Quasi periodic motions from Hipparchus to Kolmogorov.

Giovanni Gallavotti
Fisica, Università di Roma "La Sapienza"
(29 june 1999)

The evolution of the conception of motion as composed by circular uniform motions is analyzed, stressing its continuity from antiquity to our days.[1]

I. HIPPARCHUS AND PTOLEMY

Contemporary research on the problem of chaotic motions in dynamical systems finds its roots in the Aristotelian idea, often presented as kind of funny in high school, that motions can always be considered as composed by circular uniform motions,[3,4].

The reason of this conception is the perfection and simplicity of the uniform circular motion (of which the uniform rectilinear motion case must be thought as a limit case).

The idea is far more ancient than Hipparchus (from Nicea, 194-120 a.C.) from whom, for simplicity of exposition it is convenient to start. The first step is to understand clearly what the Greeks really meant for motion composed by circular uniform motions. This indeed is by no means a vague and qualitative notion, and in Greek science it acquired a very precise and quantitative meaning that was summarized in all its surprising rigor and power in the *Almagest* of Ptolemy (~100-175 d.C.).[5–7]

We thus define the *motion composed by n uniform circular motions* with angular velocities $\omega_1, \ldots, \omega_n$ that is, implicitly, in use in the *Almagest*, but following the terminology of contemporary mathematics.

A motion is said *quasi periodic* if every coordinate of any point of the system, observed as time t varies, can be represented as:

$$x(t) = f(\omega_1 t, \ldots, \omega_n t) \tag{1}$$

where $f(\varphi_1, \ldots, \varphi_n)$ is a multiperiodic function of n angles, with periods 2π and $\omega_1, \ldots, \omega_n$ are n angular velocities that are "rationally independent";[8] they were called the [velocities of the] "*motors*" of the Heavens.

We must think of such function f as a function of the positions $\varphi_1, \ldots, \varphi_n$, ("*phases*" or "*anomalies*"), of n points on n circles of radius 1 and, hence, that the state of the system is determined by the values of the n angles. Therefore to say that an observable x evolves as in (1) is equivalent to say that the motion of the system simply corresponds to uniform circular motions of the points that, varying on n circles, represent the state of the system.

We shall say, then, that the motion is *composed by n uniform circular motions* if it is quasi periodic in the sense of (1).

In reality in Greek Astronomy it is always clear that the motion of the solar system, conceived as a quasi periodic motion, is *only one among the possible motions* of a wider family that have the form

$$x(t) = f(\varphi_1 + \omega_1 t, \ldots, \varphi_n + \omega_n t). \tag{2}$$

Hence it is in a stronger sense that the motions are thought of as composed by quasi periodic motions. Indeed all the n-ples $(\varphi_1, \ldots, \varphi_n)$ of phases are considered as describing possible states of the system. This means that one thinks that the phases $(\varphi_1, \ldots, \varphi_n)$ provide a system of coordinates for the possible states of the system. The observed motion is one that corresponds, conventionally, to the initial state with phases $\varphi_1 = \varphi_2 = \ldots = 0$; but also the other states with arbitrary phases are possible and are realized in correspondence of different given initial conditions and, furthermore, *we can get close to them by waiting long enough.*

Summarizing: to say that the motions of a system are composed by n circular uniform motions, of angular velocities $\omega_1, \ldots, \omega_n$ is equivalent to say that it is possible to find a system of coordinates that describe completely the states of the system (relevant for the dynamical problem under study) in which the n coordinates are n angles and, furthermore, that in such coordinates the motion is simply a uniform circular motion of every angle, with suitable angular velocities $\omega_1, \ldots, \omega_n$. This is indeed manifestly equivalent to saying that an arbitrary observable of the system, evolving in the time, admits a representation of the type (2).

In Greek physics no methods were available (that we know of) for the computation of the angle coordinates in terms of which the motion would appear circular uniform, *i.e.* no methods were available for the computation of the coordinates φ_i and of the functions f, in terms of coordinates with direct physical meaning (*e.g.* polar or Cartesian coordinates of the several physical point masses of the system). Hence Greek astronomy did consist in the hypothesis that all the motions could have the form (2) and in deriving, then, by experimental observations the functions f and the velocities ω_i well suited to the description of the planets and stars motions, with a precision that, even to our eyes (used to the screens of digital computers), appears marvelous and almost incredible.

After Newton and the development of infinitesimal calculus it has become natural and customary to imagine dynamical problems as developing starting from initial conditions that can be quite different from those of immediate interest in every particular problem. For example it is common to imagine solar systems in which the radius of the orbits of Jupiter is double of what actually is or in which the Moon is at a distance from the Earth different from the observed one, *etc.*. Situations of this kind can be included in the Greek scheme simply by imagining that the coordinates $\varphi_1, \ldots, \varphi_n$ are not a complete system of coordinates, and other coordinates are needed to describe the motions of the same planets if they are supposed to have begun their motion in situations radically different from those which, soon or later, they would reach when starting at the given present states (and which "just" correspond to states with arbitrary values of the phase coordinates $\varphi_1, \ldots, \varphi_n$).

To get a complete description of these "other possible motions" of the system other coordinates $A_1, \ldots, A_m$ are necessary: they are, however, *constant in time* on every motion and hence they only serve to specify to which family of motions the considered one belongs. Obviously we shall have to think that the $\omega_1, \ldots, \omega_n$ themselves are functions of the A_i and, in fact, it would be convenient to take the ω_i themselves as part of the coordinates A_i, particularly when one can show that $m = n$ and that the ω_i can be independent coordinates.

Let us imagine, therefore, that the more general motion has the form:

$$x(t) = f(A_1, \ldots, A_m, \omega_1 t + \varphi_1, \ldots, \omega_n t + \varphi_n) \qquad (3)$$

where $\omega_1, \ldots, \omega_n$ are functions of $A_1, \ldots, A_m$ and the coordinates $A_1, \ldots, A_m, \varphi_1, \ldots, \varphi_n$ are a *complete* system of coordinates.

In Greek astronomy there is no mention of a relation between m and n: probably only because no mention is made of the coordinates $A_1, \ldots, A_m$ since the Greeks depended exclusively on actual observations hence they could not conceive studying motions in which the $A_1, \ldots, A_m$ (*e.g.* the radii of the orbits of the planets, the inclinations of the orbits, *etc.*) were different from the observed values.

In this respect it is important to remark that Newtonian mechanics shows that it must be $m = n = 3N =$ {number of degrees of freedom of the system}, if N is the number of bodies, even though in general it can happen that the $\omega_1, \ldots, \omega_n$ cannot be taken as coordinates in place of the A_i's because they are not always independent of each other (for instance the Newtonian theory of the two body problem gives that the three ω_i are all rational multiples of one of them, as otherwise the motion would not be periodic). Nevertheless this identity between m and n has to be considered one among the great successes of Newtonian mechanics.

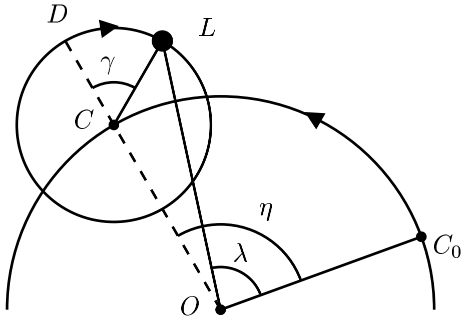
Returning to Greek astronomy it is useful to give some example of how one concretely proceeded to the determination of $\varphi_1, \ldots, \varphi_n$, of $\omega_1, \ldots, \omega_n$ and f.

A good example (other than the motion of the Fixed Stars, that is too "trivial", and the motion of the Sun that is, in a way, too "simple" to allow us to appreciate the differences between the Ptolemaic and Copernican theories) is provided by the theory of the Moon.

As first example I consider Hipparchus' theory of the Moon.

In general motions of heavenly bodies appear, in a first approximation, as uniform circular motions around the center of the Earth. Or *in average* the position of the heavenly body can be deduced by imagining it in uniform motion on a circle, (*"the oblique circle that carries the planets along"*, Dante, Par., X), with center on the Earth and rigidly attached to the sphere ("sky") of the Fixed Stars, that in turn rotates uniformly around the Earth.

This average motion was called *deferent motion*: but the heavenly body almost never occupies the average position, rather it is *slightly* away from it sometimes overtaking it and sometimes lagging behind.



h1

Fig. h1: Hipparchus' Moon theory with one deferent and one epicycle.

In the case of the motion in longitude of the Moon (*i.e.* of the projection of the lunar motion on the plane of the ecliptic) the simplest representation of these oscillations with respect to the average motion is by means of two circular motions, one on a circle of radius $CO = R$, *deferent*, with velocity ω_0 and another on a small circle of radius $CD = r$, *epicycle*, with angular velocity ω_1.

We reckon the angles from a conjunction between the Moon and the average Sun, *i.e.* when OC_0 projected on the ecliptic plane points at the position of the average Sun, which is a Sun which moves exactly on a circle with uniform motion and period one solar day (we could use instead a conjunction between the Moon and a fixed Star, with obvious changes).

The center of the epicycle rotates on the deferent with angular velocity ω_0 and the Moon L rotates on the epicycle with velocity $-\omega_1$,[4]

By using the complex numbers notation to denote a vector in the ecliptic plane and beginning to count angles from the position of apogee on the epicycle we see that the vector z that indicates the longitudinal position of the Moon is

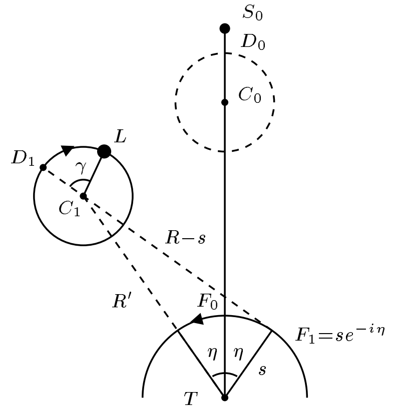
$$\begin{cases} \eta = \omega_0 t \\ \gamma = -\omega_1 t \end{cases} \Rightarrow z = R e^{i\omega_0 t} + r e^{-i(\omega_1 - \omega_0)t} \qquad (4)$$

where $\eta = \omega_0 t$ is the angle C_0OC, $\gamma = -\omega_1 t$ is the angle DCL and $2\pi/\omega_1$ is the time T that elapses between two successive returns of the Moon in apogee position on her epicycle (since $\omega_0 \sim \omega_1$ the new apogee will happen at a time T for which there exists a small angle δ such that: $\omega_0 T = 2\pi + \delta$ and $(\omega_0 - \omega_1)T = \delta \Rightarrow \omega_1 T - 2\pi = 0 \Rightarrow T = 2\pi/\omega_1 =$ *anomaly month*).

In modern language we say that the Moon, having three degrees of freedom, shall have a motion with respect to the Earth (assumed on a circular orbit) endowed with 3 periods: the month of anomaly (*i.e.* return to the apogee and "true period of revolution") of approximately 27^d, the period of rotation of the apogee of approximately 9^y (in direction concording with that of the revolution) and the period of precession (retrograde) of the node between the lunar world and the ecliptic of approximately 18.7^y. This last period, obviously, does not concern the motion in longitude which, therefore, is characterized precisely by two fundamental periods: for instance the month of anomaly and the sideral month (return to the same fixed star: note that the difference between the the two angular velocities $\omega_0 - \omega_1$ is obviously the velocity of precession of the apogee).

Hipparchus theory of the motion in longitude of the Moon yields, as we see, a quasi periodic motion with one deferent, one epicycle and two frequencies (or "motors"). It reveals itself sufficient (if combined with the theory of the motion in latitude, that we do not discuss here) for the theory of the eclipses, but it provides us with ephemerides (somewhat) incorrect when the Moon is in position of quadrature.

Ptolemy develops a more refined theory of this motion in longitude,[4,5]. Again assume that the angles in longitude are reckoned from the mean Sun S_0 starting at a conjunction, as in the previous theory. In a first version he imagines that the center of the epicycle moves at the extremity of a segment of length $R-s$ that however does not have origin on the Earth T but in a point F_1 that moves with (angular) velocity $-\omega_0$ on a small circle of radius s centered on T; we suppose that the center of the epicycle is C_1 so that the angle between C_1T and the axis of apogee (our reference x–axis on a complex z–plane) is still $\eta = \omega_0 t$, but the angle $\gamma = -\omega_1 t$ that determines the position L of the Moon on the epicycle is now reckoned from D_1:



t1

Fig. t1: The correction of Ptolemy to Hipparchus' theory of the Moon.

in formulae, if $R' = C_1T$, $\tilde{R} = C_1F_1 = R - s$ and ϑ is the angle between D_1F_1 and D_0F_0, one finds:

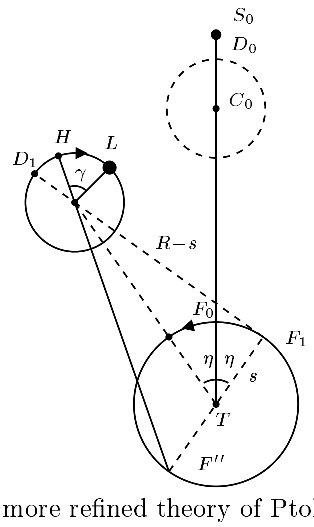
$$\begin{aligned}
\eta &= \omega_0 t,\\
\tilde{R}e^{i\vartheta} &= R'e^{i\omega_0 t} - se^{-i\omega_0 t}\\
\tilde{R} &= |R'e^{i\omega_0 t} - se^{-i\omega_0 t}|,\\
R' &\equiv R'(\omega_0 t; R, s) =\\
&= s\cos 2\omega_0 t + \tilde{R}(1 - (s/\tilde{R})^2 \sin^2 2\omega_0 t)^{1/2}\\
z &= R'(\omega_0 t; R, s)e^{i\omega_0 t} + re^{-i(\omega_1 - \vartheta)t}
\end{aligned} \tag{5}$$

which reduces to Hipparchus' moon theory if $s = 0$. It also gives the same result in conjunction and in opposition (*i.e.* when $\eta = 0, \pi$); it gives a closer Moon at quadratures (*i.e.* when $\eta = \frac{1}{2}\pi, \frac{3}{2}\pi$).

This representation reveals itself sufficient for the computation of the ephemerides also in quadrature positions, but is insufficient (although off by little) for the computation of the ephemerides in octagonal positions (*i.e.* at 45^o from the axes).

Note that, rightly so, no new periods are introduced: the motion has still two basic frequencies and (5) only has *more Fourier harmonics* with respect to (4).

The theory was therefore further refined by Ptolemy himself,[4,5], who supposed that, in the preceding representation, the computation of the angle $\gamma = -\omega_1 t$ on the epicycle should be performed not by starting from the axis C_1F_1, as in the previous case, but rather from the axis $F"H$ of the figure:



t2

Fig. t2: The more refined theory of Ptolemy.

in formulae, with R' as in (5)

$$z = R'e^{i\omega_0 t} + r\frac{R'e^{-i\omega_0 t} + se^{i\omega_0 t}}{|R'e^{-i\omega_0 t} + se^{i\omega_0 t}|}e^{-i\omega_1 t} \tag{6}$$

It is clear that with corrections of this type it is possible to obtain very general quasi periodic functions. Note that the above theory coincides with the preceding one at conjunction, opposition and quadratures and it is otherwise somewhat different (in particular at the octagonal positions).

The values that Ptolemy finds for R, r, s, so that the theoretical ephemerides conform with the experimental ones, are however such that the possible variations of the Earth–Moon distance (between $R-r-s$ and $R+r+s$) are very important and incompatible with a *not observed* corresponding variation of the apparent diameter of the heavenly body. Astronomical *distances* (as opposed to delestial longitudes and latitudes of planets) were not really measured in Greek times: but we shall see that in Kepler's theory the measurability of their value payed a major role. It is not known why the apparent diameter of the Moon did not seem to worry Ptolemy.

II. COPERNICUS

Copernicus (1473-1543)[9] (who was, indeed, very worried by the latter problem) tried to find a remedy by introducing a secondary epicycle: his model goes back to that of Hipparchus which is "improved" by imagining that the point of the epicycle in which Hipparchus set the Moon was instead the center of a smaller secondary epicycle, of radius s, on which, the Moon journeyed with angular velocity $-2\omega_0$

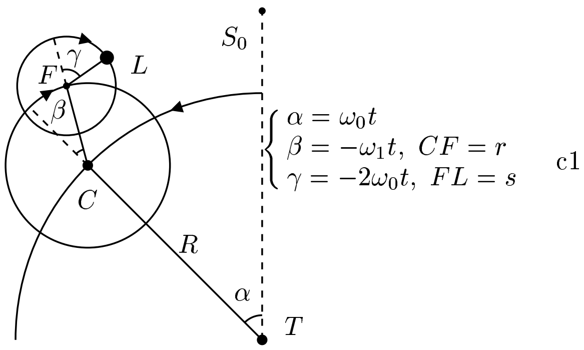


Fig. c1: Copernicus Moon theory with two epicycles.

and in formulae:

$$r_L = Re^{i\omega_0 t} + e^{-(\omega_1-\omega_0)t}(r + se^{2i\omega_0 t}) \tag{7}$$

This gives a theory of the longitudes of the Moon essentially as precise as that of Ptolemy. Note that, *again*, the same two independent angular velocities are sufficient.

Before attempting a comparison between the method of Ptolemy and that of Copernicus it is good to clarify the modern interpretation of the notions of deferent and epicycle and to clarify, also, that the motions of the Ptolemy's lunar theories are still interpretable as motions of deferents and epicycles. Which is not completely obvious since some of the axes of reference of Ptolemy *do not move of uniform circular motion*, to an extent that by several accounts, still today, Ptolemy is "accused" of having abandoned the purity of the circular uniform motions with the utilitarian scope of obtaining agreement between the experimental data and their theoretical representations[10].

I just quote here Copernicus *Commentariolus*, few lines before the statement of his famous second postulate setting the Earth away from the center of the World

"*Nevertheless, what Ptolemy and several others legated to us about such questions, although mathematically acceptable, did not seem not to give rise to doubts and difficulties*" ... "*So that such an explanation did not seem sufficiently complete nor sufficiently conform to a rational criterion*" ... "*Having realized this, I often meditated whether, by chance, it would be possible to find a more rational system of circles with which it would be possible to explain every apparent diversity; circles, of course, moved on themselves with a uniform motion*", see[11] p.108.

Therefore let us check what was, in some form, probably so obvious to Ptolemy that he did not seem to feel the necessity of justifying his alleged *deviation* from the "dogma" of decomposability into uniform motions. Namely we check that also the motions of the Ptolemaic lunar theories, as actually *all* quasi periodic motions, can be interpreted in terms of epicycles.

Consider for simplicity the case of quasi periodic motions with two frequencies ω_1, ω_2. Then the position will be

$$z(t) = \sum_{\nu_1,\nu_2}^{-\infty,\infty} \rho_{\nu_1\nu_2} e^{i(\omega_1\nu_1+\omega_2\nu_2)t} \equiv \sum_j \rho_j e^{i\Omega_j t} \tag{8}$$

by the theorem on the Fourier series, if ν_i are arbitrary integers and if j in the second sum denotes a pair ν_1, ν_2 and $\Omega_j \equiv \omega_1\nu_1 + \omega_2\nu_2$. Imagine, for simplicity, also that the enumeration with the label j of the pairs ν_1, ν_2 could be made, and is made, so that $\rho_1 >> \rho_2 \ge \rho_3 > \ldots$.

Then $r(t)$ can be rewritten as

$$r(t) = \rho_1 e^{i\Omega_1 t}\Big(1 + \frac{\rho_2}{\rho_1} e^{i(\Omega_2-\Omega_1)t} \cdot \big(\cdot\big(1 + \frac{\rho_3}{\rho_2} e^{i(\Omega_3-\Omega_2)t}(1+\ldots)\big)\Big) \tag{9}$$

which, neglecting $\rho_2, \rho_3 \ldots$, is the uniform circular motion on the deferent of radius $|\rho_1|$ with velocity angular Ω_1; neglecting only $\rho_3, \rho_4, \ldots$ it is a motion with a deferent of radius $|\rho_1|$ rotating at velocity Ω_1 on which rests an epicycle of radius $|\rho_2|$ on which the planet rotates at velocity $\Omega_2 - \Omega_1$; neglecting only ρ_j, $j \ge 4$ one obtains a motion with one deferent and two epicycles, as that used by Copernicus in the above lunar model.

If $|\rho_1|$ is not much larger than the other radii (and precisely if $|\rho_1|$ not is larger than the sum of the other $|\rho_j|$, a situation that is not met in the ancient astronomy), what said remains true except that the notion of deferent is no longer meaningful. Or, in other words, the distinction between *main* circular motion and epicycles is no longer so clear from a physical and geometrical viewpoint. The epicycle with radius larger than the sum of the radii of the other epicycles, if existent, essentially determines the average motion and is given the privileged name of "deferent". In the other cases, although the average motion still makes sense,[12], it is no longer associated with a particular epicycle, but all of them concur to define it, for an example see [AA68], p.138.

We see, therefore, the *complete equivalence* between the representation of the quasi periodic motions by means of a Fourier transform and that in terms of epicycles,[10].

Greek astronomy, thus, consisted in the search of the Fourier coefficients of the quasi periodic motions of the heavenly bodies representing them geometrically by means of uniform motions.

But Ptolemy's method is in a certain sense not systematic (see, however, below): the intricate interplay of rotating sticks that explains, or better parameterizes, the motion of the Moon is very clever and precise but it seems quite clearly not apt for obvious extensions to the cases of other planets and heavenly bodies.

Copernicus' idea, *instead*, of introducing epicycles of epicycles, as many as needed to an accurate representation of the motion, is systematic and, as seen above, coincides with the computation of the Fourier transform of the motion coordinates with coefficients ordered by decreasing absolute value. Copernicus' work (with the only exception, and such only in a rather restricted sense that it is not possible to discuss here, of some details of the motion of Mercury) is strictly coherent with this principle. set in his early project quoted above.

This is perhaps[16] the great innovation of Copernicus and not, certainly, the one he is always credited for, *i.e.* having referred the motions to the (average) Sun rather than to the Earth: that is a trivial change of coordinates, known as possible and already studied in antiquity,[14,15], by Aristarchus (of Samos, 310-235 a.C.), Ptolemy *etc.*, but set aside by Ptolemy for obvious reasons of convenience, because in the end it is from Earth that we observe the heavens (so that still today many ephemerides are referred to the Earth and not to an improbable observer on the Sun), and also because he seemed to lack an understanding of the principle of inertia (as we would say in modern language). See the *Almagest*, p.45 where allegedly Ptolemy says: "*...although there is perhaps nothing in the celestial phenomena which would count against that hypothesis* [that the Sun is the center of the World]*... one can see that such a notion is quite ridiculous.*[17]

Ptolemy, with clever and audacious geometric constructions does not compute coefficient after coefficient the first few terms of a Fourier transform of the motion. He *sees* directly series which contain infinitely many Fourier coefficients (see $R'(\omega_0 t)$ in (5) where this happens because of the square root), *i.e.* infinitely many epicycles, *most of which are obviously very small and hence irrelevant.*

We can therefore obtain the same results with several arrangements of sticks, provided that the motion that results has Fourier coefficients, I mean those which are not negligible, equal or close to those of the motion that one wants to represent: it is this absence of uniqueness that makes the method Ptolemaic appear not systematic.

It has, however, the advantage that, if applied by an astronomer like Ptolemy, it requires apparently, at equal approximation, less elementary uniform circular motions, a fact that was erroneously interpreted as meaning *less epicycles* (which, on the contrary, are very often, in the Ptolemaic constructions, infinitely many as we see in the case of (5),(6)) than usually necessary with the methods of Copernicus: a fact that was and still is considered a grave defect of the Copernican theory compared to the Ptolemaic. Ptolemy identifies 43 fundamental uniform circular motions (that combine to give rise to quasi periodic functions endowed with *infinitely many harmonics* formed with the 43 fundamental frequencies) to explain the whole system of the World: Copernicus hopes initially (in the *Commentariolus*) to be able to explain everything with 34 harmonics, only to find out in the *De Revolutionibus* that he is forced to introduce several more. See Neugebauer in[5], vol. 2, p.925- 926.[18]

One should not, however, miss stressing also that Copernicus heliocentric assumption made possible a simple and unambiguous computation of the planetary distances.[16] If the Sun is assumed as the center, and the orbits are supposed circular (to make this remark simplest) then the radii of the epicycles of the external planets (for instance) are automatically fixed to be all equal to the distance Earth–Sun. Then, knowing the periods of revolution and observing one opposition (to the Sun) of a planet and one position off conjunction at a later time, one easily deduces the distance of the planet to the Earth and to the Sun, in units of the Earth–Sun distance. In a geocentric system the radii of the epicycles are simply related to their deferents sizes and the latter are *a priori* unrelated to the Sun–Earth distance: for this reason in ancient astronomy the size of the planetary distances was a big open problem. One can "save the phenomena" by *arbitrarily scaling* deferent and epicycles radii *independently for each planet!* The possibility of reliably measuring the distances, applied by Copernicus and then by Tycho and Kepler, was essential to Kepler who could thus see that the saving of the phenomena in longitudinal observation was not the same as saving them in the radial observations, a more difficult but very illuminating task, see[23].

III. KEPLER

Today we would say that Ptolemy's theory was *nonperturbative* because it immediately represented the motions as quasi periodic functions (with infinitely many Fourier coefficients). Copernicus' is, instead, *pertubative* and it systematically generates representations of the motions by means of developments with a finite number of harmonics constructed by adding new *pure* harmonics, one after the other, with the purpose of improving the agreement with experience. The larger number of harmonics in Copernicus is simply explained because, from his point of view, harmonics multiple of others count as different epicycles, while in Ptolemy the geometric constructions associated with an epicycle sometimes introduce also harmonics that are multiples, or combinations with integer coefficients, of others already existent and produce an "apparent" saving of epicycles.

But the systematic nature of the Copernican method permitted to his successors to organize the large amount of new data of the astronomers of the Renaissance and of the Reform time. Eventually it allowed Kepler (1571-1630) to recognize that what was being painfully constructed, coefficient after coefficient, was in the simplest cases just the Fourier series of a motion that developed on an ellipse with the Sun, or the Earth in the case of the Moon, in the focus and with constant area velocity.

For reasons that escape me the History of Science usually credits Kepler to have made possible the *rejection* of the scheme of representation in terms of deferents and epicycles, in favor of motions on ellipses.

But it is instead clear that the Keplerian motions are still interpretable in terms of epicycles whose amplitudes and positions are computed with the Copernican or Ptolemaic methods (that he regarded as equivalent in a sense that reminds us of the modern theories of "equivalent ensembles" in statistical mechanics, see[20] Ch. 1-4) or, equivalently, via the modern Fourier transform. Nor it should appear as making a difference that the epicycles are, strictly speaking, infinitely many, (even though all except a small number have amplitudes, *i.e.* radii, which are completely negligible): already the Ptolemaic motions, with the their audacious constructions based on rotating sticks did require, to be representable by epicycles (*i.e.* by Fourier series) infinitely many coefficients (or *harmonics*), see (5),(6) above, in which the r.h.s. manifestly have infinitely many nonvanishing harmonics.

Only Copernican astronomy was built to have a finite number of epicycles: but their number had to be ever increasing with the increase of the precision of the approximations. Ptolemy *seemed looking* and Kepler certainly *was looking* for exact theories, Copernicus appears to our eyes doomed to look for better and better approximations.

In reality also the critique of lack of a systematic method in Ptolemy, the starting point of the Copernican theory, should be reconsidered and subject to scrutiny: indeed we do not know the theoretical foundations on which Ptolemy based the *Almagest* nor through which deductions he arrived at the idea of the *equant* and to other marvelous devices. One can even dare the hypothesis that the *Almagest* was just a volume of commented tables based on principles so well known to not even deserve being mentioned. It is difficult to imagine that Ptolemy had proceeded in an absolutely empirical manner in the invention of *anomalous* objects like "equant points" and strange epicycles (like those he uses in the theory of the Moon) and he did not feel that he was departing from the main stream based on the axiom that all motions were decomposable into uniform circular motions: it is attractive, instead, to think that he did not feel, by any means, to have violated the law of the composition of motions by circular uniform ones.

One should note that if a scientist of the stature of Copernicus in a 1000 years from now, after mankind recovered from some great disaster, found a copy the American Astronomical Almanac[19] (possibly translated from translations into some new languages) he would be astonished by the amount of details, and by the data correctness, described there and he would be left wandering how all that had been compiled: because it is *very difficult, if not impossible* to derive even the Kepler's laws, directly from it (not to mention the present knowledge on the three body problem). And he would say "surely there must be a simpler way to represent the motions of the planets, stars and galaxies", and the whole process might start anew, only to end his life (as Copernicus in "*De revolutionibus*")[11] with new tables that coincided with an appropriately updated version of the ones he found in his youth. The American Astronomical Almanac can be perhaps better compared to Ptolemy's *Planetary Hypotheses* if the latter is really due to him, as universally acccepted, while the Almagest is an earlier but more detailed version of it[17].

After the discovery of the Kepler laws the theory of gravitation of Newton (1642-1727) was soon reached,[24]. Contrary to what at times is said, far from marking the end of the grandiose Greek conception of motion as composed by circular uniform motions, Newtonian mechanics has been, instead, its most brilliant confirmation.

For example, if ϑ denotes the angle between the major semiaxis and the actual position on the orbit ("*true anomaly*"), ℓ denotes the *average anomaly*, a is the major semiaxis of the ellipse and e is its *eccentricity*, the Keplerian motion of the Mars around the Sun is described by the equations:

$$\begin{aligned} &z = p e^{i\vartheta}(1-e\cos\vartheta)^{-1},\ p = a(1-e^2)\\ &\vartheta = \ell - 2e\sin\ell + \frac{5}{4}e^2\sin 2\ell + O(e^3),\ \ell = \omega t \end{aligned} \tag{10}$$

hence

$$\begin{aligned} z &= p(1-e^2)^{1/2} e^{i\vartheta}(1+2\sum_{n=1}^{\infty}\eta(e)^n\cos n\vartheta)\\ \eta(e) &\equiv (1-(1-e^2)^{1/2})e^{-1} = \frac{1}{2}e + O(e^3) \end{aligned} \tag{11}$$

and to first order in e:

$$\begin{aligned} z &= a e^{i\omega t}(1-2e\sin\omega t)(1+2e\cos t) + O(e^2) =\\ &= a e^{i\omega t}(1+e(1+i)e^{i\omega t} + e(1-i)e^{-i\omega t} + O(e^2)) \end{aligned} \tag{12}$$

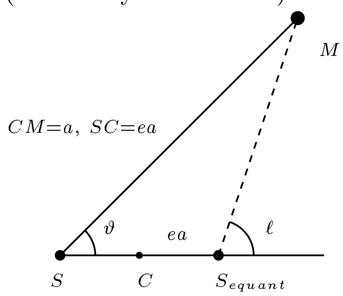
which can be described to lowest order in e, as composed by a deferent and two epicycles. Two more would be necessary to obtain an error of $O(e^3)$.

In this respect it is interesting to observe how one can arrive to an ellipse with focus on the Sun, by considering epicyclical motions. Indeed the simplest epicyclical motion is perhaps that in which one considers infinitely many pairs of epicycles, run with respective angular velocity $\pm n\omega$, with $n = 1, 2, \ldots$, and with radii decreasing in geometric progression, *i.e.* :

$$z(\vartheta) = p' e^{i\vartheta} \sum_{n=0}^{\infty} \eta'^n \cos n\vartheta \tag{13}$$

for some p', η', that leads to the ellipse in the first of the (11).

What is *less natural* in the Kepler laws, is that the time law which gives the motion on the ellipse, instead of $\vartheta \to \omega t$, is rather $\vartheta \to \omega t - 2e \sin \omega t + \ldots$. Such motion is however an old "Ptolemaic knowledge" being, at least at lowest order in e, a uniform *angular* motion around a point S_{equant} of abscissa $2ea$ from the point S with respect to which the anomaly ϑ is evaluated and of abscissa ea with respect to the center C of the circle on which the (manifestly nonuniform) motion takes place



e

Fig. e: The equant construction of Ptolemy adapted to a heliocentric theory of Mars; S is the Sun, M is Mars, C the center of the orbit and the equant point is S_{equant}.

this means that the angle ℓ in the drawing rotates uniformly and

$$\vartheta = \ell - 2e \sin \ell + e^2 \sin 2\ell + O(e^3) \tag{14}$$

Truncating the series in (13) and (14) to first order in the eccentricity we obtain (12) and hence a description in terms of one deferent, two epicycles and an equant: it is a description quite accurate of the motion of Mars with respect to the Fixed Stars Sky and it is the theory that one finds in the *Almagest*, after converting it to the inertial frame of reference fixed with the Sun.

The motion of the Earth around the Sun (or viceversa if one prefers) is similar except that the center of the deferent circle is directly the equant point, see[4] p.192, see also[20] Ch.2-4: this is usually quoted by saying the "for the Earth Ptolemy (Copernicus and Tycho) did not *bisect* the eccentricity", meaning that the center and the equant were identical and both $2\,e\,a$ away from the Sun: from[23] we deduce that this did not matter for the Earth which has a much smaller eccentricity (than Mars). Before discovering the ellipse Kepler had to redress this "anomaly" and he indeed bisected also the Earth eccentricity, see[23], making the Copernican Earth lose one more distinguishing feature with respect to the other planets.[22]

The above, however, *is not* the path followed by Kepler, see[23] where the latter is discussed in some detail.

Thus bringing the development in e to first order one reaches a level of approximation quite satisfactory for the observations to which Kepler had access, not only for the Sun but also for the more anomalous planets like Mercury, Moon and Mars: to second order however the equant becomes insufficient and Kepler realized that the ellipse had to be described at constant area velocity with respect to the focus.

We can say that *the experimental data agree within a third order error in the eccentricity with the hypothesis of an elliptical motion and with a time law based on the area law: this, within a second order error in the eccentricity, coincides with the Ptolemaic law of the equant.*

IV. MODERN TIMES

To realize better the originality of the Newtonian theory we must observe that in the approximations in which Kepler worked it was evident that the laws of Kepler were not absolutely valid: the precession of the lunar node, of the lunar perigee and of the Earth itself did require, to be explained, new epicycles: in a certain sense the Keplerian ellipses became "deferent" motions that, if run with the law of the areas, did permit us to avoid the use of equants and of other Ptolemaic "tricks". The theory of Newtonian gravitation follows after the abstraction made by Newton according to which the laws of Kepler, manifestly in contrast with certain elementary astronomical observations unless combined with suitable constructions of epicycles as Kepler himself realized and applied to the theory of the Moon,[21] were *rigorously exact* in the situation in which we *could neglect the perturbations due to the other planets*, *i.e.* if we consider the "two body problem" originally reinterpreting the Keplerian conception that the motion of a planet was due mostly to a force due to the Sun and partly to a force due to itself.

The theory of gravitation not only predicts that the motions of the heavenly bodies are quasi periodic, apparently even in the approximation in which one does not neglect the reciprocal interactions between the planets, but it *gives us the algorithms* for computing the functions $f(\varphi_1, \ldots, \varphi_n)$.

The *summa* of Laplace (1749-1827) on the *Mécanique celèste* of 1799,[25], makes us see how the description of the solar system motions, also taking account of the interactions between the planets, could be made in terms quasi periodic functions. The Newtonian mechanics allows us to compute approximately the $3N$ coordinates $\mathbf{A} = (A_1, \ldots, A_{3N})$ and the $3N$ angles $\varphi_1, \ldots, \varphi_{3N}$ and the $3N$ angular velocities $\omega_1(\mathbf{A}), \ldots, \omega_{3N}(\mathbf{A})$ in terms of which the motion simply consists of $3N$ uniform rotations of the $3N$ angles while the $\mathbf{A}$ remain constant.

Laplace makes us see that there is an algorithm that allows us to compute the A_i, ω_i, φ_i by successive approximations in a series of powers in several parameters (ratios of masses of heavenly bodies, eccentricities, ratios of the planets radii to their orbits radii *etc.*), that will be denoted here with the only symbol ε, for simplicity.

After Laplace approximately 80 years elapse during which the technique and the algorithms for the construction of the *heavenly* series are developed and refined leading to the construction of the formal structure of analytic mechanics. And Poincaré hws able to see clearly the new phenomenon that marks the first true and definitive blow to the Greek conception of motion: with a simple proof, celebrated but somehow little known, he showed that the algorithms that had obtained so many successes in the astronomy of the 1800's were in general *nonconvergent* algorithms,[26].

Few did realize the depth and the revolutionary character of Poincaré's discovery: among them Fermi that tried of deduce from the method of Poincaré the proof of the ergodicity of the motions of Hamiltonian dynamical systems that were not too special. The proof of Fermi, very instructive and witty although strictly speaking not conclusive from a physical viewpoint, remained one of the few attempts made in the first sixty years of the 1900's by theoretical physicists, to understand the importance of of Poincaré's theorem.

Fermi himself, at the end of his history, came back on the subject to reach conclusions very different from the ones of his youth (with the celebrated numerical experiment of Fermi-Pasta-Ulam).[28]

And the Greek conception of motion finds one of its last (quite improbable, *a priori*) "advocates" in Landau that, still in the 1950's, proposes it as a base of his theory of turbulence in fluids.[29] His conception that has been criticized by Ruelle and Takens (and apparently by others)[30,31] on the base of the ideas that, at the root, went back to Poincaré.

The alternative proposed by them began the modern research on the theory of the development of the turbulence and the renewed attempts at the theory of developed turbulence.

The attitude was quite different among the mathematicians who, with Birkhoff, Hopf, Siegel in particular, started from Poincaré to begin the construction of the corpus that is today called the *theory of chaos*.

But only around the middle of the 1950's it has been possible to solve the paradox consisting in the dichotomy generated by Poincaré:

(1) on the one hand the successes of classical astronomy based on Newtonian mechanics and the perturbation theory of Laplace, Lagrange, *etc.* seemed to confirm the validity of the quasi periodic conception of motions (recall for instance Laplace's theory of the World, or Gauss' "rediscovery" of Ceres,[35] and the discovery of Neptune,[32]).

(2) on the other hand the theorem of Poincaré excluded the convergence of the series used in (1)).

The fundamental new contribution came from Kolmogorov,[33,34]: he stressed the existence of two ways of performing perturbation theory,[35]. In the first way, the classical one, one fixes the initial data and lets them evolve with the equations of motion. Such equations, in all applications, depend by several small parameters (ratios of masses, *etc.*) denoted above generically by ε. And for $\varepsilon = 0$ the equations can be solved exactly and explicitly, because they reduce to a Newtonian problem of two bodies or, in not *heavenly* problems, to other *integrable* systems. One then tries to show that the perturbed motion, with $\varepsilon \neq 0$, is still quasi periodic, simply by trying to compute the periodic functions f that should represent the motion with the given initial data (and the corresponding phases φ_i, angular velocities ω_i, and the constants of motion A_i) by means of power series in ε. Such series, however, do not converge or sometimes even contain divergent terms, deprived of meaning, see[34] Sec. 5.10.

A second approach consists in fixing, instead of given initial data (note that it is in any case illusory to imagine knowing them exactly), the angular velocities (or *frequencies*) $\omega_1, \ldots, \omega_n$ of the quasi periodic motions that one wants to find. Then it is often possible to construct by means of power series in ε the functions f and the variables $\mathbf{A}, \varphi$, in terms of which one can represent quasi periodic motions, with the prefixed frequencies.

In other words, and making an example, we ask the possible question: given the system Sun, Earth, Jupiter and imagining for simplicity the Sun fixed and Jupiter on a Keplerian orbit around it, is it or not possible that *in eternity* (or also only up to just a few billion years) the Earth evolves with a period of rotation around to Sun of about 1 year, of revolution around its axis of about 1 day, of precession around the heavenly poles of about 25.500 years, *etc.*?

One shall remark that this second type of question is much more similar to the ones that the Greek astronomers asked themselves when trying to deduce from the periods of the several motions that animated a heavenly body the equations of the corresponding quasi periodic motion.

The answer of Kolmogorov is that if $\omega_1, \ldots, \omega_n$ are the n angular velocities of the motion of which we investigate the existence it will happen that for the *most part* of the choices of the ω_i there actually exists a quasi periodic motion with such frequencies and its equations can be constructed by means of a power series in ε, *convergent* for ε small.[33]

The set of the initial data that generate quasi periodic motions has a complement of measure that tends to zero as $\varepsilon \to 0$, in every bounded part of the phase space contained in a region in which the unperturbed motions are already quasi periodic.

One cannot say, therefore, whether a preassigned initial datum actually undergoes a quasi periodic motion, but one can say that near it there are initial data that generate quasi periodic motions. And the closer the smaller ε is.

By the theorem of continuity of solutions of equations of motion with respect to variations of initial data it follows that every motion can be simulated, for a long time as $\varepsilon \to 0$, by a quasi periodic motion.

But obviously there remains the problem:

(1) are there, really, initial data which follow motions that, in the long run, reveal themselves to be not quasi periodic?

(2) if yes, is it possible that in the long run the motion of a system differs substantially from that of the (abundant) quasi periodic motions that develop starting with initial data near it?

The answer to these questions is affirmative: in many systems motions that are not quasi periodic do exist and become easily visible as ε increases. Since electronic computers became easily accessible it is easy for everybody to observe personally on computer screens the very complex drawings generated by such motions (as *seen* by Poincaré, Birkhoff, Hopf *etc.*, *without* using a computer).

Furthermore quasi periodic motions although being, at least for ε small, very common and almost dense in phase space probably do not constitute an obstacle to the fact that the not quasi periodic motions evolve very far, in the long run, from the points visited by the quasi periodic motions to which the initial data were close. This is the phenomenon of *Arnold's diffusion* of which there exist quite a few examples: it is a phenomenon of wide interest. For example, if in the theory of the solar system diffusion was possible it would be conceivable the occurrence of important variations of quantities such as the radii of the orbits of the planets, with obvious (dramatic) consequences on the stability of the solar system.

In this last question the true problem is the evaluation of the time scale on which the diffusion in phase space could be observable. In systems simpler than the solar system (to which, strictly speaking, Kolmogorov's theorem does not directly apply, for some reasons that we shall not attempt to analyze here,[34] Sec. 5.10) one thinks that a sudden transition, as the intensity ε of the perturbation increases, is possible from a regime in which the diffusion times are super astronomical in correspondence of the interesting values of the parameters (*i.e.* times of several orders of magnitude larger than the age of the Universe) to a regime in which such times become so short to be observable on human scales. This is one of the central themes of the present day research on the subject,[36].

---

[1] This is, in part, a translation of the text of a conference at the University of Roma given around 1989, circulated in the form of a preprint since. The Italian text, not intended for publication, has circulated widely and I still receive requests of copies. I decided to translate it into English and make it available more widely also because I finally went into more detail in the part about Kepler, that I considered quite superficial as presented in the original text. Therefore this preprint differs from the previous, see[2], mainly (but not only) for the long new part (in footnote[23]) about *Astronomia nova*, which might be of independent interest.

[2] The original Italian text can be found and freely downloaded from `http://ipparco.roma1.infn.it`, at the page "$\leq$ 1994".

[3] A general history of astronomy is in: J. Dreyer,: *A history of astronomy*, Dover, 1953.

[4] For a simple introduction to the Ptolemaic system see: Neugebauer, O.: *The exact sciences in antiquity*, Dover, 1969. The figures of the text are taken from this volume: see p. 193–197.

[5] A critical and commented version of Ptolemy's theory, both of the *Almagest* and of the *Planetary Hypothesis*, is in: Neugebauer, O.: *A history of ancient mathematical astronomy*, part 2, Springer–Verlag, 1975.

[6] A recent edition of the *Almagest*, with comment, is: *Ptolemy's Almagest*, edited by G. Toomer, Springer Verlag, 1984.

[7] Some critiques to Ptolemy are in: R. Newton: *The crime of Claudius Ptolemy*, John's Hopkins Univ. Press, 1979.

[8] *i.e.* no linear combination of them with rational coefficients can vanish unless all coefficients vanish.

[9] Recent is the volume on Copernicus: O. Neugebauer, N. Swerdlow: *Mathematical astronomy in Copernicus' de revolutionibus*, vol. 1,2, Springer Verlag, 1984.

[10] The interpretation that the Fourier transform (and (9)) has in terms of deferent and epicyclical motions has been noted by many; the more ancient that I could retrieve is in a memory of 1874 of G. Schiaparelli, reprinted in G. Schiaparelli, *Scritti sulla storia dell' astronomia antica*, part I, tomo II, *Le sfere omocentriche di Eudosso, di Callippo e di Aristotele*, p. 11, Zanichelli, Bologna, 1926. It should be stressed that the above "reduction" of a quasi periodic motion to an epicyclic series is *not unique* and other paths can be followed: this will be very clear by the examples below[23].

[11] The work of Copernicus and Newton can be easily found in English as there are plenty of reprints; in Italian I quote the collection printed by UTET directed by L. Geymonat: N. Copernicus, *Opere*, ed. F. Barone, UTET, Torino, 1979. We find here in particular the so called *Commentariolus* that presents the plan of the Copernican work, as optimistically viewed by the young (and perhaps still naive) Copernicus himself before he really confronted himself with a work of the dimensions of the *Almagest*. Great were the difficulties that he then met, while dedicating the rest of his life to a complete realization of the program sketched in the *Commentariolus* ($\sim$1530), so that the *De revolutionibus* presents solutions quite more elaborate than those programmed in the quoted work. Nevertheless the copernican revolution appears already clearly from this brief and

illuminating work: here one finds the passage quoted in the text (p.108 of the Italian edition).

[12] And for a detailed treatment, also of the notion of average motion, see in particular (D. Boccaletti pointed out to me this bibliographic note, together with the preceding one[10]): S. Sternberg, *Celestial mechanics*, vol. 1, Benjamin, 1969.

[13] Arnold, V.I., Avez, A.: *Ergodic problems of classical mechanics*, Benjamin, New York, 1968.

[14] It is interesting, not only for the personality of Aristarchus of Samos, to read the volume: T. Heath: *Aristarchus of Samos. The ancient Copernicus*, Dover, 1981.

[15] Heath, T.L.: *Greek Astronomy*, Dover, 1991. This contains a very illuminating collection of translations of fragments from greek originals.

[16] Neugebauer assesses very lucidly Copernicus' contribution: see[4] p. 205.

[17] See Ptolemy, C.: *The Almagest*, ed. G.J. Toomer, Springer Verlag, New York, 1984. See also Theon, *Commentaires de Pappus et de Théon d'Alexandrie sur l'Almagèste*, T. II, ed. annotée par A. Rome, Biblioteca Apostolica Vaticana, Cittá del Vaticano, 1936. I often wonder whether it is possible that this passage has been contaminated by later commentators. Although certainly not much later, because it is already commented by Theon of Alexandria in the second half of the fourth century: however two centuries is a *very* long time for Science (if one thinks to what happened since Laplace). In a way this and the argument that follows it is much too rough compared to the level of the rest of the Almagest. Nevertheless if one attributes, as it seems right to do,[5] p. 900, the *Planetary Hypotheses* books to Ptolemy, then one is led to think that the passage is indeed original. This is perhaps also proved by the fact that Ptolemy does not seem to realize that the heliocentric hypothesis would have allowed a clear determination of the average radii of the orbits, missing in his work. In turn this makes us wonder which exactly was the famous heliocentric hypothesis of Aristarchus and if it went beyond a mere qualitative change of coordinates. Had it been the same as Copernicus' he could have determined the sizes of the orbits:[16] a problem in which he had a strong interest as he dedicated a book[14,15] to his determination of the distance of the Moon to the Earth, and this should have been reported by Ptolemy. From the extant information about Aristarchus' theory there is, strangely, no trace of an application of the heliocentric system to planets other than the Earth and the Sun, see[14] p. 299 and following, although it would be surprising that there was none. For a critical account of the *Planetary Hypotheses* see[5].

[18] Note that, from Newtonian mechanics and as discussed below, the motions of the 8 classical planets (the Fixed Stars Sky, Sun, Mercury, Venus, Moon, Mars, Jupiter, Saturn not counting the Earth (whose rigid motions are described by those of the Fixed Stars), or alternatively not counting the Sun and the Fixed Stars but regarding the Earth as having 6 degrees of freedom, requires a maximum of $24 = 3 \times 8$ independent "fundamental" frequencies namely three for each planet: so that both in Ptolemy and in Copernicus there must be epicycles rotating at speeds multiple of the fundamental frequencies or at least at speeds which are linear combinations with integer coefficients of the fundamental frequencies: hence the 43 frequencies cannot be rationally independent of each other; see for instance the figure on Copernicus' Moon and note that the "number of epicycles" in Ptolemy's theory could be counted differently.

[19] *The astronomical almanac for the year 1989*, issued by the National Almanac Office, US government printing office, 1989.

[20] Kepler, J.: *Astronomia nova*, reprinted french translation by J. Peyroux, Blanchard, Paris, 1979.

[21] Stephenson, B.: *Kepler's physical astronomy*, Princeton University Press, 1994.

[22] It is interesting to compare in detail the theory of Mars of Copernicus and that of Ptolemy (reduced to a heliocentric one). The first has a deferent of radius a on which a first epicycle of radius $\frac{3}{2}ea$ counter–rotates at equal speed and, *on it*, a second epicycle rotates at twice the speed; the starting configuration being the first epicycle at aphelion and the second opposite to the aphelion of the first. In other words the position z_C from the aphelion is, at average anomaly ℓ given by z_C

$$\begin{aligned} z_C &= ae^{i\ell} + \frac{3}{2}ae - \frac{1}{2}eae^{2i\ell} = \\ &= ea + ae^{i\ell}\,(1 - ie\sin\ell) \end{aligned}$$

Ptolemy has the planet on an eccentric circle, centered ea away from the Sun, whose center rotates at constant speed around the equant point which, in turn, is ea further away from the center of the orbit. Hence if ξ is the eccentric anomaly (*i.e.* the longitudinal position of the planet on the orbit as seen from the center) it is

$$z_T = ea + ae^{i\xi}$$

and from Fig. e we see that the relation between the average anomaly ℓ (*i.e.* the longitudinal position of the planet as seen from the equant point) is related to ξ by

$$\sin(\ell - \xi) = e\,\sin\ell \;\to\; \xi = \ell - e\sin\ell + O(e^3)$$

so that

$$z_T = ea + ae^{i\ell}(1 - ie\sin\ell - \frac{1}{2}e^2\sin^2\ell) + O(e^3)$$

and we see that the longitudinal difference (*i.e.* the difference of the true anomalies or longitudes from the Sun position S) is $\arg\frac{z_T}{z_C}$ or

$$\arg\Big(1 - \frac{1}{2}\frac{e^2\sin^2\ell}{e\,e^{-i\ell} + 1 - ie\sin\ell}\Big) = O(e^3\sin^3\ell)$$

However the difference in distance is $|z_C| - |z_T|$ of the order $O(\frac{1}{2}e^2\sin^2\ell)$ so that Copernicus' epicycles are equivalent to Ptolemy's equant within $O(e^3)$, or about $4'$, in longitude measurements and within $O(e^2)$, or about $40'$ in distance measurements (*i.e.* to match distances at quadrature, say, one should alter by about $40'$ the average anomaly, which means to delay the observations by about one day since Mars period is about 2 years (provided the distances could be measured accurately enough, which was not the case at Ptolemy's time).

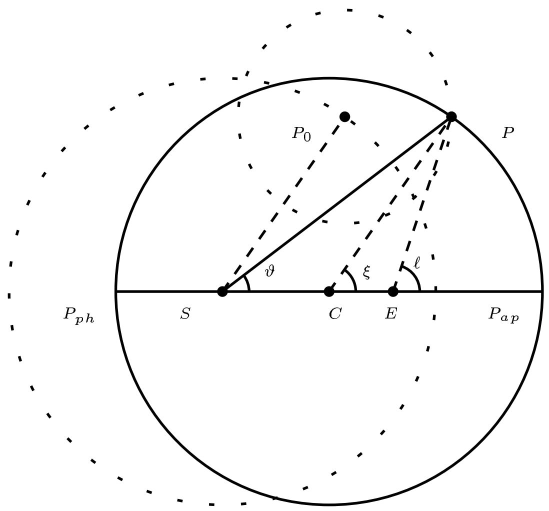
Before Tycho (relative) differences of $O(e^2)$ could not be appreciated experimentally: but their existence was derived from the theories, and used, by Kepler who discussed them at the beginning of his book in Ch.4, where the above calculation is performed for $\ell = \frac{\pi}{2}$, where the discrepancy is maximal, see[20] p. 23 (or p. 16 in the original edition).
Kepler's theory differs from both to order $O(e^2)$. He first derived a better theory for the longitudinal observations (which turned out eventually to agree with the complete theory already to $O(e^3)$) and used it to find the "correct" theory agreeing within $O(e^3)$ with the data for the distance measurements (that had become possible, see[16], after Copernicus).

[23] An account of Kepler's approach, in[20], is indeed as follows, see[21] for a very careful and detailed analysis of Kepler's discoveries from where I derive most of what follows. A key point to keep in mind in this footnote is that the resolution of the observations available to Ptolemy (and Copernicus) was of the order of $10'$ so that errors were in the order of tens of primes: this meant that one could observe first order corrections in the eccentricity of Mars but the second order corrections (of order $e^2 \simeq 10^{-2}$ or about $30'$) were barely non observable (the ensuing difficulties in interpreting the data earned Mars the name of *inobservabile sidus*, unobservable star, after Plinius). However the observations of Tycho were of the order of a few primes so that second order corrections were clearly observable, see[3] p. 385, because the third order amounts to about $3'$.
Another major point to keep in mind is, as clearly stressed in[21], that Kepler was the first to have (perhaps since Greek times) a *physical theory* to check: his language is not the one we have become accustomed to after Newton, but he had very precise laws in mind which he kept following very faithfully until the end of his work. The main one, for our purposes, was the (vituperated) law that the "speed of the motion due to the Sun is inversely proportional to the distance to the Sun", see below.
After ascertaining that the Earth and Mars orbits lie on planes through the Sun (rather than through the mean Sun, as in Copernicus and Tycho) he tried to "imitate the ancients" assigning to Mars an orbit, on an eccentric circle that I will call the *deferent*, and an equant: he noted that a very good approximation of the longitudes followed if one *abandoned* Ptolemy's theory of the center C of the orbit being half way between the Sun S and the equant E. His equant was set, to save the phenomena, a little closer to the center (with respect to the Ptolemaic equant point) at distance $e'a$ from the center C of the deferent rather than at distance ea. If z denotes the position with respect to the center C in the plane of the orbit with x–axis along the apsidal line of Mars and ζ denotes the position with respect to the Sun S (eccentric by ea away from C) and if ξ denotes the position on the deferent of the planet, called the *eccentric anomaly*, and if ω is the angular velocity with respect to the equant point E this means (using the complex numbers notation of Sec. 1, Eq. 4 with x–axis along the apsides line perihelion–aphelion)

$$z = ae^{i\xi}, \ \xi = \ell - \arcsin e' \sin \ell = \ell - e' \sin \ell + O(e'^3),$$

$$\ell = \omega t$$
$$\zeta = ea + ae^{i\xi}, \ |\zeta| = a\,((\varepsilon + \cos\xi)^2 + \sin^2\xi)^{\frac{1}{2}} =$$
$$= a\,(1 + e^2 + 2e\cos\xi)^{\frac{1}{2}}$$

which was called the *vicarious hypothesis*, illustrated in Fig. k1



k1

Fig. k1 vicarious hypothesis: here the eccentricities are $e = \frac{5}{4}0.4$ and $e' = \frac{3}{4}0.4$, $\sim$ 4 times larger than real to make a clearer picture. The dashed circles are the deferent (centered at S) and the epicycle (centered at P_0) while the planet is in P. The continuous circle is the actual orbit. The segment SP_0 is parallel to CP (partially drawn) so that $PCP_{ap} = P_0SP_{ap}$ is the eccentric anomaly while PSP_{ap} is the true anomaly and PEP_{ap} is the average anomaly, that rotates uniformly around the equant E.

The hypothesis illustrated above can be compared with the actual (later) Kepler law

$$z = a\,(\cos\xi + i\sqrt{1-e^2}\sin\xi), \ \ell = \xi + e\sin\xi,$$
$$\xi = \ell - e\sin\ell + O(e^3), \ \ell = \omega t$$

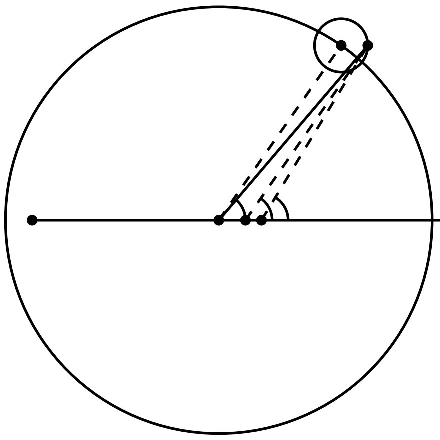
so we see that the vicarious hypothesis gave an incorrect distance Mars–Sun (because of the difference of order $O(e-e') + O(e^3) = O(e^2)$ as $e - e' \simeq 2e^2$ from the data for e'): for instance the y–coordinate was about $\frac{1}{2}e^2$ higher than it should have, see[21] p. 46, at $\xi = \frac{\pi}{2}$.
Nevertheless the vicarious hypothesis gave very accurate longitudes. Therefore the hypothesis was used by Kepler as a quick means to compute the longitudes until the very end of the research. The hypothesis was discarded because, after Copernicus and Tycho, it had become possible to compute distances from the Sun and they were incorrectly predicted by the hypothesis as the second order corrections in the eccentricity were already visible (in the case of Mars).
It is interesting to remark that *a posteriori* it is clear why the vicarious hypothesis worked so well. The relation between the longitude ϑ corresponding to a theory in which the Sun is eccentric by ea from the center of the orbit and the equant is $e'a$ further away gives a longitude ϑ as

$$\vartheta = \ell - (e + e')\sin\ell + \frac{1}{2}(ee' + e^2)\sin 2\ell + O(e^3)$$

while the final theory gives ϑ in terms of the ellipse eccentricity $\bar{e}$ as

$$\vartheta = \ell - 2\bar{e}\sin\ell + \frac{5}{4}\bar{e}^2\sin 2\ell + O(\bar{e}^3)$$

therefore a "non bisected" eccentricity (*i.e.* $e \neq e'$ with $\frac{1}{2}(e+e') = \bar{e}$ and $e = \frac{5}{4}\bar{e}$, $e = \frac{5}{4}\bar{e}, e' = \frac{3}{4}\bar{e}$) gives an agreement to order $e^3 \simeq 10^{-3}$: this is a few primes, well out of observability. As Kepler noted Tycho's and his observations gave $e + e' = 2\bar{e} = 0.18564$ and $e = 0.11332$ very close to $\frac{5}{4}\bar{e} = 0,11602$, see[21] p. 44. A realistic drawing of the vicarious hypothesis would be illustrated by Fig. k2



k2

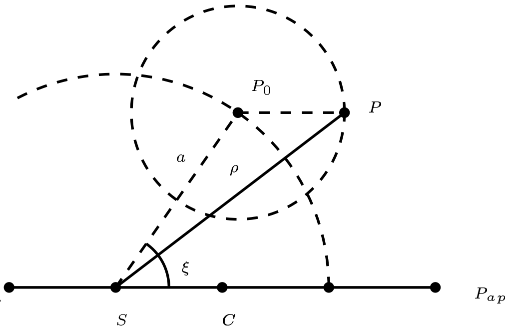
Fig. k2: same as Fig. k1 with eccentricity $e = 0.1$; only the deferent and the epicycle are drawn in *Fig. k2*

Then Kepler was assailed by the suspicion that the discrepancies in the distances that he was finding were rather due to a defective theory of the Earth motion (which was needed to convert terrestrial observation into solar ones): he was thus led to realize that the Earth too had an equant and he could check that the problem was not with an erroneous theory of the Earth motion: introducing an equant for the Earth did not affect sensibly the data for the Sun and Mars (this meant a large amount of checking, a year or so).

Returning to Mars he tried to check (again) his basic hypothesis that the velocity was inversely proportional to the distance from the Sun; he assumed that what proceeded at velocity inversely proportional to the distance from the Sun was the *center* of an epicycle of radius ea; the epicycle center had anomaly ξ on a circle of radius a around S and on it the planet was rotating at rate $-\dot{\xi}$: so that the planet was actually moving on a circle of radius a around the eccentric center C with angular velocity $\dot{\xi}$ too. The value of $\dot{\xi}$ was determined by the Kepler original law $\rho\dot{\xi} = const$, see[21] p. 101. The planet equations would be:

$$z = ae^{i\xi}, \zeta = ea + ae^{i\xi},\ \ \rho = |\zeta| = a\,(1 + 2e\cos\xi + e^2)^{\frac{1}{2}}$$

plus the dynamical law $\rho\dot{\xi} = const$: this gives both anomaly and distance differing from the vicarious hypothesis by $O(e^2)$: hence still incompatible with the observations (actually worse than the vicarious hypothesis which at least gave correctly the longitudes).



k3

Fig. k3: The first attempt to establish the law $\rho\dot{\xi} = const$. The eccentricity ($e = 0.4$) is much larger than the real value, for purposes of illustration. The deferent is a circle of radius a around S and the epicycle of radius $e\,a$ is centered at P_0, with anomaly ξ with respect to S, the planet is in P.

The successful (among others) attempt was driven by the remark that one needed to lower the y–coordinate at quadrature by $\frac{1}{2}e^2a$ (*i.e.* to eliminate a "lunula" between the vicarious hypothesis orbit and the observed orbit). In fact Kepler discovers, by chance as he reports, that the observed distance of Mars from the center of the deferent is b, *shorter* than a and precisely such that $\frac{a}{b} = \sec\vartheta$ if $\sin\vartheta = e$, or: $b = a\sqrt{1-e^2}$.

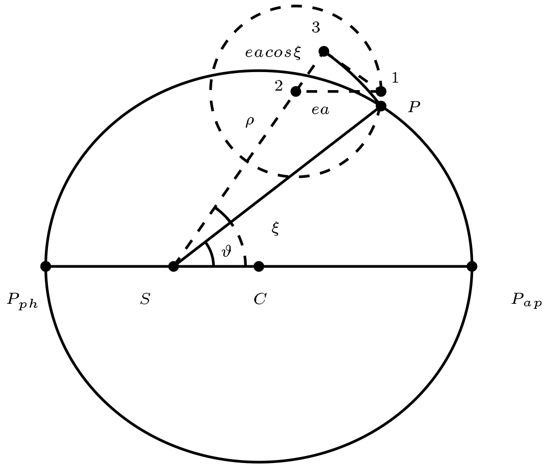
So one sees that the distance from S is at aphelion or perihelion $a\,(1 \pm e)$ from S while at quadrature on the deferent (*i.e.* at eccentric anomaly $\xi = \frac{\pi}{2}, \frac{3\pi}{2}$) it is just a, or the distance to the center C is a in the first cases and $b = a\sqrt{1-e^2} = a - \eta\,a$ with $\eta = 1 - \sqrt{1-e^2}$in the second cases.

"Therefore as if awakening from sleep" it follows that at a position with eccentric anomaly ξ the planet y coordinate is lower (in the direction orthogonal to the apsides line) by $a\,\eta\,\sin\xi$ while the x–coordinate is still $a\,\cos\xi$, see[21] p. 125: hence the orbit is an ellipse (with axes a and $b = a\sqrt{1-e^2}$ and the distance to the Sun is then easily computed to be $\rho = a\,(1 + e\cos\xi)$. Indeed the coordinates of the planet with respect to the Sun become $x = e\,a + a\,\cos\xi$, $y = a\sqrt{1-e^2}\sin\xi$ (rather than the previous $y = a\,\sin\xi$) so that $\rho = \sqrt{x^2+y^2} = a\,(1 + e\,\cos\xi)$.

Combining this with the basic dynamical law $\rho\,\dot{\xi} = const$ we deduce that the motion is over an ellipse run at constant area velocity around S (not C), as one readily checks, (see below).

Kepler's interpretation of the above relation was in the context of his attempt at a description of the motion "imitating the ancients". The eccentric anomaly ξ defines the center P' of an epicycle on a deferent circle centered at S (not at C) and with radius ea on which the planet *should* have traveled an angle $-\xi$ away from the $P'S$ axis; but in fact the actual position P was really closer to the apsides line by the (*small*) amount $a(1 - \sqrt{1-e^2})\cos\xi$ so that the distance PS was $\rho = a\,(1 + e\cos\xi)$ (as one readily checks). The law $\rho\,\dot{\xi} = const$ was quite natural as he attributed the motion around the Sun as partly due to the Sun and partly to the planet: the latter (somewhat obscurely, perhaps) was responsible for the epicyclic excursion so that

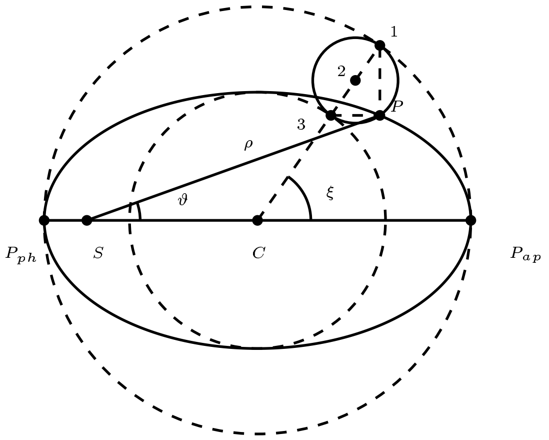
$\rho\,\dot\xi$ would be the correct variation of the eccentric anomaly which had, therefore, a physical meaning (this *is not* the common interpretation of Kepler[3,21]).



k4

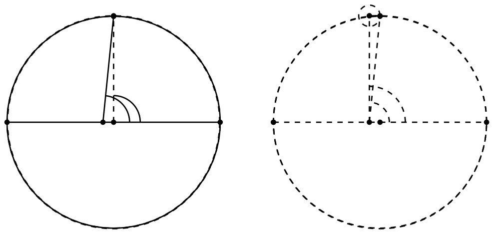
Fig. k4: Kepler's construction of the ellipse. The eccentricity is $e = 0.4$ to make a clearer drawing. The planet is in P. The segment 23 is the projection on $S3$ of the segment 31 and the segment SP is equal to $S3$. The epicycle (dashed circle) is no longer holding the planet but it determines its position. The anomaly ξ is the anomaly of the center of the epicycle. The motion verifies the law $\rho\,\dot\xi = const.$

An equivalent formulation is that the deferent is a circle centered at the center of the orbit C with radius $\frac{1}{2}(a+b)$: the planet is not on the deferent but on a (*very small*) epicycle of diameter $a-b$. If the center of the epicycle is at eccentric anomaly ξ the planet on the epicycle has traveled retrograde by 2ξ with respect to the radius from C to the epicycle center. This is described by the following figure.



k5

Fig. k5: Fabricius' interpretation, see[3] p. 402-403, of Kepler's ellipse. The ellipse has very high eccentricity ($e =$ 0.8!) to make visible the small epicycle. The planet is in P and $C\,3 = b$, $C\,1 = a$ and the center of the epicyle is 2. The eccentic anomaly is reckoned at C and the angle $1\,2\,P$ is 2ξ. Note that in Fabricius' construction the closeness of the ellipse to the deferent circle is *very small* as the epicycle has radius $O(e^2)a$; unlike what happens in the Kepler construction in which the deferent is centered at S and therefore the epicycle is much larger having a radius ea. The Fabricius' construction (done with $e = 0.1$, approximately the Mars eccentricity) makes us appreciate how subtle and refined had to be the analysis of Kepler to detect and understand the ellipse and the areal law. If one d raws on the scale of this page the Fabricius' picture and Kepler picture at quadrature, for insance, one get (of course) the same ellipse but one can see Kepler's epicycle while the Fabricius' one is not visible on this scale. The following figure *Fig. k6* (representing at quadrature Fabricius', left, and Kepler's (right) ellipses, circumscribed circle (drawn but not distinguishable on this scale) and epicycles (both drawn but only Kepler's being visible) is a quite eloquent illustration.



k6

Fig. k6: Fabricius' (left) and Kepler's (right) deferents, epicycles and ellipses with eccentricity $e = 0.1$. On this scale one does not appreciate the difference between ellipse and the deferent circles, nor the epicycle in the first drawing. The solid line and the dashed lines (indistinguishable in the scale of the picture) are respectively the ellipse and the deferent.

In other words the epicycle center moves on the deferent at speed $\dot\xi$ while the planet moves on the epicycle at speed $-2\dot\xi$. This very Ptolemaic interpretation of Kepler's ellipse, formally pointed out to Kepler by Fabricius, see[3] p. 402-403 (but it is simply impossible that Kepler had not known it immediately), is "ruined" by the very original Kepler law $\rho\,\dot\xi = const$ which implies that $\dot\xi$ is *not* constant.

No matter how audacious the last conceptual jump may look it is absolutely right and it identified the ellipse and the area law, as Kepler proved immediately, regarding visibly this fact as a final proof of his hypothesis that the area law and the inverse proportionality of the speed to the distance were absolutely correct and in fact identical. The law $\rho\,\dot\xi = const$, *i.e.* the area law, is completely out of the Copernican views and it recalls to mind the mysterious Ptolemaic lunar constructions for which we have apparently no clue on how they were derived.

That $\rho\,\dot\xi$ *is the area law* is worth noting explicitly as it is little remarked in the elementary discussions of the two body problem. If ϑ denotes the *true anomaly*, ℓ the *average anomaly* and ξ the *eccentric anomaly* then the equation of an ellipse in polar coordinates (ρ,ϑ) with center at the attraction center S can be written $\rho = p/(1-e\cos\vartheta) = a\,(1+e\cos\xi)$ with a the major semiaxis and $p = a\,(1-e^2)$. It was possibly well known, since Apollonius (?), that in such coordinates the area spanned by the radius ρ per unit time is $\frac{1}{2}\,p\,\rho\,\dot\xi$ or, as Kepler infers from his "physical conception" that "velocity [on the deferent] is inversely proportional to the distance from the Sun". This statement that is, often, interpreted as an error made by Kepler, see[3] p.388, pos-

sibly confusing the speed on the deferent with the speed around the center which is $\rho\,\dot\vartheta$: the relation $\rho^2\dot\vartheta=p\,\rho\,\dot\xi$ may explain why Kepler did not see the two laws in conflict. Although some comments on this would have helped a lot the readers, it seems unlikely that he did not notice that $\rho\,\dot\xi$ is not proportional to the area velocity unless the motion is on an ellipse with eccentric anomaly ξ, particularly after Fabricius' comments on the Ptolemaic version of the ellipse. And an error on the part of Kepler in measuring the areas is *obviously* excluded from what he writes: see[20] p. 248–251 (or p. 193–196 of the original edition).
Other interpretations of ξ are incompatible with the area law to first order in the eccentricity. Since the eccentricity of Mars is "large" and the measurements of Tycho–Brahe allowed us even to see corrections to the distances of second order in the eccentricity it was possible to realize that $\dot\xi$ was indeed inversely proportional to ρ so that $\ell=\xi+e\sin\xi$ followed. And the natural assumption that all planets (but the too close Moon and perhaps the eccentric Mercury) verified the same laws was easily checked (by Kepler) to be fully consistent with the data known at the time for the major planets.
We see that although the above Kepler approach is very original compared to Ptolemy's and Copernicus' the conclusion in[3] p.393 that "the discovery of the elliptic orbit of Mars was an absolutely new departure, as the principle of circular motion had been abandoned..." seems, after examining the methods and the ideas followed in *Astronomia nova*, a hasty conclusion to say the least. It is not even clear that Kepler himself thought so.

[24] Newton, I.: *Principia mathematica*, UTET, Torino, 1962.

[25] Recent is the reprint (economic) of the classical English translation of E. Bowditch: P. de la Place: *Celestial Mechanics*, ed. E. Bowditch, Chelsea, 1966.

[26] Poincaré, H.: *Les méthodes nouvelles de the mècanique celèste*, recenly reprinted by Blanchard, Paris, 1987.

[27] Fermi, E.: *Beweis dass ein mechanisches normalsysteme im algemeinen quasi–ergodisch ist*, Physikalische Zeitschrift, **24**, 261–265, 1923, reprinted in E., Fermi, *Collected papers*, vol I, p. 79–84, Accademia dei Lincei and Chicago University Press 1962.

[28] Fermi, E., Pasta, J., Ulam, S.: *Studies of nonlinear problems*, Los Alamos report LA-1940, 1955, reprinted in E., Fermi, *Collected papers*, vol. II, Accademia dei Lincei and Chicago University Press, 1965.

[29] Landau, L., Lifchitz, E.: *Mécanique des fluides*, MIR, Mosca, 1971.

[30] Ruelle, D., Takens, F.: *On the nature of turbulence*, Communications in mathematical Physics, **20**, 167, 1971.

[31] Ruelle, D., Takens, F.: *Note concerning our paper "On the nature of turbulence"*, Communications in Mathematical Physics, **23**, 343–344, 1971.

[32] A history of some important applications of the perturbation theory to the motion of the main planets, successive to the work of Laplace, is: M. Grosser, *The discovery of Neptune*, Dover 1979.

[33] Kolmogorov, A.N.: *Preservation of conditionally periodic movements with small change in the Hamilton function*; reprinted in: Casati, G., Ford, J.: *Stochastic behavior in classical and quantum systems*, Lecture Notes in Physics, vol. 93, Springer–Verlag, 1979.

[34] For Kolmogorov's theorem see: G. Gallavotti: *Elementary mechanics* , Springer Verlag, 1984.

[35] A classical work on the theory of unperturbed Keplerian motions is the book of Gauss, useful to whoever wishes to realize the size of even the simplest astronomical computation and thus desires to appreciate the greatness of the Greeks astronomers' work: K. Gauss, *Theory of the motion of heavenly bodies moving about the Sun in conic sections*, Dover 1971. See also the appendix Q in the second Italian edition of[34], *Meccanica elementare*, Boringhieri, Torino, 1985.

[36] A collection of modern but classical papers on the theory of quasi periodic and chaotic motions can be found in: R. McKay, J. Meiss, *Hamiltonian dynamical systems*, Hilger, 1987.

> *I know that I am mortal and the creature of a day; but when I search out the massed wheeling circles of the stars, my feet no longer touch the Earth, but, side by side with Zeus himself, I take my fill of ambrosia, the food of Gods*: quotation of Ptolemy borrowed from[15], p. lvii.

(slightly corrupted) REVTEX

Giovanni Gallavotti
Dipartimento di Fisica
Universitá di Roma "La Sapienza"
P.le Moro 2, 00185 Roma, Italy

email:
giovanni@ipparco.roma1.infn.it
tel. +39-06-49914370,
fax. +39-06-4957697
web: http://ipparco.roma1.infn.it