The Astronomical Journal, 149:174 (6pp), 2015 May doi:10.1088/0004-6256/149/5/174
© 2015. The American Astronomical Society. All rights reserved.

N-BODY SOLUTIONS AND COMPUTING GALACTIC MASSES

Donald G. Saari
University of California Irvine, CA 92697-5100, USA; dsaari@uci.edu
Received 2012 January 2; accepted 2015 February 20; published 2015 April 28

ABSTRACT

A classical approach used to determine the mass distribution of a galaxy in terms of observed rotational velocities is applied to analytic solutions of Newtonian systems of N discrete bodies (so mass distributions are known). Predictions significantly exaggerate the amount of mass distributed at larger distances from the center; e.g., rather than the actual 5% of mass on the outer edges, the method could predict over 80%. Explanations are given for the differences.

Key words: celestial mechanics – galaxies: kinematics and dynamics – methods: analytical

1. INTRODUCTION

A classical way to determine galactic mass distributions uses the circular velocities of stars. With observations of spiral galaxies, where stellar circular velocities tend to flatten or even increase with distance from the galactic center, the method predicts much larger $M(r)$ values (the total mass to distance r) than are known to exist, particularly on galactic edges (e.g., Ostriker et al. 1974; Rubin 1983, 1993). This supports the belief that huge amounts of unobserved mass must exist in a halo. However, this mathematical approach has not been evaluated to determine whether predictions could be exaggerated. The validation issue is examined here by applying the method to analytic N-body solutions.

Approaches (theoretical and computational) developed to relate mass values with circular velocities tend to be based on Newton's first and second laws (e.g., Section 2.2.1 of Binney & Tremaine 2008). In a symmetric continuum setting, for instance, a body inside a spherical shell experiences no net gravitational force from the shell; the gravitational force on a body outside a spherical shell behaves as though the shell's matter is concentrated at the center. There are also other approaches, so hereafter "continuum" refers to all models adopting these assumptions (e.g., Newton's first and second laws) to arrive at Equation (1). The "discrete" term is reserved for the standard Newtonian N-body problem.

As depicted in Figure 1, discrete N-body systems need not satisfy these assumptions. Instead, with a close approach of two bodies, each strongly influences the gravitational pull of the other, where the inner body (moving in a counterclockwise direction) tugs on the outer body. This mandatory Newtonian tugging generates larger rotational velocities, which, using continuum-type prediction approaches, would be erroneously interpreted as manifesting an added level of mass. The dragging would tend to affect bodies farther out, which suggests the possibility of exaggerated mass predictions for larger distances from the center.

With Newton's equations this pulling occurs, but does it matter? Can it be safely ignored as inconsequential, or could the dragging significantly distort the predicted mass distributions, particularly in regions farther from the center? To explore this question, a standard method is applied to analytic solutions of Newtonian N-body systems. As shown (Section 4), predictions can seriously distort actual mass distributions by significantly exaggerating the percentage of the total mass that is located on the edges of the solution. As shown, the more sharply the actual mass values decrease with increasing r, the larger the prediction errors.

It is natural to test a method by applying it to discrete N-body solutions, but this has not been previously done because of an absence of analytic solutions with suitable symmetry properties. That lacuna is partly filled here by creating, for arbitrarily large values of N, classes of solutions where each class contains a continuum (based on mass choices for individual bodies) of analytic N-body solutions. These solutions are designed to resemble (as closely as possible with discrete systems) the central properties used to develop the prediction method: solutions have highly symmetric mass distributions, orbits are circular, and each body's gravitational attraction is directed toward the center of mass.

2. STANDARD APPROACH

In symmetric settings where Newton's two laws apply, the mass up to radius r from the center of mass, $M(r)$, has the form (e.g., Section 2.2.1 of Binney & Tremaine 2008)

$$M(r) = \frac{rv_{\rm rot}^2}{G} \tag{1}$$

where $v_{\rm rot}$ is the circular velocity of a star at distance r. With the Equation (1) classical approach, the observed flattening of rotational velocities implies for large r values that

$$M(r) \approx Dr, \tag{2}$$

where D is a positive constant. As extensively described (often using Equation (1)), this $M(r)$ value is much larger than justified from the luminosity of stars and other methods. There are more sophisticated ways to estimate $M(r)$, but Equation (1) is a standard approach.

3. CENTRAL CONFIGURATIONS

To apply Equation (1) to Newtonian systems with N discrete bodies, analytic solutions are developed using "central configurations." After introducing these configurations and their properties, they are used to create any number of N-body solutions.

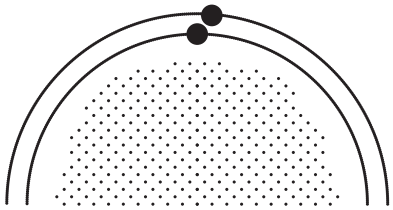

Figure 1. Discrete body interactions.

3.1. Central Configurations

N bodies define a central configuration (e.g., Wintner 1941; Saari 2005) if there is a common scalar λ (depending on distances, masses, and other N-body variables) so that each body's position and acceleration vectors satisfy

$$\boldsymbol{r}_j'' = \lambda \boldsymbol{r}_j, \quad \text{which is } \lambda \boldsymbol{r}_j = \frac{1}{m_j}\frac{\partial U}{\partial \boldsymbol{r}_j},$$

$$U = \sum_{i<k} \frac{G m_i m_k}{r_{i,k}}, \; j = 1, ..., N, \tag{3}$$

and $r_{i,k} = |\boldsymbol{r}_i - \boldsymbol{r}_k|$ is the distance between particles. In 1772, Lagrange proved, for any mass values, that the vertices of an equilateral triangle define an $N = 3$ central configuration. An example is where the Sun and Jupiter define one leg, and the Trojan asteroids populate the region around the remaining vertex of the leading and following equilateral triangles.

In an 1859 prize winning paper, Maxwell used central configurations to analyze the rings of Saturn by placing n equal masses in an evenly spaced manner on a circle. This symmetry requires the combined force acting on the jth particle to lie along the line defined by $\boldsymbol{r}_j$; i.e., for each j there is a scalar λ_j so that $\boldsymbol{r}_j'' = \lambda_j \boldsymbol{r}_j$. As all masses are equal and the particles are symmetrically spaced along the circle, an identical analysis holds for each particle, so the λ_j values agree and Maxwell's construction forms a central configuration.

Maxwell needed a massive central body to represent Saturn. Placing such a body at the center of mass of this central configuration changes the common $\lambda_j = \lambda$ value, but symmetry requires that they still agree. The central body is at $\boldsymbol{r} = \boldsymbol{0}$ (the center of mass), and the other particles are symmetrically located about 0, so it follows that $\boldsymbol{r} = \boldsymbol{r}'' = \boldsymbol{0}$. As $\lambda \boldsymbol{0} = \boldsymbol{0}$, the central body satisfies Equation (3) for any λ value. Thus, Maxwell's configuration, where a central body (with any desired mass) is symmetrically surrounded by n equal masses on a circle, defines a $N = n + 1$ body central configuration.

Moulton (1910) analyzed collinear central configurations. He proved that for any $N \geqslant 2$, any choice of masses m_j, and any ordering of the masses along a line, there exist unique (up to a common scalar multiple) spacings between adjacent particles that define a central configuration. For any specified mass choices, then, there exist $N!/2$ collinear central configurations. (To explain the denominator, there are $N!$ ways to order the masses on the line, but a reflection identifies each configuration with its reversed copy.)

The configurations designed next combine Moulton's and Maxwell's approaches to create examples with circular symmetry. For any positive integers k and n, start with k lines in the plane that pass through the origin where adjacent lines are π/k radians apart. Choose any mass values $m_j > 0$, $j = 1, ..., n$. As depicted in Figure 2 (for $k = 4$, $n = 3$), draw

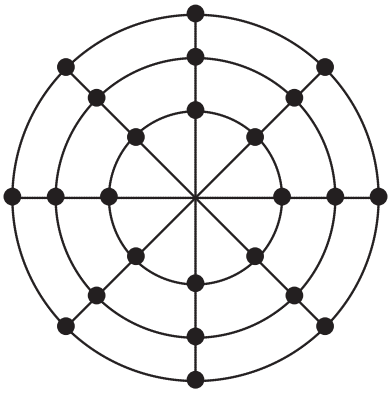

Figure 2. $k = 4$, $n = 3$ "spiderweb" central configuration.

n concentric circles and place mass m_j at each point where the jth circle intersects a line, $j = 1, ..., n$. Each circle has $2k$ masses (with two masses on each line), so this construction defines a $N = 2nk$ body configuration that resembles a spiderweb.

As asserted next (and proved in the Appendix), for any choice of k, n and masses m_j, radii for the circles exist that satisfy Equation (3) to create a "spiderweb central configuration" with $N = 2nk$ masses. In the same way Maxwell developed his model for Saturn, an associated $N = 2nk + 1$ body central configuration is created by placing a mass of any size at the origin. For $k = 1$, the spiderweb configuration is a symmetric case of Moulton's construction; for $n = 1$, it is Maxwell's construction.

Theorem 1. *For positive integers k and n and any choice of $m_j > 0$, $j = 1, ..., n$, there exist unique spacings between the concentric circles so that the configuration is a spiderweb central configuration for these spacings and any positive multiple of them.*

To indicate why Theorem 1 must be expected, symmetry requires particles on the jth circle to satisfy $\lambda_j \boldsymbol{r}_j = \frac{1}{m_j}\frac{\partial U}{\partial \boldsymbol{r}_j}$, so λ_j is a smooth function of the masses and the positioning of particles (where critical terms are the distances between circles). With fixed m_j values, the forces, which define the λ_j values, can be varied by altering the distances between circles; decreasing distances increase the λ_j magnitudes. In this manner, spacings can be selected so that $\lambda_1 = ... = \lambda_n$ to define a central configuration. Conversely, for certain distances between circles, the $\lambda_1 = ... = \lambda_n$ constraint defines linear equations in the masses that can be solved (up to a common multiple). While the primary intent of Figure 2 is to depict the construction, this configuration requires nearly equal m_1, m_2, m_3 masses.

3.2. Central Configurations and Analytic Solutions

A rotation or a scalar change of a central configuration is again a central configuration (e.g., Wintner 1941; Saari 2005), so a central configuration defines a class of configurations. For instance, specified mass choices define precisely four classes of three-body central configurations: three collinear choices (where the relative distances depend on the mass values) and the equilateral triangle (and its reflection). Each configuration can be rotated and scaled by any positive multiple; e.g., one

equilateral triangle describes the Sun–Jupiter–Trojan configuration, another identifies potential Earth–Moon–satellite docking locations, and a third describes configurations formed by bodies approaching a three-body collision (Saari 2005).

A third property (from Equation (3)) is that, with appropriate initial conditions, any coplanar N-body central configuration can be placed into a circular or elliptic orbit that preserves the configuration; the motion keeps the shape while assuming different scale or rotation values. Appropriate initial conditions for the three-body equilateral triangle configuration, then, create either a circular or elliptic orbit that maintains the equilateral triangle configuration; e.g., this motion reflects the orbit behavior of the Trojan asteroids. To analyze the rings of Saturn, Maxwell placed his central configuration model into a circular orbit. In the next section, spiderweb central configurations are placed in circular orbits.

4. PREDICTIONS OF MASS VALUES

For any choices of n, $k \geqslant 1$, let the analytic N-body solution be either the $N = 2nk$ or $N = 2nk + 1$ body spiderweb central configuration placed in a circular orbit. By construction, each particle's acceleration is pointed toward the center of mass, each body has a circular orbit, and, with large k values, the solution has a highly symmetric mass distribution. A scale change of a central configuration is again a central configuration, so the minimal distance between adjacent concentric circles is set equal to one unit.

4.1. Predicting Mass Values

Independent of the m_j values, each of these analytic Newtonian N-body solutions behaves like a rotating rigid body. For each solution, then, a constant $S > 0$ exists so that the jth particle's rotational velocity is

$$v_{rot,j} = S\, r_j, \quad j = 1, ..., N. \tag{4}$$

Incidentally, rigid body behavior can be expected to occur in portions of spiral galaxies as manifested by the nearly straight line structure of the first part of the observed rotational velocity curves. With these observed values, the rotational structure conveys the flavor of a rigid body motion where separating particles on the galactic edges are being pulled along.

While the value of S is not needed, it is derived for completeness of the exposition. The circular motion requires the kinetic energy to equal $T = \frac{1}{2}\sum_{j=1}^{N} m_j v_j^2 = \frac{1}{2}S^2\sum_{j=1}^{N} m_j r_j^2 = \frac{1}{2}S^2 I$, where $I = \sum_{j=1}^{N} m_j r_j^2$ is the polar moment of inertia. As I is a constant, it follows from the Lagrange-Jacobi equation ($I'' = 2(2\,T - U)$) that $I'' \equiv 0$ and $2T \equiv U$, so $S^2 = \frac{U}{I}$. As is already known (e.g., Winter 1941, Saari 2005), $\lambda = -\frac{U}{I}$, so λ also satisfies $|\lambda| = S^2$.

Applying Equations (1) to (4) leads to the prediction that

$$M(r) \approx \frac{S^2 r^3}{G} = \frac{|\lambda|}{G} r^3 = Dr^3 \tag{5}$$

for positive constant D. (Relative comparisons are used in the following to avoid problems caused by D changing value with n and k.) This means that the predicted mass distribution is always given by a cubic equation (Equation (5)) *even though mass values have yet to be selected.* Therefore, for all m_j choices where the actual $M(r)$ is far removed from being a cubic expression, it is impossible for Equation (5) to describe the actual mass distribution. Indeed, the incremental increase of Equation (5) is $M' \approx 3Dr^2$, so this negative assertion must hold whenever m_j values do not significantly increase for increasing j. Also, functions are compared over domains rather than a few points, so the number of rings, n, must be sufficiently large.

To illustrate, consider an $n = 100$ ring configuration where each ring has a mass equal to unity. To compute $M(r)$, determine how many rings (s of them) are required to reach distance r: $M(r)$ equals the sum of the masses of these s rings. The minimal spacing assumption (where the smallest distance is 1) means that the distance r_s required to reach the sth ring satisfies $r_s \geqslant s$; e.g., for this example with $M(r_2) = 2$, the distance r_2 is that of the first two rings, so $r_2 \geqslant 2$. More generally, if r_j is the distance to the example's jth ring, then $M(r_j) = j$. The resulting $r_j \geqslant j$ inequality requires $M(r) \leqslant r$; in other words, rather than being cubic (Equation (5)), the precise $M(r)$ distribution is, at best, linear. The more the actual $M(r)$ differs from r, the larger the r_{100} value; e.g., should computations prove that $M(r) \approx r$, then $r_{100} \approx 100$, but should computations prove that $M(r) \approx \frac{1}{2}r$, then $r_{100} \approx 200$. This difference affects the prediction errors.

A cubic equation cannot even crudely resemble a linear equation over this range of a hundred separated r_j values, which means that the cubic expression seriously distorts the true $M(r)$ distribution; this is particularly true for larger r values on the outer fringes of the solution. This assertion is illustrated by the fact that the predicted mass up to distance r relative to the mass of the first ring is $\frac{M(r)}{M(r_1)} \approx \frac{Dr^3}{Dr_1^3} = (\frac{r}{r_1})^3$; with the extreme choice of $r = r_{100}$, this means that the predicted total mass is at least $100^3 = 10^6$ times the predicted mass of the first ring, when, in reality, the multiple is only 100. (As Equation (5) might underestimate the mass of r_1, this statement does not assert that the predicted total mass is 10^6; it means that Equation (5) significantly distorts the actual mass distribution by erroneously predicting larger relative mass distributions for regions farther from the center.) Similarly, precisely 50% of the mass resides between the 50th and 100th rings, but, for $k = 3$ and the 600 body problem, $2r_{50} \approx r_{100}$ (from the approximations following the proof of Theorem 1), which means that Equation (5) erroneously predicts that approximately 87.5% (given by $\frac{D\, r_{100}^3 - D\, r_{50}^3}{Dr_{100}^3} = 1 - \frac{r_{50}^3}{r_{100}^3} \approx 1 - \left(\frac{1}{2}\right)^3 = \frac{7}{8}$) of the mass is located here.

Similarly, $M'(r) = 3Dr^2$ indicates the mass of a ring at distance r, so the prediction of the ring mass at distance r relative to that of some other ring, say the one at r_1, is $\frac{M'(r)}{M'(r_1)} \approx \left(\frac{r}{r_1}\right)^2$. With the above example and again using the extreme $r = r_{100}$, this expression erroneously predicts that the mass of the last ring is 10^4 times larger than that of the first ring,[1] while the true multiple is unity. Smaller, more realistic mass choices, where the m_j values decrease as the distance from the center of mass increases (so the $2km_j$ mass value of the jth ring differs extremely from $\frac{M'(r)}{M'(r_1)} \approx \left(\frac{r}{r_1}\right)^2$), exacerbate differences between predicated and actual distributions. Choosing,

[1] With $M(r) \approx \frac{1}{2}r$, it would be 16×10^4 larger.

for instance, $m_j = \frac{1}{2jk}$, the mass on the jth circle is $\frac{1}{j}$, so only one ring is needed to reach a mass of one unit but it takes four rings to have a mass over two units. A total mass of 20 requires s rings where $\sum_{j=1}^{s} \frac{1}{j} \approx \ln(s) \approx 20$, or $s \approx e^{20}$. Because $r \geqslant s$, the $M(r) = 20$ value requires a distance of $r \geqslant e^{20}$.

Because the true $M(r)$ is bounded by $\ln(r)$, and because a cubic equation (Equation (5)) differs drastically from a logarithmic curve, serious predicted distribution errors (demonstrating the inadequacy of Equations (1) and (5)) must be anticipated. As an illustration, Equation (5) would indicate that the total mass is exponentially larger than the mass of, say, the mass of the first four rings (given by $\left(\frac{r_{e^{20}}}{r_4}\right)^3 > \left(\frac{e^{20}}{r_4}\right)^3$), when in fact the correct multiple is less than 10. Also, only 5% of the total mass lies between the e^{19}th and e^{20}th rings. While the distribution of rings for this setting is still being examined, an estimate given by the $M(r) \leqslant \ln(r)$ expression suggests that[2] $\frac{r_{e^{19}}}{r_{e^{20}}} \approx \frac{1}{e}$, which would require Equation (5) to incorrectly predict that about 95% of the total mass is located in this fringe region; note how this error is consistent with the exaggerations of ring masses. Other examples follow by selecting other m_j values; e.g., $m_j = \frac{1}{2kj\ln(j)}$ defines the more extreme $M(r) \leqslant \ln(\ln(r))$ while $m_j = \frac{1}{2kj^2}$ requires $M(r) \leqslant M^* - \frac{1}{r}$.

Rather than being restricted to $M(r) \approx D\, r^3$, problems arise whenever the predicted $M(r)$ form significantly differs from the actual representation. This includes the Equation (2) choice of $M(r) \approx D\, r$. With $m_j = \frac{1}{2jk}$ and the actual distribution given by $M(r) \leqslant \ln(r)$, the linear $M(r) \approx Dr$ must yield seriously distorted predictions that exaggerate the relative amount of mass at a greater distance from the center.

According to the principles of scientific validation, unless new supporting arguments can be found, Equation (1) is not an appropriate way to predict $M(r)$ values for systems of N discrete bodies. Something in addition to, but probably different from, Equation (1) is needed. (Elsewhere a new approach is developed to predict and explain the full structure of rotational velocity curves for N discrete bodies including the sharp, almost linear initial growth; it predicts significantly smaller $M(r)$ values for flattened rotational velocity curves.) In particular, it follows from the above that with discrete dynamics, expressions such as Equation (1) that do not capture tugging effects are inappropriate.

4.2. More General Systems

Although the above suffices for validation purposes, it is reasonable to wonder about other settings; e.g., with three-dimensional N-body solutions or other symmetry settings. The dimensionality of physical space is not an issue; it is whether solutions (with circular rotational values) exist that cause erroneous predictions of mass distributions, such as Dr^3 or even Dr. They exist; as one of many three-dimensional analytic N-body solutions that lead to radically incorrect mass predictions, start with the $2nk + 1$ spiderweb configuration and replace the body at the origin with two bodies of equal mass that periodically oscillate (in opposite directions) along the z-axis, while the configuration rotates as described above. Similarly, central configurations can have a wide assortment of symmetry properties, ranging from collinear to spiderweb settings, so related arguments apply to a wide spectrum of symmetry and asymmetry assumptions about mass distributions.

The reason Equation (1) has these distorted predictions is captured by the difference in dynamics in Figure 1 between discrete systems and continuum approximations where the tugging required by discrete systems is not included, nor even recognized, by the approximations. In particular, a rotating central configuration requires a "tugging" on each body to keep it in a fixed position relative to the other bodies. However, the assumptions leading to Equation (1) fail to include these larger pulling velocities; this allows the Equation (1) predictions for $M(r)$ to radically differ from the precise values.

More generally, anticipate exaggerated Equation (1) predictions whenever individual rotational velocities experience tugging effects. This is due to the observed circular velocity being determined by actual $M(r)$ values *plus* the tugging of other bodies; the increased velocity caused by the pulling forces the exaggerations in Equation (1). Conversely, this effect can be minimal with a limited number of bodies and masses so small that the system can be decoupled into individual "two-body problems," such as in our planetary system where a tugging effect (e.g., as captured by the change in the perihelion of Mercury) is minimal. (Two body systems are consistent with Equation (1).) Indeed, an expression similar to Equation (1) enjoys success in determining the Sun's mass based on planetary motions. However, whenever a decoupling is inappropriate (e.g., with larger mass values and many bodies), expect Newtonian connections to exaggerate $M(r)$ predictions.

A remaining concern is whether there are settings (other than decouplings) of N discrete bodies where Equation (1) might be valid. This suggests searching for N-body solutions with the general properties used to derive Equation (1), but without the tugging effects. As indicated next, they cannot be expected to exist.

Central to Equation (1) is that all bodies have circular orbits; i.e., each body's force is directed toward the center of mass to define

$$\boldsymbol{r}_j'' = \lambda_j(t)\boldsymbol{r}_j. \tag{6}$$

The circular orbit assumption requires the moment of inertia, $I = \sum m_j r_j^2$, to have a constant value. A longstanding conjecture (Saari 1970, 2005) is that a constant I requires the system to rotate as a rigid body, which ensures Equation (1) predictions of Dr^3. A simpler conjecture (a special case) is that if each orbit in a system has a uniform circular motion, then the system rotates as a rigid body.[3] Both conjectures have been verified for any collinear N-body problem (Diacu et al. 2005; Saari 2005 (Section 2.4.2)), and for the general three-body problem (Moeckel 2005, 2008). Moreover, the conjectures can be expected to hold in general; e.g., the first one holds for generic settings of N-body systems (Schmah & Stoica 2007).

[2] Should this not be true, then $r_{e^{20}}$ has a much larger value than used with the first examples of this paragraph, which could cause even larger error multiples.

[3] Intuition supporting the simpler conjecture is that Equation (6) requires the particles to be symmetrically located with respect to each other. However, circular orbits destroy this symmetry if the angular rotation rates do not agree. If angular rotation rates agree, the particles are in a rigid body rotation.

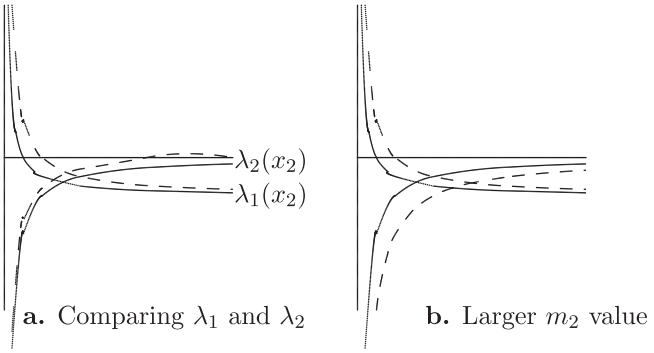

Figure 3. Comparing λ values.

(For instance, for more general N-body modelings that include perturbations.)

Thus, either the assumptions that are central in developing Equation (1) do not apply to discrete systems of N bodies, or the associated N-body solution can be expected to be a rotating rigid body causing exaggerated $M(r)$ predictions. If the basic assumptions used to derive Equation (1) fail to hold for discrete systems, then new arguments are required to justify using something resembling Equation (1) with these systems. For instance, while these arguments cast serious doubts about the derivation of Equation (2), the equation may be reasonably correct. However, to take results based on Equation (1) seriously requires developing an alternative validation argument.

This research is part of a program supported by NSF DMI-0640817 and CMMI-1016785. My thanks for comments made after several colloquia, general public, and conference presentations of these results and to G Hazelrigg for his continued interest in this project. Also, my thanks to a referee for useful suggestions.

APPENDIX

Proof of Theorem 1. The theorem is proved by Maxwell for $n = 1$ and Moulton (1910) for $k = 1$. The following supplies details for the arguments developed in Section 3.7.2 of Saari (2005) and Saari (2011).

Symmetry requires the particles on each circle to satisfy $\lambda_j \boldsymbol{r}_j = \frac{1}{m_j}\frac{\partial U}{\partial \boldsymbol{r}_j}$ where λ_j is the same for all particles on the jth circle; λ_j is a continuous function of masses and the distances between particles and circles; its sign is determined by the force direction. As described in the comments following the statement of Theorem 1, for given masses, the goal is to show that there are distances between circles where all λ_j agree.

Let r_1 be fixed. Maxwell's case is $n = 1$ where $\lambda_1 = \lambda^* < 0$. For $n = 2$, let $r_2 = r_1 + x_2$. To prove there is a unique $x_2 = \overline{x}_2$ where $\lambda_1(\overline{x}_2) = \lambda_2(\overline{x}_2)$ (defining a central configuration), properties of the $\lambda_1(x_2)$ and $\lambda_2(x_2)$ curves (roughly depicted by the solid Figure 3(a) curves) are used; they ensure that the two curves cross at an unique point where the λ_j are equal.

All particles on the second circle are attracted to the center of mass, so $\lambda_2(x_2) < 0$ for all $x_2 > 0$. The critical force terms in the analysis are determined by the distance between circles. As $x_2 \to 0$, the second circle approaches the first, creating an infinitely large attraction of particles on each circle toward its neighbor on the other circle; i.e., $\lambda_2(x_2) \to -\infty$ and $\lambda_1(x_2) \to \infty$ as $x_2 \to 0$. Because $x_2 \to \infty$ represents where the second circle is pushed away from the first, the magnitude of $\frac{1}{m_2}\frac{\partial U}{\partial \boldsymbol{r}_2}$ approaches zero, so $\lambda_2(x_2) \to 0$ in a monotonic manner as $x_2 \to \infty$. The gravitational effect that the second circle has on the first decreases to zero, reducing the problem to Maxwell's setting, so $\lambda_1(x_2) \to \lambda_1^* < 0$. The curves' properties (depicted in Figure 3(a) as solid curves) require them to cross at a unique point to share a common $\lambda < 0$ value; this completes the proof for $n = 2$.

A modification of this argument handles any number of circles. With $n \geqslant 3$, let $r_j = r_{j-1} + x_j$, $j = 2, ..., n$. The above is repeated to show, for any choice of positive $x_3, ..., x_n$, that there exists a point $\boldsymbol{x}_2(x_3, ..., x_n)$ where $\lambda_1(\boldsymbol{x}_2(x_3, ..., x_n)) = \lambda_2(\boldsymbol{x}_2(x_3, ..., x_n))$. These $(n-2)$ extra circles exert an outward pull on the first two, so $\lambda_j(x_2; x_3, ... x_n) > \lambda_j(x_2)$, $j = 1, 2$ (dashed curves in Figure 3 (a)). For fixed $x_3, ..., x_n$, $\lambda_1(x_2; x_3, ..., x_n)$ is monotonically decreasing, it approaches ∞ as $x_2 \to 0$ (the second circle comes arbitrarily close to the first), and it approaches $\lambda_1^* < 0$ as $x_2 \to \infty$ (all circles are pushed infinitely far from the first circle). For fixed $x_3, ..., x_n$ and the reasons specified for $\lambda_2(x_2)$, $\lambda_2(x_2; x_3, ..., x_n)$ approaches $-\infty$ as $x_2 \to 0$ and zero as $x_2 \to \infty$. Function $(r_1 + x_2)\lambda_2(x_2; x_3, ..., x_n) = \frac{1}{m_2}\frac{\partial U}{\partial r_2}$ is monotonically increasing and approaches a positive value (because, for fixed x_3, the third circle pulls particles outward on the second circle). Thus $\lambda_2(x_2; x_3, ..., x_n)$ is monotonic to a positive value and then decreases to approach zero.

The structure of the curves requires them to intersect to define a point $\boldsymbol{x}_2(x_3, ..., x_n)$ where $\lambda_1(\boldsymbol{x}_2(x_3, ..., x_n) = \lambda_2(\boldsymbol{x}_2(x_3, ..., x_n)$ for all $x_3, ..., x_n$. In particular, the common λ value is positive for sufficiently small x_3 values (due to the dominance of the third circle causing $\lambda_2(x_2; x_3, ..., x_n)$ to have a large positive value) and approach the negative $\lambda_1(\overline{x}_2) = \lambda_2(\overline{x}_2)$ value as $x_3 \to \infty$ (the third and other circles become infinitely far from the first two). Because of the smooth foliation formed by the $\lambda_j(\boldsymbol{x}_2(x_3, ..., x_n))$ curves and the analytic structure of the equations, $\boldsymbol{x}_2(x_3, ..., x_n)$ is a smooth function. If $x_3^* > x_3$ (the outer circles are pushed farther away from the first two), $\lambda_j(\boldsymbol{x}_2(x_3, ..., x_n) > \lambda_j(\boldsymbol{x}_2(x_3^*, ..., x_n))$. In Figure 3(a), this λ_j curve is between the corresponding solid and dashed lines, so the intersection defines a smaller common λ value; that is, the common $\lambda(\boldsymbol{x}_2(x_3, ..., x_n))$ value is monotonically decreasing in x_3.

For induction, assume that up to the $(j-1)$th circle, there is a $\boldsymbol{x}_{j-1}(x_j; x_{j+1}, ..., x_n)$ on which the first $(j-1)\lambda$ functions agree, the common $\lambda(\boldsymbol{x}_{j-1}(x_j; x_{j+1}, ..., x_n))$ has a positive value for sufficiently small x_j values, and it is monotonically decreasing to a negative value as $x_j \to \infty$. The $\lambda_j(\boldsymbol{x}_{j-1}(x_j; x_{j+1}, ..., x_n))$ function (with fixed $x_{j+1}, ..., x_n$) has the same properties as $\lambda_2(x_2; x_3, ..., x_n)$; it approaches $-\infty$ as $x_j \to 0$ and (unless $j = n$ is the last circle) monotonically increases to a positive value and then decreases to zero as $x_j \to \infty$. For sufficiently large x_{j+1} values (so the outer circles are pushed far away), $\lambda_j(\boldsymbol{x}_{j-1}(x_j; x_{j+1}, ..., x_n))$ approaches a negative valued function that approaches zero. The properties of these curves force them to intersect at a unique point. Unless $j = n$, the common intersection point has an arbitrary large positive value as $x_j \to 0$ (where the outer circles approach the inner ones) and a negative value as $x_j \to \infty$. For $j = n$, the proof is completed.

Distances between circles. Increasing a mass value, such as $m_2^* > m_2$, increases the force (with the stronger attraction of other bodies on the same circle), which moves the λ_2 curve downward as represented by the dashed line in Figure 3(b). Similarly, this larger mass value has a stronger attraction on the inner circle, which causes larger values for the λ_1 curve as depicted in Figure 3(b). Thus the separation value (where the dashed λ_j curves cross) between circles has a larger value. This captures the expected fact that smaller masses decrease separation distances while larger masses increase them.

The distances between circles for specified m_j values are difficult to compute, but they can be estimated by following the above proof; that is, compute λ_i values for specified m_j and r_j, and then adjust m_j and r_j values so that the λ_i values come closer to agreeing. With the symmetry, the particles along any specified ray can be used to determine the λ_i values; the positive x-axis is chosen. For the ith particle at distance r_i, the contribution to λ_i from other particles on the x-axis is given by $\gamma_i(0)$ where $r_i\gamma_i(0) = \sum_{j=1}^{n} \frac{-m_j}{[r_j + r_i]^2} + \sum_{j=1, j\neq i}^{n} \frac{m_j(r_j - r_i)}{|r_j - r_i|^3}$; the first summation involves particles on the negative x-axis and the second involves bodies on the positive x-axis.

With symmetry, where the force component in the y direction for a particle on the positive y-axis is canceled by its partner on the negative y-axis, the force on the ith particle by bodies on the y-axis is strictly along the x-axis. The contribution to λ_i is given by $\gamma_i(\frac{\pi}{2})$ where $r_i\gamma_i(\frac{\pi}{2}) = 2\sum_{j=1}^{n} \frac{-m_j r_i}{[r_j^2 + r_i^2]^{3/2}}$. Accompanying each line defined by θ, $0 < \theta < \frac{\pi}{2}$, is a line defined by $-\theta$ where symmetry forces components in the y direction to cancel, so the contribution to λ_i from this pair of lines is $\gamma_i(\theta)$ where

$$r_i\gamma_i(\theta) = 2\sum_{j=1}^{n} \frac{-m_j\left(r_j\cos(\theta) + r_i\right)}{\left[r_j^2 + r_i^2 + 2r_i r_j\cos(\theta)\right]^{3/2}} + 2\sum_{j=1}^{n} \frac{m_j\left(r_j\cos(\theta) - r_i\right)}{\left[r_j^2 + r_i^2 - 2r_i r_j\cos(\theta)\right]^{3/2}}. \quad (7)$$

For $k = 1$, $\lambda_i = \gamma_i(0)$; for $k = 2$, $\lambda_i = \gamma_i(0) + \gamma_i(\frac{\pi}{2})$; and for $k \geqslant 3$, $\lambda_i = \gamma_i(0) + \sum_{s=1}^{k-2} \gamma_i(\frac{s\pi}{k})$.

To illustrate with $m_j = 1$, if $x_j = r_1 = 1$ for $j = 2, ..., n$, then $\lambda_1 = 1 - \frac{1}{n^2} - \frac{1}{(n+1)^2} > 0$, and

$$\lambda_i = \frac{1}{i^3} - \sum_{k=2i-n-1}^{-n} \frac{1}{i(i-k)^2} \text{ for } 2i \leqslant n;$$
$$\sum_{k=2i-n-1, k\neq 0}^{-n} \frac{-1}{i(i-k)^2} \text{ for } 2i > n. \quad (8)$$

This expression cancels terms that are α rings to the right of r_i with terms that are α rings to the left; Equation (8) shows those terms without a canceling mate. (With decreasing $m_j > m_{j+1}$, the zeros from cancellations are replaced with negative values such as $\frac{-m_{i-\alpha} + m_{i+\alpha}}{\alpha^2}$.) These λ_i values, particularly $\lambda_1 > 0$, require a wider spacing.

Indeed, a larger $x_j = 2r_1$ leads to $\lambda_1 = -\sum_{s=1}^{n} \frac{1}{(2sr_1)^2} + \sum_{s=2}^{n} \frac{1}{[2(s-1)r_1]^2} = -\frac{1}{(2nr_1)^2}$, with $\lambda_i = \frac{-1}{(2i-1)r_i}\sum_{s=1}^{2i-1} \frac{1}{[2(n-i+s)r_1]^2}$. The relative closeness of these λ_i values, where $0 > \lambda_i > \lambda_{i+1}$ for $i = 1, ..., n = 1$, means that slightly larger separations create a central configuration. (Note the associated λ, and hence D, value is "small" for this configuration.) The actual distribution is $M(r) \approx ar$ where constant $a \approx \frac{1}{2}$; the resulting larger $r_{100} \approx 200$ increases some of the discrepancies described in Section 4.1. This change from $x_j = r_1$ to $x_j = 2r_1$ shows how the λ_i values are reasonably sensitive to changes in spacing distances. Indeed, the equations show that changes in λ_i are more sensitive to changes in the distances of circles near the ith circle.

To illustrate Equation (7) with $k = 3$ (so $\theta = \pi/3$) and $x_i = r_1 = 1$, the general expression is

$$\gamma_i\left(\frac{\pi}{3}\right) = \sum_{k=i+1}^{n} \frac{-3}{\left(k^2 + i^2 - ki\right)^{3/2}} + \sum_{k=1}^{i} \frac{k - 2i}{i\left(k^2 + i^2 - ki\right)^{3/2}} + \sum_{k=n-i+1}^{n} \frac{-(k+2i)}{i\left(k^2 + i^2 + ki\right)^{3/2}}. \quad (9)$$

Each term in each summation is negative. The first summation combines the term from the second summation of Equation (7) that is α rings beyond the ith with the term in the first summation representing the αth ring. The second and third summations are remaining not-matched terms. The λ_i values for the $k = 3$ setting and $N = 6n$ body problem is the sum of $\gamma_i(\frac{\pi}{3})$ and λ_i from Equation (8); all λ_i values are now negative and their closer values indicate that slight changes (e.g., m_2 must come slightly inwards) in distance will generate the central configuration. That is, $r_1 \approx 1$ and, while the spacing is not regular, an estimate for large n values is that $r_k \approx \frac{k}{n} r_n$, which is what one would expect from $M(r) \approx ar$ for constant $0 < a \leqslant 1$; the smaller the a value, the larger the r_{100} value, which could cause larger prediction discrepancies.

Note how changes in n and k alter the λ (and D) value. Larger k values, for instance, create a denser concentration of mass nearer the origin, which exerts an effect that is similar to increasing mass values for particles closer to the center in the collinear setting. Also, $|\lambda|$ has a larger value.

REFERENCES

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (2nd ed.; Princeton, NJ: Princeton Univ. Press)

Diacu, R., Pérez-Chavela, E., & Santoprete, M. 2005, Trans. Amer. Math. Soc., 357, 4215

Moeckel, R. 2005, Trans. Amer. Math. Soc., 357, 3105

Moeckel, R. 2008, Discrete and Continuous Dynamical Systems Series S, 1, 631

Moulton, F. R. 1910, AnMat, 12, 1

Ostriker, J. P., Peebles, P. J. E., & Yahil, A. 1974, ApJ, 193, 1

Rubin, V. C. 1983, Sci, 220, 1339

Rubin, V. C. 1993, PNAS, 90, 4814

Saari, D. G. 1970, in Periodic Orbits, Stability, and Resonances, ed. G. Giacaglia (Dordrecht: Reidel), 76

Saari, D. G. 2005, Collisions, Rings, and Other Newtonian N-body Problems (Providence, RI: American Mathematical Society)

Saari, D. G. 2011, in Expeditions in Mathematics, ed. D. Hayes, & T. Shubin (Boston, MA: Mathematical Association of America), 283

Schmah, T., & Stoica, C. 2007, Trans. Amer. Math. Society, 359, 4429

Winter, A. 1941, The Analytic Foundations of Celestial Mechanics (Princeton, NJ: Princeton Univ. Press)