Kepler and Mach's Principle

Julian Barbour

Abstract The definitive ideas that led to the creation of general relativity crystallized in Einstein's thinking during 1912 while he was in Prague. At the centenary meeting held there to mark the breakthrough, I was asked to talk about earlier great work of relevance to dynamics done at Prague, above all by Kepler and Mach. The main topics covered in this paper are: some little known but basic facts about the planetary motions; the conceptual framework and most important discoveries of Ptolemy and Copernicus; the complete change of concepts that Kepler introduced and their role in his discoveries; the significance of them in Newton's work; Mach's realization that Kepler's conceptual revolution needed further development to free Newton's conceptual world of the last vestiges of the purely geometrical Ptolemaic world view; and the precise formulation of Mach's principle required to place GR correctly in the line of conceptual and technical evolution that began with the ancient Greek astronomers.

1 Introduction

Some of the most important advances in science are associated with Prague. The meeting at which the talk on which this paper is based celebrated Einstein's breakthrough to the key ideas of general relativity (GR) in 1912 near the end of his time in the Bohemian capital. In this paper, I wish to honour Kepler and his discovery of the laws of planetary motion and Mach's critique of Newton's concepts of absolute space and time. The creation of GR is unthinkable without them. I also wish to give what I believe is the correct formulation of Mach's principle. I believe that misunderstanding about this, ironically due to Einstein, may well be holding back both cosmology and the discovery of the quantum law of the universe.

Julian Barbour
Department of Physics, University of Oxford, Oxford, UK, e-mail: julian.barbour@physics.ox.ac.uk

I begin with some background to Kepler's discovery of the laws of planetary motion and then discuss the key intuitive ideas that enabled him to find them. We shall see that Kepler's reaction to the conceptual framework he inherited from Ptolemy, Copernicus and all previous astronomers was a clear anticipation of Mach's reaction to Newton's absolute space and time. In the broadest terms, one can see the creation of GR proceeding along a line of conceptual and technical development associated with six scientists: Ptolemy, Copernicus, Kepler, Newton, Mach, Einstein.

The main topics covered in this paper are listed in the abstract, so I turn directly to their presentation.

2 Some Important Facts of Planetary Motion

Everyone knows Kepler's three laws: 1. The planets move in ellipses with the Sun at one focus. 2. The radius vector from the Sun to the planet sweeps out equal areas in equal times. 3. The period of each planet is proportional to $a^{3/2}$, where a is the major axis of its ellipse. The first two laws were discovered in 1605, the third followed in 1618.

However, what really counts for understanding the history of planetary astronomy [1] up to Kepler's discovery of his first two laws is the form they take when the eccentricity of the ellipse is relatively small, as it is for all the planets. We must start with basic facts about ellipses (Fig. 1).

Because the ellipticity ε is half the square of the eccentricity e, the magnitude of ε is small for all the planets. The planet with the largest eccentricity, $e \approx 1/5$, is Mercury; then comes Mars with $e \approx 1/11$; Jupiter and Saturn have $e \approx 1/20$; the Earth ($e \approx 1/60$) and Venus ($e \approx 1/140$) have very small eccentricities and their orbits are wonderfully circular. This is crucial for the effects that even conscientious observers, who could use only the naked eye until 1610, were likely to find. The Sun and Moon subtend about 30 arc minutes on the sky. The accuracy of Ptolemy's observations was about a third of that, $10'$. Tycho Brahe's heroic observations, mostly in the period 1576–1597, pushed the accuracy to $2'$, as Brahe claimed, or $4'$ according to Kepler's more sober estimate.

What these facts about the accuracy of naked-eye astronomy mean is that the effects due to the orbit *eccentricity*, typically with a magnitude of degrees, were readily observable, while those due to the *ellipticity* were virtually undetectable. The only planet for which this is not strictly true is Mercury, but it is close to the Sun and seldom well seen, so it played no significant role in the discovery of Kepler's laws.[2] In one of several flukes in astronomy – the nearly equal apparent diameters of the Sun and Moon and the advance of the perihelion of Mercury included – it just

[1] See [1] for a detailed discussion of the history.

[2] It did help the belated recognition of Kepler's laws. His *Rudolphine Tables* (1627) led to the correct prediction and observation of the transit of Mercury across the Sun in 1631, a year after Kepler's death. The vastly superior accuracy of the *Tables* compared with the rivals, and the laws on which they were based, could no longer be denied.

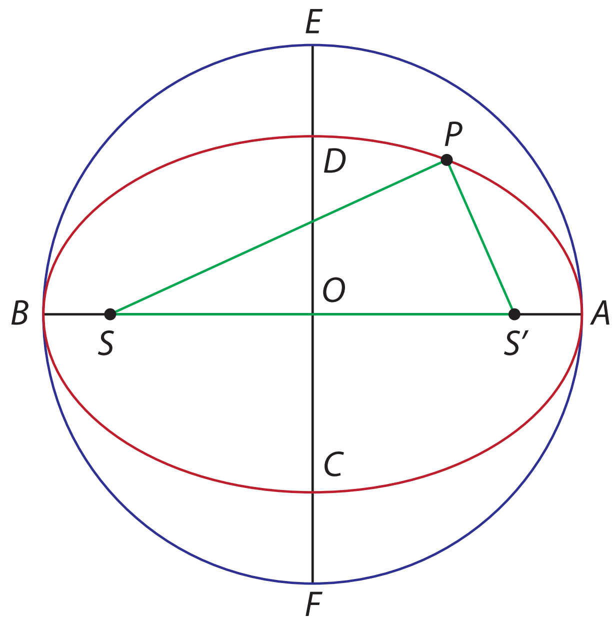


Fig. 1 The eccentricity e is OS/OB, the ellipticity ε =DE/OE is $\varepsilon = e^2/2$ and very small for the naked-eye planets.

so happens that among the remaining planets Mars was the most readily observable and has an ellipticity *just* large enough for Kepler's genius to espy it in the multitude of Brahe's observations.

To get an idea of Brahe and Kepler's achievement in the discovery of the ellipticity, Fig. 2 shows the orbit of Mercury. To the eye, it is a circle. One needs the circumscribing circle to see the difference. Fapp – for all practical purposes, to use John Bell's acronym – the planetary orbits are circles as far as naked-eye astronomy is concerned.

If the near circularity of the orbits is little known even among many astronomers, another remarkable fact is virtually unknown. It relates to a property of the *empty* focus of the planet's ellipse. If you could hover in a spacecraft just above the Sun's surface and watch a planet on the celestial sphere, you would see it move in a great circle with a decidedly non-uniform motion: first because, in accordance with Kepler's 2nd law, its physical speed in space does change, and, second, because the Sun is displaced from the centre of the orbit. This geometrical effect *doubles* the nonuniformity of the observed angular speed in the small-eccentricity approximation appropriate for the planets. If you then fly to the centre of the orbit and hover there, the geometrical distortion is eliminated, and the observed angular speed reflects the true variable speed. But a miracle happens if you journey on to the empty focus: the geometrical effect that enhanced the non-uniformity above the Sun is now

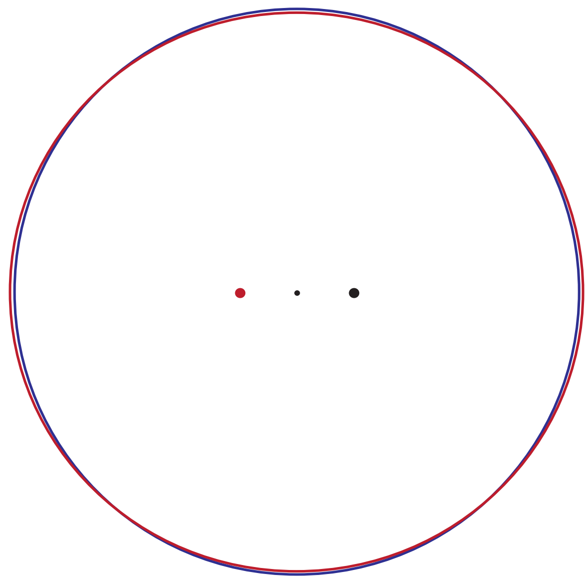


Fig. 2 The orbit of Mercury is circular to one part in 50. The figure shows the foci and centre of the orbit and the circle that most closely approximates the ellipse. The circle is inside the ellipse along the line of the apsides joining the foci and outside at the quadrants. For Mars, the planet for which Kepler discovered ellipticity, the gap between the circle and ellipse is 4.5 times less. The figure shows clearly that the effects of eccentricity are far more readily observable than those of ellipticity.

reversed: you see the planet move round its great circle with near perfect uniformity. Figure 3 illustrates the combined effect of the circularity and empty-focus effect.

The way to understand the actual process of discovery of the laws of planetary motions is through approximations to Kepler's laws. If the eccentricity e is zero, the orbits are circles and the speed uniform; if e is small, the orbits are still effectively circles but eccentric, and the speed on the circle is nonuniform though seen from the empty focus it is amazingly uniform. The major advances in the early history of planetary astronomy were largely due to these two effects. They are, respectively, very good approximations to Kepler's first two laws. It is no exaggeration to say that without them celestial dynamics could not have begun. However, they later became a source of great confusion: Copernicus and Brahe were literally going round in circles trying to make sense of the circles they imagined really were there in the sky. That's the story to which we now turn, beginning with the Greek astronomy.

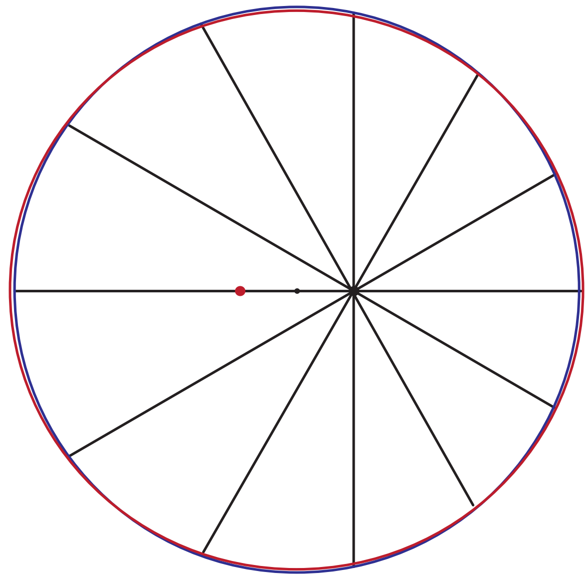


Fig. 3 Illustration of the empty-focus effect for Mercury. The Sun is left of the centre of the orbit. An observer at the empty focus on the right sees the planet move round a great circle with near perfect uniformity.

3 Greek Planetary Astronomy

Let's start with what is readily observable: the Sun. Aristotle already knew that its motion around the eclipic is non-uniform. Sometime around 150 BC (or 150 BCE as we should now say), Hipparchus, the first great astronomer of antiquity, made accurate measurements and a model to *explain* the non-uniformity. He supposed that the Sun moves on a perfect circle and with perfect uniformity and that the observed non-uniformity of its motion was due to a displacement D of the centre of the circle from the centre of the Earth, assumed to be the centre of the Universe. The ratio of D to R, the radius of the circle, was its *eccentricity*, the origin of both the technical and non-technical meanings of the word.

Hipparchus needed only two observations, the times taken by the Sun from the vernal equinox to the summer solstice and from there to the autumnal equinox, 94.5 and 92.5 days, respectively. These two data were enough to fix the two unknown parameters of his model: the magnitude of $e = D/R$ and the direction of the eccentric centre. This defined the *line of the apsides*, which joins *apogee* and *perigee* (or aphelion and perihelion in heliocentric astronomy). Hipparchus found $e = 1/24$, at that epoch more than twice the eccentricity of the Earth's heliocentric, or Sun's geocentric, orbit, which was then $\approx 1/57$.

Part of the inaccuracy was due to observational error, but the major contribution was a flaw in the theory. Having not the remotest reason to suspect non-uniformity of the motion, Hipparchus had inadvertently doubled the eccentricity. What was truly remarkable was that, when the observational accuracy had been pushed to its naked-eye limit by Islamic astronomers and Brahe, Hipparchus's incorrect model proved to be amazingly accurate in its predictions. In fact, it gives deviations from the true positions never greater than 3/4 of an arc minute, way below the detection level of naked-eye observations.[3] For this reason, Hipparchus's model, converted appropriately to heliocentric motion of the Earth, passed unscathed through the Copernican revolution. One of the gems in Kepler's work was, as we shall see, the dethronement of the Hipparchan model. It had reigned supreme for over 1700 years.

We now move on 300 years to the next great astronomer of antiquity: Claudius Ptolemy. He worked in Alexandria and had access there to, among much else, Babylonian observations made nearly a millenium earlier. He embarked on what was surely the first rationally and comprehensively planned scientific-research project in human history – the theoretical explanation through uniform circular motions of all the seven planets. Meaning *wanderers*, these included the Sun and Moon as well as Mercury, Venus, Mars, Jupiter and Saturn. Ptolemy's astronomical compendium, known through its Arabic title as *The Almagest* and written around 150 CE, was the handbook of astronomy for close on 1500 years. His epicycles tend to be mocked today as the paradigm of poor *ad hoc* science, often by scientists who should know better,[4] but they were one of the great contributions to the advance of science. On top of that, Ptolemy made what is arguably the first great discovery in the history of dynamics. To that I now turn.

Compared with the observed motion of a planet, that of the Sun is simplicity itself, being just the mirror image of the purely periodic motion of a single planet, Mother Earth, around the Sun. But the motion of a planet seen from the Earth is a compound of two incommensurate periodic motions. This compounding leads to the famous observed retrograde motions (Fig. 4).

[3] This remarkable accuracy is due to the very small eccentricity of the Earth's orbit, currently about 1/60. A solar theory developed by a Martian Hipparchus would not have survived for long because Mars has eccentricity $\approx 1/11$ (crucial for Kepler's discoveries). It is worth mentioning that the theory-independent quantity most immediately observable is always *twice* the eccentricity. For the Sun observed from the Earth, this is currently 1/30 of a radian or about 2° (four apparent solar diameters), half of which comes from the relative geometrical displacement of 1/60 and half from the physical non-uniformity described by Kepler's 2nd law. These equal contributions to observable effects are very important for understanding the history of ancient astronomy. By comparison, the observable ellipticity effects in the solar motion are about 120 times smaller at 1/30 of the apparent solar diameter, i.e., $1'$.

[4] I suspect some of them trust the scientifically very ignorant account in Koestler's *Sleepwalkers*, in which it is stated that "There is something profoundly distasteful about Ptolemy's universe; it is the work of a pedant with much patience and little originality, doggedly piling 'orb in orb'." The *Almagest* use the bare minimum of epicycles.

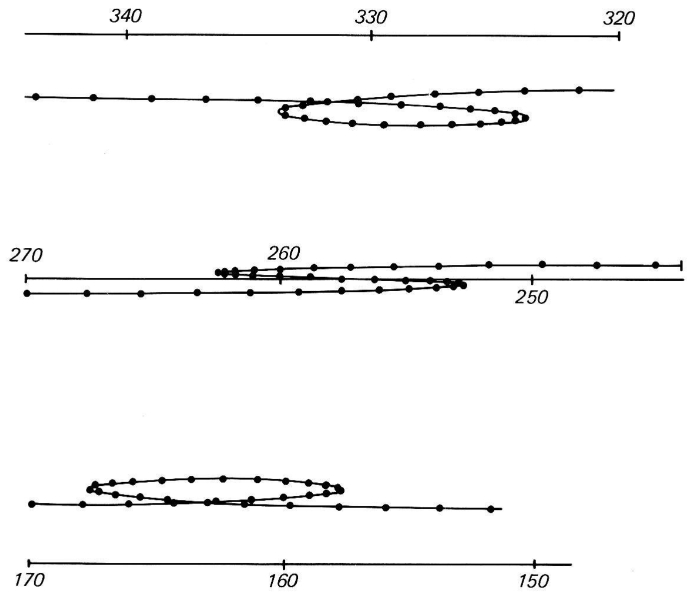


Fig. 4 The retrograde motions of Jupiter. Against the backdrop of the fixed stars, Jupiter (like all the planets) generally moves eastwards, never straying far from the ecliptic (the great circle on which the Sun moves). But about once every 13 months it executes its retrograde motions, in which the eastward motion comes to a stop and reverses into a westward motion until that halts and the eastward motion is resumed. Any attentive observer who keeps reasonably good records is likely to try to describe the complicated observed motion as a superposition of two simpler motions.

Following earlier proposals, which were probably only qualitative and may have been suggested by the great mathematician Apollonius (*circa* 255–170 BCE), Ptolemy attempted to describe them with the epicycle–deferent model (Fig. 5), which I describe in the caption only for the simpler case of the three outer planets.

In heliocentric terms, the motion of the guide point D is the *actual* motion of the planet, while the epicyclic motion is the reflection of the Earth's motion seen through the relative motion of the planet against the stars. The specific eccentricities of the various planetary orbits played a crucial role in the details of Ptolemy's theory. The key thing to understand is that the Earth's eccentricity, $e_{\rm E}$, is significantly smaller that that of the three outer planets, $e_{\rm P}$, which are moreover further from the Sun. On the sky, the observable effect O due to the orbital eccentricities is, to a first approximation,

$$O = \frac{e_{\rm P}}{e_{\rm E}} \frac{a_{\rm E}}{a_{\rm P}},$$

where $a_{\rm E}$ and $a_{\rm P}$ are the Earth's and the planet's semi-major axes, respectively. For Saturn, Jupiter and Mars $O \approx 1/30, 1/15, 1/8$, respectively. This meant that the nonuniformity in the planet's motion, represented in Ptolemy's theoretical model by the motion of the invisible guide point D, was readily observable and could not be ignored. In a first attempt to describe it, Ptolemy copied the Hipparchan solar model exactly by a simple displacement of the centre of the deferent from the terrestrial observer. For the epicycle motion, he assumed perfect uniformity around D. In heliocentric terms, this corresponds to an exactly zero-eccentricity circular orbit of the Earth. Note that the error due to circularity is tiny, around one part in 3600 and unobservable; the error due to the zero eccentricity is only 1/60. However, both of these are reduced by the ratio $a_{\mathrm{E}}/a_{\mathrm{P}}$ and escaped Ptolemy. They would in any case have been very difficult for him, with his rudimentary mathematics, to model.

What is extremely interesting is the way Ptolemy fixed the parameters of the deferent. The epicyclic motion being simply the reflection of the Earth's motion, the epicycle always points in the direction of the Sun. Ptolemy knew this; it did not prompt him to heliocentricity, but it did help him to fix the deferent parameters and to make that great discovery I mentioned.

His task was to fix the position of the invisible guide point: mission impossible you might think. But no; when the Earth is exactly between the planet and the Sun

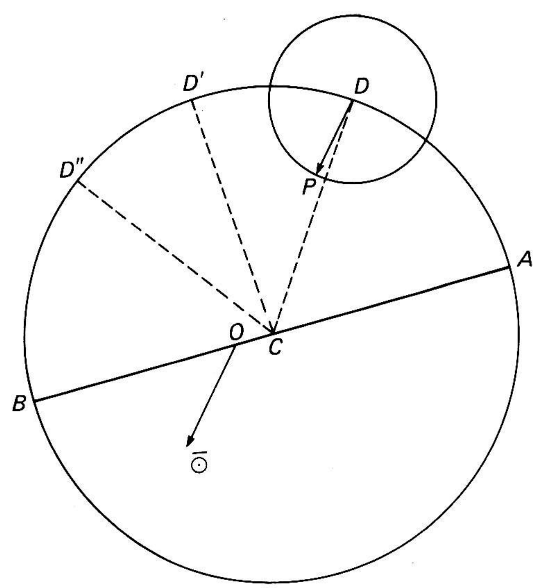


Fig. 5 Ptolemy sought to explain the regular eastward motion of, say, Jupiter, through uniform eastward motion of an invisible guide point D on a circle called the deferent. Around D a spoke of length less than the deferent radius rotated on an epicycle with perfect uniformity carrying the planet P at its tip. The centre C of the deferent is displaced from the position of the terrestrial observer O in order to explain the fact that even without the epicyclic motion the general eastward motion of Jupiter is manifestly non-uniform like that of the Sun as described by Hipparchus's model.

(so that the planet, in opposition, is due south at midnight), the epicycle, with the planet on its tip, points simultaneously towards the Sun and the Earth. Moreover, seen from the Earth, the guide point D is exactly behind the planet. Using such observations, spread necessarily over many years, Ptolemy could mimic what Hipparchus had done for the Sun. He was able to determine the eccentricity and line of the apsides of the deferent. But this was simultaneously the *heliocentric* orbit of the planet! However, because Ptolemy directly copied Hipparchus and did not suspect physical nonuniformity in the motion of D, he too found double the actual eccentricity.

Now in the solar motion there was no possibility of detecting the error. But Ptolemy's model was not yet complete. He had to fix the ratio of the epicycle and deferent radii. For this he needed just one more observation, of necessity made when the planet is not in opposition. There is an almost poetic touch worth mentioning here. When the planet is in opposition and due south a midnight, it rises at sunset, *acronychal* in Greek.[5] Ptolemy needed just one non-acronychal observation to determine the length of the epicycle.

With the model complete, Ptolemy – good scientist that he was – tested it using further non-acronychal observations. Dismay: the model failed to predict them correctly. After a long period of trial and error that, as Ptolemy admitted, had no principled basis except fidelty to observation – and hence truth – he found a deferent model that worked very well.

He discovered that it was necessary to *halve* his previous deferent eccentricity and introduce an 'equalizing point', or *equant* as it is now called. It lay on the other side of the centre of the orbit from the observer along the line of the apsides, which remained unchanged. Around the equant one had to imagine a spoke that rotated with perfectly uniform (hence equalizing) angular velocity and cut the deferent circle in its new position at the point when the guide point D must be. As before, the planet-carrying epicycle rotated with perfect uniformity about D.

With this model (somewhat modified for Venus and Mercury), Ptolemy found he could describe and predict the motion of all the planets with surprisingly good accuracy. What, in the long run, was truly significant for astronomy and dynamics, was that he had found a wonderfully good approximation to what Kepler's second law predicts. For, in heliocentric terms, Ptolemy's equant is none other than the empty focus of the planet's orbit – and I have already explained what a superb approximation that is. Because he was also working with eccentric, perfectly circular orbits, he also had an excellent approximation to Kepler's first law.

If we discount the barely observable and hence 'thankless' Mercury, Ptolemy's theory was correct for all the other planets to excellent accuracy. For Mars, with the largest eccentricity, the maximal deviation from Kepler's laws was only one part in 225. That is the measure of his achievement.

Final comments before we move on to Copernicus and Kepler. First, all of Ptolemy's work was based on measurement of *angles* between objects *that could be seen*. Second, none of his work or anything really accurate in astronomy could

[5] Acronychal and non-achronychal observations were still vital in Kepler's work.

have been done up to the invention of truly accurate clocks in the 20th century without the diurnal revolution of the stars, aka the Earth's rotation. It was the one and only clock that could be used. Time was also read off it by measurement of angles between visible objects.

4 Copernicus

It is ironic that Copernicus stumbled on his revolutionary idea by trying to undo Ptolemy's greatest discovery: the equant. Despite great admiration for Ptolemy's technical skill and achievements, Copernicus strongly disliked the equant's violation of the literally sacred principle that all the divine objects in the heavens must move with perfect uniformity in perfect circles. Ptolemy had maintained the circles (with good reason – they worked) but had discarded uniformity for the sake of truth. Copernicus, like at least one Islamic astronomer before him, sought to replace the equant device by a combination of uniform circular motions that, of necessity, was more complicated than Ptolemy's solution if fidelty to observational facts was to be maintained. While working on this project, he realized that all the retrograde motions of the planets could be understood as effects of *relative motion* if one assumes that the Earth is not at rest but moves in a circle.

There is an important point here that needs to be emphasized: Copernicus proposed a theory of *terrestrial mobility*, not heliocentricity. This was still the main point for Galileo, as shown by his famous retort "Eppur si muove." All that Copernicus needed, and said, was that the Sun must be near the centre of the circle in which the Earth moved.[6] For Copernicus, the Sun and its precise position had no physical significance. He said the Sun had a worthy place in the heavens, placed to illuminate the dance of the planets. As we shall see, the true discoverer of heliocentricity was Kepler.

Copernicus made four great contributions: first, he unavoidably, though without having an inkling of its significance, drew attention to the Sun, which was very important for Kepler; second, he explained the retrograde motions; third, as an underappreciated consequence of that, he brought to planetary astronomy a unity entirely lacking in Ptolemy's universe. Fourth, his arrangement of the solar system and the absence of observed parallax of any of the stars required the stars to be immensely farther away than Saturn.

The third contribution needs a little elaboration. Since Ptolemy worked with angles, he had no way of determining any distances. He therefore set all deferent radii equal to the nominal value unity and found the epicycle radius as a ratio to unity. Copernicus realized that this ratio, different for each planet, simply reflected the ratio of the radii of his various circles: the Earth's to that of the planets. He could use the radius of the Earth's orbit as a trigonometric base line to determine the distances to the planets. This immediately gave him a very good overall picture of the

---

[6] Copernicus actually thought that the Sun might have some slow motion of its own.

solar system. He worked out the correct order, distances and average speeds of all the planets, measured of course in terms of the radius of the Earth's orbit (now, of course, the astronomical unit) and the terrestrial day. He obtained a qualitative form of Kepler's third law and facts that Kepler subsequently found highly suggestive. It is often said that the Copernican and Ptolemaic arrangements are kinematically identical. This is not strictly true and does not do justice to Copernicus. Geometrical dispositions are part of kinematics. Ptolemy's *Almagest* did not have them or the lower bound on the distance to the stars.

A nice way to compare the respective achievements is this: when Ptolemy died, he could predict what the sky would look like – where the planets would be – as seen from Alexandria centuries after his death, but he had no idea what it would look like from Mars. When Copernicus died in 1543, he did know or, at least, knew how to calculate the positions of the Sun and planets as seen from Mars. In fact, the possibility was only literally confirmed in the space age.

Copernicus did great things, but, from a modern point of view, he bequeathed a most odd solar system to posterity. I have already mentioned the Sun's role as a mere lantern. Really strange was the location of the 'centre of the Copernican universe'. Ptolemy had discovered an equant in the deferents of all the planets, essentially because the Earth, unbeknown to him, was a spaceship that allowed him to look at the planets' positions from a whole circle in the solar system and not just from the Sun's position (as in the acronychal observations). But because the Sun's motion is merely the Earth's 180° out of phase, there was no way observation could force an equant on the solar motion. Ptolemy left the Hipparchan model unchanged.

And so did Copernicus. Even though he made Mother Earth a planet like the others and contrived makeshift substitutes for their equants, it never occurred to him that the Earth should get anything equivalent, so he simply inverted the Hipparchan model and gave the Earth an eccentricity twice what it should have. That simultaneously singled out the empty focus of the Earth's orbit as a special point. Moreover, because of the fluke of the Earth's eccentricity being so small Copernicus was misled into thinking that the lines of the apsides of the planets all converged, not on the Sun, but at the very same point that Kepler was later to identify as the Earth's empty focus. Figure 6 shows how small the mismatch was – but also that Copernicus did not propose a truly heliocentric system.

In fact, the clearest evidence of that is in the diagrams which Copernicus drew to show the orbits of the planets. *They do not show the Sun.* It was in no way an integral part of his scheme.

There were many other oddities, some very bizarre, in the Copernican cosmos, most of which arose because Copernicus simply inverted the Ptolemaic models. When *De Revolutionibus* was published in the same year 1543 that he died, Copernicus knew he had made a monumental discovery, but, like Ptolemy, his insights and methods were purely geometrical and kinematical. Kepler commented "he was unaware of his riches".

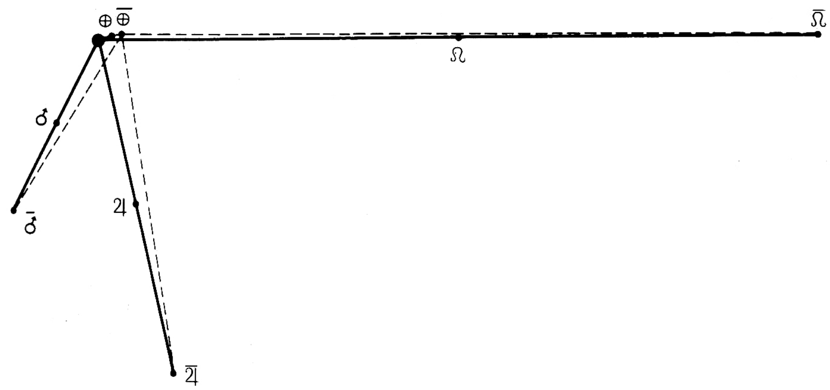


Fig. 6 Copernicus believed that the lines of the apsides of the three outer planets converged on a void point near the Sun that was actually the empty focus of the Earth's heliocentric orbit. The true Sun is the black disc on the left, where the solid lines of the apsides converge. The dashed lines show Copernicus's belief. The mismatch is small but shows that Copernicus thought solely in geometrical terms, for which such an arrangement with the Sun playing no physical role is perfectly acceptable.

5 Kepler

In this section, I want to concentrate on the huge conceptual change that Kepler (1571–1630), for whom a magnificent portrait (Fig. 7) survives, introduced and how he anticipated Mach's attitude to dynamics. I am firmly of the belief that more is still to come of it. I can only pick out the highlights. The details, which are absorbing, can be found in [1].

The best place to start is Brahe's observations of the comet of 1577, which established its distance as interplanetary. For Kepler, the supreme importance of the observations were that they 'destroyed' the crystal spheres widely believed to carry the planets. In his fascinating account of how he mastered the motion of Mars, the *Astronomia Nova* published in 1609, he repeatedly pointed out that Brahe's observations proved that the planets were not carried by spheres. The comet had passed clean through the solar system without crashing into them. They could not be there. In one of the great intuitive insights in the history of science, he proclaimed: "Henceforth the planets must find their way through the void like the birds through the air. We must philosophize about these things differently."

Crucial questions then arose. What moves the planets? How do they find their way? What if anything is directing them? Let us start with the second and third questions, which reveal Kepler's affinity with Mach – or better Mach's with Kepler. Birds find their way around the world by reference to features in the terrain and sky. But, according to the astronomy Kepler inherited from Ptolemy, Copernicus and all previous astronomers including Tycho Brahe, literally everything was controlled and directed by void points, above all the equants that Ptolemy had discovered, in empty featureless space. The difference from birds was blatant. Like them, the planets must use *visible objects*. Motion is relative to things you can observe. The only significant visible things that the planets could be guided by were the Sun and the distant stars. The guiding and determining role of visible matter is exactly what Mach was insisting on two and a half centuries later and led him to argue so persuasively against Newton's absolute space.

Moreover, if crystals spheres do not carry the planets, whence comes their motion? The planets must either have inherent motive force or be subject to it. This was a veritable change of mindset. A prominent part of the immensely long subtitle to the *Astronomia Nova* proclaimed it to be Celestial Physics. Kepler introduced forces into the heavens. True, they were Aristotelian, with the force assumed to determine the velocity it imparts and not the acceleration as in Newtonian dynamics. The astronomical data could give Kepler no hints in that direction, ironically for the same reason that Einstein three centuries later was able to subsume gravitational forces and inertia into a single geodesic law. Indeed, until very late in his work, Kepler

Fig. 7 Johannes Kepler. The epitaph that he composed for himself read "I used to measure the Heavens, now I measure the shadows of Earth. The mind belonged to Heaven, the body's shadow lies here." His grave and tombstone in Regensburg, where he died in 1630, have been lost.

believed the planets moved in circles. What more perfect and self-contained motion exists than that?

What was really important about Kepler's forces was not their mode of action but their conjectured *source*: the Sun and the planets themselves. In this key respect, Kepler's forces correctly prefigured Newton's. Their sources and controlling power resided in *physical* bodies, not void points in empty space. Here too, in identifying motion-controlling power with bodies and not space, Kepler anticipated Mach, who insisted that apparently force-free inertial motion was nothing of the sort but the outcome of an as yet unknown physical effect of all the matter in the Universe. Let me here quote Mach [2], p. 296: "The natural investigator must feel the need of further insight – of knowledge of the *immediate* connections, say, of the masses of the Universe. There will hover before him as an ideal an insight into the principles of the whole matter, from which accelerated and inertial motion result in the *same* way. The progress from Kepler's discovery to Newton's law of gravitation, and the impetus given by this to the finding of a physical understanding of the attraction in the manner in which electrical actions at a distance have been treated, may here serve as a model." That Mach sensed an affinity between himself and Kepler comes through in this quotation.

Let us return to details. Kepler had to explain two different kinds of motion: the eccentric circular motion around the Sun and the motion towards and away from the Sun during its course. To explain the circular motion, Kepler conjectured (nearly a decade before Galileo observed it!) that the Sun rotates about an axis perpendicular to the ecliptic and that what one might called ethereal 'spokes', rotating with the Sun, protruded from its equator. These, he assumed, swept the planets along in their circular motion, their strength diminishing with increasing distance from the Sun in order to explain why the more distant planets moved slower. As for the motion towards and away from the Sun, he conjectured that that it housed a powerful magnet and each planet a lesser one. The alignment of the magnetic poles would pull the planet towards the Sun on one side of the orbit and repel it on the other side.[7] These ideas are illustrated in Kepler's diagram shown in Fig. 8.

By modern standards, Kepler's forces were rather primitive and could not have survived detailed quantitative testing. What was decisive was that they focussed all of Kepler's interest on the Sun and its precise location. He was firmly convinced that the centre of the solar system did not lie at that mysterious void point that, post his discoveries, we recognize as the empty focus of the Earth's orbit, but at the entre of the relatively nearby mighty physical Sun (the distance between the Sun and the empty focus is 1/30 of the Earth's semimajor axis).

There were two ways to confirm this: first, to show that the lines of the apsides all converged *exactly* on the Sun, not the void point relatively close to it. Second, to show that the speed in orbit was not controlled by the void equant, as it was in Ptolemy's and, *de facto* despite his intense dislike of it, in Copernicus's astronomy, but by the Sun. I shall come to the crucial steps through which Kepler eventually came to his area law, which governs the speed, in a moment. First, I want to make

[7] William Gilbert's influential book on magnetism, published in 1600, strongly influenced Kepler's thinking.

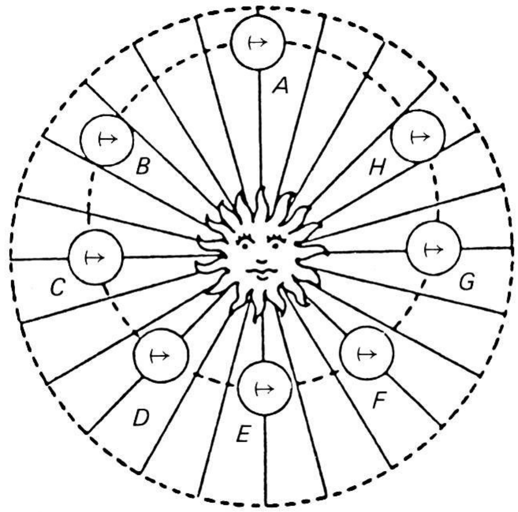


Fig. 8 This diagram encapsulates the difference between Kepler and all his predecessors. The arrival of a new mind on the astronomical scene is demonstrated nowhere more clearly than in the comparison of the diagrams in *De Revolutionibus* and the *Astronomia Nova.* The Sun is prominent by its absence in Copernicus's; in Kepler's, as here, it takes pride of place, controlling the motion of the planet through physical forces.

clear why Kepler has the credit for heliocentricity in a way that Copernicus does not. Modest as the move from the empty to the occupied focus as centre of the solar system may appear, it was a small step that anticipated and made possible Newton's giant leap of understanding in the workings of the world. It identified the turning point.

Kepler's conceptual ideas drove all the technical work done at Brahe's behest in Prague – to establish the precise motion of Mars to the same accuracy that the Dane's incomparable observations allowed. His primary tool was trigonometry, which he put to use like no one before him. Kepler was the first man who could roam truly freely in imagination through the solar system. Appropriately, he also wrote almost the first work of science fiction: a dream of a journey to the Moon. The entire thrust of his trigonometric work was to establish heliocentricity beyond gainsaying, above all to show that the knitting-needle lines of the apsides all converged bang in the middle of the Sun and not as in the Copernican scheme (Fig. 6) at the nearby void point.[8]

There is even a sense in which Kepler anticipated gauge theory: he knew perfectly well that, kinematically, all of his precise geometrical results could be exp-

[8] Kepler's work was actually a much more logically consistent and definitive proof of the Copernican cosmology than Galileo was able to muster. In fact, magnificent as his many acheivements were, the Tuscan completely failed to recognize or begin to comprehend Kepler's achievement.

resed just as adequately in the Tychonic [9] or geocentric Ptolemaic schemes as in the Copernican arrangement, but triumphantly pointed out that *in all three* the lines of the apsides meet at one point in the Sun. That was the ‘gauge-invariant’ content of his discoveries. The area law, the discovery of which we have still to discuss, had the same status.

Let us go through the most important technical advances to which Kepler was led by his intuition. I said that Plotemy’s acronychal observations were made at opposition, when the observed planet is due south at midnight and the Sun is directly behind the terrestrial observer. This is not quite true; to facilitate computations and very likely because he did not realize the importance of the difference, Ptolemy’s actual acronychal observations were made not when the true Sun was behind the observer but a substitute, a mathematically defined ‘mean Sun’ that moved around the ecliptic with perfectly uniform speed, coinciding with the true Sun only at the equinoxes. The angular distance between the true and mean Sun could be as much as those 2° corresponding to twice the Earth’s eccentricity. Brahe had continued the Ptolemaic practice of using the mean Sun; Kepler meticulously corrected his observations by interpolation to make them correspond to his beloved true Sun. This was a first useful sharpening of accuracy.

Kepler was also the first person to understand how to take into account correctly the fact that the planets do not all move in the ecliptic. This gives rise to significant effects, causing the retrogression loops to have an out-of-ecliptic component (Fig. 4); without proper understanding of and correction for the effect Kepler could never have found the ellipticity of Mars’s orbit. In fact, Kepler’s first major result is what one might call his zeroth law: each planet moves in plane that passes through the Sun and is fixed in the frame defined by the stars.

But the real gem in Kepler’s work that prepared the ground for his geatest discoveries was his finding of the *true location of the Earth*. All motion is relative. If you are trying to determine the position and motion of a distant object, you will surely make errors if you are mistaken about your own position and motion.[10] This is what Kepler understood perfectly – and he had good reason to be concerned. According to Copernicus and Brahe, the Earth’s orbit was *sui generis*: unlike those of the other planets, it had no equant. Kepler’s sense of the uniformity of nature told him that could not be true. The Earth had to have an equant. If so, that would mean it would have only *half* the eccentricity attributed to it by Copernicus and Brahe. Existing theory must be putting the Earth in the wrong place and thereby distorting the interpetation of the observations of all the extraterrestrial bodies.

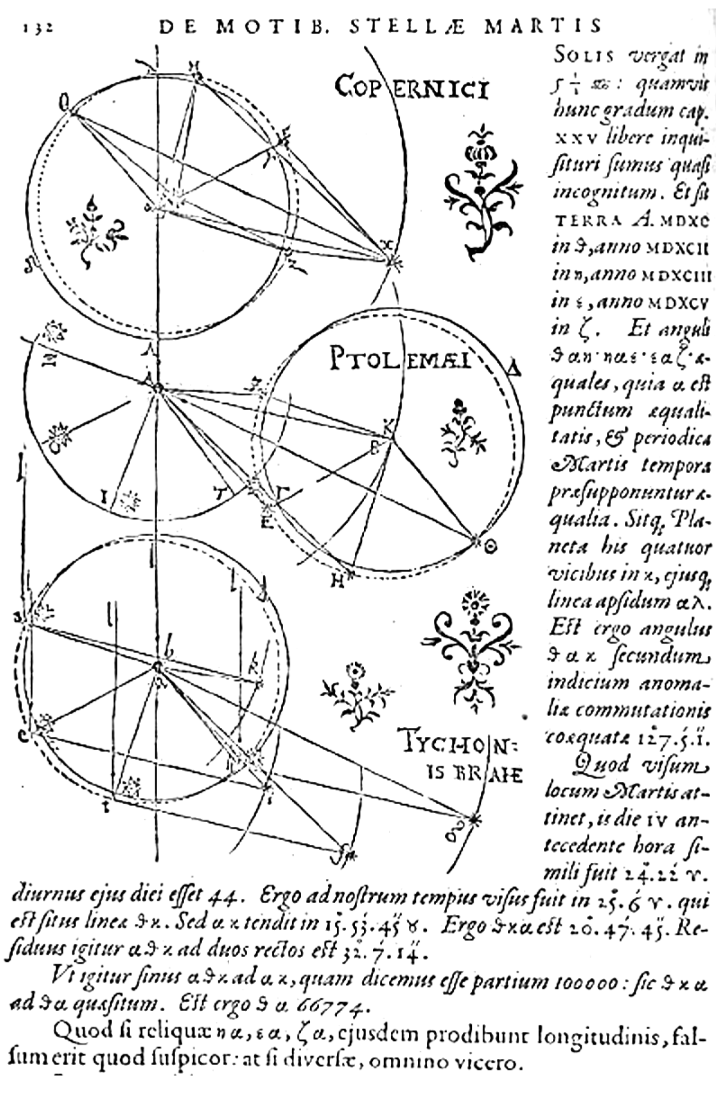
In an article written in 1930 to mark the 300th anniversary of Kepler’s death, Einstein described – with good reason – Kepler’s halving of the Earth’s eccentricity

[9] Tycho Brahe could not believe in the immensity of the Universe that, given the absence of any observable stellar parallax, followed from Copernicus’s proposal and therefore proposed that the Sun goes round the Earth while Mercury and Venus orbit it.

[10] I was told some years ago that the largest uncertainty in many high-precision tests of GR was the uncertainty in the Earth’s position that results from the perturbing influence of the asteroids, whose size is known but not, to sufficient accuracy, their densities.

as one of the most beautiful things in all of science. Kepler knew that too. He had the finest diagram in his book engraved to show the way it was done (Fig. 9).

His stroke of genius was to use a trigonometric base line formed by the Sun and Mars at times when he knew Mars was at exactly the same point in its orbit. One


132 DE MOTIB. STELLÆ MARTIS

SOLIS vergat in 5 ½ ♒ : quamvis hunc gradum cap. XXV libere inquisituri sumus quasi incognitum. Et sit TERRA A. MDXC in ϑ, anno MDXCII in η, anno MDXCIII in ε, anno MDXCV in ζ. Et anguli ϑαx · ηαε · εαζ · æquales, quia α est punctum æqualitatis, & periodica Martis tempora præsupponuntur æqualia. Sitq; Planeta his quatuor vicibus in x, ejusq; linea apsidum αλ. Est ergo angulus ϑαx secundum indicium anomaliæ commutationis coæquatæ 127.5.1. Quod visum locum Martis attinet, is die IV antecedente hora simili fuit 24.22 ♈. diurnus ejus diei esset 44. Ergo ad nostrum tempus visus fuit in 25.6 ♈. qui est situs lineæ ϑx. Sed αx tendit in 15.53.45 ♉. Ergo ϑxα est 20.47.45. Residuus igitur αϑx ad duos rectos est 32.7.14.

Ut igitur sinus αϑx ad αx, quam dicemus esse partium 100000: sic ϑxα ad ϑα quæsitum. Est ergo ϑα 66774.

Quod si reliquæ ηα, εα, ζα, ejusdem prodibunt longitudinis, falsum erit quod suspicor: at si diversæ, omnino vicero.

Fig. 9 This diagram shows the halving of the Earth's eccentricity in the Copernican cosmology and of the Sun's in the Ptolemaic and Tychonic. In the Copernican scheme, Mars is at the point x on three occasions. Knowing the relevant angles, Kepler could determine the corresponding positions of the Earth at the three points on the dashed circle. They established the true position of the Earth's orbit and that it must have half the eccentricity assumed by Copernicus.

of the great clarifications due to Copernicus's insights was that the planets traced out invariable orbits in the space defined by the Sun and fixed stars, returning to the same orbital position after completion of one orbit. Now among all data, heliocentric periods were the easiest to determine accurately; that of Mars was known to be 685 terrestrial days. Kepler searched among Brahe's 21-year treasury for Martian observations that by chance were separated by multiples of 685 days. At them, Mars must be at the same point in space. Kepler found three such observations. Acronychal observations of Mars and the theory of them, which he could trust, told Kepler the direction of Mars as seen from the Sun at all times. The direction to the Earth was also known, so Kepler could determine the angle between Mars and the Earth seen from the Sun. Brahe's observations gave him the angle between the Sun and Mars as seen from the Earth. Kepler had the one fixed Sun–Mars side of the triangle and two angles of the Sun–Mars–Earth triangle. Three such observations gave three positions of the Earth. But one only needs three known points to fix the position and size of a circle.[11] Kepler's determination of the position of the Earth's orbit finally revealed the error in Hipparchus's solar theory. The halving of the Earth's eccentricity created a firm foundation for astronomy. It was hugely important.

Having properly located the Earth, and still believing in circular orbits, Kepler set about fitting the parameters of Mars's orbit to Brahe's observations. He still made use of the Ptolemaic equant even though convinced the speed in orbit must somehow be determined relative to the Sun and not the void equant. It was a wonderful piece of work and a posthumous triumph for Ptolemy's circles and equant. But the merry-go-round just would not get everything right. Kepler tweaked here, he tweaked there, but whatever he did an occasional error of up to just 8 minutes of arc in the position of Mars would show up. He could never have done it without Brahe's observations, which were both as accurate as they could be and also, most importantly, comprehensive.[12] This is the place to quote Kepler:

"We, whom God in his goodness has given such a careful observer in Tycho Brahe, and whose observations reveal the $8'$ error of Ptolemy's calculations, should thankfully recognize the goodness of God and make use of it. That is, we should make the effort (supported by the arguments for the falsity of our assumptions) to find at last the true form of the celestial motions... These $8'$ alone reveal the need for reformulation of the whole of astronomy; they become the material of a great part of my work."

---

[11] Keplers' work on Mars began under the assumption that the Earth has an exactly circular orbit. Because the Earth's eccentricity is $\approx 1/60$, this asumption is accurate to better than one part in 7000. Even when Kepler knew the Earth's orbit could not be a perfect circle, he could assume it to be so for his work on Mars with its far larger eccentricity $\approx 1/11$. In the story of the discovery of the laws of their motion, the planets were like the characters in a good novel. Each had an individual personality determined by its eccentricity and semimajor axis. The interaction of these personalities, reflected in the observational data, is what makes the discovery of the laws of the planets' motion such an absorbing story.

[12] Ptolemy had made and used relatively few observations obtained at times and orbital positions he expected to be especially valuable for construction and testing of his observations. Brahe believed in blanket coverage of the orbits: who could know what would be relevant and revealing? This was truly prescient and of immense value to Kepler.

Now it is time to talk about the area law. Along with establishing where the lines of the apsides meet all in one place, this was the other great bonus of the 'Machian' shifting from a void point to the Sun. For the apside adjustment, the shift was from the Earth's empty focus. In the case of the area law, it was from Mars's empty focus.

Kepler was keenly aware of the value of the equant phenomenon: mathematically, in pre-calculus days, anything that involved uniform motion (angular velocity about the equant in this case) was a significant plus. Non-uniform motion was barely tractable. For this reason astronomers were still using the equant to calculate ephemerides in Newton's time three quarters of a century after Kepler, absolutely insistent on a physical interpretation of celestial motions, had done away with it.

The area law was one of the most serendipitous discoveries – of which there are so many – in science. Kepler was looking for a law, governed by the Sun, that would determine the speed in orbit of each planet. His physical intuition told him the Sun must exert a force on the planet that would be stronger, the closer the planet was to the Sun. In fact, that was clearly indicated by obsevations, which showed that the planets moved fastest when closest to the Sun. Plausibly enough, Kepler guessed that the speed would be inversely proportional to the distance from the Sun. Strong support for this came from the fact that the equant law showed the orbital speeds at aphelion and perihelion to be exactly in inverse proportion (Fig. 10).

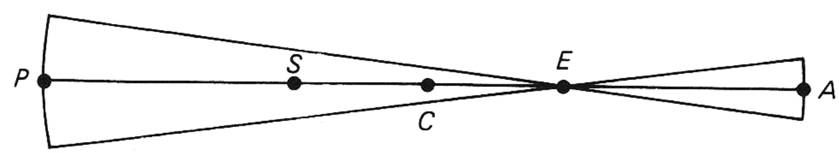


Fig. 10 The exact area law and the speed law that follows from the empty-focus effect both predict that the orbital speeds at aphelion and perihelion are inversely proportional to the distance from the Sun.

But the mathematics of the putative law, applied to the eccentric orbit of Mars, proved to be beyond Kepler's abilities. He therefore decided to replace what he regarded as the exact law by an approximation in the form of the area law! He was encouraged to this by his recollection of the way Archimedes had estimated the area of a circle by dividing it into ever smaller segments. As his work progressed and he gained an increasing number of accurate locations of Mars through his application of trigonometry to Brahe's observations, always under the key assumption that the Sun was the centre of the solar system and the controller of planetary motions, he came to realize that the area law did actually govern the speed in orbit.

Kepler's final, very tortuous breakthrough to the joint discovery of ellipticity of the orbit and the area law was in fact somewhat delayed by his enthusiasm for theory. The moment he found unambiguous evidence that Mars's orbit could not be circular, he started to speculate and initially guessed an egg-shaped orbit, i.e., fatter at one end than the other. Slowly, as he accumulated more and more accurate locations, the egg was abandoned. A chance glance at a table of logarithmic tables was what finally led him to the ellipse – more serendipity. The full truth at last came to him around Easter 1605.

6 Kepler's Significance

It would be a futile counterfactual exercise to ask how science would have developed without Brahe and Kepler's extraordinary efforts. However, it is entirely possible that, even without the discovery of the telescope and the possibility that gave for more accurate observations, decades could have passed before the discovery of the laws of planetary motion. What is absolutely certain is that Newton's *Principia* is inconceivable without Kepler's discoveries. All three of Kepler's laws were important: from the third, Newton deduced the $1/r^2$ force law for gravity; from the first, that the planet's elliptical motions could be understood as the outcome of two competing tendencies – rectilinear inertial motion and gravitationally induced deflection from it along the direction to the Sun. In many ways, Kepler's second law was actually the most important. Newton recognized this by making it the subject of his very first proposition in the *Principia*. It will be worth saying something about this.

By 1670 at the latest, Newton had most of the elements of a rudimentary dynamics, above all the notion of inertial motion that would persist forever were it not changed by the action of other bodies. He understood elastic collisons and the nature of centrifugal force. What hindered a full blossoming of dynamics was the prevailing mechanistic conception of the world due above all to Descartes. According to this view, all mechanical action took place through direct contact: collisions. The Cartesian cosmos was a terribly crowded world crammed full of pieces of matter in continual collision. Newton basically subscribed to this view. Although he had laws to describe collisions, there was little he could do with them.

The real advance almost certainly came in 1679, when Robert Hooke, newly appointed as secretary of the Royal Society, pressed Newton hard to confirm his [Hooke's] proposal "of compounding the celestiall motions of the planetts of a direct motion by the tangent & an attractive motion towards the central body." Calculations of Newton that he may well have made as a result of Hooke's letters of 1679 have survived and include the key result that became Proposition 1 in the *Principia*:

Proposition 1. *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.*

This is the theoretical explanation of Kepler's area law. It had far-reaching implications, for it told Newton that nature should be described, mathematically at least, by forces that act over distances. Huygens had coined the expression *centrifugal force*; in explicit imitation, Newton called his new forces *centripetal*. He was well aware of the revolutionary nature of what he was doing; he was proposing to give universally despised *occult* forces a decisive role in physics. He was very cautious about this and emphasized, in *hypotheses non fingo*, that he was not making any assumptions about the physical mode of action of the forces he introduced. What he did stress was that such forces, introduced mathematically, could explain at a stroke a vast number of diverse phenomena. They opened up a whole field for exploration that is still ongoing. Kepler's $8'$ led to more than the reformulation of astronomy.

7 Intermezzo: Christian Doppler

Before we move on to Mach and Einstein, brief mention should be made of Christian Doppler (Fig. 11) and the important eponymous effect that he predicted in 1842 while a professor at the Czech Polytechnic in Prague. Ironically, Doppler was seeking an explanation of the different colours of binary stars; the effect he proposed to explain the difference was physically correct but completely wrong in his application. Binary stars have different colours because, in the first place, their temperatures (and to some extent their chemical compositions) are different. Doppler suggested, without at that time any experimental support, that the observed frequency of the light emitted by the stars depended on their orbital speeds, which are different (and epoch dependent). This is true, but the effect was far too small to explain the colour differences.

Fig. 11 Christian Doppler (1803–1853).

The Dutch meteorologist Buys Ballot (1817–1890) made the first experimental confirmation of the Doppler effect in 1845 by getting a group of musicians to play a calibrated note on a train on the line between Utrecht and Amsterdam. Of course, in those days there were no police sirens that make the effect so evident today. Despite this confirmation of the effect for sound, Doppler's proposal of a dependence of the observed frequency of light on the speed of the source remained controversial for a surprisingly long time – decades. One person who helped to establish it was Ernst Mach.

8 Mach and Kinematic Residues in Dynamics

Mach (Fig. 12) was one of the great experimentalists of all times and a man of wide interests. His name is associated with three very diverse things in science: Mach bands in psychology, the Mach number in aerodynamics, and Mach's principle in the theory of gravity and inertia. Mach was twice nominated for the Nobel Prize for his discovery of shock waves (Fig. 13), but so many exciting discoveries were being made in the early 20th century that he missed the honour he deserved.

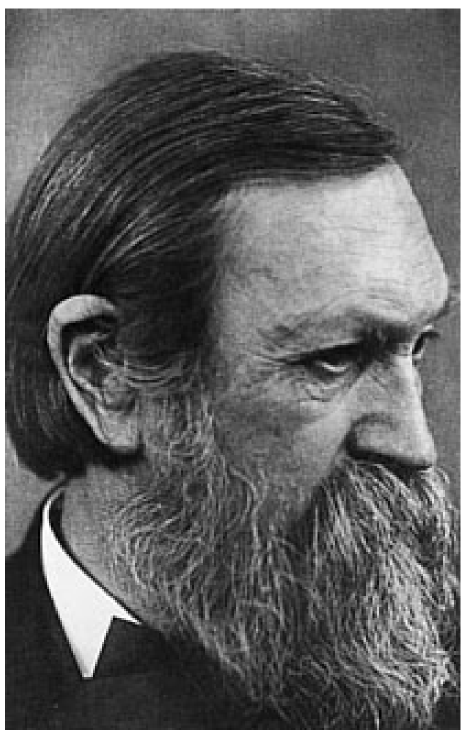


Fig. 12 Ernst Mach (1838–1916).

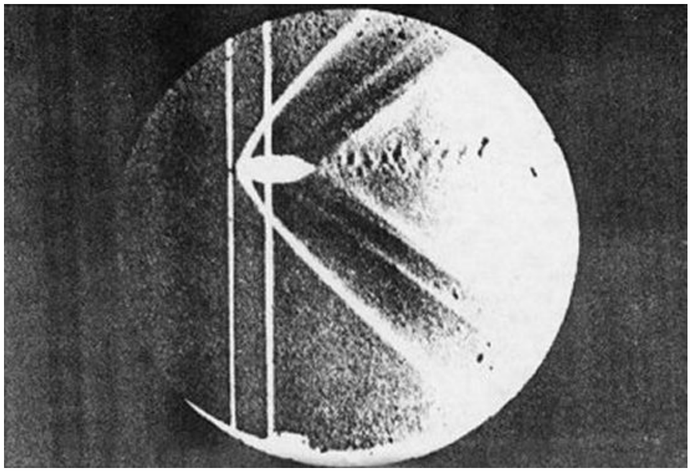


Fig. 13 Mach's flash photograph of a supersonic bullet and the shock wave generated by it.

Before we discuss Mach's critique of Newton's absolute space and time and Einstein's reaction to it, it is worth reproducing the diagram (Fig. 14) in Blackmore's valuable informative biography of Mach [3]. This shows the instrument Mach devised soon after completing his doctorate to demonstrate the Doppler effect for sound. A vertical tube AA rotates in the plane perpendicular to the page. Air forced through the column creates sound in the whistle at G. A person standing in the plane of the rotating tube hears a clearly modulated pitch of the whistle as it rotates at the end of the tube, while someone standing some distance way at right angles to the plane of rotation hears a constant pitch. The Doppler effect for sound is demonstrated in this simple way. The apparatus "became a frequently used class demonstration device throughout Central Europe for many years" [3], p. 19.

To return to our topic, visible and physical markers were key to Kepler's discoveries. The Sun had a dual role: it defined and created motion of the planets. As for the stars, Kepler, like Copernicus, declared them the ultimate frame of reference, by definition at rest. His laws did not contradict this, and the stars did not exhibit any relative motion.

The star-studded shell of this closed world retained the Sun's warmth generated at its centre, or *focus*. Kepler introduced the Latin word for hearth into scientific usage, first in optics and then in astronomy. Descartes (1596–1650) shattered Kepler's cosy 'home' when he introduced the mechanical philosophy. He, above all, marks the transition from the closed world to the infinite universe. In it, all bodies, including the stars and their constituents, move relative to each other.

Descartes actually had two diametrically opposed concepts of motion: absolute and relative. The origins of both are worth retelling (for more details, see [1]). Let us start with the first. One day, lying on his bed, he is said to have spotted a fly on the ceiling and saw he could fix its position by its two distances from the walls. The story is *ben trovato*, apposite even if invented. Cartesian coordinates, so convenient

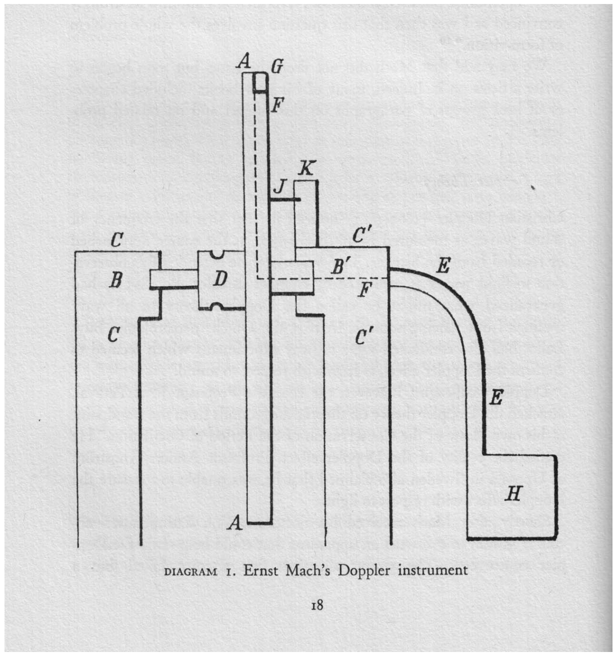


Fig. 14 Mach's device to confirm the Doppler effect for sound, reproduced from [3].

for defining straight lines, were born. The idea of rectilinear inertial motion (not yet named so) was already in circulation; in a book, *Le Monde*, ready for publication in 1632, Descartes made it the foundation of mechanics long before Newton.

Implicit here is an unchanging reified space like the ceiling on which the fly crawled: Newton's absolute frame in all but name. No longer do the Sun and stars define motion. Space does. Whereas pre-Kepler void points governed the planets' motions, now invisible space controls all motion. This is occult forces in spades!

Descartes was just about to publish *Le Monde*, which assumed correctness of the Copernican cosmology, when he heard about Galileo's condemnation by the Inquisition. In alarm – piety he claimed – Descartes hurriedly withdrew his book and thought long and hard how he could save his mechanical philosophy. Eventually he introduced a quite different definition of place and motion in his *Principles of Philosophy* (1644). He declared all motion to be relative. Now any one body has infinitely many positions and motions according to which bodies are used to define them. However, he did grant the existence of 'one true philosophical definition of position', according to which the position of any one body is defined by its envelope, i.e., the immediately adjacent matter that surrounds it. The reason for this definition, actually a throwback to the Aristotelian notion of *topos*, is to be found in Descartes's contention that the Earth is carried around the Sun by a vortex, which is thus its *immediate envelope*. The point then is that the Earth does not move relative to the vortex and therefore does not move in accordance with the true definition.

Since terrestrial mobility (and not heliocentricity) was the Inquisition's objection to Copernicus, Descartes felt he had secured his position and explicitly stated that in accordance with his proposal the Earth does not move.

But after this avowal of pure relationalism, Descartes, failing to note the contradiction, reverted to uniform rectilinear motion as the first principle of mechanics. This made no sense in a world with position and motion defined relatively in either way. It required an implicit absolute space.

Newton studied Descartes's book closely and did see the contradiction. Knowing what could be done with the law of inertia, he recoiled from the virtual impossibility of expressing it rigourously in Descartes's shifting cosmos. The prominence given to absolute space and time in the Scholium at the start of the *Principia* are a covert dismissal of Descartes, even though Newton does grant the great difficulty of distinguishing "the true motions of particular bodies from the apparent; because the parts of that immovable space in which those motions are performed do by no means come under the observation of our senses."

In fact, as Mach was later to remark, Newton's laws were never verified relative to absolute space and time but to exactly the same referents that Kepler had used: the effectively fixed stars and the time-measuring clock supplied by the diurnal revolution of the stars.

Descartes's absolute and relative are the origin of the reductionistic–holistic dichotomy. Mach the holist reacted to Newton the reductionist when he spoke of '*immediate* connections' and the ideal that hovers before the natural investigator as 'an insight into the principles of the whole matter'. The essence of reductionism is threefold: simple objects, atoms, that move in accordance with simple laws, primarily the law of inertia, in a simple background: absolute space. But if position is relative, only the totality of separations between objects is real: the world is held together by an indissoluble network of relations, and history is nothing but their evolution. As Mach said: "The universe is not twice given, with an earth at rest and an earth in motion; but only once, with its relative motions alone determinable." [Mach, p.284]

But Mach went much further than this epistemological verity. Kepler had given the Sun a dynamical role. Mach extended it to the stars or, rather, the totality of masses of the Universe. They should not only define but also control motion. That Mach envisaged this done by some as yet unkown physical mechanism is confirmed by his famous refutation of Newton's bucket argument [13] for absolute motion:

"Newton's experiment with the rotating vessel of water simply informs us that the relative rotation of the water with respect to the sides of the vessel produces no noticeable centrifugal forces, but that such forces are produced by its relative rotation with respect to the mass of the earth and the other celestial bodies. No one

[13] Newton introduced the bucket to make a serious scientific argument but simultaneously a fool of Descartes, whose mechanical philosophy relied heavily on centrifugal force. Many people writing on the absolute–relative debate and unaware of the background to the bucket argument have been misled into thinking the issue is about the difference between linear and circular motion, which is not true. In fact, I increasingly think Newton confused himself.

is competent to say how the experiment would turn out if the sides of the vessel increased in thickness and mass till they were ultimately several leagues thick."

It is well known that Mach's critique made a powerful impression on the late teenage Einstein. It was the main stimulus to his attempt to eliminate all trace of Newton's absolute space through the creation of general relativity. The idea that the totality of the masses of the universe 'work together' to create the local inertial frames in which force-free bodies move rectilinearly and uniformly is what Einstein called *Mach's principle*. For a variety of reasons, this has had a tangled history, for which several factors are responsible.

9 Einstein's Reaction to Mach

Let me start with Einstein's strange confusion of two distinct meanings of inertia: there is inertial *motion*, as defined in Newton's first law, and inertial *mass*. That they are distinct is evident: the concept of inertial mass does not enter into the statement of the first law. Mach gave a much admired operational definition of inertial mass, which he defined through the accelerations bodies impart to each other when they interact. These are inversely proportional to their *intrinsic* inertial masses. Mach's disagreement with Newton on this score was not about substance but proper formulation. What really concerned Mach was the origin of inertial motion: Newton believed absolute space governed it, Mach the totality of masses in the universe.

Reading Einstein's various comments about Mach and inertia I am forced to conclude he was the victim of semantic confusion. He does not seem to have seen any difference between the two meanings of the word inertia. His most egregious distortion of Mach is in the 1917 paper in which he laid the foundation of modern relativistic cosmology. He claims:

"In a consistent theory of relativity there can be no inertia *relatively to 'space'*, but only an inertia of masses *relatively to one another*. If, therefore, I have a mass at a sufficient distance from all other masses in the universe, its inertia must fall to zero."

Mach would have dismissed this comment as a gross distortion of his ideas; Einstein is clearly substituting a bogus issue about inertial mass for Mach's proper concern with inertial motion. Unfortunately, Einstein's 1917 comment led to several misguided attempts to implemented a Mach's principle along inappropriate lines.

However, the complexity of the Machian issue has a much more solid basis and raises a real dilemma, which is illustrated in Fig. 15. What is at stake is the very meaning of the word *relativity*. Einstein and Minkowski had in mind the observer dependence of the split of spacetime into space and time and, more generally, to the complete freedom to lay down coordinates on spacetime in any suitably continuous way. Einstein spoke of general covariance; today one speaks of four-dimensional diffeomorphism invariance. That may be called Einsteinian relativity. The most important aspect of it is the denial of simultaneity as a physically significant concept.

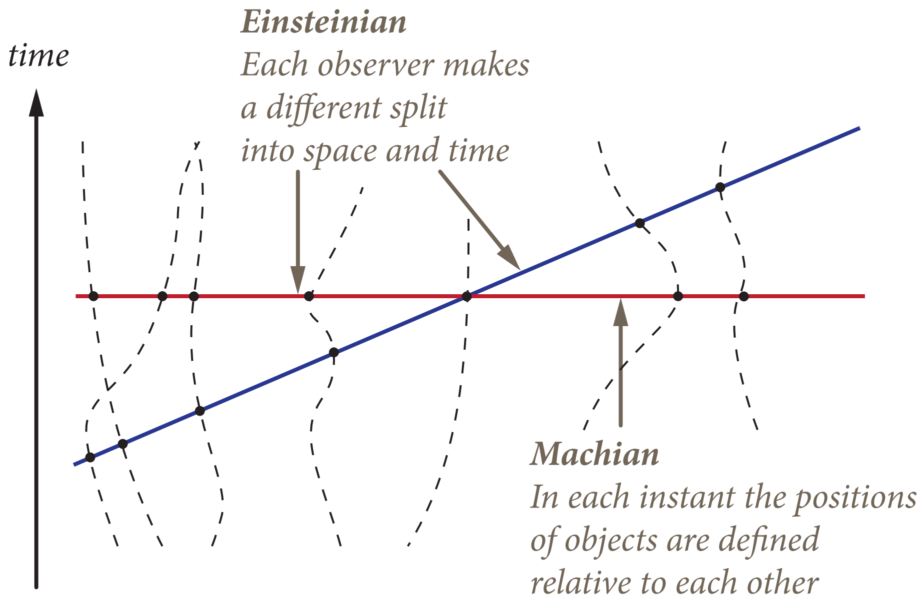


Fig. 15 The two quite different meanings of relativity. Relativity as defined by Minkowski and Einstein refers to the ambiguity in the splitting of four-dimensional spacetime into time and space. Relativity as defined by Mach means that the position of any one body is defined at a given instant by its distances to all the other bodies in the Universe at that instant. The conflict of concepts is evident: Machian relativity makes complete sense if there exists a distinguished notion of simultaneity, but that is denied as the first principle of Einsteinian relativity.

In contrast, relativity as originally formulated by Mach makes no sense without an underlying notion of simultaneity: it asserts that the position of any given object at a given instant is defined by its distance to *all* the other objects in the universe *in that instant*. Most relativists today would say that this is a hopelessly obsolete notion because Einstein and Minkowski showed that our intuitive notion of simultaneity has no counterpart in the physical world. Does this mean that Mach's principle is a dead duck?

Not necessarily. In his only article not devoted to quantum mechanics, John Bell wrote on special relativity and sought "to drive home the lesson that the laws of physics in any *one* reference frame account for all physcial phenomena, including the observations of moving observers" [4], p. 77. Attempts to find a generically distinguished frame in Minkowski space are doomed to fail on account of its high symmetry, but the case is altered in GR: gravity brings structure into spacetime. I want to use the remainder of this paper to give what I believe is the correct definition of Mach's principle and to show how Einsteinian gravity is much more compatible with Machian relativity than one might imagine. I shall even suggest that Machian relativity is the deeper principle.

However, I must first briefly recount how Einstein set out to implement the idea that Mach had espoused: that inertial motion should not be governed by absolute space but the totality of masses in the Universe (a more detailed account is given in [5]).

What was decisive for Einstein was his discovery of special relativity in 1905. This arose from his successful reconciliation of Maxwell's electrodynamics, with its only apparent need for an ether, and Galilean relativity applied to all physical phenomena. The lesson Einstein drew from his success was that uniform motion through Newton's absolute space could not be defined – it was impossible to associate any actual speed with it. Although he said nothing explicit at the time, this result already suggested to Einstein the way to implement Mach's idea: to show that the alleged absolute space had no observable effects at all, for then one could argue that it does not exist. The impossibility of determining a speed of uniform motion through space was the first step in that direction.

The decisive idea that set in motion Einstein's long search for a new theory of gravity was the equivalence principle, that 'happiest thought' of his life which occurred to Einstein in 1907. Its importance for Einstein was not so much that gravity and inertia are identical in essence but the possibility "that the principle of relativity is also satisfied for systems moving relatively to each other with acceleration" [6]. The equivalence principle suggested that this could be done for at least uniform accelerations.

Einstein's strategy from then on was clear and settled. He would attempt to extend the relativity principle ever further. The next step, clearly suggested by Mach's retort to Newton's bucket argument, called for extension to uniform circular motion, for which the magnitude of the acceleration, as in the equivalence-principle elevator, is constant in magnitude but its direction changes. From there, the logical step to complete relativity of motion was not too difficult. Einstein advanced the principle of general covariance as the physical foundation of the new theory of gravity he was seeking.

Two aspects of Einstein's approach should be noted. His principle of relativity did not in any way directly address the way in which the Universe itself behaved. It merely said that the description of its behaviour should be the same in whatever coordinate system one cared to describe it. Einstein's immediate acceptance of Kretschmann's objection that general covariance in itself had no physical content but was merely a requirement of mathematical consistency was a remarkable volte-face that has generated much argument and confusion about the foundations of general relativity and, in particular, Mach's principle. The only conclusion I wish to draw from this brief discussion is that Einstein did not attempt a direct implementation of Mach's ideas but attacked the problem indirectly. This comes out especially clearly in a comment he made in 1918 [7]:

"We want to distinguish more clearly between quantities that belong to a physical system as such ... and quantities that depend on the coordinate system. Ones initial reaction would be to require that physics should introduce in its laws only quantities of the first kind. However, the scientific development has not confirmed this conjecture. It cannot dispense with coordinate systems."

There is a clear anticipation here of the distinction, now commonplace due to developments in gauge theory, between so-called true degrees of freedom and redundant degrees of freedom. What I want to question is whether Einstein had correctly identified what are the "quantities that belong to a physical system as such". In my final section before brief conclusions, I wish to suggest that he may have made the incorrect identification.

10 The Machian Approach: Shape Dynamics

There is no doubt what Mach regarded as the true physical quantities: bodies that possess intrinsic mass and distances (in Euclidean space) between them. He most certainly did not think time had any ontological reality [2], p. 273: "It is utterly beyond our power to measure the changes of things by time. Quite the contrary, time is an abstraction at which we arrive from the changes of things." But Minkowski, followed by Einstein, had given time the same ontological status as space. This led Einstein to identify the "quantities that belong to a physical system as such" with four-dimensional spacetime intervals, whereas Mach had identified them with exclusively three-dimensional spatial entities. Let us see where such a standpoint takes us.

Let us start with one thing on which we can be sure Mach and Einstein would have agreed: if the local frames in which force-free particles move inertially are determined by the universe, there must be a sense in which the universe is a closed dynamical system, for otherwise one could never close the circle and say the whole determines the parts: local inertial frames. This underlying sense is implicit in Mach[14] and explicit in Einstein's 1917 cosmological model.

Let us then allow the notion of simultaneity and assume that the universe is a closed dynamical system. We can consider two models: an island universe of N point particles, which matches the ontology of Mach's original proposal, and a three-dimensional Riemannian geometry closed up on itself, which corresponds to closed-space vacuum GR. In the point-particle case, the difference between Newton and Mach is easily expressed in terms of configuration spaces. Newton's is R^{3N}, three coordinates for each particles. But Mach said only the inter-particle separations are real. We need to quotient R^{3N} by the Euclidean translations and rotations to obtain the $3N-6$-dimensional Machian *relative configuration space* (RCS). Absolute position and orientation are removed from the RCS. In fact, although Mach did not recognize the need, one must go a step further since *distance* persupposes an absolute scale. We also need to quotient by dilatations; this takes us to the $3N-7$-

[14] See his comment [Mach, p. 287] "Nature does not begin with elements, as we are obliged to begin with them. It is certainly fortunate for us that we can, from time to time, turn aside our eyes from the overpowering unity of the All, and allow them to rest on individual details. But we should not omit, ultimately to complete and correct our views by a thorough consideration of the things which for the time being we left out of account." How can completion come without a definite sense in which the universe is closed?

dimensional *shape space* $\mathscr{S}$. I would say that an instantaneous shape of the universe matches Mach's requirement that we grasp the 'immediate connections'.

However, that does not yet mean that we have gained 'an insight into the principles of the whole matter'. Machian histories of the universe will be curves in $\mathscr{S}$. The issue now is this: what determines these curves? The fact is that any Newtonian history can be represented as a curve in $\mathscr{S}$: one simply plots the representative points of the successive shapes. In what way would a Machian history be distinguished from an arbitrary Newtonian one plotted in $\mathscr{S}$?

A problem with Mach is that he tended to speak in general intuitive terms. It is here that a penetrating analysis by Poincaré [8], who analyzed the problem in much more precise terms, provides the guide. Poincaré asked: what defect, if any, arises from Newton's use of absolute space? His answer was that a true believer in relationalism, convinced that only inter-particle separations r_{ab} have physical significance, would pose the initial-value problem of particle dynamics in these terms: $r_{ab}, \dot{r}_{ab}$ should determine the evolution $r_{ab}(t), a, b = 1, 2, ..., N$, uniquely. This matches the formulation in terms of the particle coordinates and velocities; in accordance with Laplacian determinism, $\boldsymbol{r}_a, \dot{\boldsymbol{r}}_a, a = 1, 2, ..., N$, determine the evolution.

Poincaré pointed out that the rather natural transfer of this requirement from R^{3N} to the RCS fails. The reason is that the data $r_{ab}, \dot{r}_{ab}$ contain no information about the angular momentum $\boldsymbol{L}$ in the system, whereas this information is encoded in $\boldsymbol{r}_a, \dot{\boldsymbol{r}}_a$ (under the assumption that the masses are known). Although the presence or absence of $\boldsymbol{L}$ is undetectable in $r_{ab}, \dot{r}_{ab}$ initial data, the curves that result do encode information about $\boldsymbol{L}$. Poincaré said that this fact, reflected in the manifest presence of angular momentum in the solar system, was the true evidence for the existence of a dynamically active agent in addition to the separations r_{ab} and their rates of change $\dot{r}_{ab}$. As a convinced believer that only relative motions should have dynamical effect, Poincaré said he found this state of affairs repugnant but that it was necessary to accept the empirical evidence.

It is strange that Poincaré did not consider a Machian resolution to the problem, namely that for *for a dynamically closed universe as a whole* the relative data do determine the future uniquely. One can then attribute the failure of this requirement in subsystems of the universe precisely to the fact that the masses of the universe do determine local inertial frames of reference. Poincaré formulated his ideas in the context of the RCS, but they can be directly extended to and made more stringent in shape space. This leads me to the formulation of the Mach-Poincaré principle for particle dynamics in these terms.

Mach–Poincarè Principle. Specification of a point and direction (strong form) or point and tangent vector (weak form) in shape space $\mathscr{S}$ should determine the evolution in $\mathscr{S}$ uniquely.[15]

[15] In my mind, the great virtue of Poincaré's analysis is that he formulates requirements on the form of a dynamical theory of the Universe in terms of the *initial data* that one regards as belonging "to the physical system as such". This allows a much more precise formulation than Einstein's requirement that all coordinate systems should be on an equal footing, which is actually void of content, or that the action should satisfy certain symmetry requirements, which is also amenable to adjustment, as one sees with the passage from standard Newtonian dynamics to parametrized

It is necessary to allow for the weaker form if one is to model expansion of the Universe. For discussion of this delicate issue, see my introduction to shape dynamics [9]. The extension to dynamical geometry is relatively obvious; the shape space in this case is conformal superspace, which provides the natural framework for describing the dynamics of three-dimensional conformal geometries. I cannot describe in detail this work; the most important papers are [10, 11, 12] (see also Koslowski's contribution to this conference proceedings).

What one can say is that, if the Universe is spatially closed, there is a well-defined sense in which GR implements Mach–Poincaré principle in the weak form very well, indeed perfectly if there is no cosmological constant. For all the details I must refer the reader to the references already cited, but the key conclusions do need to be stated, at least for vacuum gravity. First, by virtue of its clearly formulated first principles shape dynamics introduces of necessity a notion of simultaneity into GR and insists that the physical enitity which is evolving is the conformal three-geometry on successive leaves of a foliation of spacetime by hypersurfaces of constant-mean-(extrinsic)-curvature (CMC). Second, the spacetime in which these hypersurfaces are embedded is completely determined by specification of a point and tangent vector in conformal superspace. This fact was first demonstrated in [13].

If the ideas of shape dynamics, which do follow very naturally from Mach's ideas, are vindicated, it will be incorrect to view shape dynamics as a rule to select certain special solutions – those that are globally hyperbolic and CMC foliable – from among the full set allowed by GR. Rather GR might have to be seen as an extension of shape dynamics beyond its physical domain. I suspect we shall have to await the quantum theory of gravity to see if this view is justified.

11 Conclusions

Galileo said "He that attempts natural philosophy without geometry is lost." He meant of course *three-dimensional* geometry, which was still Euclidean in his day, though I am sure he would have greeted Riemann's generalization with enthusiasm. The first step in the still ongoing creation of the dynamical theory of the Universe was Hipparchus's theory of the Sun's motion. It is important that all the great work in astronomy reviewed in this paper studied the evolution of *intrinsic shapes*, of which the fixed stars formed part. Absolute position, orientation and size played role at all; I have already emphasized that every conclusion drawn in astronomy was based on measurement of angles between observed physical objects. These included measurements of what was called time but was actually the diurnal rotation of the stars. Even now, with geometry curved and made dynamical, the irreducible

---

particle dynamics, which adds reparametrization invariance as a symmetry without changing the physical content of the theory. In contrast, implementation of Mach's ideas boils down (in the case of the weak Mach–Poincaré principle) to identification of the true (configurational) degrees of freedom and construction of a theory in which they and their velocities wrt an independent variable uniquely determine the evolution of the true degrees of freedom.

epistemological basis of science is observed angles. We now see the Universe as almost infinitely flexible, but we cannot do without angles.

A conformal geometry supplies the ‘immediate connections’ that Mach exhorted us to grasp. As regards “the principles of the whole matter”, I would say that as far as classical physics is concerned they are encapsulated in the weak Mach–Poincaré principle applied to a closed Universe whose possible spatial configurations are defined by conformal geometry.

Let me end with my sincere thanks to the organizers, above all Jiří Bičák, for the invitation to speak at the wonderful conference in Prague.

References

1. J.B. Barbour, *Absolute Or Relative Motion?: The discovery of dynamics*. Absolute Or Relative Motion?: A Study from a Machian Point of View of the Discovery and the Structure of Dynamical Theories. The Discovery of Dynamics (Cambridge University Press, 1989)
2. E. Mach, *The Science of Mechanics* (Open Court, 1960)
3. J.T. Blackmore, *Ernst Mach. His Life, Work, and Influence* (University of California Press, 1972)
4. J.S. Bell, *Speakable and Unspeakable in Quantum Mechanic* (Cambridge University Press, 1987)
5. J.B. Barbour, *The part played by Mach's Principle in the genesis of relativistic cosmology* (1990), p. 47
6. A. Einstein, *Berichtigungen zu der Arbeit: "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen"*, Jahrbuch der Radioaktivität und Elektronik **5**, 98 (1908)
7. A. Einstein, *Prinzipielles zur allgemeinen Relativitätstheorie*, Annalen der Physik **360**, 241 (1918)
8. H. Poincaré, *Science and Hypothesis* (Walter Scott, London, 1905)
9. J.B. Barbour, *Shape dynamics: An introduction* (Birkäuser, 2012)
10. E. Anderson, J. Barbour, B.Z. Foster, B. Kelleher, N.Ó. Murchadha, *The physical gravitational degrees of freedom*, Classical and Quantum Gravity **22**, 1795 (2005)
11. H. Gomes, S. Gryb, T. Koslowski, *Einstein gravity as a 3D conformally invariant theory*, Classical and Quantum Gravity **28**(4), 045005 (2011)
12. H. Gomes, T. Koslowski, *The link between general relativity and shape dynamics*, Classical and Quantum Gravity **29**(7), 075009 (2012)
13. J. Barbour, N.Ó. Murchadha, *Conformal superspace: the configuration space of general relativity*, ArXiv e-prints [arXiv:1009.3559 [gr-qc]] (2010)