Snippets of Physics

10. Thermodynamics of Self-Gravitating Particles

T Padmanabhan

The statistical mechanics of a system of particles interacting through gravity leads to several counter-intuitive features. We explore one of them, called Antonov instability, in this installment.

T Padmanabhan works at IUCAA, Pune and is interested in all areas of theoretical physics, especially those which have something to do with gravity.

Suppose we put a large number (N) of particles, each of mass m and interacting through a two-body potential $U(\mathbf{x}-\mathbf{y})$ into a spherical box of radius R. We will arrange matters such that the particles move randomly to start with and bounce off the surface of the sphere elastically. Let the total energy of the system be E which, of course, will remain a constant. We are interested in the behaviour of the system at late times, when the particles will have had sufficient time to interact with each other and exchange energy.

The result will clearly depend on the nature of the interaction, specified by $U(\mathbf{x}-\mathbf{y})$ as well as the other parameters. If $U(\mathbf{x}-\mathbf{y})$ is a short range potential representing intermolecular forces and if E is sufficiently high, then the system will relax towards a Maxwellian distribution of velocities and nearly uniform density in space. (The velocity distribution will have a characteristic temperature $T \simeq 2E/3N$ and we are assuming that this is higher than the 'boiling point' of the 'liquid' made of these particles. If not, the eventual equilibrium state will be a mixture of matter in liquid and vapour state. Also note that we use units with $k_B = 1$ throughout.) All this is part of standard lore in statistical mechanics.

Keywords
Thermodynamics, gravitation.

What happens if the $U(\mathbf{x}-\mathbf{y})$ is due to gravitational interaction of the particles? What are the different phases in which matter can exist in such a case? I will discuss some of the peculiar effects that arise in this context.

To do this, let us begin by quickly reviewing the way one introduces the equilibrium configuration in statistical mechanics. Consider a system described by a distribution function $f(\mathbf{x}, \mathbf{p}, t)$ such that $f \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{p}$ denotes the total mass in a small phase space volume. We assume that the evolution of the distribution function is given by some equation (usually called the Boltzmann equation) of the form $\mathrm{d}f/\mathrm{d}t = C(f)$. The term $C(f)$ on the right hand side describes the effect of collisions. While the precise form of $C(f)$ can be quite complicated, we can usually assume that the collisional evolution of f, driven by $C(f)$, satisfies two reasonable conditions:
(a) The total mass and energy of the system are conserved and (b) the mean field entropy, defined by

$$S = -\int f \,\ln\, f\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{p} \tag{1}$$

does not decrease (and in general increases).[1] For any such system, we can obtain the equilibrium form of f by extremising the entropy while keeping the total energy and mass constant using two Lagrange multipliers. This is a standard exercise in statistical mechanics and the resulting distribution function is the usual Boltzmann distribution governed by:

$$f(\mathbf{x}, \mathbf{v}) \propto \exp\left[-\beta\left(\frac{1}{2}v^2 + \phi\right)\right];\ \phi(\mathbf{x}) = \int \mathrm{d}^3\mathbf{y}\, U(\mathbf{x}, \mathbf{y})\rho(\mathbf{y}). \tag{2}$$

Integrating over velocities, we get the closed system of equations for the density distribution:

$$\begin{aligned} \rho(\mathbf{x}) &= \int \mathrm{d}^3\mathbf{v}\, f = A \exp(-\beta\phi(\mathbf{x})); \\ \phi(\mathbf{x}) &= \int \mathrm{d}^3\mathbf{y}\, U(\mathbf{x}, \mathbf{y})\rho(\mathbf{y}). \end{aligned} \tag{3}$$

The final result is quite understandable: It is just the Boltzmann factor for the density distribution: $\rho \propto \exp(-\beta V)$,

[1] For those who are unfamiliar with this expression, here is a recap: In the standard derivation of Boltzmann distribution, one extremises the function $S = -\sum n_i \ln n_i$ of the occupation numbers n_i subject to the constraint on total energy and number. In the continuuum limit one works with f rather than n_i and the summation over i becomes an integral over the phase space leading to (1).

where V is the potential energy at a given location due to the distribution of particles. One could have almost written this down 'by inspection'! (See Appendix for more details)

One could have almost written down equation (3) 'by inspection'.

Everything that we have said so far is independent of the nature of the potential U (except for one important caveat which we will discuss right at the end). In the case of gravitational interaction, (3) becomes:

$$\rho(\mathbf{x}) = A\exp(-\beta\phi(\mathbf{x}));\ \phi(\mathbf{x}) = -G\int\frac{\rho(\mathbf{y})\mathrm{d}^3\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}. \quad (4)$$

The integral equation (4) for $\rho(\mathbf{x})$ can be easily converted to a differential equation for $\phi(\mathbf{x})$ by taking the Laplacian of the second equation – leading to $\nabla^2\phi = 4\pi G\rho$ – and using the first equation. We then get, for the spherically symmetric case, the isothermal sphere equation:

$$\nabla^2\phi = \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}\right) = 4\pi G\rho_c e^{-\beta[\phi(r)-\phi(0)]}. \quad (5)$$

The constants β and ρ_c (the central density) have to be fixed in terms of the total number (or mass) of the particles and the total energy. Given the solution to this equation, which represents an extremum of the entropy, all other quantities can be determined. As we shall see, this system shows several peculiarities.

To analyse (5), it is convenient to introduce length, mass and energy scales by the definitions

$$L_0 \equiv (4\pi G\rho_c\beta)^{1/2}, \quad M_0 = 4\pi\rho_c L_0^3, \quad \phi_0 \equiv \beta^{-1} = \frac{GM_0}{L_0} \quad (6)$$

All other physical variables can be expressed in terms of the dimensionless quantities $x \equiv (r/L_0)$, $n \equiv (\rho/\rho_c)$, $m = (M(r)/M_0)$, $y \equiv \beta[\phi - \phi(0))]$, where $M(r)$ is the mass inside a sphere of radius r. These variables satisfy the equations:

$$y' = m/x^2; \quad m' = nx^2; \quad n' = -mn/x^2. \tag{7}$$

In terms of $y(x)$ the isothermal equation, (5), becomes

$$\frac{1}{x^2}\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\frac{\mathrm{d}y}{\mathrm{d}x}\right) = \mathrm{e}^{-y} \tag{8}$$

with the boundary condition $y(0) = y'(0) = 0$.[2] Let us consider the nature of the solutions to this equation.

[2] We have assumed that the system is spherically symmetric; it turns out that this is indeed the extremal solution.

By direct substitution, we see that $n = (2/x^2), m = 2x, y = 2\ln x$ satisfy (7) and (8). This solution, however, is singular at the origin and hence is not physically admissible. The importance of this solution lies in the fact that – as we will see – all other (physically admissible) solutions tend to this solution [1, 2] for large values of x. This asymptotic behavior of all solutions shows that the density decreases as $(1/r^2)$ for large r implying that the mass contained inside a sphere of radius r increases as $M(r) \propto r$ at large r. Of course, in our case, the system is enclosed in a spherical box of radius R with a given mass M.

Equation (8) is invariant under the transformation $y \to y + a$; $x \to kx$ with $k^2 = \mathrm{e}^a$. This invariance implies that, given a solution with some value of $y(0)$, we can obtain the solution with any other value of $y(0)$ by simple rescaling. Therefore, only one of the two integration constants needed in the solution to (8) is really nontrivial. Hence it must be possible to reduce the degree of the equation from two to one by a judicious choice of variables. One such set of variables is:

$$v \equiv \frac{m}{x}; \quad u \equiv \frac{nx^3}{m} = \frac{nx^2}{v}. \tag{9}$$

In terms of v and u, (5) becomes

$$\frac{u}{v}\frac{\mathrm{d}v}{\mathrm{d}u} = -\frac{(u-1)}{(u+v-3)}. \tag{10}$$

The boundary conditions $y(0) = y'(0) = 0$ translate into the following: v is zero at $u = 3$, and $(\mathrm{d}v/\mathrm{d}u) = -5/3$ at (3,0). (You can prove this by examining the behaviour of (7) near $x = 0$ retaining up to necessary order in x; try it out!) The solution $v\,(u)$ to equation (10) can be easily obtained numerically: it is plotted in *Figure* 1 as the spiralling curve. The singular points of this differential equation are given by the location in the uv plane at which both the numerator and denominator of the right hand side of (10) vanish. Solving $u = 1$ and $u + v = 3$ simultaneously, we get the singular point to be $u_s = 1$, $v_s = 2$. Using (9), we find that this point corresponds to the asymptotic solution $n = (2/x^2), m = 2x$. It is obvious from the nature of the equation that the solution curve will spiral around the singular point asymptotically approaching the $n = 2/x^2$ solution at large x.

The nature of the solution shown in *Figure* 1 allows us to put interesting bounds on various physical quantities

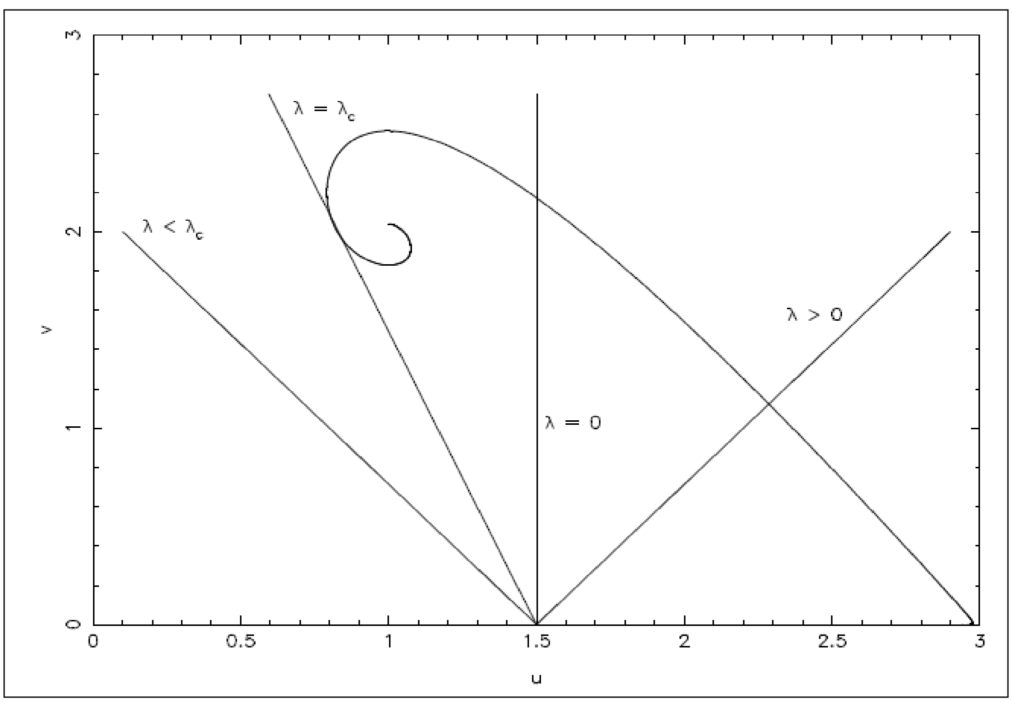

Figure 1. Bound on RE/GM² for the isothermal sphere.

including energy. To see this, we shall compute the total energy E of the isothermal sphere. The potential and kinetic energies are

$$\begin{aligned} U &= -\int_0^R \frac{GM(r)}{r}\frac{\mathrm{d}M}{\mathrm{d}r}\mathrm{d}r = -\frac{GM_0^2}{L_0}\int_0^{x_0} mnx\mathrm{d}x, \\ K &= \frac{3}{2}\frac{M}{\beta} = \frac{3}{2}\frac{GM_0^2}{L_0}m(x_0) = \frac{GM_0^2}{L_0}\frac{3}{2}\int_0^{x_0} nx^2\mathrm{d}x, \end{aligned} \tag{11}$$

where $x_0 = R/L_0$ is the boundary and the expression for K follows from the velocity dependence of f in (2). The total energy is, therefore,

$$\begin{aligned} E &= K + U = \frac{GM_0^2}{2L_0}\int_0^{x_0} \mathrm{d}x(3nx^2 - 2mnx) \\ &= \frac{GM_0^2}{2L_0}\int_0^{x_0} \mathrm{d}x\frac{\mathrm{d}}{\mathrm{d}x}\{2nx^3 - 3m\} \\ &= \frac{GM_0^2}{L_0}\left\{n_0x_0^3 - \frac{3}{2}m_0\right\}, \end{aligned} \tag{12}$$

where $n_0 = n(x_0)$ and $m_0 = m(x_0)$. The dimensionless quantity (RE/GM^2) is given by

$$\lambda \equiv \frac{RE}{GM^2} = \frac{1}{v_0}\left\{u_0 - \frac{3}{2}\right\}. \tag{13}$$

Note that the combination (RE/GM^2) *is a function only of* (u, v) *at the boundary.* Let us now consider the constraints on λ. Suppose we specify some value for λ by specifying R, E and M. Then such an isothermal sphere *must* lie on the curve

$$v = \frac{1}{\lambda}\left(u - \frac{3}{2}\right); \qquad \lambda \equiv \frac{RE}{GM^2} \tag{14}$$

which is a straight line through the point $(1.5, 0)$ with a slope λ^{-1}. On the other hand, since *all* isothermal spheres must lie on the $u-v$ curve, *an isothermal sphere* *can exist only if the line in equation (14) intersects the $u - v$ curve.*

For large positive λ (positive E) there is just one intersection. When $\lambda = 0$, (zero energy) we still have a unique isothermal sphere. (For $\lambda = 0$, (14) represents a vertical line through $u = 3/2$.) When λ is negative (negative E), the line can cut the $u - v$ curve at more than one point; thus more than one isothermal sphere can exist with a given value of λ. (Of course, specifying M, R, E individually will remove this non-uniqueness). But as we decrease λ (more and more negative E) the line in (14) will slope more and more to the left; and when λ is smaller than a critical value λ_c, the intersection will cease to exist. So we reach the key conclusion that *no isothermal sphere can exist if* (RE/GM^2) *is below a critical value* λ_c. This fact[3] follows immediately from the nature of the $u-v$ curve and (14). The value of λ_c can be found from the numerical solution and turns out to be about -0.335.

[3] This derivation is due to the author [3,1]. It is surprising that Chandrasekhar, who has worked out the isothermal sphere in *u–v* coordinates as early as 1939, missed discovering the energy bound shown in *Figure* 1. Chandrasekhar [2] has the *u–v* curve but does not over-plot lines of constant λ. If he had done that, he would have discovered Antonov instability decades before Antonov did [4].

What does this result mean? To understand its implications, consider constructing such a system with a given mass M, radius R and an energy $E = -|E|$ which is negative. (The last condition means that the system is gravitationally bound.) In this case, $\lambda = RE/GM^2 = -R|E|/GM^2$ is a negative number but let us assume that it is above the critical value; that is, $\lambda > \lambda_c$. Then we know that an isothermal sphere solution exists for the given parameter values. By construction, this solution is the local extremum of the entropy and could represent an equilibrium configuration if it is also a global maximum of entropy.

But for the system we are considering, it is actually quite easy to see that there is no global maximum for entropy. This is because, for a system of point particles interacting via Newtonian potential, there is no lower bound to the gravitational potential energy. If we take

But for the system we are considering, it is actually quite easy to see that there is no global maximum for entropy.

an amount of mass $m < M$ and form a compact core of radius r inside the spherical cavity, then by decreasing r one can supply arbitrarily large amount of energy to the rest of the particles. Very soon, the remaining particles will have very large kinetic energy compared to their gravitational potential energy and will essentially bounce around inside the spherical cavity like a non-interacting gas of particles. The compact core in the center will continue to shrink thereby supplying energy to the rest of the particles. It is easy to see that such a core–halo configuration can have arbitrarily high values for the entropy. All this goes to show that the isothermal sphere cannot be a *global* maximum for the entropy. (This was the caveat in the calculation we performed to derive the isothermal sphere equation; we tacitly assumed that the extremum condition can be satisfied for a finite value of entropy.)

If we increase the radius of the spherical box (with some fixed value for $E = -|E|$), the parameter λ will become more and more negative and for sufficiently large R, we will have a situation with $\lambda < \lambda_c$. Now the situation gets worse. The system does not even have a local extremum for the entropy and will evolve directly towards a core–halo configuration. This is closely related to a phenomenon called Antonov instability [4, 3].

In real life, of course, there is always some short distance cut-off because of which the core cannot shrink to an arbitrarily small radius. In such a case, there is a global maximum for entropy achieved by the (finite) core–halo configuration which could be thought of as the final state in the evolution of such a system. It will be highly inhomogeneous and, in fact, is very similar to a system which exists as a mixture of two phases. This is one key peculiarity introduced by long range attractive interactions in statistical mechanics.

Suggested Reading

[1] T Padmanabhan, *Physics Reports*, Vol.188, p.285, 1990.

[2] S Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Dover ,1939.

[3] T Padmanabhan, *Astrophys. Jour. Supp.*, Vol.71, p.651, 1989.

[4] V A Antonov, *Vest. Leningrad Univ.*, Vol.7, p.135, 1962. Translation: *IAU Symposium*, Vol.113, p.525, 1985.

Address for Correspondence
T Padmanabhan
IUCAA, Post Bag 4
Pune University Campus
Ganeshkhind
Pune 411 007, India.
Email:
paddy@iucaa.ernet.in
nabhan@iucaa.ernet.in

Appendix

In the text, I derived equation (3) from the expression for entropy in equation (1). Given the peculiarities of gravitating systems one may wonder how trustworthy this approach is. Here I describe briefly a more basic derivation of the expression in (3).

The equilibrium state is the one that maximizes the entropy of the system. When we study the system in the microcanonical ensemble, this entropy S is the logarithm of the volume $g(E)$ of the phase space available to the system if the total energy is E. That is:

$$\mathrm{e}^S = g(E) = \frac{1}{N!}\int \mathrm{d}^{3N}x\ \mathrm{d}^{3N}p\ \delta(E-H), \tag{15}$$

where H is the Hamiltonian for the system of N particles given by the sum of the kinetic energies $(\mathbf{p}_i^2/2m)$ $(i = 1, 2, ..., N)$ and the potential energy of pairwise interaction. The Dirac delta function tells us that the system is confined to the boundary of a $3N$ dimensional sphere in momentum space given by the equation

$$\sum_{i=1}^{N}\mathbf{p}_i^2 = 2m\left[E - \frac{1}{2}\sum_{i\neq j}U(\mathbf{x}_i, \mathbf{x}_j)\right] \equiv l^2. \tag{16}$$

Obviously, the momentum integration in (15) will lead to a term proportional to l^{3N-1}; so, performing the momentum integrations and using $N \gg 1$, we get

$$\mathrm{e}^S = g(E) \propto \frac{1}{N!}\int \mathrm{d}^{3N}x\left[E - \frac{1}{2}\sum_{i\neq j}U(\mathbf{x}_i, \mathbf{x}_j)\right]^{\frac{3N}{2}}. \tag{17}$$

The integral in (17) is impossible to evaluate for any realistic potential but there is a standard approximation using which we can map this problem to a more tractable one.

Let us divide the the spatial volume V into J (with $J \ll N$) cells of equal size, large enough to contain many particles but small enough for the potential to be treated as a constant within each cell. (We will assume that such an intermediate scale exists, which usually does.) Instead of *integrating* over all the particle coordinates $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)$ we shall *sum* over the number of particles n_a in the cell centered at $\mathbf{x}_a$ (where $a = 1, 2, ..., J$) thereby approximating the integral by a discrete sum. Using the obvious fact that the integration over $(N!)^{-1}\mathrm{d}^{3N}x$ can be replaced by

$$\sum_{n_1=1}^{\infty}\left(\frac{1}{n_1!}\right)\sum_{n_2=1}^{\infty}\left(\frac{1}{n_2!}\right)...\sum_{n_J=1}^{\infty}\left(\frac{1}{n_J!}\right)\left(\frac{V}{J}\right)^N \tag{18}$$

(subject to the constraint that $\sum n_a = N$) and approximating the $n!$ s by Stirling's approximation: $\ln n! = n\ln(n/e)$ we get, after some straightforward algebra the result:

$$\mathrm{e}^S \approx \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}...\sum_{n_J=1}^{\infty}\ \exp S[\{n_a\}], \tag{19}$$

where

$$S[\{n_a\}] = \frac{3N}{2}\ln\left[E - \frac{1}{2}\sum_{a\neq b}^{J} n_a U(\mathbf{x}_a,\mathbf{x}_b)n_b\right] - \sum_{a=1}^{J} n_a \ln\left(\frac{n_a J}{\mathrm{e}V}\right). \tag{20}$$

If the number of particles in each cell is sufficiently large – as usually is the case – this is not a drastic approximation. We are, of course, interested in the configuration of $\{n_a\}$ for which the summand in (19) reaches the maximum value, subject to the constraint on the total number. This works in standard statistical mechanics because, in most cases of interest, the largest term actually dominates the sum and the error involved in ignoring the rest is small. That is, to a high order of accuracy $S = S[n_{a,\max}]$, where $n_{a,\max}$ is the solution to the variational problem $(\delta S/\delta n_a) = 0$ with the sum of particle numbers n_a being kept equal to N. Imposing this constraint by a Lagrange multiplier and using the expression (20) for S, we obtain the equation satisfied by $n_{a,\max}$:

$$\frac{1}{T}\sum_{b=1}^{J} U(\mathbf{x}_a,\mathbf{x}_b)n_{b,\max} + \ln\left(\frac{n_{a,\max}J}{V}\right) = \text{constant}, \tag{21}$$

where we have *defined* the temperature T through the relation:

$$\frac{1}{T} = \frac{3N}{2}\left(E - \frac{1}{2}\sum_{a\neq b}^{J} n_a U(\mathbf{x}_a,\mathbf{x}_b)n_b\right)^{-1} = \beta \tag{22}$$

with $n_a = n_{a,\max}$. To see that this is not as strange as it looks, you only need to note, from (20) that this β is also equal to $(\partial S/\partial E)$; therefore one can think of T as the correct thermodynamic temperature. We can now return back to the continuum limit of (21) by writing $n_{a,\max}(J/V) = \rho(\mathbf{x}_a)$ and replacing the sum over particles by integration with the measure J/V. In this continuum limit, the extremum solution in (21) is given precisely by (3).