Suppose, however, that we instead use the actual orbital equal time points A, B, C′, D′, etc. and make these points the basis for our focal triangles. The situation is shown in Fig. 4.

The beginning of the polygonal path is done as before. The time Δt to cover the orbit arc AB or the chord AB is the basis for the equal times to cover the polygon sides. The body moves along chord **AB** with speed $v_{B-} = AB/\Delta t$, exactly as before. Now, however, the impulse at B is different and adjusted so that after the next Δt the body is at C′, where the orbit time to go from B to C′ is equal to the orbit time to go from A to B. The velocity change of the body is at C′, where the orbit time to go from B to C′ is equal to the orbit time to go from A to B. The velocity change $\Delta \mathbf{v}$ is now given by $\Delta \mathbf{v} = \mathbf{cC}'/\Delta t$. The point C′ is, in general, different from the former point C, which is also shown in Fig. 4. Since cC′ is not, in general, parallel to SB it will no longer be the case that the area of focal triangle SAB equals the area of focal triangle SBC′.

However, the actual orbital area swept out from A to B equals the orbital area swept out from B to C′, etc. In the limit as $\Delta t \to 0$, cC′ will become parallel to SB, the infinitesimal focal triangles will become equal in area, and the polygon will become the curved orbit. Thus, in the limit, the areas of all infinitesimal focal sectors traversed in equal infinitesimal times will be equal one to the other. From this it follows that equal areas are tracked out in equal times along the curved orbit.

Why did Newton choose his polygonal path rather than this alternate path? One can only speculate. Perhaps Newton thought it was neater to have the focal triangles equal in area for all Δt rather than equal in area only in the limit. However this may be, for either polygonal method the points on the polygon remain on the actual curved orbit as we pass to the limit, and both methods are exact.

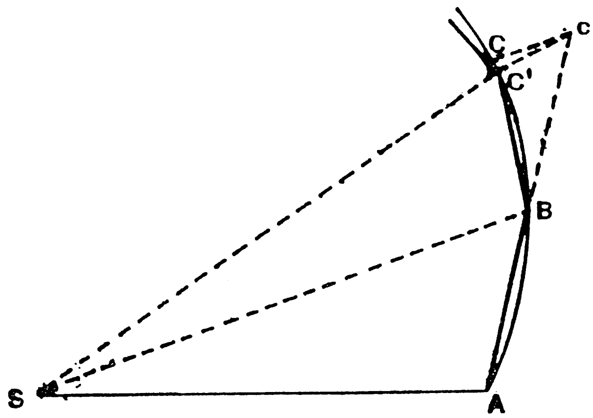

Fig. 4. The orbital time along arc $\overset{\frown}{BC}$ equals the orbital time along arc $\overset{\frown}{AB}$, and also equals the chord time along chords **AB**, **BC′**, etc.

3. Connection of this paper with the earlier paper (September 1991) published in the *American Journal of Physics*

Newton's concept of force is not easy to understand, especially since it was developed over a period of time. One should consult the paper "Motive Force and Centripetal Force in Newton's Mechanics," which was published in the September 1991 issue of the *American Journal of Physics*, to obtain an account of Newton's development of his force concept.

We here quote from the abstract of that paper:

> Newton's concept of force was developed over a period of time, starting with a collisional model in his *Waste Book*, and culminating in his measure of the centripetal force at a point found in *De Motu* and again in *Principia*, Book I, Proposition VI. Newton kept developing his force concept, adding to it and making it many faceted. This rich concept of force has often posed difficulties for historians of science [for example, R.S. Westfall, *Force in Newton's Physics* (MacDonald and American Elsevier, New York, 1971) and J. Herivel, *The Background to Newton's Principia* (Oxford U.P., New York, 1965)] and others. These difficulties are related to at least six things as follows: (1) lack of understanding that Newton's "motive force" is an abbreviation for "motive quantity of a force," so that Newton's "motive force" and his "force" are not two different types of force. Motive force (motive quantity) is used to quantify force; (2) lack of understanding of Newton's model for instantaneous impulses; (3) lack of understanding that although Newton's motive force has directionality Newton added it as a scalar in his analysis of uniform circular motion; (4) lack of understanding of Newton's polygon model; (5) lack of understanding that Newton's measure of the motive value of a force, the motive force, was always used to obtain a comparative measure of forces in his early mechanics; and (6) lack of understanding that Newton's Prop. VI measure of centripetal force at a point permitted Newton to bring the time dimension into his motive force concept, thus developing it into something very close to our modern definition of force. This paper analyzes Newton's development of his motive force concept, and discusses the difficulties that some scholars have had with various aspects of this concept.

The 1991 paper starts with Section I, entitled "Newton's Motive Force Concept and Instantaneous Impulses." The section starts out with Newton's Definition VIII: "*The motive quantity of a centripetal force is the measure of the same, proportional to the motion which it generates in a given time*." The paper continues with "The quantity of motion of a moving body was defined by Newton as follows: Definition II: The quantity of motion is the measure of the same, arising from the velocity and quantity of matter conjointly."

The word "velocity" in Newton's definition of quantity of motion is the scalar speed of the body. Both quantity of motion and motive force have directionality but both were added as scalars by Newton. The exception to this was instantaneous motive forces, which Newton added by the parallelogram law of addition.

To understand Newton's work, particularly the early work in the *Waste Book*, the modern reader can be aided by referring to modern symbols and modern concepts, but he or she also has to make the necessary effort to understand the concepts of Newton's 17th-century mechanics. In the *Waste Book*, Newton thought of force as the producer of change of motion. Forces were compared by comparing the motion changes they produced for equal time intervals.

Much of the confusion that has existed about Newton's concept of force is due to the use of the terms "motive force" and "centripetal force." Strictly speaking, motive force is the *motive quantity* of a force as defined by Newton in Definition VIII. In modern terms, Newton's motive quantity of a force is proportional to the scalar magnitude of the change in momentum, $\Delta m\mathbf{v}$, or its equivalent, the scalar magnitude of the impulse $\mathbf{J}$. Newton said "These *quantities* of forces, we may, for the sake of *brevity*, call by the names of motive, accelerative, and absolute forces." The force F is measured by the motive force, proportional to $|\Delta m\mathbf{v}|$, but is not equal to the motive force because the time interval has to be taken into account. The force F is proportional to the motive force divided by the time interval. In modern terms $F = \lim(|\Delta m\mathbf{v}|/\Delta t)$. When two steady (constant in magnitude) forces are compared for equal time intervals, the steady forces are proportional to the motive quantities of those forces for equal time intervals.

Although Newton defined two other quantities of a force, absolute and accelerative, it is clearly the motive quantity of a force that plays by far the central role in Newtonian mechanics. The absolute quantity of a force is a measure of the strength of the force source and the accelerative quantity of a force is simply the motive quantity divided by the mass.

The concepts of quantity of motion and motive force originated in the consideration of collisional forces. A collisional force is an impulsive force that acts for a very short interval of time and that changes the quantity of motion of a body. Newton idealized actual collisional forces by reducing the time interval of the collision to zero (or an infinitesimal). This idealization was Newton's "instantaneous impulse." Motive force was measured by the change in the quantity of motion (what we would now call the magnitude of the change in linear momentum). Directionality was introduced by having the action of the motive force in the direction of the change in motion. Motive force for collisional forces depended on the change in motion and did not depend on whether this change was accomplished all at once or over a time interval. If it was accomplished all at once it was an instantaneous impulse.

Newton's instantaneous impulse would be today described as an infinite force (modern concept of force) that discontinuously changes the linear momentum of a body in zero time. In terms of Newton's mechanics there was no mention of an infinite force because Newton was thinking of the finite change of motion (proportional to the finite motive force). Newton's instantaneous impulses involve a finite motive force (finite impulse in modern terms) which acts for an instant. For Newton, the motive force of the instantaneous impulse is associated with a finite Δt, so the force measured by the instantaneous impulse is finite. In modern terms, the force associated with the instantaneous impulse is the average force. When Δt is shrunk to zero, the average force becomes the instantaneous force at a point.

References

Dirac, P.A.M., 1958. The Principles of Quantum Mechanics, 4th Edition. Oxford Univ. Press, London.

Herivel, J., 1965. The Background to Newton's Principia. Oxford Univ. Press, London.

Newton, I., 1687. Philosophia Naturalis Principia Mathematica, 1st Edition. English transl. A. Motte, 1729; rev. F. Cajori, Univ. of California Press, Berkeley, LA, 1934. All page references are to the revised edition.