COMAP

R4R

Refresh for Relevance

WWW.COMAP.ORG/R4R

The Consortium for Mathematics and Its Applications (COMAP)
175 Middlesex Turnpike, Suite 3B
Bedford, MA 01730

Published and distributed by

COMAP, Inc.
www.comap.org

2025 Disclaimer

This reprinted publication contains some information that may be outdated or no longer accurate, such as web links and contact details. For the most up-to-date information about COMAP and its publications, please visit: www.comap.org.

The text of this publication, or any part thereof, is made available as part of the Refresh for Relevance (R4R) project. More information about the project can be found at

www.comap.org/R4R

UMAP

Modules in Undergraduate Mathematics and Its Applications

Published in cooperation with

The Society for Industrial and Applied Mathematics,

The Mathematical Association of America,

The National Council of Teachers of Mathematics,

The American Mathematical Association of Two-Year Colleges,

The Institute for Operations Research and the Management Sciences, and

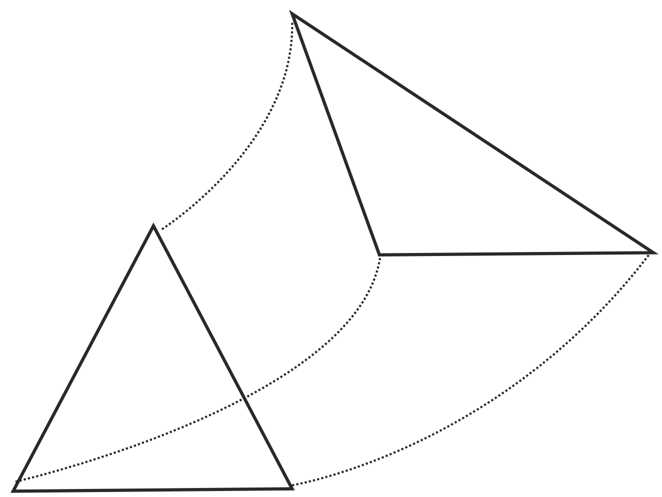
The American Statistical Association.

COMAP

Module 789

Liouville's Theorem in Dynamics

Paul Atsusi Isihara
Benjamin Noonan



Application Field: Physics and Astronomy

COMAP, Inc., Suite 210, 57 Bedford Street, Lexington, MA 02420 (781) 862–7878

INTERMODULAR DESCRIPTION SHEET:	UMAP Unit 789
TITLE:	Liouville's Theorem in Dynamics

AUTHOR:	Paul Atsusi Isihara
	Dept. of Mathematics
	Wheaton College
	Wheaton, IL 60187
	`Paul.A.Isihara@wheaton.edu`
	
	Benjamin Noonan
	`Ben.J.Noonan@wheaton.edu`
	Nathaniel Stapleton
	`Nat.J.Stapleton@wheaton.edu`

MATHEMATICAL FIELD:	Calculus of variations, dynamical systems
APPLICATION FIELD:	Physics, astronomy
TARGET AUDIENCE:	Undergraduate math majors who have completed a basic calculus and physics sequence.
ABSTRACT:	The purpose of this Module is to give an elementary introduction to Liouville's theorem in dynamics. Well-known to physicists, and in particular, fundamental within statistical mechanics, this theorem has many applications, including the focusing of charged particle beams by accelerators and the determination of information about galactic systems. We begin the Module by providing pertinent background material on Euler's equation in the calculus of variations (Section 2) and on Hamiltonian dynamics and phase space (Section 3). We then derive Liouville's theorem in two dimensions (Section 4). An example of three falling balls is given to help visualize an important fact related to Liouville's theorem, namely, that an energy-conserving flow through phase space is incompressible (Section 5). We conclude the Module with further discussion of two of the theorem's applications (Section 6).
PREREQUISITES:	A basic knowledge of physics and of multivariable calculus.

UMAP/ILAP Modules 2002–03: Tools for Teaching, 49–67.

COMAP, Inc., Suite 210, 57 Bedford Street, Lexington, MA 02420
(800) 77-COMAP = (800) 772-6627, or (781) 862-7878; `http://www.comap.com`

Liouville's Theorem in Dynamics

Paul Atsusi Isihara
Dept. of Mathematics
Wheaton College
Wheaton, IL 60187
Paul.A.Isihara@wheaton.edu

Benjamin Noonan
Ben.J.Noonan@wheaton.edu

Nathaniel Stapleton
Nat.J.Stapleton@wheaton.edu

Table of Contents

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS (UMAP) PROJECT

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications, to be used to supplement existing courses and from which complete courses may eventually be built.

The Project was guided by a National Advisory Board of mathematicians, scientists, and educators. UMAP was funded by a grant from the National Science Foundation and now is supported by the Consortium for Mathematics and Its Applications (COMAP), Inc., a nonprofit corporation engaged in research and development in mathematics education.

Paul J. Campbell	Editor
Solomon Garfunkel	Executive Director, COMAP

1. Introduction

We give a concrete introduction to an important theorem in dynamical systems known as *Liouville's theorem*. Well-known to physicists and fundamental in the field of statistical mechanics, this theorem has been widely applied—for instance, to study the focusing of charged particle beams by accelerators and to determine the potential function of the galactic gravitational field from the distribution of stars [Marion and Thornton 1995].

We begin by providing pertinent background on Euler's equation in the calculus of variations (**Section 2**) and on Hamiltonian dynamics and phase space (**Section 3**). We then derive Liouville's theorem in two dimensions (**Section 4**), the more general $2N$-dimensional case being proved in a similar manner. The dynamics of three freely-falling balls helps to visualize an important fact related to Liouville's theorem, namely, the incompressibility of an energy-conserving flow through phase space (**Section 5**). We conclude with further mention of the theorem's applications (**Section 6**).

2. Euler's Equation from the Calculus of Variations

Hamilton's principle, a fundamental insight in basic classical dynamics, posits that

> an object moves in such a way that the time integral of the difference between its kinetic and potential energies is minimized
>
> [Marion and Thornton 1995].

We explain how the calculus of variations is used to minimize such integrals.

First, consider an integral of the form

$$J = \int_{t_1}^{t_2} f\{y(t), \dot{y}(t); t\}\, dt,$$

where y and $\dot{y} = dy/dt$ both depend on t. Our goal is to find a function $y = y_0(t)$ that minimizes J. For example, let us take $f\{y, \dot{y}; t\} = \sqrt{1 + (\dot{y})^2}$, $t_1 = 0, t_2 = 1$, so that

$$J = \int_0^1 \sqrt{1 + (\dot{y})^2}\, dt.$$

If we specify further that $y(0) = 0$ and $y(1) = 1$, then J gives the length of the graph of a differentiable function $y = f(t)$ that passes through the points $(0, 0)$ and $(1, 1)$. In this case, we know that the function that minimizes J must be $y_0(t) = t$, since the shortest distance between two points is a straight line.

To arrive at this conclusion by means of a calculus-of-variations approach, we begin by setting

$$y = y(\alpha, t) = y_0(t) + \alpha\eta(t).$$

Here $\eta(t)$ is an arbitrary differentiable function defined on $t_1 \leq t \leq t_2$ that satisfies

$$\eta(t_1) = \eta(t_2) = 0.$$

For every real number α, the function y agrees with the unknown optimizing function $y_0(t)$ at the endpoints $t = t_1$ and $t = t_2$, but it may vary from $y_0(t)$ in the interval $t_1 < t < t_2$. Note that y and $\dot{y}$ are now functions of both t and α, with $\partial y/\partial\alpha = \eta(t)$ and $\partial\dot{y}/\partial\alpha = \eta'(t)$. The integral J must also depend on α:

$$J = J(\alpha) = \int_{t_1}^{t_2} f\{y(\alpha,t), \dot{y}(\alpha,t); t\}\, dt,$$

and we have

$$\begin{aligned} \frac{\partial J}{\partial\alpha} &= \int_{t_1}^{t_2} \frac{\partial}{\partial\alpha}\left[f\{y,\dot{y};t\}\right]\, dt \\ &= \int_{t_1}^{t_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial\alpha} + \frac{\partial f}{\partial\dot{y}}\frac{\partial\dot{y}}{\partial\alpha}\right) dt \\ &= \int_{t_1}^{t_2} \left(\frac{\partial f}{\partial y}\,\eta(t) + \frac{\partial f}{\partial\dot{y}}\eta'(t)\right)\, dt. \end{aligned}$$

Using integration by parts, we get

$$\frac{\partial J}{\partial\alpha} = \int_{t_1}^{t_2} \frac{\partial f}{\partial y}\eta(t)\, dt + \eta(t)\frac{\partial f}{\partial\dot{y}}\bigg|_{t_1}^{t_2} - \int_{t_1}^{t_2} \eta(t)\frac{d}{dt}\left[\frac{\partial f}{\partial\dot{y}}\right]\, dt.$$

Because $\eta(t_1) = \eta(t_2) = 0$, the second term drops out; and after simplification, we have

$$\frac{\partial J}{\partial\alpha} = \int_{t_1}^{t_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dt}\left[\frac{\partial f}{\partial\dot{y}}\right]\right)\eta(t)\, dt.$$

Since J has a minimum at $\alpha = 0$, $\partial J/\partial\alpha$ must equal zero when $\alpha = 0$. Since the function $\eta(t)$ is arbitrary, we must have

$$\frac{\partial f}{\partial y} - \frac{d}{dt}\left[\frac{\partial f}{\partial\dot{y}}\right] = 0. \qquad \textit{(Euler's equation).}$$

Euler's equation gives a necessary condition for J to have a minimum.

For the arclength example, $f = \sqrt{1 + (\dot{y})^2}$ and $\partial f/\partial y = 0$, so Euler's equation specifies that

$$\frac{d}{dt}\left[\frac{\partial f}{\partial\dot{y}}\right] = \frac{d}{dt}\left[\frac{\dot{y}}{\sqrt{1+(\dot{y})^2}}\right] = 0,$$

or

$$\frac{\dot{y}}{\sqrt{1+(\dot{y})^2}} = c \qquad \text{(where } c \text{ is a constant).}$$

Solving for $\dot{y}$, we find that

$$\dot{y} = \sqrt{\frac{c^2}{1-c^2}} = \text{constant},$$

and hence y must be linear.

Exercise

1. Use Euler's equation to find the function y that minimizes

$$J = \int_{t_1}^{t_2} \left(\tfrac{1}{2}\dot{y}^2 + y\right) dt,$$

where y satisfies the initial conditions $y(0) = y_0$, $\dot{y}(0) = \dot{y}_0$.

3. Hamiltonian Dynamics and Phase Space

If an object's motion is constrained to one dimension, the position of the object at time t is given by an ordinary real valued function $y(t)$. In Newtonian dynamics, the motion of the object is analyzed by consideration of the total force acting on the object and Newton's Second Law:

$$\text{Force} = \text{mass} \cdot \text{acceleration} = m\ddot{y}(t).$$

In Hamiltonian dynamics, the object's motion is obtained by consideration of total energy rather than total force. For simplicity, we assume that the object's kinetic energy T is a function of the object's position $y(t)$ and momentum $m\dot{y}$ (the object's mass m is a constant). The object's potential energy, on the other hand, depends only on its position $y(t)$. Thus, we may write

$$T = T(y, \dot{y}), \qquad U = U(y).$$

In particular, we assume that neither T nor U has a direct dependence on time t. Hamilton's principle tells us that the object "moves in such a way that the time integral of the difference between its kinetic and potential energies is minimized" [Marion and Thornton 1995]. In other words, the integral $J = \int_{t_1}^{t_2} L\, dt$, where $L = T - U$, is minimized (L is called the *Lagrangian*). Taking $f\{y, \dot{y}; t\} = L(y, \dot{y}, t) = T(y, \dot{y}) - U(y)$, Euler's equation implies that

$$\left(\frac{\partial T}{\partial y} - \frac{dU}{dy}\right) - \frac{d}{dt}\left[\frac{\partial T}{\partial \dot{y}}\right] = 0,$$

which leads to the equality

$$\frac{\partial T}{\partial y} = \frac{dU}{dy} + \frac{d}{dt}\left[\frac{\partial T}{\partial \dot{y}}\right].$$

The total derivative of L indicates how L changes as the object moves through time. We compute this total derivative as follows:

$$\begin{aligned}
\frac{dL}{dt} &= \frac{\partial T}{\partial y}\,\dot{y} + \frac{\partial T}{\partial \dot{y}}\,\ddot{y} - \frac{dU}{dy}\,\dot{y} \\
&= \left(\frac{dU}{dy} + \frac{d}{dt}\left[\frac{\partial T}{\partial \dot{y}}\right]\right)\dot{y} + \frac{\partial T}{\partial \dot{y}}\,\ddot{y} - \frac{dU}{dy}\,\dot{y} \\
&= \frac{d}{dt}\left[\frac{\partial T}{\partial \dot{y}}\right]\dot{y} + \frac{\partial T}{\partial \dot{y}}\,\ddot{y} \\
&= \frac{d}{dt}\left[\frac{\partial T}{\partial \dot{y}}\,\dot{y}\right].
\end{aligned}$$

Hence, the object moves through time in such a way that

$$\frac{d}{dt}\left(L - \frac{\partial T}{\partial \dot{y}}\,\dot{y}\right) = 0,$$

or equivalently,

$$L - \frac{\partial T}{\partial \dot{y}}\,\dot{y} = \text{constant} = -\mathcal{H}.$$

The $\mathcal{H}$ in the last equality is the value of the object's *Hamiltonian*. The Hamiltonian is dependent on the object's dynamical conditions (kinetic energy and potential energy). Different objects may therefore have different values for the Hamiltonian, but the same object maintains a single value of the Hamiltonian throughout its motion.

We will now show that for the simple dynamics under consideration in this Module, the Hamiltonian $\mathcal{H}$ represents total energy (that is, it is the sum of the kinetic energy T and potential energy U). We assume that the kinetic energy of an object is given by $T = \frac{1}{2}m(\dot{y})^2$, and therefore

$$\frac{\partial T}{\partial \dot{y}} = m\dot{y}.$$

It follows that

$$\dot{y}\,\frac{\partial T}{\partial \dot{y}} = m\dot{y}^2 = 2T.$$

Hence,

$$-\mathcal{H} = L - 2T = T - U - 2T.$$

Our desired result follows by elementary algebra:

$$\mathcal{H} = T + U.$$

Since the value of $\mathcal{H}$ remains constant when describing the motion of an object, the total energy of the object must be conserved.

Whereas the Lagrangian is expressed as a function of y, $\dot{y}$, and t, the Hamiltonian should be expressed as a function of q, p and t, where

$$q = y, \qquad p = \frac{\partial L}{\partial \dot{y}}.$$

(Later, when defining Hamilton's equations, we shall see that the use of "conjugate variables" p and q gives rise to a nice symmetry in expressions for $\partial\mathcal{H}/\partial q$ and $\partial\mathcal{H}/\partial p$.)

In our case, $L = \frac{1}{2}m(\dot{y})^2 - mgy$, so that $p = \partial L/\partial\dot{y} = m\dot{y}$. Thus, the Hamiltonian may be expressed as

$$\mathcal{H} = \mathcal{H}(q,p,t) = \frac{p^2}{2m} + mgq.$$

(Here there is no direct dependence of the Hamiltonian on time t. In general, the Hamiltonian of an isolated dynamical system without energy dissipation is independent of time.)

This Hamiltonian arises in the idealized motion of ball with mass m falling freely with constant gravitational acceleration $-g$ (i.e. we neglect air resistance). For simplicity, we take $m = 1$ and $g = 1$. Let $y(t)$ be the height of a ball at time t. In Newtonian mechanics, by anti-differentiating the acceleration, we obtain both the height and the velocity of the ball at arbitrary time t given the initial velocity v_0 and the initial position y_0:

$$\frac{d^2y}{dt^2} = -1, \qquad \frac{dy}{dt} = -t + v_0, \qquad y(t) = -\frac{1}{2}t^2 + v_0 t + y_0.$$

Thus, we may describe the motion of the ball as tracing the right half of a vertical parabola in which the height y of the ball is on the vertical axis and the time t is on the horizontal axis (see **Figure 1**).

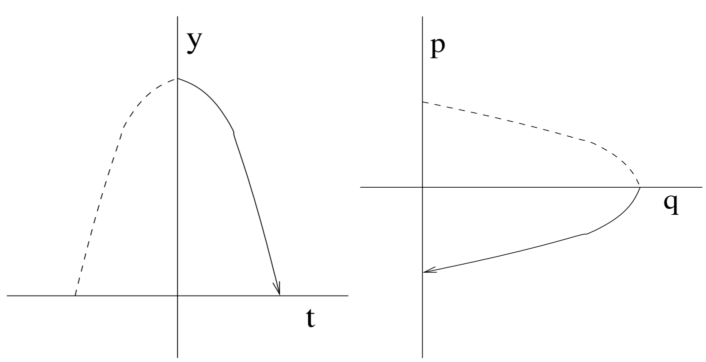


Figure 1. Freely-falling ball's motion described in terms of time t vs. height y (left) and in terms of height q vs. momentum p (right).

In Hamiltonian dynamics, we describe the motion of the ball using the coordinates $\big(q(t), p(t)\big)$, where $q(t)$ is the position of the ball at time t (i.e., $q(t) = y(t)$) and $p(t)$ is, in this simple case, the momentum of the ball at time t

(i.e., $p(t) = m\dot{y} = \dot{y}$ since $m = 1$). The motion of the ball in the q-p plane traces the bottom half of a horizontal parabola (see **Figure 1** and **Exercise 3**). The q-p plane is called a *two-dimensional phase space*, and it is the setting in which we discuss Liouville's theorem in the next section.

Since the Hamiltonian $\mathcal{H}$ is the sum of the kinetic and potential energies ($\mathcal{H} = p^2/2 + q$), one can verify that (**Exercise 4**)

$$\frac{\partial \mathcal{H}}{\partial q} = -\frac{dp}{dt}, \qquad \frac{\partial \mathcal{H}}{\partial p} = \frac{dq}{dt}. \qquad \textit{(Hamilton's equations)}.$$

These equations are fundamental in Hamiltonian dynamics; we use them in the next section when we derive Liouville's theorem.

Exercises

2. Consider a ball with mass m that falls freely under the influence of constant gravity (i.e., without air resistance). Let $y(t)$ $(t \geq 0)$ be the height of the ball at time t. The kinetic energy of the ball is

$$T(\dot{y}) = \tfrac{1}{2} m (\dot{y})^2,$$

and its potential energy is

$$U(y) = mgy,$$

where $-g$ is the constant acceleration due to gravity.

a) Find a simple differential equation for y that is obtainable from Euler's equation.

b) Check that the the total energy $E = T + U$ is conserved, by computing dE/dt and simplifying using the answer to part **a)**.

3. Consider a ball with unit mass that at $t = 0$ is dropped from rest at a height of 1, so that $(q_0, p_0) = (1, 0)$. Assuming constant acceleration $-g = -1$, show that in the two-dimensional q-p phase plane, the freely-falling ball's motion is along the horizontal parabola $q = 1 - \frac{1}{2}p^2$.

4. Verify Hamilton's equation for the case of a freely-falling ball. That is, show that

$$\frac{\partial \mathcal{H}}{\partial q} = -\frac{dp}{dt}, \qquad \frac{\partial \mathcal{H}}{\partial p} = \frac{dq}{dt}.$$

4. Derivation of Liouville's Theorem in Two Dimensions

Intuitively, Liouville's theorem in two dimensions says that

> a large collection of phase points will always occupy the same amount of area, no matter how the shape of the area they occupy may change.

In other words, the density of a large collection of phase points does not change with time, even though the region in phase space formed by these phase points usually will change with time. (see **Figure 2**).

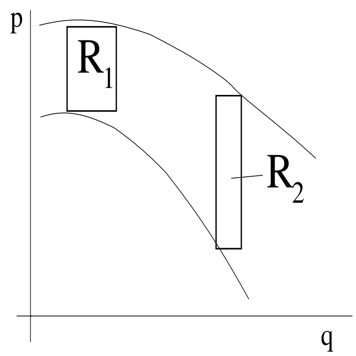


Figure 2. The phase points forming region R_1 at time t_1 form region R_2 at time t_2. The shape of the region changes but not the amount of area or the density of phase points.

We now introduce the concept of a *phase space density function* ρ, which may be evaluated at any phase point (q,p) and at any time t. For a small region in phase space, the phase density is the number of phase points in that region divided by the area of the region. Since the density ρ is dependent on q, p, and t, we write $\rho = \rho(q,p;t)$. Furthermore, we are interested in how the density changes as we move with the flow of points through phase space. In other words, we must compute the change in value of the density ρ as we move with a phase point along its path $\big(q(t),p(t)\big)$ through phase space. Since we allow q and p to change with time, we write $\rho = \rho\big(q(t),p(t);t\big)$.

Liouville's theorem in two dimensions asserts that the total derivative $d\rho/dt$ is zero. This means that the density will remain constant in the following sense. If at time t_0 we evaluate the density of an object at its phase point $\big(q(t_0),p(t_0)\big)$, and at any later time t we evaluate the density at the object's new phase point $\big(q(t),p(t)\big)$, we will obtain the same number.

By the chain rule for derivatives, we have

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial q}\,\dot{q} + \frac{\partial\rho}{\partial p}\,\dot{p} + \frac{\partial\rho}{\partial t}.$$

Hence, one way of proving that $d\rho/dt$ equals zero is to show that

$$\frac{\partial\rho}{\partial t} = -\left(\frac{\partial\rho}{\partial q}\,\dot{q} + \frac{\partial\rho}{\partial p}\,\dot{p}\right),$$

where $\partial\rho/\partial t$ is the rate of change of phase density with time if we keep fixed the point of evaluation (q,p). To compute $\partial\rho/\partial t$, we consider the flow of phase points through a small rectangular area element of phase space. We position the bottom left corner of the rectangle at (q,p) and give the area element a horizontal length dq and vertical length dp (see **Figure 3**), so that the area of the rectangle is $dq\ dp$. Phase points flow in and out of this rectangle, with the net

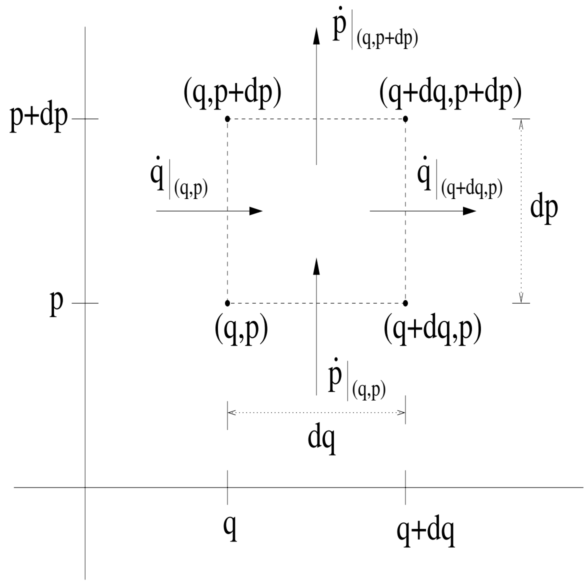


Figure 3. Diagram for computation of $\partial\rho/\partial t$.

rate of change (i.e., the increase in the number of phase points per unit time) for this area element being given by $\partial\rho/\partial t \; dq \; dp$.

The number of phase points flowing in through the bottom of the rectangle per unit time is given by the expression $\rho\dot{p} \mid_{(q,p)} dq$ (the notation $|_{(q,p)}$ indicates that both ρ and $\dot{p}$ are to be evaluated at the phase point (q,p)). This expression is the product of $\dot{p}$ (the vertical component of the flow velocity at the point (q,p)) with the horizontal length dq and the density function ρ. Similarly, the amount flowing in the left is $\rho\dot{q}\; dp$, where we keep in mind that both the density ρ and speed $\dot{q}$ are evaluated at the point (q,p).

Assuming that the vertical component of velocity of the phase points is constant along the top of the rectangle, the amount flowing out the top is $\rho\dot{p} \mid_{(q,p+dp)} dq$. Analogous to the fact that the linear approximation $f(x+h) \approx f(x) + f'(x)h$ becomes exact as h approaches zero, the linear approximation

$$\rho\dot{p} \mid_{(q,p+dp)} \approx \rho\dot{p} \mid_{(q,p)} + \frac{\partial}{\partial p}\left[\rho\dot{p}\right] \mid_{(q,p)} \; dp$$

becomes exact as dp approaches zero. It follows that the number of phase points flowing out the top is given by

$$\rho\dot{p} \mid_{(q,p)} dq + \frac{\partial}{\partial p}\left[\rho\dot{p}\right] \mid_{(q,p)} \; dp\; dq$$

where the point of evaluation is (q,p). The number of phase points out the right can be computed using the same reasoning and is therefore given by

$$\rho\dot{q}\; dp \; + \frac{\partial}{\partial q}\left[\rho\dot{q}\right] \; dq\; dp,$$

where the point of evaluation, (q,p), has been assumed.

The total change in number of phase points in the area element per unit time is equal to the rate phase points flow in minus the rate that they flow out. Hence, we obtain the following expression for $\partial\rho/\partial t \; dq \; dp$:

$$\frac{\partial \rho}{\partial t}\, dp\, dq = \rho\dot{p}\, dq \; + \rho\dot{q}\, dp \; - \left(\rho\dot{p}\, dq \; + \frac{\partial}{\partial p}\left[\rho\dot{p}\right]\; dp\, dq\right) - \left(\rho\dot{q}\, dp + \frac{\partial}{\partial q}\left[\rho\dot{q}\right]\; dq\, dp\right).$$

The first term in the right side of the equation represents the amount that flows in the bottom, the second term represents the amount that flows in the left, the third represents the amount that flows out the top, and the last term represents the amount that flows out the right. Cancelling and simplifying leads to

$$\begin{aligned}
\frac{\partial \rho}{\partial t}\, dp\, dq &= -\left(\frac{\partial}{\partial p}\left[\rho\dot{p}\right] + \frac{\partial}{\partial q}\left[\rho\dot{q}\right]\right)\; dq\, dp, \\
\frac{\partial \rho}{\partial t} &= -\left(\frac{\partial}{\partial p}\left[\rho\dot{p}\right] + \frac{\partial}{\partial q}\left[\rho\dot{q}\right]\right), \\
\frac{\partial \rho}{\partial t} &= -\left(\frac{\partial \rho}{\partial p}\,\dot{p} + \rho\,\frac{\partial \dot{p}}{\partial p} + \frac{\partial \rho}{\partial q}\,\dot{q} + \rho\,\frac{\partial \dot{q}}{\partial q}\right).
\end{aligned}$$

The last step uses the product rule for partial derivatives.

Next, because $\dot{p} = -\partial\mathcal{H}/\partial q$ and $\dot{q} = \partial\mathcal{H}/\partial p$ in Hamiltonian dynamics, we may substitute, so that

$$\frac{\partial \rho}{\partial t} = -\left(\frac{\partial \rho}{\partial p}\,\dot{p} + \rho\,\frac{\partial}{\partial p}\left[-\frac{\partial \mathcal{H}}{\partial q}\right] + \frac{\partial \rho}{\partial q}\,\dot{q} + \rho\,\frac{\partial}{\partial q}\left[\frac{\partial \mathcal{H}}{\partial p}\right]\right).$$

The second and fourth terms cancel due to the equality of mixed partials, so we get

$$\frac{\partial \rho}{\partial t} = -\left(\frac{\partial \rho}{\partial p}\,\dot{p} + \frac{\partial \rho}{\partial q}\,\dot{q}\right).$$

Substituting this expression for $\partial\rho/\partial t$ into the total derivative $d\rho/dt$ given at the onset results in

$$\begin{aligned}
\frac{d\rho}{dt} &= \frac{\partial \rho}{\partial q}\,\dot{q} + \frac{\partial \rho}{\partial p}\,\dot{p} + \frac{\partial \rho}{\partial t} \\
&= \frac{\partial \rho}{\partial q}\,\dot{q} + \frac{\partial \rho}{\partial p}\,\dot{p} - \left(\frac{\partial \rho}{\partial p}\,\dot{p} + \frac{\partial \rho}{\partial q}\,\dot{q}\right) \\
&= 0.
\end{aligned}$$

This proves Liouville's theorem and confirms that the density does not change if at the initial time t_0 we evaluate it at the point $(q(t_0), p(t_0))$ and at any later time t we evaluate it at the point $(q(t), p(t))$.

There is a simple relationship between the density $\rho(q,p;t)$ and a probability distribution $f(q,p;t)$. Let N be the total number of phase points in the phase space. Then we have

$$f(q,p;t) = \frac{\rho(q,p;t)}{N}.$$

In other words, the number of phase points in an area $dq\,dp$ is given by $\rho\,dq\,dp$, whereas the probability of finding a phase point in the same area is $f\,dq\,dp$. It is then simple to show (**Exercise 5**) that Liouville's theorem may be expressed in its probability distribution form as $df/dt = 0$. The two-dimensional probability distribution surface $z = f(p,q;t)$ moves with the flow of phase points in such a way that the height of the surface above each particular phase point remains constant as that phase point moves through phase space.

Exercise

5. Prove the probability distribution form of Liouville's theorem. That is, show that $df/ft = 0$, where $f(p,q;t)$ is a probability distribution function for a two-dimensional phase space. (Josiah Willard Gibbs (1839–1903) of Yale University is given credit for being the first to derive explicitly the general probability distribution form of Liouville's theorem [Binney and Tremaine 1987]. Gibbs was also the first to recognize that this theorem could be applied to astrophysics. We discuss this idea further in **Section 6**.)

5. A Simple Illustration Using Three Freely-Falling Balls

Let $\mathbf{F}(q,p)$ be the vector field that gives the "velocity" of a phase point (q,p) in two-dimensional phase space. In other words,

$$\mathbf{F}(q,p) = \left\langle \frac{dq}{dt}, \frac{dp}{dt} \right\rangle.$$

Hence we may write

$$\mathbf{F}(q,p) = \left\langle \frac{\partial \mathcal{H}}{\partial p}, -\frac{\partial \mathcal{H}}{\partial q} \right\rangle.$$

The divergence of $\mathbf{F}$ is zero, as we can see from the following:

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{\partial}{\partial q}\left[\frac{\partial \mathcal{H}}{\partial p}\right] + \frac{\partial}{\partial p}\left[-\frac{\partial \mathcal{H}}{\partial q}\right] \\ &= \frac{\partial^2 \mathcal{H}}{\partial q\,\partial p} - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial q} = 0. \end{aligned}$$

A vector field whose divergence is zero at each point is said to be *incompressible*. This means that a collection of phase points forming a region with a certain area at time t_0, will move through phase space in such a way that the area occupied by these points will be the same at any later time t. This result may also be obtained as a special case of Liouville's theorem in which the distribution f is uniform (or constant) on an area region A in q-p phase space (**Exercise 7**).

We may illustrate the incompressibility of phase flow in the following way. Let us consider three freely-falling balls A, B, and C whose initial positions in phase space are

$$A_0 = (1, 0), \qquad B_0 = (1.5, -1), \qquad C_0 = (2.5, -1).$$

We may think of balls B and C as having been released from heights $y = 2$ and $y = 3$, respectively, one second before ball A is released from rest at height $y = 1$. Note that the area of the triangle $\triangle A_0B_0C_0$ is $(0.5)(1)(1) = 0.5$ square units (see **Figure 4**).

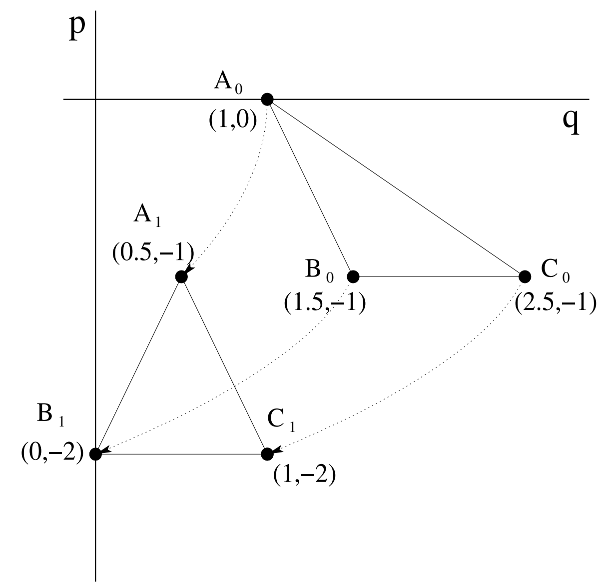


Figure 4. A special case of Liouville's theorem implies that the areas of triangles $A_0B_0C_0$ and $A_1B_1C_1$ must be the same.

A phase point initially at (q_0, p_0) will be at the new phase point

$$(q(t), p(t)) = \left(-\tfrac{1}{2}t^2 + p_0t + q_0,\ p_0 - t\right)$$

after an amount of time t. Hence, after one second, the phase points that form $\triangle A_0B_0C_0$ will form $\triangle A_1B_1C_1$, where

$$\begin{aligned} A_1 &= (-0.5 + 1, -1) = (0.5, -1) \\ B_1 &= (-0.5 - 1 + 1.5, -2) = (0, -2) \\ C_1 &= (-0.5 + 2.5 - 1, -2) = (1, -2). \end{aligned}$$

With a little thought about the physical motion of freely-falling balls, one realizes that all of the representative phase points on or inside $\triangle A_0B_0C_0$ at time

$t = 0$ must be on or inside $\triangle A_1B_1C_1$ at time $t = 1$. Furthermore, no point outside $\triangle A_0B_0C_0$ at time $t = 0$ will be on or inside $\triangle A_1B_1C_1$ at time $t = 1$. Since the flow of phase points is incompressible ($\nabla \cdot \mathbf{F} = 0$), the areas of the two triangles must be the same. Indeed, both triangles have an area of 0.5 square units as can be seen in **Figure 4**.

Exercises

6. Find $\triangle A_2B_2C_2$, which is formed after two seconds by the points of the phase space initially in $\triangle A_0B_0C_0$, and show that the area of $\triangle A_0B_0C_0$ equals the area of $\triangle A_2B_2C_2$.

7. Let $A(t_0)$ be an area region formed by a collection of phase points at time t_0. Let $A(t)$ be the area region occupied by these phase points at any later time t. Use the probability distribution form of Liouville's theorem to prove that area($A(t_0)$) = area($A(t)$).

6. Further Applications

6.1 Charged-Particle Accelerators

Particle accelerators attempt to create subatomic particles and interactions; in addition, they are used in a remarkable variety of practical applications (see [Wilson 2001]), including

- manufacturing of such diverse products as computer disks, shrink-wrap, automobile tires and telephone cables;
- purification of food stuffs, drinking water and surgical tools; and
- in medical procedures such as diagnostic imaging systems and radiation therapy techniques.

A simple case occurs when the beam circulates around an accelerator with constant energy. A cross section of such a beam may be represented by an ensemble of particles in two-dimensional q-p phase space, forming an elliptical region. Because the particles are charged, they can be focused by an arrangement of magnets, much as a light beam can be focused by geometric lenses. Since the magnetic force is conservative, by Liouville's theorem the shape of the elliptical-region in phase space may change during focusing but its area remains constant (**Figure 5**). An important quantity called the *emittance* of a beam is proportional to the area of the region occupied by the points in phase space. Hence, the emittance remains constant during the focusing of a constant energy beam.

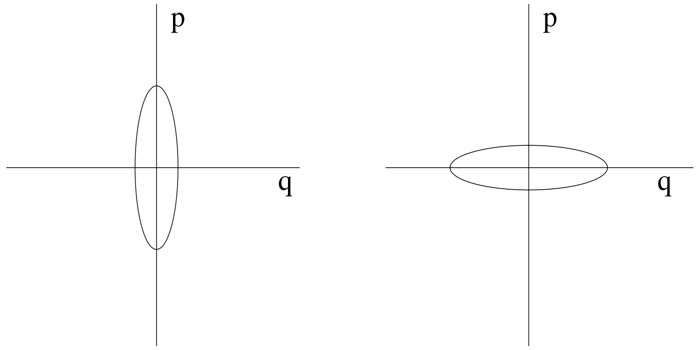


Figure 5. A cross section of particles in a constant-energy beam may be represented by an elliptic region in phase space. Liouville's theorem says that the area of the region occupied in phase space remains constant. Hence, the beam's emittance, which is proportional to the area, is also constant.

6.2 Galactic Dynamics

Two years after Liouville's death in 1882, Josiah W. Gibbs (1839–1903), a mathematical physicist at Yale University, recognized that the probability distribution form of Liouville's theorem could be applied to astronomy. Liouville's theorem arises within galactic dynamics in at least three different settings [Binney and Tremaine 1987]:

6.2.1 Analyzing Motion of Stars

A star moving within a galaxy may be represented by a phase point moving within a phase space with three position and three momentum coordinates. In this context, the number density function $\rho = \rho(q(t), p(t), t)$ with $q(t) = \big(q_1(t), q_2(t), q_3(t)\big)$ and $p(t) = \big(p_1(t), p_2(t), p_3(t)\big)$ is used to specify the number $\rho\, dq\, dp$ of stars within a small volume dq and momentum range dp centered at (q, p). Liouville's theorem asserts that $d\rho/dt = 0$, or, in other words, the number density remains the same around the phase point representing a given star.

6.2.2 Describing Macroscopic Properties

A galactic system with N stars may also be represented by a $6N$-dimensional phase space called a Γ-*space*; a point $(w_1, w_2, ..., w_N)$ in this phase space is called a Γ-*point*. The ith coordinate w_i of a Γ-point describes the position and momentum of a particular star. Different Γ-points may correspond to the same set of macroscopic galactic properties (density distribution, velocity distribution, number of binary stars etc.) Collectively, Γ-points giving rise to the same set of macroscopic properties are called an *ensemble.* The N*-particle distribution function* $f^{(N)} = f^{(N)}(w_1, ..., w_N, t)$ is then used to obtain the probability that a Γ-point belonging to an ensemble is found in a region D of Γ-space at time t. (One finds the probability by integrating $f^{(N)}$ over D.) The evolution of a galactic ensemble is therefore described by the evolution of its distribution $f^{(N)}$, the latter being constrained by the probability distribution form of Liouville's theorem ($df^{(N)}/dt = 0$).

6.2.3 Modeling the Dynamics in a Cluster of Galaxies

On a massive scale, the dynamical motion of thousands of galaxies comprising a gigantic cluster may be modeled in a similar way to the motion of stars within a galaxy, hence creating a third setting for Liouville's theorem to be applied. Clusters continue to be an important site for the investigation into the existence, nature and distribution of enigmatic dark matter. Thirteen different clusters, including the Coma cluster and Abel 2142, have been studied by the Chandra X-ray observatory [Harvard-Smithsonian Center for Astrophysics 2003] since it was first deployed in 1999 by the space shuttle Columbia.

7. Solutions to the Exercises

1. Euler's equation implies that $\ddot{y} = 1$. It follows by antidifferentiation that $y = y_0 + \dot{y}_0 t + \frac{1}{2}t^2$.

2. **a)** According to Hamilton's principle, the ball falls so that the integral of the difference between its kinetic and potential energy is minimized. Letting $f(y, \dot{y}; t) = \frac{1}{2}m\dot{y}^2 - mgy$, Euler's condition becomes

$$\begin{aligned} -mg - \frac{d}{dt}(m\dot{y}) &= 0, \\ mg + m\ddot{y} &= 0, \\ \ddot{y} &= -g. \end{aligned}$$

b) $E = \frac{1}{2}m\dot{y}^2 + mgy$, so $dE/dt = m\dot{y}\ddot{y} + mg\dot{y} = m\dot{y}(-g) + mg\dot{y} = 0$.

3. The motion in the q-p phase plane will be along the parabola $q = -\frac{1}{2}p^2 + 1$, since

$$\begin{aligned} \frac{dy}{dt} &= p = -t, \\ y &= q = -\tfrac{1}{2}t^2 + 1 = -\tfrac{1}{2}p^2 + 1. \end{aligned}$$

4. Assume that the ball has mass m and constant acceleration $-g$. The Hamiltonian $\mathcal{H}$ is given by $\mathcal{H} = p^2/2m + mgq$. Hence, $\partial\mathcal{H}/\partial q = mg$ and $\partial\mathcal{H}/\partial p = p/m$. We also have that

$$\begin{aligned} \text{Force } &= -mg = \frac{dp}{dt}, \\ \text{Velocity } &= \frac{dq}{dt} \quad = \frac{p}{m}. \end{aligned}$$

It follows that $\partial\mathcal{H}/\partial q = -dp/dt$ and $\partial\mathcal{H}/\partial p = dq/dt$.

5.
$$\frac{df}{dt} = \frac{1}{N}\frac{d\rho}{dt} = 0.$$

6. $A_2 = (-1, -2)$, $B_2 = (-2.5, -3)$, $C_2 = (-1.5, -3)$. Area of $\triangle A_2B_2C_2$ is 0.5.

7. Let area$\big(A(t_0)\big) = k$ and let $f(p, q, t_0) = 1/k$ for all $(p, q) \in A(t_0)$ and zero otherwise. By Liouville's theorem, $f(p, q, t) = 1/k$ for all $(p, q) \in A(t)$ and zero otherwise. Since the total probability is equal to one, area $A(t)) = k$, and we conclude that area$(A(t)) =$ area$(A(t_0))$.

References

Binney, James and Scott Tremaine. *Galactic Dynamics*. Princeton, NJ: Princeton University Press 1987.

Harvard-Smithsonian Center for Astrophysics. 2003. Chandra Images by Category. `http://chandra.harvard.edu/photo/category/galaxyclusters.html` .

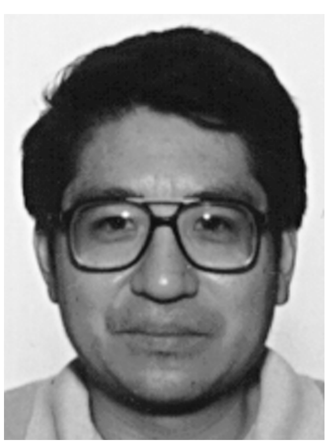
Marion, Jerry B., and Stephen Thornton. 1995. *Classical Dynamics of Particles and Systems*. 4th ed. San Diego, CA: Harcourt Brace Jovanovich.

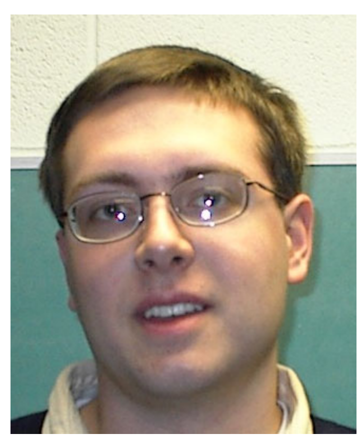
Wilson, E. 2001. *An Introduction to Particle Accelerators.* New York: Oxford University Press.

Acknowledgments

The authors thank Dr. Akira Isihara, a distinguished theoretical physicist in statistical mechanics, for inspiring the writing of this Module, and a referee for helpful corrections.

About the Authors





Paul Isihara, Professor of Mathematics at Wheaton College (IL), has a special interest in doing expository and creative research projects with undergraduates.

Nat Stapleton (mathematics ’04) and Benjamin Noonan (physics ’04) worked on this Module while participants in the 2003 Wheaton College Summer Science Research Program.