Galactic spiral structure

BY CHARLES FRANCIS1* AND ERIK ANDERSON2

125 Elphinstone Road, Hastings, TN34 2EG, UK
2800 Morton Street, Ashland, OR 97520, USA

We describe the structure and composition of six major stellar streams in a population of 20,574 local stars in the New Hipparcos Reduction with known radial velocities. We find that, once fast moving stars are excluded, almost all stars belong to one of these streams. The results of our investigation have led us to re-examine the hydrogen maps of the Milky Way, from which we identify the possibility of a symmetric two-armed spiral with half the conventionally accepted pitch angle. We describe a model of spiral arm motions that matches the observed velocities and compositions of the six major streams, as well as the observed velocities of the Hyades and Praesepe clusters at the extreme of the Hyades stream. We model stellar orbits as perturbed ellipses aligned at a focus in coordinates rotating at the rate of precession of apocentre. Stars join a spiral arm just before apocentre, follow the arm for more than half an orbit, and leave the arm soon after pericentre. Spiral pattern speed equals the mean rate of precession of apocentre. Spiral arms are shown to be stable configurations of stellar orbits, up to the formation of a bar and/or ring. Pitch angle is directly related to the distribution of orbital eccentricities in a given spiral galaxy. We show how spiral galaxies can evolve to form bars and rings. We show that orbits of gas clouds are stable only in bisymmetric spirals. We conclude that spiral galaxies evolve toward grand design two-armed spirals. We infer from the velocity distributions that the Milky Way evolved into this form about 9 billion years ago (Ga).

Keywords: astrometry; celestial mechanics; stars: kinematics; stars: statistics; Galaxy: kinematics and dynamics; Galaxy: solar neighbourhood

1. Background

Stellar orbits are not elliptical because the gravitating mass of the galaxy is distributed in the disc and the halo (Binney & Tremaine 1987, ch. 3). In addition, orbits oscillate in the direction perpendicular to the disc. Orbits are expected to precess from both these causes, generating a rosette. It is usually assumed that, in time, an equilibrium state will be attained in which the distribution is well mixed. Spiral structure is usually explained using the density wave hypothesis of Lin et al. (1969), according to which stars move through the arms, which consist of dense regions analogous to regions of heavy traffic on a motorway. A simple

*Author for correspondence (charles.e.h.francis@googlemail.com).

analogy with patches of heavy traffic fails because a wave effect would require that stars slow down when they approach a dense region, but the gravity of the dense region would cause them to speed up. Following Kalnaj (1973), density wave theory is usually explained by means of a diagram (such as figure 1) constructed by enlarging and rotating ovals. The orbit is an epicyclic approximation in coordinates rotating at a rate $\Omega - K/2$, where $\Omega$ and $K$ are the frequencies of circular motion and the radial oscillation (see the electronic supplementary material, appendix A).

Density wave theory could apply to gas in laminar flow, but not to observed turbulent gas motions, or to mature stars for which the increase in density in figure 1 represents only a small proportion of the orbit. If, as one expects, $K$ is similar in magnitude to $\Omega$ (see §7), then coordinates rotate at near half orbital velocity and figure 1 shows a single spiral twice, not a bisymmetric spiral. Orbits that appear separate in these coordinates, and that are treated as a laminar flow in density wave theory, actually cross in physical space. We show in the electronic supplementary material (appendix A) that an entire class of rotating coordinates in which orbits are closed has been overlooked in Lindblad’s epicycle theory, including the most natural solution. Other models of spiral arms include the stochastic self-propagating star formation hypothesis proposed by Mueller & Arnett (1976), which suggests a mechanism by which young stars may form transitory spiral segments, but does not apply to enduring grand design two-armed spirals.

The established models are not borne out by the analysis of kinematic data in the solar neighbourhood. After the removal of fast moving stars, whose orbits are highly eccentric and/or significantly inclined to the disc, and which principally belong to the halo or the thick disc, far from being well mixed or obeying a laminar flow, the remaining population of thin disc stars divides into six major streams with distinct motions, and containing stars of all ages. We will describe an alternative mechanism, which does not depend on epicycles, and which also results in a spiral structure. We will show that this structure is dynamically stable, and that the observed stream motions are precisely those the structure predicts.
2. Stellar streams

Stellar streams consist of large populations of stars with similar motions. The existence of moving groups was first established from astronomical investigations dating as far back as 1869 (Eggen 1958). They were thought to consist of previously clustered coeval stars that have been gradually dispersed by the dynamic processes of tidal forces, differential galactic rotation and encounters with other stars. Increasingly comprehensive star catalogues published in the 1950s opened the way for more thorough analyses. Beginning in 1958, Eggen produced a series of seminal studies of stellar streams using RA:DE proper motion ratios in conjunction with radial velocities. Eggen’s investigations showed significantly increased membership counts and spatial extents of stellar streams, leading him to hypothesize a more protracted process of dissolution for star clusters. In Eggen’s scenario, as star clusters dissolve during their journeys around the Galaxy, they are stretched into tube-like formations, which were subsequently called superclusters.

The investigation of stellar streams received a major boost with the arrival of the precision astrometry afforded by the Hipparcos mission. Dehnen (1998), using transverse velocities derived from Hipparcos, produced maps of the local stellar velocity distribution showing that streams contain a significant proportion of late type stars. A wide range of stellar ages was identified within superclusters, challenging Eggen’s hypothesis of common origin (e.g. Chereul et al. 1998, 1999). The search for other types of dynamical mechanisms to account for streams has been ongoing. Candidates include migrations of resonant islands (Sridhar & Touma 1996; Dehnen 1998) and transient spiral waves (De Simone et al. 2004; Famaey et al. 2005) in which streams originate from perturbations in the gravitational potential associated with spiral structure.

3. The local standard of rest

The local standard of rest (LSR) is defined to mean the velocity of a circular orbit at the Solar radius from the Galactic centre. The definition idealizes an axisymmetric galaxy in equilibrium, ignoring features such as the bar, spiral arms and perturbations due to satellites. An accurate estimate of the LSR is required to determine parameters such as the enclosed mass at the solar radius and the eccentricity distribution that is of importance in understanding galactic structure and evolution.

As is customary in kinematic analyses of the stellar population, we denote velocity in the direction of the galactic centre by \( U \), in the direction of rotation by \( V \), and perpendicular to the galactic plane by \( W \). The solar motion relative to the LSR is \((U_0, V_0, W_0)\). The usual way to calculate the LSR is to calculate the mean velocity of a stellar population, and to correct \( V_0 \) for asymmetric drift. The method assumes a well-mixed distribution. However, as seen in §4, the observed kinematic distribution is highly structured, and divides into six populations each with distinct motion and stellar composition. Ignoring the possibility of perturbations to the galactic plane, motions of thin disc stars in the \( W \)-direction may be treated as a low amplitude oscillation due to the gravity

of the disc, and as independent of orbital motion in the $U-V$ plane. It is thus not unreasonable to calculate $W_0$ as the mean motion of a population. However, in the absence of knowledge of the causes for streams, there is no way to relate the statistical properties of their motion to $U_0$ and $V_0$.

Francis & Anderson (2009) studied a population of 20,574 Hipparcos stars with complete kinematic data, described in the electronic supplementary material (appendix B). We observed a deep minimum in the velocity distribution at a particular value of $(U, V)$ and argued that such a minimum might be expected at the LSR as a consequence of disc heating. Heating is the process by which scattering events cause the random velocities of stars to increase with age (e.g. Jenkins 1992). In thermal equilibrium in a well-mixed population, one would expect that the modal magnitude of random peculiar velocity denotes disc temperature. Circular motion represents an absolute zero temperature and can be expected to be rare for mature orbits. As a result, the distribution in velocity space can be expected to have a minimum at circular motion. In this paper, we will show that the true cause of the minimum is the perturbation of orbits due to the spiral structure. We will use the value of the solar motion found from the minimum in the velocity distribution, $(U_0, V_0, W_0) = (7.5 \pm 1.0, 13.5 \pm 0.3, 6.8 \pm 0.1) \text{km s}^{-1}$. We will use an adopted Solar transverse orbital velocity of 225 km s$^{-1}$ and a distance to the Galactic centre of 7.4 kpc, consistent with recent determinations (Reid 1993; Layden et al. 1996; Eisenhauer et al. 2005; Bica et al. 2006; Nishiyama et al. 2006) and the proper motion of Sgr A* determined by Reid & Brunthaller (2004) on the assumption that Sgr A* is stationary at the Galactic barycentre.

4. Stream properties

Famaey et al. (2005) described six kinematic groups: three streams, Hyades/Pleiades, Sirius and Hercules, a group of young giants, high velocity stars and a smooth background distribution (figure 2). We smoothed the velocity distribution by replacing each discrete point with a two-dimensional Gaussian function and finding the sum. The choice of smoothing parameter depends on the density of stars in the plot, and the required visual balance between overall structure and detail. Too large a smoothing parameter obscures structure, while two small a value confuses random fluctuations with structure. A standard deviation of 0.75 km s$^{-1}$ gives a clear contour plot (figure 3) and shows the major features of the distribution.

Streams are seen as dense regions in the velocity plots (figures 3–9). It is not possible to give precise criteria for stream membership from purely statistical data, as there is some overlap. Here we seek only a broad description. We distinguish the Hyades and Pleiades streams, since the velocity distributions show separate peaks, and, as we will see, these streams contain different distributions of stellar types and ages. There is a large and well-dispersed stream centred at $(U, V) = (25, -23) \text{km s}^{-1}$, noted by Dehnen (1998). Our estimate of its position is in good agreement with Chakrabarty (2007), who identified a clump in the velocity distributions at $(U, V) = (20, -20) \text{km s}^{-1}$. We have called it the Alpha Ceti stream, after the brightest star we identified with this motion.
Figure 2. $U$–$V$ plot showing groups identified by Famaey et al. (2005). These represent only a small proportion of the true membership of the streams. The calculated position of the LSR is shown for clarity.

Figure 3. The distribution of $U$- and $V$-velocities using Gaussian smoothing with a standard deviation of 0.75 km s$^{-1}$, showing the Hyades, Pleiades, Sirius, Hercules, Alpha Lacertae and Alpha Ceti streams.

Our analysis shows that the Pleiades stream consists largely of new-born stars, originating in our own spiral arm and with low eccentricities and typical orbits near apocentre, and also contains mature orbits with slightly greater eccentricity. We distinguish it from a stream with orbits close to pericentre, which contains
young as well as old stars. We have called this the Alpha Lacertae stream. It appears that Famaey’s young giants belong to the Alpha Lacertae stream. Famaey found a total stream membership of over 25 per cent, but the velocity distribution is highly structured by colour. When this is taken into consideration one sees that streams represent the bulk of the population.

The bluest stars, with $B - V < 0.04$ (approx. B–A0), reflect recent star formation in the Pleiades and Alpha Lacertae streams (figure 4). The overlap between these streams and the Hyades stream makes it difficult to ascertain the earliest Hyades stars. There are Hyades candidates of type B2, having a maximum age of about 20 Myr, stronger candidates at type B4, a maximum age of about 60 Myr, and clear signs of a Hyades population at type B9, an age of about
Figure 8. $U-V$ distribution: 2107 dwarfs, $0.56 \leq B-V < 0.8$ (approx. G), smoothing $\sigma = 1.5$.

Figure 9. $U-V$ distribution: 480 GC-S dwarfs with ages 9–13 billion years (Gyr), smoothing $\sigma = 3$.

400 Myr. Apart from a single star of type B7, the earliest clear indication of the Sirius stream is for stars of type B8, an age about 300 Myr. The Hercules and Alpha Ceti streams also contain members as early as B8, but become well populated at type F0, corresponding to an age of about 2.5 Gyr.

For $0.04 < B - V < 0.16$ (approx. A1–A5), the velocity distribution is concentrated in the Pleiades and Sirius streams (figure 5). The Hyades stream becomes more prominent than the Sirius stream for $0.16 < B - V < 0.4$ (approx. A6–A9) (figure 6), and dominates the velocity distributions (by density, not by total population) for dwarves with $0.4 < B - V < 0.56$ (approx. F) (figure 7) and $0.56 < B - V < 0.8$ (approx. G).

A clear indication of the stability of stream motions is given by the velocity distribution for old stars, using isochrone ages given by G-CS II (figure 9). There are known problems with isochrone ageing for very young stars; we found that a number of stars with young kinematics had been assigned ages greater than 13 Gyr. In other respects, G-CS II isochrone ages appear to be at least broadly reasonable, in accordance with their positions on the H-R diagram. For stars aged 9–13 Gyr, there is little indication of Sirius or Pleiades streams. The Hyades stream shows a sharp peak. The Hercules and Alpha Ceti streams, which are diffuse but contain more stars, are also prominent.

5. The eccentricity distribution

For an elliptical orbit, the eccentricity vector is defined as the vector pointing towards pericentre and with magnitude equal to the orbit’s scalar eccentricity. It is given by

$$e = \frac{|v|^2 r}{\mu} - \frac{(r \cdot v)v}{\mu} - \frac{r}{|r|},$$

where $v$ is the velocity vector, $r$ is the radial vector, and $\mu = GM$ is the standard gravitational parameter for an orbit about a mass $M$ (e.g. Arnold 1989; Goldstein 1980). For a Keplerian orbit, the eccentricity vector is a constant of the motion. Stellar orbits are not strictly elliptical, but the orbit will approximate an ellipse at each part of its motion, and the eccentricity vector remains a useful measure (the Laplace–Runge–Lenz vector, which is the same up to a multiplicative factor, is also used to describe perturbations to elliptical orbits). We smoothed the eccentricity distribution by replacing each discrete point with a two-dimensional Gaussian function and finding the sum. Standard deviation, $\sigma$, is used as a smoothing parameter. A standard deviation of 0.005 gave a clear contour plot (figure 10). In a well-mixed population, eccentricity vectors will be spread smoothly in all directions, with an overdensity at apocentre and underdensity at pericentre, because of the increased orbital velocity at pericentre and because stars at apocentre come from a denser population nearer the galactic centre. This is not seen in the plot. In practice, the distribution is concentrated at particular values corresponding to stream motions.

6. A model of spiral structure

In polar coordinates $(r, \theta)$ an equiangular spiral is given, for positive real $a$ and $b$, by

$$r = ae^{b\theta}.$$
The constant pitch angle of the spiral is $\phi = \arctan b^{-1}$. We demonstrate here that an equiangular spiral structure can be constructed from elliptical orbits by enlarging an ellipse by a constant factor, $k$, centred at the focus and rotating it by a constant angle, $\tau$, with each enlargement (figure 11). Orbits of different sizes align in a spiral pattern, leading to an overdensity of stars, which creates the spiral arms. The pitch angle of the spiral depends only on $k$ and $\tau$, not on the eccentricity of the ellipse. For a given pitch angle, ellipses with a range of eccentricities can be fitted to the spiral, depending on how narrow one wants to make the spiral structure and what proportion of the circumference of the ellipse one wants to lie within it. In general terms, ellipses with greater eccentricity fit with spirals with greater pitch angles.

In a practical model for galactic spiral structure, stellar orbits are approximately elliptical and are gravitationally aligned to a spiral arm. Unaligned orbits lying between the arms will be drawn to one arm or the other, and orbits will precess due to the distributed matter distribution of the galaxy, such that they become aligned. Once alignment of the orbit with the spiral arm is achieved, it will be maintained by perturbations to the orbit due to the gravity of the arm.
A star close to apocentre will approach the inside of the arm on account of the pitch angle. If it has greater eccentricity than that of stars in the arm, the gravity of the arm will draw it closer, causing a reduction in eccentricity. If it has lower eccentricity than the arm stars, it will pass through the arm. Because of the curve of its orbit, it will spend more time in the gravitational field on the outside of the arm, and will be drawn back towards the arm, with a net increase in eccentricity.

The mechanism binding stars to the arm is explained in more detail in the electronic supplementary material (appendix E). It reinforces the spiral, showing that spiral arms are stable dynamical structures (they may eventually be destroyed by the growth of a bar and/or a ring). Orbital precession will mean that the spiral pattern rotates, but the winding problem is resolved because, for a wide range of orbits, orbital alignment is determined by the gravitational field of the arm (the classic winding problem is inapplicable because spiral structure does not depend directly on orbital velocity at different radii). The evolution of bisymmetric spirals from flocculent and multi-armed spirals is explained in §12. The model predicts trailing spirals, not leading spirals, because orbital velocity and gravitational field strength owing to a centrally concentrated mass distribution are lower near apocentre, so that the gravitational field of the arm is of greater influence in perturbing the orbit near apocentre. Thus, the alignment of orbits to the spiral arm proceeds from the outside towards the inside, not the other way about. This agrees with the long-established observational result (de Vaucouleurs 1958). The few exceptional candidates for leading spirals are thought to be induced by special mechanisms of tidal interactions with companion galaxies (Väisänen et al. 2008).

An animation of a galaxy formed from aligned rosettes with similar parameters to the Milky Way is described in the electronic supplementary material (appendix C). The animation clearly shows stars crossing an arm at the same part of their outward motion (Hyades stream), as well as the differing velocities of stars in the arm.
Figure 12. An equiangular spiral with a pitch angle of 11°, constructed by repeatedly enlarging an ellipse with eccentricity 0.3 by a factor 1.05 and rotating it through 15° with each enlargement. Lower eccentricity ellipses produce a narrower structure. Ellipses with eccentricity greater than about 0.25 have more than half their circumference within the spiral region. Ellipses with eccentricity greater than about 0.35 produce probably too broad a structure to model a spiral arm with this pitch angle, but give a good fit for spirals with greater pitch angle.

7. Precession of apocentre and spiral pattern speed

For an orbit in the thin disc, motion perpendicular to the disc may be treated as an independent oscillation superimposed on an orbit in the plane of the disc. To a good approximation, this oscillatory motion does not cause the orbit to precess (since the oscillation is perpendicular to the centripetal force and to orbital motion). The mass distribution in the halo is generally assumed spherical. By Newton’s shell theorem it can be treated as a central mass, which reduces as orbital radius reduces. The effect is to reduce the curvature of the orbit at pericentre, such that pericentre regresses during the orbit. The result is less obvious for matter in the disc, because the gravitational effect of nearby matter in a uniform ring outweighs the net effect of the farther parts of the ring. The reduction in curvature of the orbit near pericentre owing to lower enclosed mass is offset by the increase in density of nearby matter. We used a numerical simulation for a galaxy with a central core plus a disc with an exponentially decreasing surface density to establish that orbits also regress (figure 12).

Knowledge of the mass distribution of the Milky Way is not sufficiently precise to choose a specific model from which an exact rate of precession may be calculated, but it is not large. The distribution of dark matter in the halo and in the thin disc affects the rate of precession of stellar orbits, but it has no direct impact on spiral structure. If we assume that the rate of precession is constant, or approximately so, for orbits at different radii, and for eccentricities within the range of binding by the spiral arm, then we may choose coordinates rotating at the rate of precession of apocentre. In these coordinates, orbits do not precess,
and may be taken to be approximately elliptical. The spiral arm structure applies as before, and we find that, in non-rotating coordinates, spiral pattern speed is equal to the rate of precession of apocentre.

We require only that the rate of precession is an approximate constant for a stable spiral arm, because gravitational binding to the arm outweighs the consequence of small changes in the rates of precession at different orbital radii and for different eccentricities. If the matter distribution is such that orbital precession is not constant for different radii, the pitch angle of the spiral will alter over time. Eccentricities will adjust to the pitch angle, and stability may be achieved at altered values of pitch angle and eccentricity (we may conjecture that the rings of Saturn comprise very tightly wound spirals built from orbits that owe their low eccentricity to mass distribution). For the purpose of this paper, we will use coordinates rotating at spiral pattern speed, and we will assume that orbits can be approximated by ellipses in these coordinates.

8. Fitting the model to the Milky Way

It is straightforward to observe spiral structure in other galaxies, but extremely difficult to observe it within our own galaxy, as recently illustrated by observations of the Spitzer telescope showing that stellar concentrations are not found at the positions where two arms were thought to be (Benjamin 2008). There have been two principal methods for locating spiral arms. The usual four-armed spiral is derived principally from the distribution of ionized hydrogen (Georgelin & Georgelin 1976; Russeil 2003), but, in fact, the distribution is so sparse and irregular that it is difficult to be certain that anything has really been fitted. We will see in §9 that ionized hydrogen is not expected to give a good fit to spiral structure in this model. The neutral hydrogen distribution was famously mapped by Oort et al. (1958), and more recently by Levine et al. (2006). Levine et al. (2006) fitted (slightly irregular) four-armed spirals, but comment that other fits are possible.

The four-armed spirals fitted by Georgelin & Georgelin, Russeil & Levine, and Blitz & Heiles have a pitch angle of about 10–15°, corresponding to orbital eccentricities in the range greater than about 0.25. This is not consistent with the eccentricity distribution for the Milky Way, in which the modal value is a little above 0.1 (figure 13), suggesting a much lower pitch angle than is used in four-armed spirals. We found good visual fits to the hydrogen maps of Oort et al. (1958) and of Levine et al. (2006) for bisymmetric spirals with a 8.2kpc bar and pitch angles in the range 5.3 ± 0.5° (figure 14), in agreement with 5.1° and 5.3° found from HII regions for the two-armed logarithmic model by Hou et al. (2009). There is a subjective element in the quality of such a fit, but the two-armed spirals seem to us to better follow the line of the hydrogen clouds, while the more open four-armed spirals appear to follow clouds bridging the true line of the arms. We also fitted to the map of Levine et al. by maximizing the mass per unit length lying on a bisymmetric spiral, finding maxima at pitch angles of 4.9° and 5.7°. As the model is predicted to give a ragged arm in gas motions, this did not appear to be better than the visual fitting method.

Figure 13. The eccentricity vector of an orbit regresses for a central core plus disc. Regression has been exaggerated by increasing the mass of the disc relative to the core (by comparison with the Milky Way). The simulation used a central mass of 35 billion solar masses, a disc density $0.3e^{-R/3}$ billion solar masses per kpc$^2$, initial radius 8 kpc and initial velocity 190 km s$^{-1}$.

We constructed a symmetric two-armed spiral with pitch angle $5.44^\circ$ from ellipses using an angular increment $\tau = 30^\circ$ for each 105 per cent enlargement (electronic supplementary material figure S19, appendix D). For the calculated value of the LSR, current solar eccentricity is 0.138. The Sun is 16.3$^\circ$ before pericentre (as determined by the current eccentricity vector), at which point it should lie near the inner edge of the arm, and be heading outwards through the arm. We were not able to make a meaningful map showing the positions of stars with velocities in the arm with respect to the Sun, because the data from Hipparcos are not sufficiently comprehensive over large enough distances, and because the data for which we have radial velocities are strongly weighted to the northern hemisphere, but, in agreement with typical estimates, there is some indication in the data that we are in the arm, about 100–150 pc from the inner rim, and too far to be able to detect the outer rim.

9. Young stars

Star formation has been a central problem for galactic dynamics. There is not enough mass in the disc for gas clouds to collapse under gravity and form protostars. Depending on the width of the arms, orbital alignment in spiral arms results in an increase in stellar density by a factor of about 5. Gas clouds follow orbits following spiral arms according to the same laws as those governing stars (figure 15). Gas in the arm seeks to gain velocity as it approaches pericentre, and also to follow paths crossing within the arm. Thus, motions are complicated by collisions between clouds and resulting turbulence. The increase in surface density of gas is about half as much (figure 14). When clouds of atomic hydrogen with spiral arm motions meet with clouds crossing the arm it is to be expected
that higher densities obtain, the height of gas is increased (figure 14) and that greater turbulence is created, with pockets of high density generating molecular gas clouds from which protostars form.

Since outgoing gas has generally lower density than ingoing gas, outgoing motions normally terminate at the arm. In regions where outgoing gas clouds of greater than normal density meet regions where gas in the arm is less dense, the pattern of the clouds deviates from the stellar spiral, as seen in figure 14. We therefore do not expect the distribution of ionized hydrogen or of star-forming regions to depict spiral structure accurately, as assumed by Georgelin & Georgelin (1976) and Russeil (2003).

We may expect that star formation typically initiates with a build-up of gas and dust on the inside of the arm, and continues through the arm. The less massive outgoing gas clouds add a component of radial velocity to those in the arm, with the result that stars form with motions found in the Pleiades stream. The modal value of eccentricity in the Pleiades stream is approximately 0.065, and is increasing (see the electronic supplementary material, figure S21, appendix E). Figure 16a shows an orbit starting at apocentre with eccentricity 0.074. In figure 16b, initial eccentricity is lower, 0.033. The orbits hug the outside of the arm, while eccentricity continues to increase. This is consistent with observations on many spiral galaxies showing that the brightest spots, groups of young stars, lie on the outside of the arms (e.g. M51, M74, M83, M101).

The resulting orbits achieve higher than normal eccentricity, seen in young stars in the Sirius stream and Hyades cluster. The orbits do not at first align with the arm. A typical orbit with eccentricity 0.075 at apocentre meets the arm before pericentre, with an eccentricity characteristic of the Sirius stream. Thereafter the trajectory continues to meet the spiral at pericentre for a number of orbits, while pericentre regresses so as to improve the alignment of the orbit with the arm. For eccentricity 0.034 at apocentre, the orbit does not rejoin the arm, but eccentricity continues to increase and pericentre advances, toward alignment with the other arm. In both cases it may take several more orbits before alignment with an arm is achieved.

10. The Hercules and Alpha Ceti streams

It is clear that the Hercules stream, and the higher eccentricities seen in the Alpha Ceti stream, do not fit with the pattern of typical spiral arm motions. Also observed on the eccentricity distribution for mature orbits (see the electronic supplementary material, figure S23, appendix E) are small numbers of stars with eccentricities up to about 0.3 approaching the semi-latus rectum on the
Figure 16. Orbits of young stars passing through apocentre with eccentricities 0.074 and 0.033, typical of the Pleiades stream.

inward arm of the orbit, and extending from well before pericentre to just after pericentre. These motions can also be picked out on the velocity plot (see the electronic supplementary material figure S22, appendix E), and have increased prominence on the distributions for late type stars (figure 8) and old stars (figure 9), suggesting that these are also stable orbits in coordinates rotating with the arms.

It is possible to align orbits with eccentricities in the region of 0.3 with spiral arms such that the orbit follows the locus of one arm for a period after apocentre, and with the other arm during the inner part of the orbit (figure 17). Since these streams consist of predominantly older stars, we may conclude they are stable orbits formed during more turbulent motions in the early galaxy, and that young stars rarely join these motions.

11. Formation of bars and rings

When the mass of the arm is increased in the numerical simulation, the orbit continues to follow the arm after pericentre, eventually leaving with an increased eccentricity. The density of the arm increases with the density of the disc towards the galactic centre, and, depending on the mass distribution of the galaxy, the spiral structure breaks down. If, as may be expected, orbits depart from the arm
at a similar point, then a symmetrical interlinked pattern of opposing orbits is formed (figure 18; e.g. UGC12646, ESO 325-28, NGC 2665, NGC 619). It seems unlikely that a structure of interlinked rings is stable. If eccentricities continue to increase and the structure collapses, a bar will be formed.

According to this model, spiral arms do not start at the bar, but extend inwards beyond its ends. This is observed in a number of barred spirals (e.g. M92, M109, NGC 1300) and in the Oort hydrogen map (figure 14), in which the spiral structure of both arms continues a quarter turn beyond the bar. Once a
bar is formed, it perturbs nearby orbits such that the arms are destroyed shortly after stars pass the ends of the bar. This is likely to cause the bar to grow. As with spiral arms, bar pattern speed will depend on the eccentricity of orbits in the bar. Since eccentricities in the bar are very high, we may expect that bar pattern speed does not match spiral pattern speed.

In time, orbits approaching the bar will also be perturbed and will not rejoin the spiral arm. We have not performed a dynamical analysis, but conjecture that these perturbations lead to the inner ring observed in galaxies like M95, NGC 4314, NGC 1433, IC 5240, IC 5340. It is also possible that inner rings as well as a bar can evolve from interlinked rings.

An inner ring formed from a bar will be oval. A circular ring may be formed when the two spiral arms are close enough to perturb each other’s orbits such that stars in the inner arm are drawn toward the outer arm on each side of the spiral (NGC 4725, M81, NGC 6902, NGC 7217, NGC 3124). This process may be encouraged if the inner part of the spiral arm has already been sucked into an inner ring or bar. Some galaxies display both types of ring (NGC 1543, IC 1438, NGC 7187). If the ring has sufficient density, stars following the line of the arm and meeting the ring will become gravitationally bound to the ring. The gravitational attraction of the ring will allow it to contain stars with a range of orbital velocities, which will tend to smooth the mass distribution in the ring such that a stable structure may be formed. One would then expect a gap to develop between the ring and the arms (NGC 210, NGC 5701, NGC 1291), or the arms to disappear entirely, where stars from the arms have joined the ring (NGC 7742, PGC 54599, NGC 4553, NGC 4419, NGC 1543, NGC 7020). In the case of PGC 54599, Hoag’s Object, one can see a residual arm. The natural evolution of Hoag’s Object from a spiral galaxy would explain its symmetrical form and its lack of features characterizing rings resulting from galactic collision. NGC 4725 is sometimes described as a one-armed spiral, but in fact appears to be a ring forming from a two-armed spiral at an earlier stage of evolution than Hoag’s Object. One arm has almost been destroyed, while the other is still clearly apparent.

12. Evolution of bisymmetric spirals

Starting from an unstructured initial condition, there is reason to think that mutual gravitational attraction will cause stars to form groups, spiral segments (flocculence), and, ultimately, spiral arms, but no strong reason to expect that these structures will reduce to a bisymmetric spiral. Gas will also be attracted into spiral segments, but when gas clouds meet, they combine to form larger clouds, and add more quickly to the mass of the segment, so that the process of reducing spiral segments to a small number of arms is faster for gaseous galaxies.

In the Milky Way, the density of gas in a spiral arm is greater by a factor of about 2 than that of gas between the arms (figure 14). This is sufficient to maintain spiral structure when gas in the arms meets with outgoing gas. In a symmetric three-armed spiral, the density of gas in the arm would be reduced by about 33 per cent, and the density of outgoing gas distributed evenly in the disc would be increased by about 67 per cent. Thus, outgoing gas meeting the arm
would outweigh the ingoing gas, and would tend to remove gas from the arm. Thus, a two-armed gaseous spiral can be stable, whereas multiarmed gaseous spirals cannot.

Outgoing gas applies a pressure to the inside of a spiral arm with an inverse proportionality to radius (figure 15). If one gaseous arm advances compared with the bisymmetric position, the pressure due to gas from the other arm will be reduced. At the same time, pressure on the retarded arm due to outgoing gas from the advanced arm will be increased. Thus gas motions provide a mechanism to maintain the symmetry of two-armed spirals. The same pressure means that gaseous arms will advance with respect to non-gaseous arms and spiral segments. Eventually the arms will combine, so that a two-armed spiral is formed.

Thus gas establishes the pattern of the major spiral arms in flocculent and multi-arm spirals representing earlier stages in the development of grand design spirals. The lower the gas content, the longer the process of evolution. This pattern of evolution is consistent with the observations of Thornley (1996), who found a low-lying spiral structure in four nearby flocculent spirals using near-infrared imaging. We also note that the marked lack of grand design spirals in the Hubble ultra-deep field, in which all galaxies are young, is indicative that this form takes some time to evolve. The stability of orbits with high eccentricities in the Hercules and Alpha Ceti streams depends on the prior formation of a bisymmetric spiral (figure 17). Francis & Anderson (2009) found that a sharp change in velocity components at age 9 ± 1 Gyr (previously seen by Quillen & Garnett (2000)) is caused by an increased membership of these streams for stars of greater age. We infer that a grand design two-armed spiral was formed in the Milky Way about 9 billion years ago (Ga).

13. Conclusion

After studying the velocity distributions for local stars we have concluded that the observed stellar streams reflect the spiral structure of the Milky Way. We have presented a straightforward model of equiangular spiral arms constructed from elliptical orbits aligned at a focus. This model applies in coordinates rotating at the spiral pattern speed, which is equal to the mean rate of orbital precession. We have shown by qualitative argument and numerical simulation describing perturbations to elliptical orbits, that, for a range of arm densities, spiral structure is dynamically stable, up to destruction by a bar and/or a ring. We have shown that, for a two-armed equiangular spiral with pitch angle set to match the distribution of neutral hydrogen, the observed eccentricity and velocity distributions are a good fit to the predictions of the model after taking expected perturbations into account. We have accounted for all stellar streams in the observed local velocity distributions. We find that the Sun follows a very typical orbit aligned to the Orion arm, which is a major spiral arm containing Perseus and Sagittarius sectors. We have calculated that its current eccentricity is 0.138. This is a little higher than the modal value, 0.11, for stars in the arm, giving a typical orbital period of about 300 Myr—longer than usually estimated because of the greater eccentricity. We have seen how spiral structure can evolve to form the rings and bars found in many
Galactic spiral structure

3445

galaxies, and that gas motions determine that flocculent galaxies evolve toward bisymmetric spirals. We have found that the Milky Way evolved to this form about 9 Ga.

It is perhaps worth remarking that the model has made genuine predictions, and not merely been retrodictively fitted to data. Having made a prediction of a galactic structure, we searched images to find examples of the configuration. The interlinked ring structure of figure 18 was recognized by the astronomer (E.A.) among the authors, but it was not known to the mathematician (C.F.), who produced the figure from the numerical solution of perturbed orbits. The same was true of the prediction that young stars are to be found on the outside of spiral arms. Nor did we know of galaxies where the spiral arms are separate from the ring. We have not made any predictions of galactic structures for which we were unable to find examples.

The compiled data used in this paper can be downloaded from http://data.rggravity.net/lsr/.

References


