On a mechanism that structures galaxies

D. Lynden-Bell
Institute of Astronomy, The Observatories, Cambridge CB3 0HA

Received 1978 September 20; in original form May 9

Summary. By considering the interaction of a single stellar orbit with a weak \( \cos 2\phi \) potential it is shown that in the central regions of galaxies with slowly rising rotation curves, the elongations of the orbits will align along any potential valley and oscillate about it. This effect is more pronounced for elongated orbits. In such regions any pair of orbits will naturally align under their mutual gravity and so a bar will form. The gravity of this bar will drive a spiral structure in the outer parts of the galaxy where differential rotation is too strong to allow the orbits to be caught by the bar. The spiral structure carries a torque which slowly drains angular momentum from the bar, gradually making its outline more eccentric and slowing its pattern speed. In the outer parts of the bar only the more eccentric orbits align with the potential valley; the rounder ones form a ring or lens about the bar. As the pattern speed slows down, the corotation resonance and outer Lindblad resonance, which receive the angular momentum, move outwards. The evolution of the system is eventually slowed down by the weakness of these outer resonances where the material is rather sparse.

1 Introduction

Most authors who have tackled the problem of eccentric orbits resonating with the spiral structure near the inner Lindblad resonance have come to the conclusion that such orbits will anti-align with the valley in the spiral potential thus tending to eliminate it [15].* However, James & Sellwood [6] and Huntley, Sanders & Roberts [7] have recently found barred forms growing in computer experiments on galaxies which have a double inner Lindblad resonance. This has prompted me to reinvestigate the behaviour of orbits near this resonance and this has exposed the importance of the former assumption that the resonant frequency was decreasing as the angular momentum of the orbit increases. While this assumption is in general correct it is not always true, and is often incorrect in the central regions of galaxies where the velocity curve is rising. This is where bars and lenses are seen to be. These regions,

*Contopoulos & Mertzanides [15] found an exception when the amplitude of the spiral wave increased very strongly inwards; then the opposite occurs.
which normally have double (or even quadruple) inner Lindblad resonances, are motors that drive both bars and spiral structure in galaxies. The structures of many galaxies fall into place if it is assumed that this is the driving motor.

In Section 2 we give a brief account of the treatment of resonant orbits described in more mathematical detail in my Saas Fee lectures [2]; the only difference is that we here envisage that the quantity called \( s \) in those notes may sometimes be negative. In order to describe strongly elliptical orbits in the plane of a galaxy it is necessary to have an axially symmetrical model potential in which all the unperturbed orbits may be explicitly calculated. For our model we use the Isochrone [8–10], \( \psi = GM(b + \sqrt{R^2 + b^2})^{-1} \), whose pertinent properties are illustrated in Figs 1 and 3. Having demonstrated those regions of the galaxy in which the elliptical orbits will align with a potential valley, we proceed to argue the importance of this result for galactic structure and galactic evolution Section 3.

2 Resonant orbit theory

In a central potential a typical orbit is a rosette with an angle between consecutive apocentres of between \( \pi \) and \( 2\pi \). If we view any one orbit from rotating axes, we can always choose the rotation of our axes \( \Omega_i \) so that with respect to them the angle between apocentres is \( \pi \). In those axes the orbit will be a bisymmetric figure like a centred ellipse. If \( \Omega \) is the mean angular velocity of the star about the galaxy in inertial axes, and \( \kappa \) is the radial angular frequency of its motion in and out, then the condition that the axes are chosen in this way is \( \Omega_i = \Omega - \frac{1}{2}\kappa \). In general each orbit will need a different \( \Omega_i \), but it was Lindblad’s idea that for near-circular orbits \( \Omega_i \) did not vary much over quite a large region of galaxy and that this allowed non-axially symmetrical structures made up of oval orbits to last a long time. This is indeed the case for the Isochrone; compare the weak variations of the \( \Omega - \kappa/2 \) curve of Fig. 1 with the rapid change of \( \Omega \). It is interesting that if we consider the same problem for less round orbits, then \( \Omega - \kappa/2 \) decreases as they become more oval. Thus one could artificially keep \( \Omega - \kappa/2 \) even more constant by populating those radii in which \( \Omega - \kappa/2 \) for near-circular orbits is larger than average, by orbits that are eccentric.

We now consider the interaction of an orbit with a weak bar-like potential trough \( \psi_2 \) which rotates with some pattern speed \( \Omega_p \) which is close to the \( \Omega_i \) of the orbit concerned.

![Figure 1. The angular velocity of the lobes of near-circular orbits, \( \Omega_i = \Omega - \kappa/2 \); the angular velocity of the stars around the orbits \( \Omega \), and the circular velocity \( V \) plotted as functions of radius for the isochrone model. Units have been chosen so that \( GM = 1 \) and \( b = 1 \).](https://academic.oup.com/mnras/article-abstract/187/1/101/1084199/fig1)
By this we mean
\[
\frac{\Omega_i - \Omega_p}{\Omega} \ll 1.
\]

Since both \(\Omega_i\) and \(\Omega_p\) are significantly less than \(\Omega\) over the central regions that concern us (see Fig. 1), this condition is satisfied by most orbits there. Taking axes rotating with \(\Omega_p\) we see the star pursuing its orbit at the rapid rate \(\Omega - \Omega_p\); but the star’s orbit is an almost closed oval — indeed we may regard it as a closed oval rotating at the slow rate \(\Omega_i - \Omega_p\). Since \(\Omega - \Omega_p\) is so rapid there is no time for a weak perturbing potential \(\psi_2\) to affect the star’s fast motion around the oval, which is therefore adiabatically invariant and

\[
\frac{1}{2\pi} \int p \, dq = 2J_f
\]

is approximately constant. Since the orbit has half a turn about the galactic centre for each radial oscillation in and out \(J_f = \frac{3}{2} h + J_R\) where \(J_R\) is

\[
\frac{1}{2\pi} \int R \, dR
\]

and \(h\) is the angular momentum. However, \(\psi_2\) will provide a persistent weak couple on the oval as a whole because it rotates only slowly. Hence the oval will change to another oval of the same \(J_f\) but different angular momentum \(h\). If the potential trough lies behind the elongation of the oval, see Fig. 2, then \(h\) will be decreased and as a result \(\Omega_i(J_f, h)\) will change. If \(\Omega_i\) decreases as \(h\) decreases (\(J_f\) held constant) then the motion of the elongation of the oval would be slowed down. In such circumstances ovals with unperturbed \(\Omega_i\) close to \(\Omega_p\) would be trapped to oscillate with their elongation along that of the bar-shaped potential. However, normally over most regions of a galaxy \(\Omega_i\) increases as \(h\) decreases, so the elongation of the orbit behaves like a donkey accelerating when held back and slowing down when urged forward. Such orbits can be trapped, but only at right angles to the valley in the potential. Thus in such a case the gravity of the trapped orbits will tend to cancel the potential that traps them and no big perturbations can build up.

![Figure 2](https://example.com/image2.png)

**Figure 2.** The torque \(T\) tries to align the lobes of the orbit with the potential trough, but in the normal region the orbits are donkeys accelerating forwards when tugged back and vice versa. Only in the abnormal regions do they slow down when pulled back. This leads to orbits librating about the potential trough. (In potentials that are more centrally condensed than the Isochrone the elliptic orbits in rotating axes turn back to make figures of eight at each end.)
Now let us return to the abnormal situation in which \( \Omega_l \) decreases as \( h \) decreases. In this case each orbit vibrates about the state in which it is aligned along the potential valley, so an assembly of many such orbits in which those of least energy are more populous will concentrate about the potential valley. The gravity of this concentration will add to the potential valley — trap yet more orbits and lead to a strong bar structure. To discover what regions of a galaxy will lead to a barring of near-resonant orbits we must calculate \( \Omega_l(J_t, h) \) for our equilibrium model and discover in what regions it behaves abnormally by decreasing as \( h \) decreases. It is quite simple to do this for the Isochrone model, for we have

\[
J_R = \frac{1}{2\pi} \oint \hat{\mathbf{R}} \, dR = \frac{1}{2\pi} \oint [2(E + \psi) - h^2 R^{-2}]^{1/2} \, dR
\]

where

\[
\psi = \frac{GM}{b + \sqrt{R^2 + b^2}}
\]

One can calculate this integral using \( \psi \) as a substitution variable in place of \( R \) and after tedious algebra one obtains

\[
J_R = GM(-2H)^{-1/2} - \frac{1}{2}[h + (h^2 + 4GMb)^{1/2}] \quad (1)
\]

where I have written \( H \) in place of \( E \) to emphasize that we could reorganize this equation to give us the Hamiltonian \( H(J_R, h) \) as a function of the action variables \( J_R \) and \( h \).

Now by the normal property of action variables the angular frequencies are given by

\[
\kappa = \left( \frac{\partial H}{\partial J_R} \right)_h, \quad \Omega = \left( \frac{\partial H}{\partial h} \right)_{J_R} \quad (2)
\]

Hence from equations (1) and (2) it is simple to calculate \( \Omega_l = \Omega - \kappa/2 \). For the Isochrone, Fig. 3 illustrates contours of constant \( \Omega_l \) in the \( J_t, h \) plane (remember \( J_t = J_R + \frac{1}{2}h \)). To the left of the dashed line we have the abnormal situation in which \( \Omega_l \) decreases as \( h \) decreases. For the Isochrone one may define an eccentricity \( e \) of the symmetrically centred oval orbit by

\[
\frac{1}{2}(1 - e^2/2)(1 - e^2)^{-1/2} = \frac{1}{2}J_R/h = J_t/h.
\]

This formula gives the correct eccentricity for the small elliptical orbits close to the centre where the Isochrone is almost simple harmonic. At all radii \( e = 0 \) are circles and \( e = 1 \) are straight lines or parabolas. (However, the \( e \) thus defined is not that appropriate to the eccentricity \( e^* \) of Kepler’s elliptical orbits about a focus that occur at large radii [10–11]. For them the appropriate one is \( e^* \), given by \( (1 - e^{*2})^{-1/2} = (1 + J_R/h) \).) Notice that if \( h \) decreases at constant \( J_t \) then \( e \) always increases. Thus the orbits become more oval as their angular momentum is decreased by a back-dragging torque.

Notice that for both Keplerian and \( V = \) constant velocity curves the abnormal region in which resonant orbits align hardly occurs. Thus it is associated with regions in which \( V \) is rising, although that is neither a necessary nor a sufficient condition.

3 Consequences of aligning orbits

Consider two orbits whose \( \Omega - \kappa/2 \) differ but little and let both be in the aligning regions in which \( \Omega - \kappa/2 \) decreases as \( h \) decreases (at constant \( J_t \)). Then, under the influence of the gravity of the stars in the other orbit each orbit will try to align with the other. In practice
they will oscillate about one another, but the net effect will be a gravity trough deeper than either would make on its own rotating with a pattern speed $\Omega_p$. This trough will be deep enough to catch orbits whose $\Omega - \kappa/2$ differs from $\Omega_p$ by slightly more, so more orbits join the bar and yet more will be caught by the increased gravity trough. This is I believe to be the detailed mechanism by which bars are made in galaxies. Once the bar is made, its perturbations on the outer parts can drive a spiral structure which carries a torque out to the corotation region and probably even onwards to outer Lindblad resonance. The constant torque on the bar causes it to lose angular momentum, but as those orbits lose angular momentum they become more eccentric as discussed at the end of the last section. Under the action of this torque $\Omega - \kappa/2$ of each orbit slows down, and hence the pattern speed of the bar slows down. This is so because the bar only forms in the abnormal region in which $\Omega - \kappa/2$ decreases as $h$ decreases.

We are thus led to a satisfying picture of barred spiral evolution in which the bar forms naturally and becomes steadily thinner and slower as the spiral structure drains angular momentum from it to be deposited at the outer resonances. The increased eccentricity of the bar ensures that it is all the more capable of driving the spiral structure as the process goes on. The length of the bar will not change very much.

It is notable that many galaxies with bars have central lenses or rings about the bars. The bar-like potential has very little grip on an almost circular orbit and near the ends of the visible bar it is the eccentric orbits from further in that make it, although these will be helped by untrapped near-circular orbits and Contopoulos’s mechanism, as explained in the Appendix.

This elegant picture comes with one drawback. The spiral structure of the normal spirals does not appear to be bar-driven, yet it looks not unlike the spiral structure of the outer
parts of barred galaxies. If that fact were taken as fundamental then one could alter the theory as follows: asymmetries in a galaxy moving inward with the group velocity generate the spiral structure in all galaxies, but in those with a gradually rising velocity curve the disturbance, instead of making a spiral, will catalyse a bar within the inner Lindblad resonance. Spiral structure in all galaxies would then be a reflection of the asymmetries found at large radii by Bosma [12]. At present I prefer the view that the bars are the motors for the well-organized structures seen in the barred spirals.

One prediction of this theory of bars which may be evaluated from Bosma’s work is that bars occur when the rotation curve rises slowly to its maximum. Only then is the abnormal aligning orbit region reasonably extended. In Bosma’s picture-compilation [12] all four of the SB galaxies have slowly rising velocity curves and the only other galaxies with fairly slow rises are the SAB galaxies and the small spirals M33 and NGC 4244.

It is interesting to remark that the little Lindblad ellipses at corotation and the large ring orbits at outer Lindblad resonances are both anti-aligning, so the orbits, like wild asses [4], will hover around those orientations corresponding to the mountains of the potential.

If we contemplate the possibility that the orbits at outer Lindblad resonances are so oriented and that they have been given much angular momentum by the spiral structure over their life, then we get an interesting prediction. For the evolution of these orbits the outer Lindblad adiabatic invariant \( \frac{1}{2} h - J_R \) has been conserved, yet \( h \) has steadily increased. Hence \( \frac{1}{2} - \frac{J_R}{h} \) has steadily decreased. Thus \( J_R / h \) has steadily changed towards the value \( \frac{1}{2} \). If the medium at outer Lindblad resonance were gaseous the orbits would get more eccentric until eventually one would cross itself in the rotating axes. That first occurs when the axial ratio of the ring is 3:2 with the long axis perpendicular to the bar.

Acknowledgments

I thank Drs Sancisi and Kormendy [16] for telling me of their interesting research on barred galaxies, and Drs James and Sanders for communicating their papers to me in advance of publication.

References

Appendix

There are many other phenomena on which this paper has not dwelt which must nevertheless be included in any complete theory of barred spirals, and at Professor Contopoulos's behest I discuss a few important ones below.

First, among these are the effects of nearly trapped and just trapped orbits of large libration. In the barring case their effects are easiest seen by analogy with the motion of a frictionless ball rolling in a channel over hills and valleys. Some balls have sufficient energy to be free; they speed up in the valley and contribute most density on the hills. So also do those that are just trapped — they oscillate rapidly across the valley that caught them. In our discussions above we have assumed that both of these are outweighed by the population of eccentric orbits that are well and truly caught in the bar. They contribute most where they always are — in the valley of the bar itself. Thus, in summary, when the well-caught orbits are barring, the weakly caught and uncaught orbits contribute more weakly in the opposite direction.

Secondly, in a normal or anti-barring region all the above effects are reversed so that the weakly trapped and untracked orbits actually help the bar even though well-trapped orbits then hinder it.

Thirdly, as Contopoulos has emphasized, the behaviour of the nearly circular orbits can be somewhat different for the non-axially symmetrical potential itself can then alter the shape of the orbit faster than the couple can slew the line of apsides. Due to this effect near-circular orbits with \( \Omega - \kappa/2 \) far from \( \Omega_p \) can yet be made exactly periodic. It turns out that these orbits can be regarded as a limiting case of the free untrapped orbits and their contribution to the baring potential therefore behaves in the same way. A theory of bars based around the dominance of these last effects has been advanced by Contopoulos and Merzanides [13–15].