© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

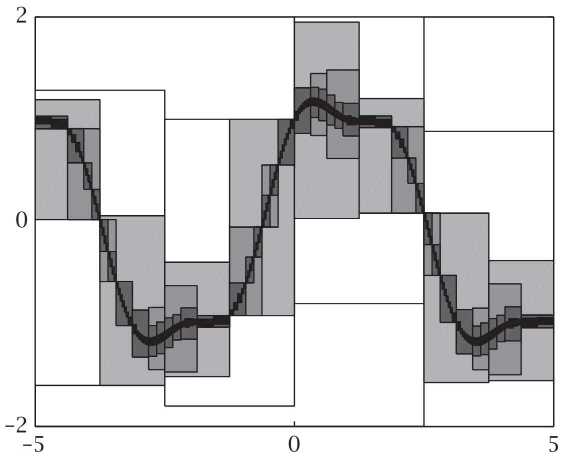

Figure 1 Successively tighter enclosures of a graph.

2 Interval Analysis

The inclusion principle (2) enables us to capture continuous properties of a function, using only a finite number of operations. Its most important use is to explicitly bound discretization errors that naturally arise in numerical algorithms.

As an example, consider the function $f(x) = \cos^3 x + \sin x$ on the domain $\boldsymbol{x} = [-5, 5]$. For any decomposition of the domain $\boldsymbol{x}$ into a finite set of subintervals $\boldsymbol{x} = \bigcup_{i=1}^{n} \boldsymbol{x}_i$, we can form the set-valued graph consisting of the pairs $(\boldsymbol{x}_1, F(\boldsymbol{x}_1)), \dots, (\boldsymbol{x}_n, F(\boldsymbol{x}_n))$. As the partition is made finer (that is, as $\max_i \operatorname{diam}(\boldsymbol{x}_i)$ is made smaller), the set-valued graph tends to the graph of f (see figure 1). And, most importantly, every such set-valued graph contains the graph of f.

This way of incorporating the discretization errors is extremely useful for quadrature, optimization, and equation solving. As one example, suppose we wish to compute the definite integral $I = \int_0^8 \sin(x + \mathrm{e}^x)\,\mathrm{d}x$.

A MATLAB function `simpson` that implements a simple textbook adaptive Simpson quadrature algorithm produces the following result.

```
% Compute integral I with tolerance 1e-6.
>> I = simpson(@(x) sin(x + exp(x)), 0, 8)
I =
   0.251102722027180
```

A (*very naive*) set-valued approach to quadrature is to enclose the integral I via

$$I \in \sum_{i=1}^{n} F(\boldsymbol{x}_i)\operatorname{diam}(\boldsymbol{x}_i),$$

which, for a sufficiently fine partition, produces the integral enclosure

$$I \in 0.3474001726^{66}_{49}.$$

Thus, it turns out that the result from `simpson` was completely wrong! This is one example of the importance of rigorous computations.

3 Recent Developments

There is currently an ongoing effort within the IEEE community to standardize the implementation of interval arithmetic. The hope is that we will enable computer manufacturers to incorporate these types of computations at the hardware level. This would remove the large computational penalty incurred by repeatedly having to switch rounding modes—a task that central processing units were not designed to perform efficiently.

Further Reading

Alefeld, G., and J. Herzberger. 1983. *Introduction to Interval Computations*. New York: Academic Press.

Kulisch, U. W., and W. L. Miranker. 1981. *Computer Arithmetic in Theory and Practice*. New York: Academic Press.

Lerch, M., G. Tischler, J. Wolff von Gudenberg, W. Hofschuster, and W. Krämer. 2006. filib++, a fast interval library supporting containment computations. *ACM Transactions on Mathematical Software* 32(2):299–324.

Moore, R. E., R. B. Kearfott, and M. J. Cloud. 2009. *Introduction to Interval Analysis*. Philadelphia, PA: SIAM.

Rump, S. M. 1999. INTLAB—INTerval LABoratory. In *Developments in Reliable Computing*, edited by T. Csendes, pp. 77–104. Dordrecht: Kluwer Academic.

Tucker, W. 2011. *Validated Numerics: A Short Introduction to Rigorous Computations*. Princeton, NJ: Princeton University Press.

II.21 Invariants and Conservation Laws

Mark R. Dennis

As important as the study of *change* in the mathematical representation of physical phenomena is the study of *invariants*. Physical laws often depend only on the *relative* positions and times between phenomena, so certain physical quantities do not change; i.e., they are *invariant*, under continuous translation or rotation of the spatial axes. Furthermore, as the spatial configuration of a system evolves with time, quantities such as total energy may remain unchanged; that is, they are

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

conserved. The study of invariants has been a remarkably successful approach to the mathematical formulation of physical laws, and the study of continuous symmetries and conservation laws—which are related by the result known as *Noether's theorem*—has become a systematic part of our description of physics over the last century, from the atomic scale to the cosmic scale.

As an example of so-called *Galilean invariance*, Newton's force law keeps the same form when the velocity of the frame of reference (i.e., the coordinate system specified by x-, y-, and z-axes) is changed by adding a constant; this is equivalent to adding the same constant velocity to all the particles in a mechanical system. Other quantities *do* change under such a velocity transformation, such as the kinetic energy $\frac{1}{2}m|\boldsymbol{v}|^2$ (for a particle of mass m and velocity $\boldsymbol{v}$); however, for an evolving, nondissipative system such as a bouncing, perfectly elastic rubber ball, the total energy is constant in time—that is, energy is conserved.

The development of our understanding of fundamental laws of dynamics can be interpreted by progressively more sophisticated and general representations of space and time themselves: ancient Greek physical science assumed absolute space with a privileged spatial point (the center of the Earth), through static Euclidean space where all spatial points are equivalent, through CLASSICAL MECHANICS [IV.19] where all inertial frames, moving at uniform velocity with respect to each other, are equivalent according to Newton's first law, to the modern theories of special and general relativity. The theory of relativity (both general and special) is motivated by Einstein's principle of covariance, which is described below. In this theory, space and time in different frames of reference are treated as coordinate systems on a four-dimensional pseudo-Riemannian manifold (whose mathematical background is described in TENSORS AND MANIFOLDS [II.33]), which manifestly combines conservation laws and continuous geometric symmetries of space and time. In special relativity (described in some detail in this article), this manifold is flat *Minkowski space-time*, generalizing Euclidean space to include time in a physically natural way. In general relativity, described in detail in GENERAL RELATIVITY AND COSMOLOGY [IV.40], this manifold may be curved, depending in part on the distribution of matter and energy according to EINSTEIN'S FIELD EQUATIONS [III.10].

In quantum physics, the description of a system in terms of a complex vector in Hilbert space gives rise to new symmetries. An important example is the fact that physical phenomena do not depend on the overall phase (argument) of this vector. Extension of Noether's theorem here leads to the conservation of electric charge, and extension to Yang–Mills theories provides other conserved quantities associated with the nuclear forces studied in contemporary fundamental particle physics. Other phenomena, such as the Higgs mechanism (leading to the Higgs boson recently discovered in high-energy experiments), are a consequence of the breaking of certain quantum symmetries in certain low-energy regimes. Symmetry and symmetry breaking in quantum theory are discussed briefly at the end of this article.

Spatial vectors, such as $\boldsymbol{r} = (x, y, z)$, represent the spatial distance between a chosen point and the origin, and of course the vector between two such points $\boldsymbol{r}_2 - \boldsymbol{r}_1$ is independent of translations of this origin. Similarly, the scalar product $\boldsymbol{r}_2 \cdot \boldsymbol{r}_1$ is unchanged under rotation of the coordinate system by an orthogonal matrix $\boldsymbol{R}$, under which $\boldsymbol{r} \to \boldsymbol{R}\boldsymbol{r}$.

Continuous groups of transformations such as translation and rotation, and their matrix representations, are an important tool used in calculations of invariants. For example, the set of two-dimensional matrices $\left(\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right)$, representing rotations through angles θ, may be considered as a continuous one-parameter Abelian group of matrices generated by the MATRIX EXPONENTIAL [II.14] $e^{\theta \boldsymbol{A}}$, where $\boldsymbol{A}$ is the generator $\left(\begin{smallmatrix}0 & -1 \\ 1 & 0\end{smallmatrix}\right)$. The generator itself is found as the derivative of the original matrix with respect to θ, evaluated at $\theta = 0$. Translations are less obviously represented by matrices; one approach is to append an extra dimension to the position vector with unit entry, such as $(1, x)$ specifying one-dimensional position x; a translation by X is thus represented by

$$\begin{pmatrix}1 & 0 \\ X & 1\end{pmatrix} = \exp\left(X\begin{pmatrix}0 & 0 \\ 1 & 0\end{pmatrix}\right). \qquad (1)$$

When a physical system is invariant under a one-parameter group of transformations, the corresponding generator plays a role in determining the associated conservation law.

1 Mechanics in Euclidean Space

It is conventional in classical mechanics to define the positions of a set of interacting particles in a vector space. However, we do not observe any unique origin to the three-dimensional space we inhabit, which we therefore take to be the *Euclidean space* $\mathbb{E}^3$; only relative positions between different interacting subsystems

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

(i.e., positions relative to the common center of mass) enter the equations of motion. The entire system may be translated in space without any effect on the phenomena.

Of course, *external forces* acting on the system may prevent this, such as a rubber ball in a linear gravity field. (In such situations the source of the force, such as the Earth as the source of gravity, is not considered part of the system.) The gravitational force may be represented by a potential $V = gz$ for height z and gravitational acceleration g; the ball's mass m times the negative gradient, $-m\nabla V$, gives the downward force acting on the ball. The contours of V, given by $z =$ const., nevertheless have a symmetry: they are invariant to translations of the horizontal coordinates x and y. Since the gradient of the potential is proportional to the gravitational force—which, by virtue of Newton's law equals the rate of change of the particle's linear momentum—the horizontal component of momentum does not change and is therefore conserved even when the particle bounces due to an impulsive, upward force from the floor. The continuous, horizontal translational symmetry of the system therefore leads to conservation of linear momentum in the horizontal plane. In a similar argument employing Newton's laws in cylindrical polar coordinates, the invariance of the potential to rotations about the z-axis leads to the conservation of the vertical component a body's angular momentum, as observed for tops spinning frictionlessly.

2 Noether's Theorem

The Lagrangian framework for mechanics (CLASSICAL MECHANICS [IV.19 §2]), which describes systems acting under forces defined by gradients of potentials (as in the previous section), is a natural mathematical setting in which to explore the connection between a system's symmetries and its conservation laws. Here, a mechanical system evolving in time t is described by n generalized coordinates $q_j(t)$ and their time derivatives $\dot{q}_j$, for $j = 1, \dots, n$, where the initial values $q_j(t_0)$ at time t_0 and final values $q_j(t_1)$ at t_1 are fixed. The *action* of the system is the functional

$$S[\{q_j\}] = \int_{t_0}^{t_1} L(\{q_j\}, \{\dot{q}_j\}, t)\, \mathrm{d}t,$$

where $L(\{q_j\}, \{\dot{q}_j\}, t)$ is the *Lagrangian*; this is a function of the coordinates, their corresponding velocities, and maybe time, specified here by the total kinetic energy minus the total potential energy of the system (thereby capturing the forces as gradients of the potential energy). Using the CALCULUS OF VARIATIONS [IV.6], the functions $q_j(t)$ that satisfy the laws of mechanics are those that make the action stationary, and these satisfy Lagrange's equations of motion

$$\frac{\partial L}{\partial q_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_j} = 0, \quad j = 1, \dots, n. \tag{2}$$

The argument of the time derivative in this expression, $\partial L/\partial \dot{q}_j$, is called the *canonical momentum* p_j for each j. The set of equations (2) involves the combination of partial derivatives of the Lagrangian with respect to the coordinates and velocities, together with the total derivative with respect to time. By the chain rule, this total derivative affects explicit time dependence in L and the implicit time dependence in each q_j and $\dot{q}_j$. Many of the conservation laws involving Lagrangians involve such an interplay of explicit and implicit time dependence.

Any transformation of the coordinates q_j that does not change the Lagrangian is a symmetry of the system. If L does not have explicit dependence on a coordinate q_j, then the first term in (2) vanishes: $\mathrm{d}p_j/\mathrm{d}t = 0$, i.e., the corresponding canonical momentum is conserved in time. In the example from the last section of a particle in a linear gravitational field, the coordinates can be chosen to be Cartesian x, y, z, or cylindrical polars r, ϕ, z; L is independent of x and y, leading to conservation of horizontal momentum, and also ϕ, leading to conservation of angular momentum about the z-axis. The theorem is proved for symmetries of this type in CLASSICAL MECHANICS [IV.19 §2.3]: if a system is homogeneous in space (translation invariant), then linear momentum is conserved, and if it is isotropic (independent of rotations, such as the Newtonian gravitational potential around a massive point particle exerting a central force), then angular momentum is conserved (equivalent to Kepler's second law of planetary motion for gravity).

Since it is the equations (2) that represent the physical laws rather than the form of L or S, the system may admit a more general kind of symmetry whose transformation adds a time-dependent function to the Lagrangian L. If, under the transformation, the Lagrangian transforms $L \to L + \mathrm{d}\Lambda/\mathrm{d}t$ involving the total time derivative of some function Λ, the action transforms $S \to S + \Lambda(t_1) - \Lambda(t_0)$. Thus the transformed action is still made stationary by functions satisfying (2), so transformations of this kind are symmetries of the system, which are also continuous if Λ also depends on a continuous parameter s so that its time derivative is

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

zero when $s = 0$. It is not then difficult to see that the quantity

$$\sum_{j=1}^{n} p_j \frac{\partial q_j}{\partial s}\bigg|_{s=0} - \frac{\partial \Lambda}{\partial s}\bigg|_{s=0}, \quad (3)$$

defined in terms of the generators of the transformation on each coordinate and the Lagrangian, is constant in time. This is Noether's theorem for classical mechanics.

An important example is when the Lagrangian has no explicit dependence on time, $\partial L/\partial t = 0$. In this case, under an infinitesimal time translation $t \to t + \delta t$, $L \to L + \delta t \mathrm{d}L/\mathrm{d}t$, so here Λ is $L\delta t$, with L evaluated at t, and δt plays the role of s. Under the same infinitesimal transformation, $q_j \to q_j + \delta t \dot{q}_j$, so the relevant conserved quantity (3) is $\sum_j p_j \dot{q}_j - L$, which is the Hamiltonian of the system, which is equal to the total energy in many systems of interest. It is apparently a fundamental law of physics that the total energy in physical processes is conserved in time; energy can be in other forms such as electromagnetic, gravitational, or heat, as well as mechanical. Noether's theorem states that the law of conservation of energy is equivalent to the fact that the physical laws of the system, characterized by their Lagrangian, do not change with time.

The vanishing of the action functional's integrand (i.e., the Lagrangian L) is equivalent to the existence of a first integral for the system of Lagrange equations, which is interpreted in the mechanical setting as a constant of the motion of the system. In this sense, Noether's theorem may be applied more generally in other physical situations described by functionals whose physical laws are given by the corresponding Euler–Lagrange equations. In the case of the Lagrangian approach applied to fields (i.e., functions of space and time), Noether's theorem generalizes to give a continuous density ρ (such as mass or charge density) and a flux vector $\boldsymbol{J}$ satisfying the continuity equation $\dot{\rho} + \nabla \cdot \boldsymbol{J} = 0$ at every point in space and time.

3 Galilean Relativity

Newton's first law of motion can be paraphrased as "all inertial frames, traveling at uniform linear velocity with respect to each other, are equivalent for the formulation of mechanics"—that is, without action of external forces, a system will behave in the same way regardless of the motion of its center of mass. The behavior of a mechanical system is therefore independent of its overall velocity; this is a consequence of Newton's second law, that force is proportional to acceleration. According to pre-Newtonian physics, forces were thought to be proportional to *velocity* (as the effect of friction was not fully appreciated), and it was not until Galileo's thought experiments in friction-free environments that the proportionality of force to *acceleration* was appreciated. In spite of Galilean invariance, problems involving circular motion do in fact seem to require a privileged frame of reference, called *absolute space*. One example due to Newton himself is the problem of explaining, without absolute space, the meniscus formed by the surface of water in a spinning bucket; such problems are properly overcome only in general relativity.

With Galilean relativity, absolute position is no longer defined: events occurring at the same position but at different times in one frame (such as a moving train carriage) occur at different positions in other frames (such as the frame of the train track). However, changes to the state of motion, i.e., accelerations, have physical consequences and are related to forces. This is an example of a *covariance principle*, whose importance for physical theories was emphasized by Einstein. According to this principle, from the statement of physical laws in one frame of reference (such as the laws of motion), one can derive their statement in a different frame of reference from the application of the appropriate transformation rule between reference frames. The statement in the new frame should have the same *mathematical form* as in the previous frame, although quantities may not take the same values in different frames.

Transformations between different inertial frames are represented mathematically in a similar way to the translations of (1); events are labeled by their positions in space and time, such as (t, x) in one frame and (t, x') in another moving at velocity v with respect to the first. Since $x' = x - vt$, the transformation from (t, x) to (t, x') is represented by the matrix $\left(\begin{smallmatrix} 1 & 0 \\ -v & 1 \end{smallmatrix}\right)$. This Galilean transformation (or *Galilean boost*) differs from (1) in that time t is here appended to the position vector, since the translation from the boost is time-dependent. Galilean boosts in three spatial dimensions, together with regular translations and rotations, define the *Galilean group*. It can be shown that the Lagrangian of a free particle follows directly from the covariance of the corresponding action under the Galilean group.

Infinitesimal velocity boosts generate a Noetherian symmetry on systems of particles interacting via forces that depend only on the positions of the others. Consider N point particles of mass m_k and position $\boldsymbol{r}_k$ such that V depends only on $\boldsymbol{r}_k - \boldsymbol{r}_\ell$ for $k, \ell = 1, \dots, N$. Under

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

an infinitesimal boost by $\delta\boldsymbol{v}$ for each k, $\boldsymbol{r}_k \to \boldsymbol{r}_k + t\delta\boldsymbol{v}$, $\dot{\boldsymbol{r}}_k \to \dot{\boldsymbol{r}}_k + \delta\boldsymbol{v}$, and

$$L \to L + \tfrac{1}{2}|\delta\boldsymbol{v}|^2 \sum_k m_k + \delta\boldsymbol{v} \cdot \sum_k m_k \dot{\boldsymbol{r}}_k.$$

The resulting Noetherian conserved quantity (3) is

$$\boldsymbol{C} = t \sum_k m_k \dot{\boldsymbol{r}}_k - \sum_k m_k \boldsymbol{r}_k. \qquad (4)$$

This quantity is constant in time; its value is minus the product of the system's total mass with the position of its center of mass when $t = 0$. Thus, $\dot{\boldsymbol{C}} = t \sum_k m_k \ddot{\boldsymbol{r}}$, that is, t times the sum of the forces on each of the N particles, which is zero by assumption.

4 Special Relativity

Historically, the physics of fields, such as electromagnetic radiation described by MAXWELL'S EQUATIONS [III.22], was developed much later than the mechanics of Galileo and Newton. The simplest wave equations, in fact, suggest a different invariant space-time from the Galilean framework above, which was first investigated by Hendrik Lorentz and Albert Einstein at the turn of the twentieth century.

Consider the time-dependent wave equation $\nabla^2\varphi - c^{-2}\ddot{\varphi} = 0$ for some scalar function $\varphi(t, \boldsymbol{r})$ and ∇^2 the LAPLACE OPERATOR [III.18]. If all observers agree on the same value of the wave speed c (as proposed by Einstein for c the speed of light, and verified by experiments), this suggests that the transformation between moving inertial frames should be in terms of *Lorentz transformations* (or *Lorentz boosts*) in the x-direction,

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma(v) \begin{pmatrix} 1 & -v/c \\ -v/c & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \qquad (5)$$

where (ct', x') denotes position in space and time (multiplied by the invariant speed c) in a frame moving along the x-axis at velocity v with respect to the frame with space-time coordinates (ct, x), and

$$\gamma(v) = (1 - v^2/c^2)^{-1/2}.$$

Requiring $\gamma(v)$ to be finite, positive, and real-valued implies $|v| < c$. According to the Lorentz transformation with $x = 0$, $t' = \gamma(\boldsymbol{v})t$, suggesting that $\gamma(v)$ may be interpreted as the rate of flow of time in the transformed frame moving at speed v with respect to time in the reference frame.

According to Einstein's theory of *special relativity*, *all* physical laws should be invariant to translations, rotations, and Lorentz boosts, and Galilean transformations and Newton's laws arise in the low-velocity limit. Rotations and Lorentz boosts, together with time reflection and spatial inversion, define the *Lorentz group*; the *Poincaré group* is the semidirect product group of the Lorentz group with space-time translations. Special relativity is therefore also based on the principle of covariance, but with a different set of transformations between frames of reference. All of the familiar results of CLASSICAL MECHANICS [IV.19] are recovered when the boost transformations are restricted to $v \ll c$.

In special relativity different observers measure different time intervals between events, as well as different spatial separations. For a relativistic particle, different observers will disagree on the particle's relativistic energy E and momentum $\boldsymbol{p}$. Nevertheless, according to the principle of special relativity, all observers agree on the quantity

$$E^2 - |\boldsymbol{p}|^2c^2 = m^2c^4,$$

so m is an invariant on which all observers agree, called the *rest mass* of the particle. In the frame in which the particle is at rest, $E = mc^2$ arises as a special case. Otherwise, $E = m\gamma(v)c^2$, so in special relativity the moving particle's energy is explicitly related to the flow of time in the particle's rest frame with respect to that in the reference frame.

The Lagrangian formalism can be generalized to special relativity, with all observers agreeing on the form of the Euler–Lagrange equations; instead of being equal to its kinetic energy, the Lagrangian of a free particle is $-mc^2/\gamma(v)$.

The set of space-time events $(ct, \boldsymbol{r})$ described by special relativity is known as *Minkowski space* or Minkowski space-time (discussed in TENSORS AND MANIFOLDS [II.33]). It is a flat four-dimensional manifold with one time coordinate and three space coordinates (i.e., a manifold with a $3 + 1$ pseudo-Euclidean metric), and the inertial frames with perpendicular spatial axes are analogous to Cartesian coordinates in Euclidean space. Time intervals and (simultaneous) spatial distances between events are no longer separately invariant; only the *space-time interval*

$$s^2 = |\Delta\boldsymbol{r}|^2 - c^2\Delta t^2$$

between two events with spatial separation $\Delta\boldsymbol{r}$ and time separation Δt is invariant to Lorentz transformations and hence takes the same value in all inertial frames. s^2 may be positive, zero, or negative: if positive, there is a frame in which the pair of events are simultaneous; if zero, the events lie on the trajectory of a light ray; if negative, there is a frame in which the events occur at the same position. Since nothing can travel faster than

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

c, only events with $s^2 \leqslant 0$ can be causally connected; the events that may be affected by a given space-time point are those within its *future light cone*.

Being a manifold, the symmetries of Minkowski space are easily characterized, and its continuous symmetries are generated by vector fields, called *Killing vector fields*: dynamics in the space-time manifold is not changed under infinitesimal pointwise translation along these fields. The symmetries are translations in four linearly independent space-time directions, rotation about three linearly independent spatial axes, and Lorentz boosts in three linearly independent spatial directions (which are equivalent to rotations in space-time that mix the t-direction with a spatial direction). There are therefore ten independent symmetry transformations in Minkowski space-time (generating the Poincaré group), such that, for special relativistic systems not experiencing external forces, relativistic energy, momentum, angular momentum, and the analogue of $\boldsymbol{C}$ in (4) are conserved.

Both electromagnetic fields described by MAXWELL'S EQUATIONS [III.22] and relativistic quantum matter waves (strictly for a single particle) described by the DIRAC EQUATION [III.9] can be expressed as Lagrangian field theories; the combination of these is called *classical field theory*. As described above, Noether's theorem relates continuous symmetries of these theories to continuity equations of relativistic 4-currents. For instance, space-time translation symmetry ensures that the relativistic rank-2 *stress–energy–momentum* tensor, which describes the space-time flux of energy (including matter), is divergence free.

5 Other Theories

Einstein's theory of *general relativity* is a yet more general approach, which admits transformations between any coordinate systems on space-time (not simply inertial frames). This general formulation now applies to an arbitrary coordinate system, such as rotating coordinates (resolving the paradoxes implicit in Newton's bucket problem). This is formulated on a possibly curved, pseudo-Riemannian manifold where local neighborhoods of space-time events are equivalent to Minkowski space and free particles follow geodesics on the manifold. The geometry of the manifold—expressed by the Einstein curvature tensor—is proportional, by EINSTEIN'S FIELD EQUATIONS [III.10], to the stress–energy–momentum tensor, and gravitational forces are fictitious forces as freely falling particles follow curved space-time geodesics. The symmetries of general relativistic space-time manifolds are much more complicated than those for Minkowski space but are still formulated in terms of Killing vector fields that generate transformations that keep equations invariant.

In quantum physics (in a Galilean or special relativistic setting), the ideas described in this article are very important in *quantum field theory*, required to describe the quantum nature of electromagnetic and other fields, and systems involving many quantum particles. In quantum field theory, all energies in the system are quantized, often around a minimum-energy configuration. However, many field theories involve potentials whose minimum energy is not the most symmetric choice of origin; for instance, a "Mexican hat potential" of the form $|\boldsymbol{v}|^4 - 2a|\boldsymbol{v}|^2$ for some field vector $\boldsymbol{v}$ and constant $a > 0$ is rotationally symmetric around $\boldsymbol{v} = 0$ but has a minimum for any $\boldsymbol{v}$ with $|\boldsymbol{v}| = \sqrt{a}$. A minimum-energy excitation of the system *breaks* this symmetry by choosing an appropriate minimum energy $\boldsymbol{v}$. Many phenomena in the quantum theory of condensed matter (such as the Meissner effect in superconductors) and fundamental particles (such as the Higgs boson) arise from this type of symmetry breaking.

In classical mechanics, the time direction is privileged (it is common to all observers), so its space-time structure cannot be phrased simply in terms of manifolds. Nevertheless, the Galilean group of transformations naturally has a fiber bundle structure, with the base space given by the one-dimensional Euclidean line $\mathbb{E}^1$ describing time and the fiber being the spatial coordinates given by $\mathbb{E}^3$.

Galilean transformations can naturally be built into this fiber bundle structure, known as *Newton–Cartan space-time*, by defining the paths of free inertial particles (i.e., straight lines) in the connections of the bundle. This provides a useful mathematical framework to compare the physical laws and symmetries of Newtonian and relativity theories, and interesting connections exist between general relativity and Newtonian gravity in the Newton–Cartan formalism.

Further Reading

Lorentz, H. A., A. Einstein, H. Minkowski, and H. Weyl. 1952. *The Principle of Relativity*. New York: Dover.

Misner, C. W., K. S. Thorne, and J. A. Wheeler. 1973. *Gravitation*. New York: W. H. Freeman.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Neuenschwander, D. E. 2010. *Emmy Noether's Wonderful Theorem*. Baltimore, MD: Johns Hopkins University Press.
Penrose, R. 2004. *The Road to Reality* (chapters 17–20 are particularly relevant). London: Random House.

II.22 The Jordan Canonical Form
Nicholas J. Higham

A canonical form for a class of matrices is a form of matrix—usually chosen to be as simple as possible—to which all members of the class can be reduced by transformations of a specified kind. The Jordan canonical form (JCF) is associated with similarity transformations on $n \times n$ matrices. A similarity transformation of a matrix A is a transformation from A to $X^{-1}AX$, where X is nonsingular. The JCF is the simplest form that can be achieved by similarity transformations, in the sense that it is the closest to a diagonal matrix.

The JCF of a complex $n \times n$ matrix A can be written $A = ZJZ^{-1}$, where Z is nonsingular and the Jordan matrix J is a block-diagonal matrix

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}$$

with diagonal blocks of the form

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}.$$

Here, blanks denote zero blocks or zero entries. The matrix J is unique up to permutation of the diagonal blocks, but Z is not. Each λ_k is an eigenvalue of A and may appear in more than one Jordan block. All the EIGENVALUES [I.2 §20] of the Jordan block J_k are equal to λ_k. By definition, an eigenvector of J_k is a nonzero vector x satisfying $J_k x = \lambda_k x$, and all such x are nonzero multiples of the vector $x = [1\ 0\ \cdots\ 0]^{\mathrm{T}}$. Therefore J_k has only one linearly independent eigenvector. Expand x to a vector $\tilde{x}$ with n components by padding it with zeros in positions corresponding to each of the other Jordan blocks J_i, $i \neq k$. The vector $\tilde{x}$ has a single 1, in the rth component, say. A corresponding eigenvector of A is $Z\tilde{x}$, since $A(Z\tilde{x}) = ZJZ^{-1}(Z\tilde{x}) = ZJ\tilde{x} = \lambda_k Z\tilde{x}$; this eigenvector is the rth column of Z.

If every block J_k is 1×1 then J is diagonal and A is similar to a diagonal matrix; such matrices A are called *diagonalizable*. For example, real symmetric matrices are diagonalizable—and moreover the eigenvalues are real and the matrix Z in the JCF can be taken to be orthogonal. A matrix that is not diagonalizable is *defective*; such matrices do not have a complete set of linearly independent eigenvectors or, equivalently, their Jordan form has at least one block of dimension 2 or greater.

To give a specific example, the matrix

$$A = \frac{1}{2}\begin{bmatrix} 3 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

has a JCF with

$$Z = \begin{bmatrix} 0 & \frac{1}{2} & 1 \\ -1 & -\frac{1}{2} & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad J = \left[\begin{array}{c|cc} \frac{1}{2} & 0 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right].$$

As the partitioning of J indicates, there are two Jordan blocks: a 1×1 block with eigenvalue $\frac{1}{2}$ and a 2×2 block with eigenvalue 1. The eigenvalue $\frac{1}{2}$ of A has an associated eigenvector equal to the first column of Z. For the double eigenvalue 1 there is only one linearly independent eigenvector, namely the second column, z_2, of Z. The third column, z_3, of Z is a generalized eigenvector: it satisfies $Az_3 = z_2 + z_3$.

The JCF provides complete information about the eigensystem. The *geometric multiplicity* of an eigenvalue, defined as the number of associated linearly independent eigenvectors, is the number of Jordan blocks in which that eigenvalue appears. The *algebraic multiplicity* of an eigenvalue, defined as its multiplicity as a zero of the characteristic polynomial $q(t) = \det(tI - A)$, is the number of copies of the eigenvalue among all the Jordan blocks. For the matrix (1) above, the geometric multiplicity of the eigenvalue 1 is 1 and the algebraic multiplicity is 2, while the eigenvalue $\frac{1}{2}$ has geometric and algebraic multiplicities both equal to 1.

The *minimal polynomial* of a matrix is the unique monic polynomial ψ of lowest degree such that $\psi(A) = 0$. The degree of ψ is certainly no larger than n because the CAYLEY–HAMILTON THEOREM [IV.10 §5.3] states that $q(A) = 0$. The minimal polynomial of an $m \times m$ Jordan block $J_k(\lambda_k)$ is $(t - \lambda_k)^m$. The minimal polynomial of A is therefore given by

$$\psi(t) = \prod_{i=1}^{s} (t - \lambda_i)^{m_i},$$

For general queries, contact webmaster@press.princeton.edu