Reading the Master: Newton and the Birth of Celestial Mechanics

Bruce Pourciau

Dedicated to J. Bruce Brackenridge

One factor that has remained constant through all the twists and turns of the history of physical science is the decisive importance of the mathematical imagination.
—Freeman J. Dyson

I. In January of 1684, the young astronomer Edmund Halley travelled from Islington up to London for a meeting of the Royal Society. Later, perhaps over tea and chocolate at a nearby coffee house, he chatted casually about natural philosophy and other topics with Sir Christopher Wren and Robert Hooke. Talk soon turned to celestial motions, and Halley later reconstructed the conversation [22, p. 26]:

I, having from the consideration of the sesquialter proportion of Kepler concluded that the centripetal force [to the Sun] decreased in the proportion of the squares of the distances reciprocally, came one Wednesday to town, where I met with S' Christ. Wren and M' Hook, and falling in discourse about it, M' Hook affirmed that upon that principle all the Laws of the celestiall motions were to be demonstrated, and that he himself had done it. I declared the ill success of my attempts; and S' Christopher to encourage the Inquiry said that he would give M' Hook or me 2 months time to bring him a convincing demonstration thereof, and besides the honour, he of us that did it, should have from him a present of a book of 40 shillings. M' Hook then said that he would conceal [his] for some time that other triing and failing, might know how to value it, when he should make it publick. ...I remember S' Christopher was little satisfied that he could do it, and though M' Hook then promised to show it him, I do not yet find that in that particular he has been as good as his word.

The two month deadline passed. Wren and Halley waited through the summer, but still the promised proof from Hooke never came. Finally, in August, Halley would wait on Hooke no longer. He carried the question to Cambridge and the Lucasian Professor of Mathematics, Isaac Newton.

Newton's secretary and attendant has painted a portrait, daubed with colorful and concrete detail, of the eccentric Cambridge professor Halley had finally decided to approach [12, p. xiii–xiv]:

I cannot say, I ever saw him laugh, but once...I never knew him take any Recreation or Pastime, either in Riding out to take ye Air, Walking, Bowling
or any other Exercise whatever, thinking all Hours lost, \textit{y} was not spent in his Studyes, to \textit{w}E he kept so close... so intent, so serious upon [them], \textit{y} he eat very sparingly, nay, oft times he has forgot to eat at all, so \textit{y} going into his Chamber I have found his Mess untouch'd, of \textit{w}E when I have reminded him, [he] would reply, Have I; \& then making to \textit{y} Table, would eat a bit or two standing, for I cannot say, I ever saw Him sit at Table by himself... He very rarely went to Dine in \textit{y} Hall unless upon some Publick Days, \& then, if He has not been minded, would go very carelessly, \textit{w}E Shooes down at Heels, Stockins unty'd, Suplice on, \& his Head scarcely comb'd... At some seldom Times when he design'd to dine in \textit{y} Hall [he] would turn to \textit{y} left hand, \& go out into \textit{y} street, where making a Stop, when he found his mistake, [he] would hastily turn back \& then sometimes instead of going into \textit{y} Hall, would return to his Chamber again... .

... in his Garden, \textit{w}E was never out of Order, ... he would, at some seldom Times, take a short Walk or two, not enduring to see a Weed in it... . When he has some Times taken a turn or two [he] has made a sudden Stand, turn'd himself about, run up \textit{y} Stairs [\&] like another A[r]chimedes, with an \textit{ευρήκα} fall to write on his Desk standing, without giving himself the Leasure to draw a Chair to sit down on... .

In a letter from 1727 [22, p. 27], Abraham de Moivre set the scene as Halley, having arrived in Cambridge, posed the crucial question to the reclusive mathematician:

... after they had been some time together, the D' asked [Newton] what he thought the Curve would be that would be described by the Planets supposing the force of attraction towards the Sun to be reciprocal to the square of their distance from it. S' Isaac replied immediately that it would be an Ellipsis. The Doctor struck with joy and amazement asked him how he knew it. Why saith he I have calculated it... .

Witness the birth of celestial mechanics: the embryonic question has been answered—

\textit{every orbital motion subject to an inverse-square force lies on a conic having focus at the force center}

—not with a guess, but with a \textit{mathematical demonstration}!

Semester after semester, at every college and university, we give our students the same answer Newton gave to Halley, our demonstrations—so different from Newton's—blessed by the glories of vector calculus, and in this way we honor Newton and celebrate the emergence of celestial dynamics. In the present article, we honor Newton in the way of Abel, who counsels us to read the masters. We shall place the original argument from Newton's \textit{Principia} next to a modern counterpart, delighting in the stark contrasts. One delightful difference: Newton's argument requires that we first answer the \textit{converse} to Halley's question—

\textit{What force law maintains a conic motion orbiting about the focus?}

—and again, reading the master, we shall juxtapose the \textit{Principia}'s very geometric proof of this reversal with its demonstration by vector calculus. In this mix of old
and new, of geometry and analysis, some insights and surprises make their way to the surface:

- The mathematics of the *Principia* is geometric *analysis*, both analysis in the sense of ‘taking apart’ as well as analysis in the sense of *calculus*. Newton’s geometry is calculus—limits, derivatives, integrals, acceleration, curvature—masked as geometry.
- While less precise than their vector calculus descendants, the *Principia’s* definitions have a concrete, visceral character that informs our geometric and physical intuition.
- The first ten sections of the *Principia* (apart from the statement of the Third Law) contain no physics, only mathematics. Newton may write of ‘forces,’ but he calculates accelerations. His concentration on acceleration and shape reminds us that force and mass take no part in the mathematics of the one-body problem, which occupies the leading sections of the *Principia*.
- In contrast to force, curvature is deeply involved with the *Principia’s* orbital dynamics, yet apart from rare oblique sightings, the dependence on curvature remains hidden.
- Asked who should receive credit for answering Halley’s question with a demonstration rather than a guess, historians of science bow to Newton. Asked for evidence to back up their claim, the historians open the *Principia* and point to a *two-sentence* argument. We confirm that Newton’s little sketch, given air and sun, blossoms into a cogent proof.
- Reading the masters—Archimedes, Newton, Euler, Gauss, Riemann, . . . —can mean entering a foreign paradigm, an unfamiliar mathematical world where alien values, language, definitions, tools, strategies, and assumptions frustrate our attempts to understand. And so it is with the *Principia*. But with persistence and prayer, even the *Principia* sends up her secrets. As we slowly learn to navigate in Newton’s world, we deepen our understanding of the *Principia’s* paradigm as well as our own.

It may seem odd to have placed our conclusions here in the introduction, but with these closing remarks now out of the way, we can read on unburdened by the western need to fret and fuss about the point of it all. As the Taoist philosopher Chuang Tzu suggests [19, p. 126], we can now lean back and float with the current, “going under with the swirls and coming out with the eddies, following along the way the water goes, and never thinking . . .”

2. We begin with Newton’s generalized answer to Halley—that every orbit produced by an inverse-square force must lie on a conic—in this section giving a contemporary proof and in the next exploring the *Principia’s* original argument. But we should first agree on some technical vocabulary, so that we can be more precise. Any smooth map \( r = r(t) \) from an open interval \( J \) into euclidean 3-space is a *motion*. Every motion \( r \) has a velocity \( v = \dot{r} \) and an acceleration \( a = \ddot{r} \). For the magnitude of a vector, we choose the same letter in nonbold italic: thus, for example, \( r = |r|, \) \( v = |v|, \) and \( a = |a| \). (We tacitly assume that \( r \) and \( v \) (the *speed*) never vanish.) We say the motion \( r \) has an *inverse-square acceleration* provided for some nonzero \( \lambda \),

\[
a = -\frac{\lambda}{r^2} U
\]

for all \( t \) in \( J \). Here \( U \) stands for the unit direction vector \( r/r \). More generally, whenever the cross-product \( r \times a \) vanishes identically, we call \( r \) an *orbital motion*. 
If the origin S has some significance—it might be the focus of a conic or the pole of a spiral, for instance—an orbital motion may be labelled a *motion about S*. A sentence that would be typical of the *Principia*, "A body is urged by a centripetal force continually directed toward an immovable center S," becomes briefer in our language: "Given a motion about S."

Assuming that Mars traversed an ellipse with its position vector sweeping out equal areas in equal times, Kepler made predictions in his *Astronomica nova* of 1609 that matched the careful observations of Tycho Brahe. In Propositions I and II (Section II, Book I) of the *Principia*, Newton uses this area principle to characterize orbital motions in general [11, p. 40 and 42]:

**PROPOSITION I  THEOREM I**

The areas which revolving bodies describe by radii drawn to an immovable centre of force do lie in the same immovable planes, and are proportional to the times in which they are described.

**PROPOSITION II  THEOREM II**

Every body that moves in any curved line described in a plane, and by a radius drawn to a point either immovable, or moving forwards with an uniform rectilinear motion, describes about that point areas proportional to the times, is urged by a centripetal force directed to that point.

Today of course we translate these propositions into the language of vectors:

**NEWTON’S AREA THEOREM**  For any motion \( \mathbf{r} = \mathbf{r}(t) \), the following are equivalent:

(a) \( \mathbf{r} \) is orbital
(b) the (massless) angular momentum \( \mathbf{h} = \mathbf{r} \times \mathbf{v} \) is constant
(c) \( \mathbf{r} \) is planar and sweeps out area at a constant rate

The proof is simple, especially once we agree that the area swept out is

\[
\frac{1}{2} \int_{t_0}^{t} |\mathbf{r} \times \mathbf{v}| \, dt,
\]

the only slippery step being to show \( \mathbf{r} \) is planar when \( \mathbf{h} \) vanishes everywhere, but in this case the derivative \( \dot{\mathbf{U}} \) vanishes everywhere (recall \( \mathbf{U} = \mathbf{r}/r^3 \)), indicating that the motion lies on a fixed ray from the origin. That \( \dot{\mathbf{U}} \) remains zero follows from a simple fact:

\[
\dot{\mathbf{U}} = \frac{\mathbf{h} \times \mathbf{r}}{r^3} \tag{1}
\]

Halley's question and Newton's answer involve the relationship between the acceleration of the motion and the shape of the orbit. Moving from acceleration to shape, we define the trajectory of a motion \( \mathbf{r} = \mathbf{r}(t) \) to mean the subset \( \{ \mathbf{r}(t) : t \in J \} \) of 3-space. An orbit is then just the trajectory of an orbital motion. If a trajectory lies on a conic, say, or a spiral, we would have a conic or spiral motion. The Principian sentence, "A body, urged by a centripetal force continually directed toward an immovable center S, moves in a conic section with focus at S," now turns into "Consider a conic motion about S." Of course conics hold some special interest for us here, and we recall the following definition: a conic is the locus of points whose distance from a given point S (the *focus*) is some positive constant e
(the *eccentricity*) times the distance from a given line (the *directrix*). Perhaps we should put this definition in vector dress, so it will feel more comfortable when vector calculus comes to call. If we let \( \mathbf{r} \) be the position vector from the focus, \( d \) the distance from the directrix to the focus, and \( \mathbf{e} \) (the *eccentricity vector*) a vector of length \( e \) which points perpendicularly toward the directrix, then the definition tells us that

\[
\mathbf{r} = e \left( d - \mathbf{r} \cdot \frac{\mathbf{e}}{e} \right),
\]

and with the notation \( \mathbf{U} = \mathbf{r}/r \) and \( l = de \), this formula turns into the *vector conic equation*:

\[
\mathbf{r} \cdot (\mathbf{e} + \mathbf{U}) = l. \tag{2}
\]

The constant \( l \) is called the *semi-latus rectum* of the conic. Given a positive constant \( l \) and a nonzero vector \( \mathbf{e} \), the vector conic equation defines a conic with semi-latus rectum \( l \), eccentricity \( e = |\mathbf{e}| \), axis along \( \mathbf{e} \), and focus at the origin. When \( \mathbf{e} = \mathbf{0} \), then (2) describes a circle of radius \( l \) about the origin, and if \( l = 0 \), we have a ray from the origin.

At this point, we have the vocabulary and background to explore a contemporary version of Newton’s answer to Halley. Suppose we have a motion \( \mathbf{r} = \mathbf{r}(t) \) with an inverse-square acceleration, so that for some nonzero number \( \lambda \),

\[
\mathbf{a}(t) = -\frac{\lambda}{r^2} \mathbf{U}(t)
\]

for all \( t \) in some open interval \( J \). Crossing with the angular momentum \( \mathbf{h} = \mathbf{r} \times \mathbf{v} \), we have

\[
\mathbf{a} \times \mathbf{h} = -\frac{\lambda}{r^2} \mathbf{U} \times \mathbf{h}
\]

\[
= -\frac{\mathbf{r} \times \mathbf{h}}{r^3}
\]

which becomes, using (1),

\[
\mathbf{a} \times \mathbf{h} = \lambda \dot{\mathbf{U}}.
\]

Now antidifferentiate, remembering that \( \mathbf{h} \) is constant because \( \mathbf{r} \) is orbital:

\[
\mathbf{v} \times \mathbf{h} = \lambda \mathbf{U} + \mathbf{c}
\]

\[
= \lambda (\mathbf{U} + \mathbf{e})
\]
for some constant vectors \( c \) and \( e = \frac{1}{\lambda} c \). If we dot with \( r \), we find
\[
\frac{1}{\lambda} r \cdot (v \times h) = r \cdot (e + U),
\]
and then permuting the entries in the scalar triple product uncovers the vector conic equation (2):
\[
\frac{h^2}{\lambda} = r \cdot (e + U).
\]
When the constant vector \( h \) vanishes, this reduces to \( U = -e \), and the motion must then lie on a fixed ray from the origin. If \( h \) does not vanish, but \( e \) does, we conclude \( r = h^2/\lambda \), so the orbit lies on a circle centered at the origin. Supposing neither \( h \) nor \( e \) vanishes, we have seen that the vector conic equation (2) defines a conic with focus at the origin. And that seals it:

**NEWTON’S SHAPE THEOREM.** Apart from motion on a ray from the center, every motion with an inverse-square acceleration must be a conic motion about the focus.

A second proof of the Shape Theorem is quick but sly. Assume again that
\[
a(t) = \frac{-\lambda}{r^2} U(t)
\]
Then of course \( h \) remains constant, but (surprise!) so does the vector \( L = \frac{1}{\lambda} v \times h - U \). To check, compute the derivative:
\[
\dot{L} = \frac{1}{\lambda} a \times h - \frac{h \times r}{r^3} = \frac{1}{\lambda} \left( \frac{-\lambda}{r^2} U \right) \times h - \frac{h \times U}{r^2} = 0
\]
Now just dot \( r \) with \( L + U \),
\[
r \cdot (L + U) = \frac{1}{\lambda} r \cdot (v \times h) = \frac{h^2}{\lambda},
\]
and we recognize the vector conic equation (2). That’s all there is to it.

The sly part of this proof is (un)clear: why would one expect the vector \( \frac{1}{\lambda} v \times h - U \) to be constant? The secret lies in a formula for the eccentricity vector \( e \). Given any conic motion \( r = r(t) \), if we differentiate the vector conic equation,
\[
r \cdot (e + U) = l,
\]
and solve for the (constant) eccentricity vector \( e \), we obtain the

**ECCENTRICITY FORMULA.** For any motion \( r = r(t) \) satisfying the vector conic equation (2),
\[
e = \frac{l}{h^2} v \times h - U. \tag{3}
\]
Of course we began with an inverse-square motion, not a conic motion, but if we had had a conic motion, then the vector \((l/h^2)v \times h - U\), representing as it does the eccentricity vector, would have been a priori constant. Knowing that \( \lambda \) turns out to be \( h^2/l \) (see our first proof), it seems natural then to suspect that \( L = (1/\lambda)v \times h - U \) should be constant in the case of inverse-square acceleration. If you do not like this sneaky proof of the Shape Theorem, blame Laplace. The vector \( L \), sometimes called the Laplace-Runge-Lenz vector, has the history of its rediscoveries etched in its name.
Now that we have seen two contemporary proofs, let us drift back in time, back to the 1680s, to examine Newton's original argument for the Shape Theorem in the *Principia*.

3. Only with some nervousness, do we open Newton's monumental work *Philosophiae Naturalis Principia Mathematica*. It had a reputation in 1687; it has a reputation still—a reputation for being impenetrable. In the latter half of the eighteenth century and on into the nineteenth, this reputation fed a cottage industry of writing notes and commentaries devoted entirely to 'understanding' the *Principia*. (The industry may have declined, but it still produces excellent commentaries from time to time: witness [5] and [6], just out in 1995.) Always formal, terse, and crabbed in his scholarly work, Newton took these stylistic tendencies to their limit in the *Principia*. Why? A decade earlier, his theory of colors had been attacked by Leibniz, Hooke, Linus, Lucas, as well as others, and Newton had detested the controversy. In a shrill letter to Henry Oldenburg, who was then Secretary of the Royal Society, Newton despairs, "I see I have made myself a slave to Philosophy, but if I get free of Mr. Linus's business I will resolutely bid adieu to it eternally, excepting what I do for my private satisfaction or leave to come out after me. For I see a man must either resolve to put out nothing new or become a slave to defend it." [7, p. 198] Of course, Newton did not "leave [the *Principia*] to come out after [him]," but he did choose to limit his readership and therefore his potential critics by composing in an icy, mathematical style, ultimately producing 500 pages of dense Latin text—definitions, axioms, lemmas, theorems, propositions, demonstrations, scholia, and figures, all fixed in place, a massive ordered regiment of abstract formality. According to a close friend of Newton's [2, p. 168], controversy of any kind

made sr Is[aac] very uneasy; who abhorred all Contests...And for this reason, mainly to avoid being baited by little Smatterers in Mathematicks, he told me, he designedly made his Principia abstruse; but yet so as to be understood by able Mathematicians, who he imagined, by comprehending his Demonstrations, would concurr with him in his Theory.

Yet even the most able mathematicians of the day struggled with the *Principia*. The confident young mathematician Abraham de Moivre happened to be visiting the Duke of Devonshire when Newton arrived to present the Duke with a copy of the new work [21, p. 471]:

[de Moivre] opened the book and deceived by its apparent simplicity persuaded himself that he was going to understand it without difficulty. But he was surprised to find it beyond the range of his knowledge and to see himself obliged to admit that what he had taken for mathematics was merely the beginning of a long and difficult course that he had yet to undertake. He purchased the book, however; and since the lessons he had to give forced him to travel about continually, he tore out the pages in order to carry them in his pocket and to study them during his free time.

Prepared by its scary reputation, we cannot conjure up the initial poise of de Moivre as we open the *Principia*, but prepared for some hard work, let us take a look at Newton's argument for the Shape Theorem. Actually, to do this in the proper order, we should close the *Principia* for the moment and begin nearer the
beginning, returning to Halley’s call on Newton in 1684. Earlier we have read de Moivre’s description of their meeting [22, p. 27]:

... after they had been some time together, the Df asked him what he thought the Curve would be that would be described by the Planets supposing the force of attraction towards the Sun to be reciprocal to the square of their distance from it. Sf Isaac replied immediately that it would be an Ellipsis. The Doctor struck with joy and amazement asked him how he knew it. Why saith he I have calculated it....

But stopping here is a rude interruption, for de Moivre continues [7, p. 283],

... whereupon D’ Halley asked him for his calculation without any farther delay, S’ Isaac looked among his papers but could not find it, but he promised him to renew it, & sent it.

It would be three months before Newton made good his promise, but idleness had not caused the delay, for he not only renewed his calculation for the ellipse, but embedded that calculation in a nine-page tract, “De motu Corporum in gyrum” (“On the Motion of Bodies in Orbit”), which Halley received in November.

It is in “De motu” then that we should look for Newton’s original demonstration of the Shape Theorem, that an inverse-square force implies conic orbits. Thumbing through its pages, we pass a line of definitions, hypotheses, theorems, corollaries, and problems until we stop at a familiar-looking claim [12, VI p. 49]:

Scholium The major planets orbit, therefore, in ellipses having a focus at the centre of the Sun ... exactly as Kepler supposed.

The Shape Theorem (at least for ellipses)! Eagerly we anticipate the proof—hunched over the scholium, eyes narrowed, pencil poised—but then the adrenaline seeps away as we scan down the page to find... nothing. Newton has left the Shape Theorem, his answer to Halley, as a bald claim, completely unsupported! Because the scholium directly follows

Problem 3 A body orbits in an ellipse: there is required the law of centripetal force tending to a focus of the ellipse.

we would guess that Newton must have viewed the Shape Theorem as a trivial corollary of his solution to Problem 3, or, more generally, of what we shall call

**NEWTON’S ACCELERATION THEOREM.** Every conic motion about the focus has an inverse-square acceleration.

Not understanding how the Shape Theorem would follow trivially from the Acceleration Theorem, we turn from “De motu” to the *Principia*, expecting the fuller development there to enlighten us.

Halley’s question in August of 1684 had reseeded Newton’s interest in celestial mechanics, and “De motu” was just the first little sprout. In January of 1685, he wrote Flamsteed, the Astronomer Royal, “Now that I am upon this subject, I would gladly know ye bottom of it before I publish my papers.” [7, p. 286] What understatement: between November of 1684 and April of 1687, Newton came to
“know ye bottom of it,” and the nine-page treatise exploded into a five hundred page masterpiece.

Now remember that “De motu” had left the Shape Theorem unproved. And the 1687 Principia? No better! In Section III of Book I, Newton demonstrates Propositions XI–XIII, which, taken together, establish the Acceleration Theorem and then follows with the Shape Theorem dressed as a corollary [11, p. 61] to this trio of propositions:

Cor. I From the three last Propositions it follows, that if a body P goes from place P with any velocity in the direction of any right line PR, and at the same time is urged by the action of a centripetal force that is inversely proportional to the square of the distance of the places from the center, the body will move in one of the conic sections, having its focus in the center of force.

But again, no proof. Worse yet, no one complained—not Halley, not Leibniz, not Huygens, not de Moivre—until, in October of 1710, twenty-three years after the publication of the Principia, Johann Bernoulli finally pointed out the obvious: Corollary 1 needed a demonstration. By this time, however, perhaps getting an early wind of Bernoulli’s criticism, Newton had already decided to fill the gap, instructing his editor, in a letter dated 11 October 1709, to slip the following argument [13, p. 5–6] into the second edition (1713) of the Principia:

Nam datis umbilico et puncto contactus & positione tangenti, describi potest Sectio conica quae curvaturam datam ad punctum illud habebit. Datur autem curvatura ex data vi centripeta: et Orbis duo se mutuo tangentes eadem vi describi non possunt.

For the third edition (1726), Newton added to this shockingly brief sketch the word ‘velocity’ in two places, resulting in [11, p. 61]

NEWTON’S ARGUMENT FOR THE SHAPE THEOREM
For the focus, the point of contact, and the position of the tangent, being given, a conic section may be described, which at that point shall have a given curvature. But the curvature is given from the centripetal force and velocity of the body being given; and two orbits, touching one the other, cannot be described by the same centripetal force and the same velocity.

Brevity may be the soul of wit, but it may be the seed of confusion as well. No one laughs when a fundamental proposition of celestial mechanics is followed by a two-sentence sketch which fails to persuade. At least Newton’s plan, although strikingly different from what we saw in Section 2, seems both familiar and clear—to prove that every solution to a given initial-value problem has a particular form, we exhibit a solution of that form and then invoke a uniqueness principle—but connecting all the dots in the outline may be another story, especially when some of the dots themselves are missing.

Expanding Newton’s sketch in a natural way, we arrive at what we take as his intended strategy:

NEWTON’S STRATEGY FOR PROVING THE SHAPE THEOREM
1. Suppose given any motion $\mathbf{r} = \mathbf{r}(t)$ with an inverse-square acceleration. At some time $t_0$, note the position $\mathbf{r}_0$, velocity $v_0$, and curvature $\kappa_0$ of the motion $\mathbf{r}$.  

2. Construct a conic \( \mathcal{C} \), having focus at the origin, that passes through the tip of \( \mathbf{r}_0 \) with tangent parallel to \( \mathbf{v}_0 \) and curvature \( \kappa_0 \).

3. On the conic \( \mathcal{C} \), put a motion \( \mathbf{r} = \mathbf{r}(t) \) about the focus that leaves the tip of \( \mathbf{r}_0 \) with velocity \( \mathbf{v}_0 \). (Newton never mentions this step, which involves making sure the position vector sweeps out area at a uniform rate, but it's a simple matter, and one that he probably took for granted.)

4. From Propositions XI–XIII (the Acceleration Theorem), infer that \( \mathbf{r} = \mathbf{r}(t) \), a conic motion about the focus, must have an inverse-square acceleration.

5. Thus both \( \mathbf{r} \) and \( \mathbf{r} \) have inverse-square accelerations, but even better, the matching of position, velocity, and curvature is steps (2) and (3) forces \( \mathbf{r} \) and \( \mathbf{r} \) to share the same proportionality constant.

6. Finally, noting that \( \mathbf{r} \) and \( \mathbf{r} \) now both solve the same initial-value problem, invoke a uniqueness principle to conclude that \( \mathbf{r} = \mathbf{r} \), proving that our given inverse-square motion \( \mathbf{r} \) must be a conic motion about the focus as desired.

As we begin to check whether this six-step strategy unfolds further into a convincing proof, we can see already that step (2) will block us, unless we know a little about the curvature of conics. For a motion \( \mathbf{r} = \mathbf{r}(t) \), the curvature \( \kappa \) is \( |\mathbf{T}|/\nu \) and the radius of curvature \( \rho \) is \( 1/\kappa \), where \( \mathbf{T} \) is the unit tangent \( \nu/\nu \). From the velocity and the acceleration, we can easily find the curvature from a well-known formula:

\[
\rho = \frac{\nu^3}{|\mathbf{a} \times \nu|}.
\tag{4}
\]

To calculate the radius of curvature for a conic, we start with any motion \( \mathbf{r} = \mathbf{r}(t) \) satisfying the vector conic equation (2),

\[
\mathbf{r} \cdot (\mathbf{e} + \mathbf{U}) = l,
\]

differentiate twice to get

\[
\mathbf{a} \cdot (\mathbf{e} + \mathbf{U}) + \nu \cdot \frac{\mathbf{h} \times \mathbf{r}}{r^3} = 0,
\]

and insert our formula (3) for the eccentricity vector \( \mathbf{e} \) to see that

\[
\frac{l}{h^2} \mathbf{a} \cdot (\mathbf{v} \times \mathbf{h}) + \nu \cdot \frac{\mathbf{h} \times \mathbf{r}}{r^3} = 0.
\]

Sliding the entries in the scalar triple products gives back

\[
|\mathbf{a} \times \nu| = \frac{1}{l} \left( \frac{h}{r} \right)^3,
\]

which leads to

\[
\rho = \frac{\nu^3}{|\mathbf{a} \times \nu|} = l \left( \frac{\nu}{h} \right)^3,
\]

or, rephrasing, to the

**CONIC CURVATURE LEMMA.** For any conic motion with semi-latus rectum \( l \),

\[
\rho = \frac{l}{|\mathbf{U} \times \mathbf{T}|^3}.
\tag{5}
\]

Newton cast this lemma more elegantly [12, III p. 159]: If the line perpendicular to the conic at \( P \) meets the focal axis at \( N \), then \( \rho \) varies as \( PN^3 \). (The equivalence to
our lemma follows from a geometric fact about conics: the projection of PN onto SP is the semi-latus rectum.) This lovely property is just one of several striking results on curvature obtained by Newton in his 1671 tract on series and fluxions. "The problem [of curvature]," he wrote in this tract, "has the mark of exceptional elegance and of being pre-eminently useful in the science of curves."[12, III p. 151] From an insight in his Waste Book made around December of 1664 (over twenty years before the Principia), we have evidence that Newton also recognized the fundamental place of curvature in the study of orbital motions: "If the body b moved in an Ellipsis, then its force in each point (if its motion in that point bee given) may be found by a tangent circle of equall crookedness [read curvature] with that point of the Ellipsis." [22, p. 14] It is perhaps surprising then that curvature plays no role in the 1687 Principia. However, in the 1690s Newton made radical plans for revising the first edition, plans that would have made curvature the centerpiece of his celestial mechanics. Sadly, this radical revision never made it into print, and in the end Newton contented himself with relatively minor changes, squeezing some curvature methods into the second (1713) and third (1726) editions as tucked on corollaries. For more on the role of curvature in Newton's celestial mechanics, see [3, 4, 10, and 17].

Now that we know something about the curvature of conics, we can begin to connect all the dots in a proof of the Shape Theorem inspired by Newton's two-sentence argument in the Principia. We follow the six-step strategy above, for it seems to be the only plausible interpretation of what Newton had in mind.

**Step 1:** We give ourselves any motion \( \mathbf{r} = \mathbf{r}(t) \) with an inverse-square acceleration: for some nonzero \( \lambda \), suppose \( \mathbf{r} \) solves the initial-value problem

\[
\ddot{r}(t) = \frac{\lambda}{r^2} \mathbf{U}(t), \quad r(t_0) = r_0, \quad \dot{r}(t_0) = v_0
\]

on the open interval \( J \). If \( r_0 \times v_0 = 0 \), then the motion lies on a fixed ray through the origin, but apart from this special case, we need to prove that \( \mathbf{r} \) is a conic motion about the focus. Since \( \mathbf{r} \) is an orbital motion, the orbit lies in a fixed plane and the angular momentum remains fixed at \( \mathbf{h}_0 = r_0 \times v_0 \).

**Step 2:** In this fixed plane, we now construct a conic that "fits" the orbit of \( \mathbf{r} \). Let \( \rho_0 \) be the radius of curvature of \( \mathbf{r} \) at \( \mathbf{r}(t_0) = r_0 \). Put

\[
I = \rho_0 |\mathbf{U}_0 \times \mathbf{T}_0|^3
\]

\[
\mathbf{e} = \frac{l}{h_0^2} v_0 \times h_0 - \mathbf{U}_0
\]

where \( \mathbf{U}_0 = r_0/r_0 \), \( \mathbf{T}_0 = v_0/u_0 \), and \( \mathbf{h}_0 = r_0 \times v_0 \). (As \( r_0 \) and \( v_0 \) are not parallel, \( h_0 \neq 0 \) and \( \mathbf{e} \) is well-defined.) The vector-conic equation (2)

\[
r \cdot (\mathbf{e} + \mathbf{U}) = I
\]

now defines a particular conic \( \mathcal{C} \). One easily checks that \( \mathcal{C} \) has a focus at the origin, and that \( \mathcal{C} \) passes through the tip of \( r_0 \) with its tangent parallel to \( v_0 \) and its radius of curvature equal to \( \rho_0 \).

**Step 3:** At this point, we would like to apply Newton's Acceleration Theorem to our constructed conic, but the Acceleration Theorem applies only to conic motions, indeed only to conic motions about the focus, not to mere conic loci. Therefore, on
the conic locus $C$ we now place a motion about the focus. (To put it differently, we must parameterize the conic locus $C$ in a way that keeps the acceleration vector pointed at the focus.) By the Area Theorem, to make a motion about the focus, we need only make a motion whose position vector from the focus sweeps out area at a constant rate, and intuitively we can do this by arranging for the area swept out to be our parameter. More precisely: Using arc-length measured from the tip of $r_0$, let $r_1 = r_1(s)$ be the unit-speed motion on $C$ having initial velocity $T_0$. The real function

$$a(s) = t_0 + \int_0^s \frac{1}{h_0} |r_1(s) \times \dot{r}_1(s)| \, ds$$

is smooth and strictly increasing. (Note that $h_0 = |r_0 \times v_0| \neq 0$ and $|r_1(s) \times \dot{r}_1(s)| \neq 0$ for all $s$, because tangents to $C$ never pass through the focus.) Take the (smooth) inverse $a^{-1} = a^{-1}(t)$, and use it to define a motion $r = r(t)$ on $C$ by

$$r(t) = r_1[a^{-1}(t)].$$

This constructed conic motion $r$ is also a motion about the focus $S$, for it has constant angular momentum $h_0 = r_0 \times v_0$. Moreover, $r(t_0) = r_0$ and $\dot{r}(t_0) = v_0$.

We haven’t done anything here, by the way, that Newton couldn’t do. You can find him geometrically constructing motions about the focus, on given conic loci, in the Principia, Book I, Section VI [11, p. 109–116]. Such constructions are even implicit in Newton’s proof of the Area Theorem in Propositions I and II, at the very beginning of the Principia. In his two-sentence argument for the Shape Theorem, Newton fails to mention the problem of putting an orbital motion on his constructed conic, but at the Principia’s level of rigor, this is a trivial omission. Refer to [15 and 16] for some discussion of this point.

**Step 4:** We apply the Acceleration Theorem (Propositions XI–XIII, Section III, Book I) to $r = r(t)$, our newly minted conic motion about the focus, and conclude that $r$ has an inverse-square acceleration: for some nonzero $\mu$,

$$\ddot{r}(t) = \frac{\mu}{r^2} U(t)$$

for all $t$.

**Step 5:** To prove that $\mu = \lambda$, we return to the curvature matching we did in Step 2. By design, both our constructed motion $r$ and our given motion $\tilde{r}$ share the same radius of curvature at the tip of $r_0$, namely $\rho_0$. For the conic motion $r$, by (4),

$$\rho_0 = \frac{v_0^3}{|a_0 \times v_0|} = \frac{v_0^3}{|\frac{\mu}{r_0^2} U_0 \times v_0|} = \frac{h_0^2}{\mu |U_0 \times T_0|^3}$$

Similarly, for the given motion $\tilde{r}$,

$$\rho_0 = \frac{h_0^2}{\lambda |U_0 \times T_0|^3}.$$

It follows that $\mu = \lambda$.

**Step 6:** We now have two solutions, the constructed conic motion $r$ and the given inverse-square motion $\tilde{r}$, to the initial-value problem

$$\ddot{r}(t) = \frac{\lambda}{r^2} U(t), \quad r(t_0) = r_0, \quad \dot{r}(t_0) = v_0$$
on the interval $J$. By standard uniqueness theorems (equivalent to Propositions XLI and XLII, Section VIII, Book I, Principia) for differential equations, we conclude that $r = f$ on $J$, and it follows that our given inverse-square motion must be a conic motion about the focus, as expected.

This completes a “Newtonian” proof of the Shape Theorem—that every motion having an inverse-square acceleration is a conic motion about the focus—a proof springing from Newton’s two-sentence argument in the Principia. Is this proof the contemporary version of what Newton had in mind? Probably, but the sheer brevity of his sketch leaves room for other views. On this issue, read [15, 16, 20, and 23].

Of course, our “completed” Newtonian demonstration is really anything but complete, since in step four, to ensure that our constructed conic motion had an inverse-square acceleration, we called on the unproved reversal of the Shape Theorem:

**NEWTON’S ACCELERATION THEOREM.** Every conic motion about the focus has an inverse-square acceleration.

We now intend to study the original argument for the Acceleration Theorem and then contrast the original with what we might do today, but as we return with this intention to the Principia (and specifically to Propositions XI, XII, and XIII in Book I), we must first page back to Proposition VI in order to understand how Newton measures orbital acceleration.

4. In May of 1686, just one month after the Principia was presented to the Royal Society, Halley sent news to Newton of the plans for printing and publication, but his cheerful letter ended with a sour lemon [21, p. 446]: “There is one thing more I ought to informe you of,” he wrote,

that Mr Hook has some pretensions upon the invention of ye rule of the decrease of Gravity, being reciprocally as the squares of the distances from the Center. He sais you had the notion from him...how much of this is so, you know best, as likewise what you have to do in this matter, only Mr Hook seems to expect you should make some mention of him, in the preface...

“Now is not this very fine?” sneered back Newton [21, p. 448],

Mathematicians that find out, settle & do all the business must content themselves with being nothing but dry calculators & drudges & another that does nothing but pretend & grasp at all things must carry away all the invention...And why should I record a man for an Invention who founds his claim upon an error therein & on that score gives me trouble? He imagines he obliged me by telling me his Theory, but I thought myself disobliged by being upon his own mistake corrected magisterially & taught a Theory wch every body knew & I had a truer notion of then himself.

In his fury at Hooke’s pretensions, Newton struck back with his pen, literally striking out almost every reference to Hooke in the entire Principia.

Even so, Hooke did in fact make one significant contribution to the Principia, for he was the first to see orbital motions as the geometric signature of a central attraction that pulls the orbiting body away from its linear inertial path. In November of 1679, as the new Secretary of the Royal Society, Hooke had asked
Newton to [22, p. 22] "please...continue your former favors to the Society by communicating what shall occur to you that is Philosophicall," and he added,

for my own part I shall take it as a great favor...if you will let me know your thoughts of [my hypothesis] of compounding the celestial motions of the planets of a direct [straight] motion by the tangent & an attractive motion towards the centrall body.

Hooke had this hypothesis as early as 1670, a time when Newton's eyes were still clouded by thoughts of "outward endeavors" and "Cartesian vortices." Still, Hooke's physical insight could take him only so far. In his hands, the hypothesis remained just that: a guess, a guess rooted in physical intuition and mechanical experiment, yet still a guess. But in Newton's hands, the hands of a soaring mathematical imagination, Hooke's hypothesis rose to an aerie of definitions, lemmas, and propositions. Look, for example, at the figure Newton draws to illustrate his proof of Propositions I and II (Section II, Book I), where we see, for the very first time,

the mathematical equivalence of central attraction and the area law, and you behold, in its central attraction and deviations from the tangent, the risen form of Hooke's hypothesis.

Later, in Proposition VI, Newton fashions from Hooke's inward deviation a formula for measuring the acceleration of an orbital motion. (In the *Principia*,

14 NEWTON AND THE BIRTH OF CELESTIAL MECHANICS [January
accelerations for general motions are never even defined.) If a particle in orbital motion falls freely toward the acceleration center S, Newton may have reasoned that the particle could be thought of as instantaneously in free fall from the tangent down to its position on the orbit. In a given time \( t \), suppose a particle moves along its orbit from P to Q. If there had been no acceleration during this time interval, the particle would have proceeded instead along the tangent at constant speed \( v \) to a location L. The deviation QL, nearly parallel to SP, would be like the "distance fallen toward S," which we would expect to be approximately \( \frac{1}{2}at^2 \), where \( a \) gives the acceleration at P. This suggests

\[
\frac{QL}{t^2} \to \frac{1}{2}a
\]

as \( t \to 0 \). Sanding top and bottom, Newton could now have shaped the measure QL/t² to fit squarely into his geometric approach. First nudge L just a bit along the tangent to the position R, making the deviation QR exactly parallel to SP.

Because time varies as the area in orbital motions, replace \( t \) by the area of the "sector" PSQ, and the sector in turn by the approximating triangle PSQ, in the process turning \( t \) into the product SP \cdot QT—no need to keep tabs on constant factors, such as the missing 1/2 here, for Newton works with proportions, not equations—and the measure QL/t² into QR/(SP \cdot QT)². The limit of this ratio, as Q \to P, gauges the acceleration at P. In the Principia, this measure of acceleration appears as Corollary I to Proposition VI (Section II, Book I) [11, p. 48]. With this corollary, Newton later derives acceleration laws from orbit shapes.

Cor I. If a body P revolving about the center S describes a curved line APQ, which a right line ZPR touches in any point P; and from any other point Q of the curve, QR is drawn parallel to the distance SP, meeting the tangent in R; and QT is drawn perpendicular to the distance SP; the centripetal force will be inversely as the solid SP² \cdot QT²/QR, if the solid be taken of that magnitude which it ultimately acquires when the points P and Q coincide.
Before we leave the topic of acceleration, we should take a moment to discuss the role of force and mass in the early sections of the *Principia*. The word ‘force’ appears, as it does above in Corollary I, in many of the definitions, axioms, corollaries, and propositions of the *Principia*, but in the first ten sections, where Newton attends to the one-body problem, force, and mass as well, exist literally in name only, playing no part in the *mathematics*. He may talk of ‘force,’ but Newton calculates accelerations. The Cartesians, Huygens and Leibniz among them, claimed that Newton, by introducing gravity, and therefore action at a distance, brought Aristotelian ‘occult qualities’ back into physics. But he should plead innocent to this charge. In the *Principia’s* work on orbital motions, ‘force’ and ‘gravity’ become merely convenient words, as Newton stresses the relations and laws, with no comment on *causes*. The cause of gravity comes up only in a General Scholium on the final pages of the *Principia* [11, p. 547]: “But hitherto I have not been able to discover the cause of those properties of gravity from phenomena,” wrote Newton,

and I frame no hypotheses; for whatever is not deduced is to be called an hypothesis; and hypotheses, whether metaphysical or physical, whether of occult qualities or mechanical, have no place in experimental philosophy. ...

And to us it is enough that gravity does really exist, and act according to the laws which we have explained, and abundantly serves to account for all the motions of the celestial bodies, and of our sea.

Wouldn’t Newton, that lover of geometry and curvature, have been delighted with Einstein’s view that geometry, indeed the curvature of spacetime, is the very cause of gravity?

After this interlude on Newton’s measure of acceleration, we remain in the past, looking for the original proof of the Acceleration Theorem in the *Principia*.

5. Wasting no time after Corollary I to Proposition VI, Newton attacks a series of problems with his new measure of acceleration. In Propositions VII through XIII, he calculates the acceleration law for circular motions about any given point, semicircular motions about a point infinitely remote, spiral motions about the pole, elliptical motions about the center, and then, in a stately section all their own, elliptical, hyperbolic, and parabolic motions about the focus. Taken together, this final triumphant trio of propositions (XI, XII, and XIII) establishes the Acceleration Theorem: *Every conic motion about the focus has an inverse-square acceleration*.

Newton could have proved the Acceleration Theorem in a single proposition covering general conic motions, but “... because of the dignity of the Problem . . .”, he writes, “I shall confirm the . . . cases by particular demonstrations.” [11, p. 57] These “particular demonstrations” naturally offer the same argument with minor variations, so we may safely choose one of the propositions to represent all three. Turn then to the most celebrated page of the *Principia* and to Newton’s analysis for Proposition XI:

**PROPOSITION XI PROBLEM VI**

*If a body revolves in an ellipse; it is required to find the law of the centripetal force tending to the focus of the ellipse.*
In the ellipse, Newton draws conjugate diameters DK and PG, with DK parallel to the tangent RPZ. (The midpoints of parallel chords in an ellipse lie on a line, called a diameter of the ellipse, and the parallel chords are then called the ordinates of the diameter. Two diameters with the property that each bisects every chord parallel to the other are said to be conjugate diameters.) From Q he drops three lines: QR parallel to the focal radius SP, QT perpendicular to SP, and Qx completing the parallelogram QxPR. He then extends Qx until it meets PG at v and draws PF perpendicular to DK.

Newton’s analysis requires the services of three lemmas, one of his own and two well known to Apollonius of Perga. (For the two Apollonian lemmas, see [1, I p. 15 and VII p. 31] or [18, p. 151 and p. 169].)

**NEWTON’S LEMMA.** PE = AC

**LEMMA 1.** All parallelograms circumscribed about any conjugate diameters of an ellipse have equal area.

**LEMMA 2.** In an ellipse, the squares of the ordinates of any conjugate diameter are proportional to the rectangles under the segments which they make on the diameter.

As we have seen in the previous section, Newton measures the acceleration of an orbital motion by computing the limit of the ratio

\[
\frac{QR}{(SP \cdot QT)^2}
\]

as Q → P. To infer an inverse-square acceleration for this case of elliptical motion about the focus, he must therefore prove that QR/QT² has a limit independent of P. In fact, as we now show, Newton’s argument reveals that QT²/QR tends to the *latus rectum* of the ellipse.
Because QR is PX and (by Newton’s Lemma) PE is AC, the similarity of the triangles PXV and PEC implies
\[ QR = \frac{PV \cdot AC}{PC}. \]

On the other hand, Newton’s Lemma (again) and the similarity of the triangles QXT and PEF give
\[ QT = \frac{QX \cdot PF}{AC} = \frac{QX \cdot BC}{CD}, \]
where the second equality follows from Lemma 1, which assures us that PF \cdot CD = BC \cdot AC. We infer
\[ \frac{QT^2}{QR} = \frac{QX^2 \cdot BC^2}{CD^2} \cdot \frac{PC}{PV \cdot AC} = \frac{1}{2} \frac{QX^2 \cdot PC}{PV \cdot CD^2}, \]
where we have replaced 2BC^2/AC by L. (Following Apollonius, Newton calls 2BC^2/AC the latus rectum.) If now Q \to P, this last expression has the same limit as
\[ \frac{1}{2} \frac{L \nu G}{PC}, \]
for Qx/Qx tends to one and Lemma 2 implies
\[ \frac{Qx^2}{PV \cdot \nu G} = \frac{CD^2}{PC^2}. \]
But \( \nu G \to 2PC \), so that \( \frac{1}{2} L(\nu G/PC) \), and thus also QT^2/QR, must tend to L. This completes Newton’s analysis for Proposition XI: Every elliptical motion about the focus has an inverse-square acceleration.

6. We have been “going under with the swirls and coming out with the eddies, following along the way the water goes,” but now just one quick swirl remains: to return from the Principia to the present, from Newton’s original work on the Acceleration Theorem to the delightful contrast of a contemporary argument.

Any conic motion \( r = r(t) \) about the focus must satisfy the vector-conic equation (2),
\[ r \cdot (e + U) = l, \]
for some positive constant \( l \) and constant vector \( e \). Since \( r \) is an orbital motion, \( h = r \times v \) is a constant vector. Since \( r \) is a conic motion,
\[ L = \frac{l}{h^2} v \times h - U \]
is a second constant vector (equal to the eccentricity vector \( e \) by (3)). Differentiating \( L \) yields
\[ 0 = \frac{l}{h^2} a \times h - \frac{h \times r}{r^3}, \]
and taking lengths we uncover an inverse-square acceleration,
\[ a = \frac{h^2}{l} \frac{1}{r^2}, \]
proving again

**NEWTON’S ACCELERATION THEOREM.** Every conic motion about the focus has an inverse-square acceleration.
ACKNOWLEDGMENTS. I am indebted to J. Bruce Brackenridge, David Cook, Nathaniel Grossman, Michael Nauenberg, Alan Parks, D. T. Whiteside, Curtis Wilson, and the referees for thoughtful comments—historical, mathematical, and stylistic. In most cases, I took their advice, and where I did not, I probably should have.

REFERENCES


Bruce Pourciau
Department of Mathematics
Lawrence University
Appleton, Wisconsin 54912
Bruce.H.Pourciau@Lawrence.edu