Chapter 1

Differential Geometry and Mathematical Physics

Andrew Hone and Steffen Krusch
School of Mathematics, Statistics & Actuarial Science
University of Kent, Canterbury CT2 7NF, UK

The chapter will illustrate how concepts in differential geometry arise naturally in different areas of mathematical physics. We will describe manifolds, fibre bundles, (co)tangent bundles, metrics and symplectic structures, and their applications to Lagrangian mechanics, field theory and Hamiltonian systems, including various examples related to integrable systems and topological solitons.

1. Manifolds

Manifolds are a central concept in mathematics and have natural applications to problems in physics. Here we provide an example-led introduction to manifolds and introduce important additional structures.

A simple, yet non-trivial, example of a manifold is the 2-sphere S^2, given by the set of points

$$S^2 = \left\{ (x_1, x_2, x_3) : \sum_{i=1}^{3} x_i^2 = 1 \right\}.$$

Often, we label the points on the sphere by polar coordinates:

$$x_1 = \cos\phi \sin\theta, \quad x_2 = \sin\phi \sin\theta, \quad x_3 = \cos\theta,$$

$$\text{where } 0 \leq \phi < 2\pi,\ 0 \leq \theta \leq \pi.$$

However, there is the following problem: we cannot label S^2 with a single coordinate system such that nearby points have nearby coordinates, and every point has unique coordinates.

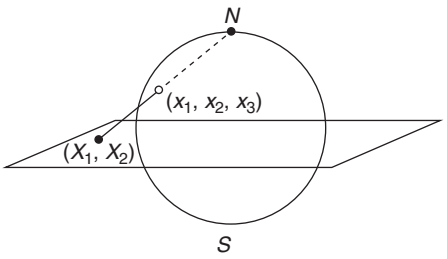


Fig. 1. Stereographic projection.

Polar coordinates suffer both problems: the point $(1,0,0)$ has coordinates $\theta = \pi/2$ and $\phi = 0$, whereas the nearby point $(\sqrt{1-\epsilon^2}, -\epsilon, 0)$ for small ϵ has coordinates $\theta = 0$ and $\phi \approx 2\pi - \epsilon$, which is not nearby. Furthermore, the north pole $(0,0,1)$ is given by $\theta = 0$ and ϕ is arbitrary, which is clearly not unique. An alternative is to use stereographic coordinates which are constructed as follows. Let N be the north pole given by $(0,0,1)$. Given any point $P \in S^2 \setminus \{N\}$, draw a line through N and P. This line intersects the x_1x_2-plane at the point $(X_1, X_2, 0)$, see Fig. 1. A short calculation shows that these coordinates are given by $X_1 = \frac{x_1}{1-x_3}, X_2 = \frac{x_2}{1-x_3}$. Stereographic coordinates give every point other than the North pole N a unique set of coordinates. However, close to the North pole nearby points do not have nearby coordinates. In order to proceed, we have to use more than one coordinate system, and this gives rise to the definition of a manifold.

Definition 1.1. M is an m-dimensional (differentiable) manifold if

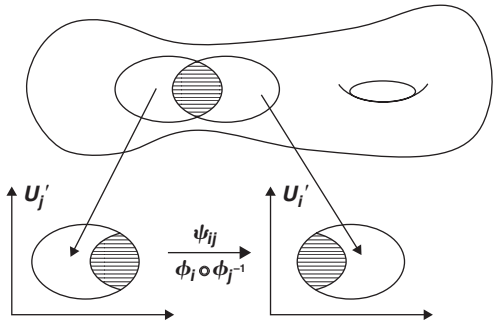
- M is a topological space.
- M comes with family of charts $\{(U_i, \phi_i)\}$ known as an *atlas.*[a]
- $\{U_i\}$ is family of open sets covering M: $\bigcup_i U_i = M$.
- ϕ_i is homeomorphism from U_i onto an open subset U_i' of $\mathbb{R}^m$.
- For each $U_i \cap U_j \neq \emptyset$, the map

$$\psi_{ij} = \phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$$

is C^∞. The ψ_{ij} are called *crossover maps.*

The above definition is illustrated in Fig. 2.

[a]Two atlases are *compatible* if their union is again an atlas. Compatibility introduces an equivalence relation with equivalence classes known as *differential structures.*



$$\psi_{ij} = \phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$$

Fig. 2. A manifold and two overlapping charts.

Example 1.2 (The 2-sphere S^2). The projection from the North pole gives

$$X_1 = \frac{x_1}{1 - x_3}, \quad X_2 = \frac{x_2}{1 - x_3},$$

with open sets $U_1 = S^2 \setminus \{N\}$ and $U_1' = \mathbb{R}^2$. Then the first chart is given by

$$\phi_1 : \ U_1 \to \mathbb{R}^2$$
$$(x_1, x_2, x_3) \mapsto (X_1, X_2).$$

Similarly, the projection from the South pole results in

$$Y_1 = \frac{x_1}{1 + x_3}, \quad Y_2 = \frac{x_2}{1 + x_3},$$

with open sets $U_2 = S^2 \setminus \{S\}$ and $U_2' = \mathbb{R}^2$. Then the second chart is

$$\phi_2 : \ U_2 \to \mathbb{R}^2$$
$$(x_1, x_2, x_3) \mapsto (Y_1, Y_2).$$

The *crossover map* $\psi_{21} = \phi_2 \circ \phi_1^{-1}$ can then be calculated as

$$\psi_{21} : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2; (X_1, X_2) \mapsto (Y_1, Y_2) = \left(\frac{X_1}{X_1^2 + X_2^2}, \frac{-X_2}{X_1^2 + X_2^2} \right). \quad (1.1)$$

Example 1.3 (The 3-sphere S^3). The 3-sphere is given by the set of points

$$S^3 = \left\{ (x_1, x_2, x_3, x_4) : \sum_{i=1}^{4} x_i^2 = 1 \right\}.$$

It can also be labelled by stereographic coordinates, which work the same way as for S^2, e.g. $X_i = x_i/(1-x_4)$ for $i = 1, 2, 3$ define the projection from the North pole N. However, S^3 can also be identified with the Lie group SU(2), i.e. complex 2×2 matrices which satisfy

$$U\, U^\dagger = U^\dagger U = 1 \quad \text{and} \quad \det U = 1. \tag{1.2}$$

Upon setting

$$U = \begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix},$$

all the conditions in (1.2) are seen to be satisfied provided $|z_1|^2 + |z_2|^2 = 1$, which corresponds to the equation for S^3 e.g. if we set $z_1 = x_1 + \mathrm{i}x_2$, $z_2 = x_3 + \mathrm{i}x_4$. Hence the 3-sphere can be given a group structure.

1.1. *Maps between manifolds*

Let M be an m-dimensional manifold with charts $\phi_i : U_i \to \mathbb{R}^m$ and N be an n-dimensional manifold with charts $\psi_j : \tilde{U}_j \to \mathbb{R}^n$. Let $f : M \to N$ be a map between the two manifolds, so a point $p \mapsto f(p)$. This has a coordinate presentation

$$F_{ji} = \psi_j \circ f \circ {\phi_i}^{-1} : \mathbb{R}^m \to \mathbb{R}^n, x \mapsto \psi_j(f(\phi_i^{-1}(x))),$$

where $x = \phi_i(p)$ ($p \in U_i$ and $f(p) \in \tilde{U}_j$). Using the coordinate presentation, all the rules of calculus in $\mathbb{R}^n$ work for maps between manifolds. If the presentations F_{ji} are differentiable in all charts, then f is said to be differentiable.

2. Fibre Bundles

Often manifolds can be built up from smaller manifolds. An important example is the Cartesian product of two manifolds, $E = M \times F$. *Fibre bundles* are manifolds which look like Cartesian products *locally*, but not *globally*. This concept is very useful for physics: non-trivial fibre bundles

occur in general relativity, and also in field theories due to boundary conditions "at infinity."

Definition 2.1. A *fibre bundle* (E, π, M, F, G) consists of the following properties:

- A manifold E called the *total space*, a manifold M called the *base space* and a manifold F called the *fibre* (or typical fibre).
- A surjection $\pi : E \to M$ called the *projection.* The inverse image of a point $p \in M$ is called the fibre at p, namely $\pi^{-1}(p) = F_p \cong F$.
- A Lie group G called the *structure group* which acts on F on the left.
- A collection of open sets $\{U_i\}$ covering M and diffeomorphisms $\phi_i : U_i \times F \to \pi^{-1}(U_i)$, such that $\pi \circ \phi_i(p, f) = p$. Each map ϕ_i is called a *local trivialization*, since ϕ_i^{-1} maps $\pi^{-1}(U_i)$ to $U_i \times F$.
- Transition functions $t_{ij} : U_i \cap U_j \to G$, such that $\phi_j(p, f) = \phi_i(p, t_{ij}(p)f)$, so $t_{ij} = \phi_i^{-1} \circ \phi_j$ holds for fixed p.

The transition functions t_{ij} satisfy the conditions

$$\begin{array}{ll} t_{ii}(p) = e & \forall p \in U_i, \\ t_{ij}(p) = t_{ji}^{-1}(p) & \forall p \in U_i \cap U_j, \\ t_{ij}(p) \cdot t_{jk}(p) = t_{ik}(p) & \forall p \in U_i \cap U_j \cap U_k, \end{array}$$

to be consistent with the fact that $t_{ij} \in G$. If all the transition functions are the identity map e, then the fibre bundle is the *trivial bundle* $E = M \times F$. In general, the transition functions of two local trivializations $\{\phi_i\}$ and $\{\tilde{\phi}_i\}$ for fixed $\{U_i\}$ are related via

$$\tilde{t}_{ij}(p) = g_i^{-1}(p) \cdot t_{ij}(p) \cdot g_j(p),$$

where for fixed p, we define $g_i : F \to F$, $g_i = \phi_i^{-1} \circ \tilde{\phi}_i$. For the trivial bundle, $t_{ij}(p) = g_i^{-1}(p) \cdot g_j(p)$.

Example 2.2 ($U(1)$ bundles over S^2 and magnetic monopoles). Consider a fibre bundle with fibre $U(1)$ and base space S^2. Let $\{U_N, U_S\}$ be an open covering of S^2 where

$$U_N = \{(\theta, \phi) : 0 \leq \theta < \pi/2 + \epsilon, 0 \leq \phi < 2\pi\},$$
$$U_S = \{(\theta, \phi) : \pi/2 - \epsilon < \theta \leq \pi, 0 \leq \phi < 2\pi\}.$$

The intersection $U_N \cap U_S$ is a strip which is basically the equator. The local trivializations are $\phi_N^{-1}(u) = (p, e^{i\alpha_N})$, $\phi_S^{-1}(u) = (p, e^{i\alpha_S})$, where $p = \pi(u)$.

Possible transition functions are $t_{NS} = e^{in\phi}$, where $n \in \mathbb{Z}$, and define different bundles, denoted P_n. The fibre coordinates in $U_N \cap U_S$ are related via

$$e^{i\alpha_N} = e^{in\phi} e^{i\alpha_S}.$$

If $n = 0$ this is the trivial bundle $P_0 = S^2 \times S^1$. For $n \neq 0$ the $U(1)$ bundle P_n is twisted. P_n is an example of a *principal bundle* because the fibre is the same as the structure group $G = U(1)$. In physics, P_n is interpreted as a magnetic monopole of charge n.

Example 2.3 (Hopf bundle). The Hopf map $\pi : S^3 \to S^2$, with $S^3 = \left\{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\right\}$, is defined by

$$\xi_1 = 2(x_1x_3 + x_2x_4), \quad \xi_2 = 2(x_2x_3 - x_1x_4), \quad \xi_3 = x_1^2 + x_2^2 - x_3^2 - x_4^2,$$

which implies $\xi_1^2 + \xi_2^2 + \xi_3^2 = 1$. It turns out that with this choice of coordinates S^3 can be identified with P_1, a non-trivial $U(1)$ bundle over S^2, known as the Hopf bundle.

2.1. *Tangent vectors*

Given a curve $c : (-\epsilon, \epsilon) \to M$ and a function $f : M \to \mathbb{R}$, we define the tangent vector $X[f]$ at $c(0)$ by taking the the directional derivative of $f(c(t))$ along $c(t)$ at $t = 0$, namely,[b]

$$X[f] = \left.\frac{df(c(t))}{dt}\right|_{t=0}.$$

In local coordinates (x^μ), with $n = \dim M$, this becomes

$$\sum_{\mu=1}^{n} \frac{\partial f}{\partial x^\mu} \left.\frac{dx^\mu(c(t))}{dt}\right|_{t=0}.$$

Hence tangent vectors act on functions via

$$X[f] = X^\mu \left(\frac{\partial f}{\partial x^\mu}\right),$$

where above there is an implicit sum over repeated indices. (This is the Einstein summation convention, and henceforth we shall always use this.)

[b]Alternatively, tangent vectors can be defined as equivalence classes of curves.

Vectors are independent of the choice of coordinates, hence we write

$$X = X^\mu \frac{\partial}{\partial x^\mu} = \tilde{X}^\mu \frac{\partial}{\partial y^\mu}$$

for another set of coordinates (y^μ), with X^μ and $\tilde{X}^\mu$ related via $\tilde{X}^\mu = X^\nu \frac{\partial y^\mu}{\partial x^\nu}$. It is very useful to define the pairing

$$\left\langle dx^\nu, \frac{\partial}{\partial x^\mu} \right\rangle = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu,$$

which leads us to 1-*forms* (or covectors) $\omega = \omega_\mu dx^\mu$, also independent of choice of coordinates, and extends by linearity to the pairing between vectors (or vector fields) and 1-forms:

$$i_X \omega := \langle \omega, X \rangle = \omega_\mu X^\mu.$$

Now, under a change of local coordinates, we have

$$\omega = \omega_\mu dx^\mu = \tilde{\omega}_\nu dy^\nu \implies \tilde{\omega}_\nu = \omega_\mu \frac{\partial x^\mu}{\partial y^\nu}. \tag{2.1}$$

This can be generalized further to tensors $T^{\mu_1 \cdots \mu_q}{}_{\nu_1 \ldots \nu_r}$ of type (q, r), with q upper (contravariant) and r lower (covariant) indices.

2.2. *Tangent bundle*

Let M be an n-dimensional manifold. At each point $p \in M$, tangent vectors can be added and multiplied by real numbers, so the set of all tangent vectors forms an n-dimensional vector space, denoted T_pM. For a set of local coordinates (x^μ) at p, a basis of T_pM is given by $\partial/\partial x^\mu$, $1 \leq \mu \leq n$. The union of all tangent spaces forms the tangent bundle

$$TM = \bigcup_{p \in M} T_pM.$$

TM is a $2n$-dimensional manifold with base space M and fibre $\mathbb{R}^n$. It is an example of an important kind of fibre bundle known as a *vector bundle.*

Similarly, at each point $p \in M$ the 1-forms form the cotangent space ${T_p}^*M$, which is the dual space to T_pM. The union of all cotangent spaces forms the cotangent bundle

$$T^*M = \bigcup_{p \in M} T^*_pM,$$

another $2n$-dimensional manifold which is a vector bundle with base space M and fibre $\mathbb{R}^n$. More generally, at each point p, tensors of type (q, r) are

elements of the tensor product $T_pM \otimes \cdots \otimes T_pM \otimes T_p^*M \otimes \cdots \otimes T_p^*M$ with q copies of T_pM and r copies of ${T_p}^*M$, and the corresponding tensor bundle is defined similarly.

Example 2.4 (The tangent bundle TS^2). To construct the tangent bundle of S^2 we use the two stereographic projections as our charts. The coordinates $(X_1, X_2) \in U_1'$ and $(Y_1, Y_2) \in U_2'$ are related via (1.1). Given $u \in TS^2$ with $\pi(u) = p \in U_1 \cap U_2$, then the local trivializations ϕ_1 and ϕ_2 satisfy $\phi_1^{-1}(u) = (p, V_1^\mu)$ and $\phi_2^{-1}(u) = (p, V_2^\mu)$. The local trivialization is

$$t_{12} = \frac{\partial(Y_1, Y_2)}{\partial(X_1, X_2)} = \frac{1}{(X_1^2 + X_2^2)^2} \begin{pmatrix} X_2^2 - X_1^2 & -2X_1X_2 \\ -2X_1X_2 & X_1^2 - X_2^2 \end{pmatrix}.$$

The reader may check that $t_{21}(p) = t_{12}^{-1}(p)$.

Given a smooth map $f : M \to N$ between manifolds, we can define a map between tangent spaces,

$$\begin{aligned} f_* : T_pM &\to T_{f(p)}N \\ V &\mapsto f_*V, \end{aligned}$$

which is called the *pushforward* of f. If $g \in C^\infty(N)$ then $g \circ f \in C^\infty(M)$, and the action of the vector f_*V on g is given by

$$f_*V(g) = V(g \circ f).$$

Similarly, there is a map between cotangent spaces,

$$\begin{aligned} f^* : {T^*}_{f(p)}N &\to {T^*}_pM \\ \omega &\mapsto f^*\omega \end{aligned}$$

called the *pullback*. The pullback can be defined via the pairing between vectors and 1-forms: $\langle f^*\omega, V\rangle_M = \langle \omega, f_*V\rangle_N$.

2.3. *Sections*

Definition 2.5. Let (E, M, π) be a fibre bundle. A section $s : M \to E$ is a smooth map which satisfies $\pi \circ s = id_M$. Here, $s|_p$ is an element of the fibre $F_p = \pi^{-1}(p)$. The space of sections is denoted by $\Gamma(E)$.

A local section is defined on a subset $U \subset M$, only. Not all fibre bundles admit global sections (defined on the whole of M).

Example 2.6 (Vector and tensor fields). A vector field V on a manifold M is a section of the tangent bundle: $V \in \Gamma(TM)$, associating a tangent vector to each point in M. In local coordinates $x = (x^\mu)$ this means $V = V^\mu(x)\partial/\partial x^\mu$, where the components $V^\mu(x)$ define a map from (part of) $\mathbb{R}^n \to \mathbb{R}^n$. Similarly, a differential 1-form is a section of T^*M, and a tensor field of type (q, r) is a section of the corresponding tensor bundle.

Example 2.7 (Wave functions). The wave function $\psi(\mathbf{x}, t)$ of a single particle in quantum mechanics can be thought of as a section of a complex line bundle $E = \mathbb{R}^{3,1} \times \mathbb{C}$.

Vector bundles always have at least one section, the null section s_0 with

$$\phi_i^{-1}(s_0(p)) = (p, 0)$$

in any local trivialization. A principal bundle E only admits a global section if it is trivial: $E = M \times G$. A section in a principal bundle can be used to construct the trivialization of the bundle, using the fact that there is a right action which is independent of the local trivialization: $ua = \phi_i(p, g_i a)$ for any $a \in G$.

2.4. *Associated vector bundle*

Given a principal fibre bundle P with base M, $\pi : P \to M$ with $\pi^{-1}(p) \cong G$, and a k-dimensional representation $\rho : G \to \mathrm{Aut}(V)$ of G on a vector space V, the *associated vector bundle* $E = P \times_\rho V$ is defined by identifying the points

$$(u, v) \quad \text{and} \quad (ug, \rho(g)^{-1}v) \in P \times V,$$

where $u \in P$, $g \in G$, and $v \in V$. The projection $\pi_E : E \to M$ is defined by $\pi_E(u, v) = \pi(u)$, which is well defined because

$$\pi_E(ug, \rho(g)^{-1}v) = \pi(ug) = \pi(u) = \pi_E(u, v).$$

The transition functions of E are given by $\rho(t_{ij}(p))$ where $t_{ij}(p)$ are the transition functions of P. Conversely, a vector bundle naturally induces a principal bundle associated with it.

2.5. *Metric tensor*

Manifolds can carry further structure, an important example being a *metric.* A metric g is a $(0, 2)$ tensor which, at each point $p \in M$, must satisfy

the following conditions, for all $U, V \in T_pM$:

(1) $g_p(U,V) = g_p(V,U)$;
(2) $g_p(U,U) \geq 0$, with equality only when $U = 0$.

The metric g provides an inner product on each tangent space T_pM, and (M,g) is called a Riemannian manifold.[c] The metric gives an isomorphism between vector fields $X \in \Gamma(TM)$ and 1-forms $\eta \in \Gamma(T^*M)$ via

$$g(\cdot, X) = \eta_X.$$

In physics notation $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$ lower and raise indices, and one writes

$$g = ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu.$$

If M is a submanifold of N with metric g_N and $f : M \to N$ is the embedding map, then the induced metric g_M is

$$g_{M\mu\nu}(x) = g_{N\alpha\beta}(f(x)) \frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\beta}{\partial x^\nu}.$$

2.6. *Connection on the tangent bundle*

The "derivative" of a vector field $V = V^\mu \frac{\partial}{\partial x^\mu}$ with respect to x^ν is

$$\frac{\partial V^\mu}{\partial x^\nu} = \lim_{\Delta x^\nu \to 0} \frac{V^\mu(\dots, x^\nu + \Delta x^\nu, \dots) - V^\mu(\dots, x^\nu, \dots)}{\Delta x^\nu}.$$

This does not work as the first vector is defined at $x + \Delta x$ and the second at x. We need to transport the vector V^μ from x to $x + \Delta x$ "without change." This is known as *parallel transport*, and is achieved by specifying a *connection* $\Gamma^\mu{}_{\nu\lambda}$, namely the parallel transported vector $\tilde{V}^\mu$ is given by

$$\tilde{V}^\mu(x + \Delta x) = V^\mu(x) - V^\lambda(x)\Gamma^\mu{}_{\nu\lambda}(x)\Delta x^\nu.$$

The covariant derivative of V with respect to x^ν is

$$\lim_{\Delta x^\nu \to 0} \frac{V^\mu(x + \Delta x) - \tilde{V}^\mu(x + \Delta x)}{\Delta x^\nu} = \frac{\partial V^\mu}{\partial x^\nu} + V^\lambda \Gamma^\mu{}_{\nu\lambda}.$$

[c]Condition (2) defines a Riemannian metric. If g is non-degenerate but of indefinite signature, it is called pseudo-Riemannian; this is required for special and general relativity.

In the case where a manifold comes equipped with a (pseudo-) Riemannian metric g, it can be required that the metric g is covariantly constant, meaning that if two vectors X and Y are parallel transported along any curve, then the inner product $g(X,Y)$ remains constant. The condition

$$\nabla_V(g(X,Y)) = 0,$$

yields the *Levi-Civita connection.* The Levi-Civita connection is specified locally by the *Christoffel symbols*, written as

$$\Gamma^{\kappa}{}_{\mu\nu} = \tfrac{1}{2}g^{\kappa\lambda}\left(\partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}\right). \tag{2.2}$$

Example 2.8 (General relativity). The Christoffel symbols do not transform like a tensor. However, from them, one builds the Riemann curvature tensor:

$$R^{\kappa}{}_{\lambda\mu\nu} = \partial_\mu\Gamma^{\kappa}{}_{\nu\lambda} - \partial_\nu\Gamma^{\kappa}{}_{\mu\lambda} + \Gamma^{\eta}{}_{\nu\lambda}\Gamma^{\kappa}{}_{\mu\eta} - \Gamma^{\eta}{}_{\mu\lambda}\Gamma^{\kappa}{}_{\nu\eta}.$$

Important contractions of the curvature tensor are the *Ricci tensor Ric*,

$$Ric_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu},$$

and the *scalar curvature* $\mathcal{R}$,

$$\mathcal{R} = g^{\mu\nu}Ric_{\mu\nu}.$$

Now, we have the ingredients for *Einstein's field equations of general relativity*, namely

$$Ric_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}\mathcal{R} = 8\pi G T_{\mu\nu},$$

where G is the gravitational constant and $T_{\mu\nu}$ is the energy–momentum tensor which describes the distribution of matter in spacetime.

2.7. *Yang–Mills theory and fibre bundles*

An example of Yang–Mills theory is given by the Lagrangian density

$$\mathcal{L} = \tfrac{1}{8}\mathrm{Tr}\left(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}\right) + \tfrac{1}{2}\left(D_\mu\Phi\right)^\dagger D^\mu\Phi - U(\Phi^\dagger\Phi), \tag{2.3}$$

where $U(\cdot)$ is a potential energy density,

$$D_\mu\Phi = \partial_\mu\Phi + \mathbf{A}_\mu\Phi, \quad \text{and} \quad \mathbf{F}_{\mu\nu} = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu].$$

Here Φ is a two-component complex scalar field (Higgs field), while $\mathbf{A}_\mu$ is called a gauge field and is $\mathfrak{su}(2)$-valued, i.e. $\mathbf{A}_\mu$ are anti-Hermitian 2×2 matrices; $\mathbf{F}_{\mu\nu}$ is known as the field strength (also $\mathfrak{su}(2)$-valued). This Lagrangian is Lorentz invariant: indices are raised and lowered with the Minkowski metric $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$, and the Lagrangian remains the same under linear coordinate transformations which preserve this quadratic form.

The Lagrangian (2.3) is also invariant under local *gauge transformations:* Let $\mathbf{g} \in \mathrm{SU}(2)$ be a spacetime dependent gauge transformation with

$$\Phi \mapsto \mathbf{g}\Phi, \quad \text{and} \quad \mathbf{A}_\mu \mapsto \mathbf{g}\mathbf{A}_\mu\mathbf{g}^{-1} - \partial_\mu\mathbf{g}\cdot\mathbf{g}^{-1}.$$

The covariant derivative $D_\mu\Phi$ transforms as

$$D_\mu\Phi \mapsto \partial_\mu(\mathbf{g}\Phi) + \left(\mathbf{g}\mathbf{A}_\mu\mathbf{g}^{-1} - \partial_\mu\mathbf{g}\cdot\mathbf{g}^{-1}\right)\mathbf{g}\Phi = \mathbf{g}D_\mu\Phi$$

Hence $\Phi^\dagger\Phi \mapsto (\mathbf{g}\Phi)^\dagger\mathbf{g}\Phi = \Phi^\dagger\mathbf{g}^\dagger\mathbf{g}\Phi = \Phi^\dagger\Phi$, and similarly for $(D_\mu\Phi)^\dagger D^\mu\Phi$. Finally, $\mathbf{F}_{\mu\nu} \mapsto \mathbf{g}\mathbf{F}_{\mu\nu}\mathbf{g}^{-1}$, so $\mathrm{Tr}\,(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu})$ is also gauge invariant.

Yang–Mills theory can also be described in a more mathematical language: The gauge field $\mathbf{A}_\mu$ corresponds to a connection on a principal SU(2) bundle. The field strength $\mathbf{F}_{\mu\nu}$ corresponds to the curvature of the principal SU(2) bundle. The Higgs field Φ is a section of the associated $\mathbb{C}^2$ vector bundle. The action of $\mathbf{g} \in \mathrm{SU}(2)$ on Φ and $\mathbf{A}_\mu$ is precisely what is expected for an associated fibre bundle. Surprisingly, mathematicians and physicists derived the same result very much independently!

3. Forms and Integration

3.1. *Differential forms*

Differential r-forms are $(0, r)$ tensor fields which are completely antisymmetric, that is, they acquire a minus sign under interchange of any two indices. A basis for $\Omega^r(M)$, the set of r-forms on an n-dimensional manifold M, is given in local coordinates at each point by

$$\{dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \,|\, 1 \leq \mu_1 < \cdots < \mu_r \leq n\}.$$

Setting $\Omega^0(M) = C^\infty(M)$, for $0 \leq k, l \leq n$ the *wedge product* gives a bilinear map

$$\wedge : \Omega^k \times \Omega^l \to \Omega^{k+l},$$

satisfying associativity, i.e. $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$, and the property

$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha.$$

The *exterior derivative* $d : \Omega^k \to \Omega^{k+1}$ can be defined as follows: Given

$$\omega = \frac{1}{k!} \omega_{\mu_1 \dots \mu_k} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k} \in \Omega^k,$$

its exterior derivative $d\omega \in \Omega^{k+1}$ is

$$d\omega = \frac{1}{k!} \left(\frac{\partial}{\partial x_\nu} \omega_{\mu_1 \dots \mu_k} \right) dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k}.$$

From this definition it is an easy exercise to show that $d^2\omega = 0$ for any differential k-form ω.

Example 3.1 (Symplectic form). A symplectic form ω is a 2-form which satisfies

(1) ω is closed, i.e. $d\omega = 0$;
(2) ω is non-degenerate: at each point $p \in M$, $\omega(U, V) = 0$ for all $V \in T_pM$ implies $U = 0$.

In local coordinates, ω can be written as

$$\omega = \tfrac{1}{2} \omega_{\mu\nu} dx^\mu \wedge dx^\nu.$$

Like a metric g, a symplectic form also provides an isomorphism between vector fields and 1-forms via

$$\begin{aligned} \Gamma(TM) &\to \Gamma(T^*M) \\ V &\mapsto \omega(\cdot, V), \end{aligned} \tag{3.1}$$

and a manifold M equipped with a symplectic form is called a *symplectic manifold*.

3.2. *Integration on manifolds*

Recall that, under change of coordinates, 1-forms transform according to (2.1). Two charts define the same orientation provided that

$$\det \left(\frac{\partial x^\mu}{\partial y^\nu} \right) > 0. \tag{3.2}$$

A manifold is orientable if for any overlapping charts U_i and U_j there exist local coordinates x^μ for U_i and y^μ for U_j such that (3.2) holds. The invariant

volume element on a (pseudo-)Riemannian manifold M is given by

$$\Omega = \sqrt{|g|}dx^1 \wedge \cdots \wedge dx^m \quad \text{where } |g| = |\det(g_{\mu\nu})|. \tag{3.3}$$

Given Ω and an orientation, a function $f : M \to \mathbb{R}$ can be integrated over M. For one chart,

$$\int_{U_i} f\Omega = \int_{\phi_i(U_i)} f(\phi_i^{-1}(x))\sqrt{|g(\phi_i^{-1}(x)|}dx^1 dx^2 \dots dx^m.$$

For the whole of M, one takes a *partition of unity*, which is a family of differentiable functions $\epsilon_i(p)$, $1 \leq i \leq k$ such that

(1) $0 \leq \epsilon_i(p) \leq 1$;
(2) $\epsilon_i(p) = 0$ if $p \notin U_i$;
(3) $\epsilon_1(p) + \cdots + \epsilon_k(p) = 1$ for any point $p \in M$.

Then the integral of f over the whole manifold M, covered by k charts, is

$$\int_M f\Omega = \sum_{i=1}^{k} \int_{U_i} f(p)\epsilon_i(p)\Omega.$$

A fundamental result on integration is the generalized Stokes theorem.

Theorem 3.2. *If ω is an r-form and R an $r+1$-dimensional region in M with boundary ∂R, then*

$$\int_R d\omega = \int_{\partial R} \omega.$$

Example 3.3 (Green's theorem). For $\omega = p\,dx + q\,dy$ in $\mathbb{R}^2$,

$$d\omega = (\partial_x q - \partial_y p)dx \wedge dy.$$

Hence, integrating around a closed curve $\mathcal{C}$ enclosing a region $R \subset \mathbb{R}^2$ gives

$$\oint_{\mathcal{C}} (pdx + qdy) = \iint_R (\partial_x q - \partial_y p)dxdy,$$

which is Green's theorem in the plane.

Example 3.4 (Stokes theorem and divergence theorem). In $\mathbb{R}^3$ with $\omega = f_1\,dx + f_2\,dy + f_3\,dz$, we have

$$d\omega = (\partial_y f_3 - \partial_z f_2)dy \wedge dz + (\partial_z f_1 - \partial_x f_3)dz \wedge dx + (\partial_x f_2 - \partial_y f_1)dx \wedge dy,$$

which gives rise to the traditional Stokes theorem in vector form:

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint_S (\nabla \wedge \mathbf{f}) \cdot \mathbf{n}\, dS.$$

On the other hand, if $\omega = f_1\,dy \wedge dz + f_2\,dz \wedge dx + f_3\,dx \wedge dy$ then

$$d\omega = (\partial_x f_1 + \partial_y f_2 + \partial_z f_3)\ dx \wedge dy \wedge dz,$$

which yields the divergence theorem:

$$\iiint_V \nabla \cdot \mathbf{f}\, dxdydz = \iint_S \mathbf{f} \cdot \mathbf{n}\, dS.$$

When a manifold has a metric, there is another important operation on forms called the *Hodge star*. The totally antisymmetric tensor is defined by

$$\epsilon_{\mu_1\mu_2\ldots\mu_n} = \begin{cases} +1 & \text{if } (\mu_1\mu_2\ldots\mu_n) \text{ is an even permutation of } (12\ldots n), \\ -1 & \text{if } (\mu_1\mu_2\ldots\mu_n) \text{ is an odd permutation of } (12\ldots n), \\ 0 & \text{otherwise.} \end{cases}$$

The Hodge star is a linear map $* : \Omega^r(M) \to \Omega^{n-r}(M)$ which acts on a basis vector in $\Omega^r(M)$ according to

$$*(dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r}) = \frac{\sqrt{|g|}}{n!}\epsilon^{\mu_1\ldots\mu_r}{}_{\nu_{r+1}\ldots\nu_n}\, dx^{\nu_{r+1}} \wedge \cdots \wedge dx^{\nu_n}.$$

The invariant volume element in (3.3) is $\Omega = *1$.

Example 3.5 (Hodge star on $\mathbb{R}^3$). In the Euclidean space $\mathbb{R}^3$,

$$*1 = dx \wedge dy \wedge dz, *dx = dy \wedge dz, *dy = dz \wedge dx, *dz = dx \wedge dy,$$

$$*dy \wedge dz = dx, *dz \wedge dx = dy, *dx \wedge dy = dz, *dx \wedge dy \wedge dz = 1.$$

Now suppose that (M, g) is Riemannian, $\dim M = n$ and ω is an r-form. Then applying the Hodge star twice gives

$$** \omega = (-1)^{r(n-r)}\omega,$$

which shows that $*$ is an isomorphism. Given the coordinate expressions

$$\omega = \frac{1}{r!}\omega_{\mu_1\ldots\mu_r}\, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \quad \text{and} \quad \eta = \frac{1}{r!}\eta_{\mu_1\ldots\mu_r}\, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r},$$

in terms of the volume form (3.3) we have

$$\omega \wedge *\eta = \cdots = \frac{1}{r!}\omega_{\mu_1\ldots\mu_r}\eta^{\mu_1\ldots\mu_r}\,\Omega.$$

Thus an inner product on r-forms is defined via

$$(\omega, \eta) = \int_M \omega \wedge *\eta, \tag{3.4}$$

and this product is symmetric and positive definite.

Example 3.6 (Ginzburg–Landau potential). Ginzburg–Landau vortices on $\mathbb{R}^2$ are minimals of the potential energy functional

$$V(\phi, a) = \frac{1}{2}\int_{\mathbb{R}^2}\left(da \wedge *da + \overline{d_a\phi} \wedge *d_a\phi + \frac{\lambda}{4}(1-\bar{\phi}\phi)^2 * 1\right), \tag{3.5}$$

where $\phi : \mathbb{R}^2 \to \mathbb{C}$ is a complex scalar field, $a = a_1\,dx^1 + a_2\,dx^2 \in \Omega^1(\mathbb{R}^2)$ is the gauge potential 1-form, $d_a\phi = d\phi - \mathrm{i}a\phi$, and $*$ is the Hodge isomorphism. In usual physics notation,

$$V = \frac{1}{2}\int\left(\frac{1}{2}f^{ij}f_{ij} + \overline{D^i\phi}D_i\phi + \frac{\lambda}{4}(1-\overline{\phi}\phi)^2\right)dx^2,$$

where $D_i\phi = \partial_i\phi - \mathrm{i}a_i\phi$ and $f_{12} = \partial_1 a_2 - \partial_2 a_1$.

Given the exterior derivative $d : \Omega^{r-1}(M) \to \Omega^r(M)$ we can define the adjoint exterior derivative $d^\dagger : \Omega^r(M) \to \Omega^{r-1}(M)$ via

$$d^\dagger = (-1)^{nr+n+1} * d * .$$

If (M, g) is compact, orientable and without boundary, and $\alpha \in \Omega^r(M)$, $\beta \in \Omega^{r-1}(M)$ then

$$(d\beta, \alpha) = (\beta, d^\dagger\alpha),$$

so $d^\dagger$ is the adjoint with respect to the inner product (3.4). The Laplacian $\Delta : \Omega^r(M) \to \Omega^r(M)$ is defined by

$$\Delta = (d + d^\dagger)^2 = dd^\dagger + d^\dagger d.$$

For a function $f : M \to \mathbb{R}$, it is a useful exercise to show that in coordinates the Laplacian is given by

$$\Delta f = -\frac{1}{\sqrt{|g|}}\,\partial_\nu(\sqrt{|g|}g^{\mu\nu}\partial_\mu f).$$

An r-form ω_r is called harmonic if $\Delta\omega_r = 0$, and the set of harmonic r-forms is denoted by $\mathrm{Harm}^r(M)$.

Theorem 3.7 (Hodge decomposition).

$$\Omega^r(M) = d\Omega^{r-1}(M) \oplus d^\dagger\Omega^{r+1} \oplus \mathrm{Harm}^r(M),$$

that is, any $\omega_r \in \Omega^r(M)$ *can be decomposed as*

$$\omega_r = d\alpha_{r-1} + d^\dagger\beta_{r+1} + \gamma_r$$

with $\Delta\gamma_r = 0$.

In fact, $\mathrm{Harm}^r(M) \cong H^r(M)$, where $H^r(M)$ is the de Rham cohomology group (closed r-forms modulo exact r-forms).

Example 3.8 (Maxwell's equations). The four Maxwell equations for the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ can be written in vector form as

$$\nabla \cdot \mathbf{E} = \rho, \quad \nabla \wedge \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{j}, \quad \nabla \cdot \mathbf{B} = 0, \quad \text{and} \quad \nabla \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0,$$

where, in terms of the electromagnetic potential $(A_0, \mathbf{A})$,

$$\mathbf{E} = -\nabla A_0 - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \wedge \mathbf{A}.$$

In differential geometry notation, the potential is encoded into a 1-form A, and the electromagnetic field tensor is the 2-form $F = dA$. With a 1-form current j, the Maxwell equations are just

$$d^\dagger F = j \quad \text{and} \quad dF = 0.$$

3.3. *Complex manifolds*

If $z = x + iy$ and $f = u + iv$ then $f(x, y)$ is *holomorphic* in z provided the Cauchy–Riemann equations are satisfied:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

A complex manifold is a manifold such that the crossover maps ψ_{ij} are all holomorphic. Examples of complex manifolds are $\mathbb{C}^n$, S^2, T^2, $\mathbb{C}P^n$, and $S^{2n+1} \times S^{2m+1}$.

An *almost complex structure* is a $(1,1)$ tensor field J which acts linearly on T_pM in terms of real coordinates as

$$J_p\frac{\partial}{\partial x^\mu}=\frac{\partial}{\partial y^\mu},\quad J_p\frac{\partial}{\partial y^\mu}=-\frac{\partial}{\partial x^\mu} \tag{3.6}$$

with ${J_p}^2=-\mathrm{id}_{T_pM}$.[d] In terms of complex coordinate vectors we have

$$J_p\frac{\partial}{\partial z^\mu}=\mathrm{i}\frac{\partial}{\partial z^\mu},\quad J_p\frac{\partial}{\partial \bar{z}^\mu}=-\mathrm{i}\frac{\partial}{\partial \bar{z}^\mu}.$$

(multiplication by $\mathrm{i}=\sqrt{-1}$).

A *Hermitian metric* is a Riemannian metric which satisfies

$$g_p(J_pX,J_pY)=g_p(X,Y),$$

i.e. g is compatible with J_p. The vector J_pX is orthogonal to X:

$$g_p(J_pX,X)=g_p(J_p^2X,J_pX)=-g_p(J_pX,X)=0.$$

For a Hermitian metric $g_{\mu\nu}=0$ and $g_{\bar{\mu}\bar{\nu}}=0$, e.g.

$$g_{\mu\nu}=g\left(\frac{\partial}{\partial z^\mu},\frac{\partial}{\partial z^\nu}\right)=g\left(J_p\frac{\partial}{\partial z^\mu},J_p\frac{\partial}{\partial z^\nu}\right)=g\left(\mathrm{i}\frac{\partial}{\partial z^\mu},\mathrm{i}\frac{\partial}{\partial z^\nu}\right)=-g_{\mu\nu}.$$

Given the metric and an (almost) complex structure, define ω via

$$\omega_p(X,Y)=g_p(J_pX,Y),\quad X,Y\in T_pM.$$

Then ω is an antisymmetric tensor field, and invariant under J_p :

$$\omega(X,Y)=-\omega(Y,X),\quad \omega(J_pX,J_pY)=\omega(X,Y).$$

Moreover, ω is a real form and can be written as

$$\omega=-\mathrm{i}g_{\mu\bar{\nu}}\,dz^\mu\wedge d\bar{z}^\nu.$$

Also, $\omega\wedge\cdots\wedge\omega$ ($\dim_{\mathbb{C}}M$ times) provides a volume form for M. If $d\omega=0$ then g is called a *Kähler metric*. For a Kähler manifold, the metric g is related to the antisymmetric Kähler form ω which can be interpreted as a symplectic 2-form.

[d]Compare this with the matrix appearing in the right-hand side of (4.9) below.

Remark 3.9. As we shall see later, topological solitons of Bogomolny type usually have a "moduli space" of static solutions which is a smooth manifold with a natural Kähler metric given by the kinetic energy.

4. Geometry in Classical Mechanics

The language of differential geometry is extremely useful for formulating classical mechanics. There are two main approaches: Lagrangian mechanics and Hamiltonian mechanics; and these two approaches lead to two different geometrical settings: symplectic geometry and Poisson geometry.

4.1. *Lagrangian mechanics*

The traditional formulation of Lagrangian mechanics involves two manifolds, namely the configuration space Q together with its tangent bundle TQ. Suppose that Q has dimension n, and local coordinates $(\mathbf{q}, \mathbf{v}) = (q^1, q^2, \ldots, q^n, v^1, v^2, \ldots, v^n)$ are given on TQ. Then the motion of a system is specified by the Lagrangian, which is a function

$$L: \quad TQ \to \mathbb{R},$$

so locally $L = L(\mathbf{q}, \mathbf{v})$. If the velocity $\mathbf{v} = \dot{\mathbf{q}}$ is the tangent vector to a path γ in Q parametrized by time t (with dot denoting d/dt), then we can write $L = L(\mathbf{q}, \dot{\mathbf{q}})$ and consider the function L along the path γ. The action $\mathcal{S}$ associated with L is given by integrating it between two fixed, arbitrary times $t_0 < t_1$ to obtain

$$\mathcal{S} = \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt,$$

and the Principle of Least Action selects those paths which are stationary under variation of the action, i.e. $\delta\mathcal{S} = 0$. By a standard result of the calculus of variations [9], this yields the Euler–Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) - \frac{\partial L}{\partial q^j} = 0, \quad j = 1, \ldots, n. \tag{4.1}$$

Example 4.1 (Harmonic oscillator). With $Q = \mathbb{R}$, a harmonic oscillator of mass m and frequency ω is specified by the Lagrangian $L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2$.

Example 4.2 (Simple pendulum). For $Q = S^1$ with angular coordinate $q \in (-\pi, \pi)$, the Lagrangian for a simple pendulum of mass m and length ℓ experiencing gravitational acceleration g is $L = \frac{1}{2}m\ell^2\dot{q}^2 - mg\ell(1-\cos q)$.

Example 4.3 (Newton's second law). The preceding two examples are particular cases of a Lagrangian of the form $L = \frac{1}{2}\sum_j m_j(\dot{q}^j)^2 - V(\mathbf{q})$ (kinetic energy minus potential energy), with the Euler–Lagrange equations (4.1) being

$$m_j\ddot{q}^j = -\frac{\partial V}{\partial q^j}, \quad j = 1, \dots, n,$$

corresponding to Newton's second law with conservative forces described by the potential V.

Example 4.4 (Geodesic flow). Given a Riemannian manifold Q with metric g given in local coordinates as $ds^2 = g_{jk}\,dq^j\,dq^k$, consider $\mathcal{S} = \int_{t_0}^{t_1} L\,dt$ with the purely kinetic Lagrangian

$$L = \tfrac{1}{2}g(\dot{\mathbf{q}}, \dot{\mathbf{q}}) = \tfrac{1}{2}g_{jk}(\mathbf{q})\dot{q}^j\dot{q}^k \tag{4.2}$$

(the summation convention is assumed throughout). Then (4.1) produces

$$\ddot{q}^j + \Gamma^j_{kl}\dot{q}^k\dot{q}^l = 0, \tag{4.3}$$

where Γ^j_{kl} are the Christoffel symbols (2.2). These geodesic equations can also be obtained as the paths of minimum length [5], by taking the alternative action $\int_{t_0}^{t_1} ds = \int_{t_0}^{t_1} \sqrt{g(\dot{\mathbf{q}}, \dot{\mathbf{q}})}\,dt$. The latter interpretation no longer applies in the pseudo-Riemannian setting, since g is not positive definite, e.g. in the case of general relativity with $n = 4$, where g has the Lorentzian signature $(+,-,-,-)$, and equations (4.3) describe a free particle moving through spacetime [15].

4.2. *Hamiltonian mechanics: canonical case*

The Euler–Lagrange equations (4.1) are ordinary differential equations of second order for the generalized coordinates $\mathbf{q}$. An alternative formulation of mechanics, due to Hamilton, is framed in terms of differential equations of first order. The canonical version of Hamiltonian mechanics can be derived from the Lagrangian setting by performing a Legendre transformation, defining the generalized momenta $\mathbf{p} = (p_1, p_2, \dots, p_n)$ according to

$$p_j = \frac{\partial L}{\partial v^j}, \quad j = 1, \dots, n, \tag{4.4}$$

and introducing

$$H(\mathbf{q},\mathbf{p}) = \langle \mathbf{p}, \mathbf{v} \rangle - L(\mathbf{q},\mathbf{v}), \tag{4.5}$$

with $\langle , \rangle$ denoting the standard scalar product (so in components $H = p_j v^j - L$). In general, the Hamiltonian H can only be found as a function of $\mathbf{p}$ (and $\mathbf{q}$) if (4.4) can be inverted to find $\mathbf{v} = \mathbf{v}(\mathbf{q},\mathbf{p})$; so the Hessian matrix $(\frac{\partial^2 L}{\partial v^j \partial v^k})$ must be non-singular, in order to apply the implicit function theorem.

Theorem 4.5. *Given momenta p_j and velocities $v^j = \dot{q}^j$ related by* (4.4)*, and the corresponding Legendre transformation* (4.5) *between functions L and H, the Euler–Lagrange equations* (4.1) *hold for $\mathbf{q}$ if and only if Hamilton's canonical equations*

$$\dot{q}^j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \tag{4.6}$$

are satisfied for $j = 1, \ldots, n$.

Proof. From (4.5) we have on the one hand

$$dH = v^j\, dp_j - \frac{\partial L}{\partial q^j}\, dq^j + \left(p_j - \frac{\partial L}{\partial v^j} \right) dv^j,$$

where the last term in brackets vanishes by (4.4), while on the other hand

$$dH = \frac{\partial H}{\partial q^j}\, dq^j + \frac{\partial H}{\partial p^j}\, dp^j.$$

Comparing these two expressions yields

$$\frac{\partial H}{\partial q^j} = -\frac{\partial L}{\partial q^j}, \quad v^j = \frac{\partial H}{\partial p^j},$$

and with $v^j = \dot{q}^j$ the result follows. □

Example 4.6 (Natural Hamiltonian). Applying the Legendre transformation to the Lagrangian in Example 4.3, we see that the Hamiltonian is

$$H = \sum_j \frac{(p_j)^2}{2m_j} + V(\mathbf{q}),$$

where the momenta are $p_j = m_j \dot{q}^j$ (mass times velocity) for $j = 1, \ldots, n$. This is an example of a natural Hamiltonian, given as a sum of kinetic and potential energy, with the kinetic term being quadratic in momenta.

Example 4.7 (Geodesic Hamiltonian). The most general kinetic term, quadratic in momenta, arises from Example 4.4. In that case, the Legendre transformation produces the Hamiltonian for geodesic flow, which is

$$H = \tfrac{1}{2} g^{jk}(\mathbf{q}) p_j p_k, \tag{4.7}$$

where g^{jk} are the components of the co-metric tensor ($g^{jk} g_{k\ell} = \delta^j_\ell$).

Geometrically, the formula (4.4) defines a map $TQ \to T^*Q$, which sends the vector $v^j \partial/\partial q^j$ to the Poincaré 1-form $\alpha = p_j dq^j$. This yields the symplectic form

$$\omega = d\alpha = dp_j \wedge dq^j, \tag{4.8}$$

which endows the phase space $M = T^*Q$ with the natural structure of a symplectic manifold. Any symplectic manifold is of even dimension, $2n$ say, and the Darboux theorem says that in the neighbourhood of any point there exist coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ (called canonical coordinates and momenta, or Darboux coordinates) such that the symplectic form is given by the expression (4.8).

By contraction, the symplectic form defines the isomorphism (3.1) taking vector fields to one-forms. Thus to any function $F \in C^\infty(M)$ we can associate a Hamiltonian vector field $V_F \in \Gamma(TM)$, defined by $\omega(\cdot, V_F) = dF$. Equations (4.6) give the flow of the Hamiltonian vector field V_H, which can be written with vector notation as

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \dfrac{\partial H}{\partial \mathbf{q}} \\ \dfrac{\partial H}{\partial \mathbf{p}} \end{pmatrix}. \tag{4.9}$$

The symplectic form is preserved under Hamiltonian flow, in the sense that $\varphi_t^* \omega = \omega$ where $\varphi_t : M \to M$ is the time t flow map generated by V_H. It follows that for each k, $1 \leq k \leq n$, the $2k$-form $\omega \wedge \omega \wedge \cdots \wedge \omega$ (k times) is preserved by the flow; the particular case $k = n$ (preservation of the phase space volume form) is known as Liouville's theorem, and is a key result in statistical mechanics [16].

The inverse of the map (3.1) defines a bivector field J (a contravariant antisymmetric 2-tensor), called the Poisson tensor, which maps 1-forms to vector fields, so in particular $V_H = J(\cdot, dH)$, and in terms of Darboux coordinates the components of J are given by the matrix appearing on the right-hand side of (4.9). The Poisson tensor can be used to define the

Poisson bracket $\{,\}$ between functions, which is a bilinear skew-symmetric bracket given by

$$\{F, G\} = J(dF, dG) \tag{4.10}$$

for $F, G \in C^\infty(M)$. This yields the canonical Poisson bracket relations between the Darboux coordinates, namely

$$\{q^j, p_k\} = \delta^j_k, \quad \{q^j, q^k\} = 0 = \{p_j, p_k\}. \tag{4.11}$$

After quantization, the latter relations become the canonical commutation relations discovered by Born [6].[e]

The problem of explicitly integrating differential equations is a difficult one, and in general is a hopeless task. However, for a Hamiltonian system, Poisson brackets provide an algebraic way of determining first integrals (constants of motion) due to the following result.

Proposition 4.8. *A function F is a first integral of the Hamiltonian flow generated by H if and only if it is in involution with H, i.e. $\{F, H\} = 0$.*

Proof. For any function F the time evolution is given by

$$\frac{dF}{dt} = i_{V_H} dF = J(dF, dH) = \{F, H\}. \tag{4.12}$$

Hence $\frac{dF}{dt} = 0$ if and only if the Poisson bracket of F with H vanishes. □

From the skew-symmetry of the bracket, it follows that H itself is a constant of motion for the Hamiltonian flow that it generates. Thus the solution of a Hamiltonian system is restricted to a level set H =constant, of codimension one in phase space. If there are additional first integrals then the motion is further restricted, and it may even be possible to integrate it completely.

Definition 4.9 (Complete integrability). A Hamiltonian system on a symplectic manifold M of dimension $2n$ is said to be completely integrable if it admits n functions $H = H_1, H_2, \ldots, H_n$ which are functionally independent ($dH_1 \wedge dH_2 \wedge \cdots \wedge dH_n \neq 0$) and in involution with respect to the Poisson bracket on M, i.e. $\{H_j, H_k\} = 0$ for all j, k.

[e]In quantum mechanics, each pair of classical position and momentum variables q, p is replaced by a pair of operators Q, P on a Hilbert space of states, satisfying $[Q, P] = \mathrm{i}\hbar 1$ (see e.g. [12, 21]).

The above definition is also referred to as integrability in the sense of Liouville, or Liouville integrability, due to another result known as Liouville's theorem, with a more contemporary proof and extension due to Arnold [1].

Theorem 4.10 (Liouville–Arnold). *If a Hamiltonian system satisfies the conditions of Definition* 4.9, *then Hamilton's equations for H can be solved by quadratures. Furthermore, each compact level set of the n functions $H_1, H_2, \ldots, H_n$ is diffeomorphic to an n-dimensional torus T^n.*

A key ingredient of the proof is the construction of action-angle coordinates in the neighbourhood of each level set: these are action variables I_j (functions of $H_1, H_2, \ldots, H_n$ only) and canonically conjugate angles θ_j (coordinates on a torus T^n). In these coordinates, the Hamiltonian is a function of the action variables only, and Hamilton's equations become

$$\dot{\theta}_j = \frac{\partial H}{\partial I_j}, \quad \dot{I}_j = 0, \;\; j = 1, \ldots, n,$$

with the solution giving straight line motion on the torus:

$$\theta_j(t) = \frac{\partial H}{\partial I_j}\, t + \theta_j(0), \quad I_j = \text{constant}.$$

So for a completely integrable system, the motion is quasiperiodic on each compact level set of the first integrals.

Example 4.11 (Kepler problem). The Hamiltonian for a body of mass m moving in three dimensions in an attractive central force obeying the inverse square law is

$$H = \frac{|\mathbf{p}|^2}{2m} - \frac{\kappa}{|\mathbf{q}|}, \quad \kappa > 0,$$

where $(\mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3$. Defining the angular momentum vector

$$\mathbf{L} = \mathbf{q} \times \mathbf{p}$$

with components L_j, it can be verified that the canonical bracket (4.11) for the positions and momenta leads to the relations

$$\{L_j, L_k\} = \epsilon_{jk\ell} L_\ell, \tag{4.13}$$

from which it can be verified that

$$\{H, |\mathbf{L}|^2\} = \{H, L_3\} = \{|\mathbf{L}|^2, L_3\} = 0,$$

so this is a completely integrable system.

The existence of first integrals is often associated with symmetries of a system; this connection can be made precise using Noether's theorem (traditionally in the Lagrangian setting). For the Kepler problem, the conservation of angular momentum is a consequence of rotation invariance, described by the Lie group SO(3), and the same is true if the potential $-\kappa/|\mathbf{q}|$ is replaced by any rotation-invariant function $V(|\mathbf{q}|)$. However, the inverse square law is special: it has an extra hidden symmetry [11], leading to an additional conserved vector, namely the Laplace–Runge–Lenz vector $\mathbf{p} \times \mathbf{L} - m\kappa\mathbf{q}/|\mathbf{q}|$.

4.3. *Hamiltonian mechanics: General case*

Rather than starting with a Lagrangian and proceeding via a Legendre transformation, there is a more general formulation of Hamiltonian mechanics which takes the Poisson bracket as the starting point. The development of this point of view was actually inspired by the infinite-dimensional case (Hamiltonian partial differential equations), which will be described below in due course.

Definition 4.12 (Poisson bracket). Given an algebra $\mathcal{F}$ over $\mathbb{R}$, the map

$$\{,\}: \ \mathcal{F} \times \mathcal{F} \to \mathcal{F}$$

is called a Poisson bracket if the following properties hold $\forall F, G, H \in \mathcal{F}$:

- **Skew-symmetry:** $\{F, G\} = -\{G, F\}$;
- **Bilinearity:** $\{\lambda F + \mu G, H\} = \lambda\{F, H\} + \mu\{G, H\}\ \forall \lambda, \mu \in \mathbb{R}$;
- **Jacobi identity:** $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$;
- **Derivation (Leibniz rule):** $\{F, GH\} = \{F, G\}H + G\{F, H\}$.

Observe that, as stated, the second property of the above is just linearity in the first argument, but together with the first property this implies bilinearity of the bracket (i.e. linearity in both arguments), and the first three properties endow $\mathcal{F}$ with the structure of a Lie algebra. A manifold M equipped with a Poisson bracket on the functions $\mathcal{F} = C^\infty(M)$ is called a *Poisson manifold*. On a Poisson manifold M, a Hamiltonian vector field V_F is associated to each function by taking $V_F = \{\cdot, F\}$. If M has dimension d, then in local coordinates $\mathbf{x} = (x^1, \ldots, x^d)$ the equations of motion for the Hamiltonian H are

$$\dot{\mathbf{x}} = J(\cdot, dH), \tag{4.14}$$

where the Poisson tensor J is a bivector field given locally by

$$J = \frac{1}{2} J^{jk} \frac{\partial}{\partial x^j} \wedge \frac{\partial}{\partial x^k}, \quad J^{jk} = \{x^j, x^k\}.$$

Every symplectic manifold is also a Poisson manifold, with Poisson tensor defined by the inverse of (3.1), but otherwise the dimension d need not be even.

Example 4.13 (Lie–Poisson bracket for $\mathfrak{so}(3)$). Take $M = \mathbb{R}^3$ with coordinates $\mathbf{L} = (L_1, L_2, L_3)$ and bracket (4.13). The Poisson tensor has rank 2 at all points of $\mathbb{R}^3$ except the origin, where it vanishes. This is an example of a Lie–Poisson bracket: given any Lie algebra $\mathfrak{g}$ with basis $(X_j)_{j=1}^d$ satisfying $[X_j, X_k] = c_{jk\ell} X_\ell$ for structure constants $c_{jk\ell}$, a linear Poisson bracket on the dual space $\mathfrak{g}^*$ with coordinates (x_j) is defined by $\{x_j, x_k\} = c_{jk\ell} x_\ell$.

In general, a Poisson manifold is foliated by symplectic leaves: define an equivalence relation on the points of M by saying that $x \sim y$ if x and y are connected by piecewise smooth curves, each component of which is an integral curve of a Hamiltonian vector field; then each equivalence class is an immersed submanifold $N \subset M$, a symplectic manifold whose dimension is the rank of the Poisson tensor at any point of N [23]. For instance, in the preceding example, the symplectic leaves consist of two-dimensional spheres together with the origin (dimension zero); in this case they correspond to coadjoint orbits: the orbits of $G = \mathrm{SO}(3)$ acting on $\mathfrak{g}^* = \mathfrak{so}(3)^* \simeq \mathbb{R}^3$.

Another feature of Poisson manifolds, not arising in the symplectic setting, is that there can be non-constant functions whose differential is in the kernel of J.

Definition 4.14 (Casimir function). A function $C \in C^\infty(M)$ on a Poisson manifold M is called a Casimir function if $\{C, F\} = 0 \ \forall F \in C^\infty(M)$.

Example 4.15. In Example 4.13, the function

$$|\mathbf{L}|^2 = (L_1)^2 + (L_2)^2 + (L_3)^2 \tag{4.15}$$

is a Casimir for the $\mathfrak{so}(3)$ bracket (4.13). The level sets of this function coincide with the orbits of $\mathrm{SO}(3)$ acting on $\mathbb{R}^3$.

From Proposition 4.8 it is clear that Casimir functions provide first integrals for a Hamiltonian system. However, the algebra of Casimir functions

can be very complicated, and it is not clear what should be the correct generalization of Definition 4.9 in the case of Poisson manifolds. Nevertheless, we can adopt the following definition from [22], which is sufficient to describe integrability in many situations, particularly in an algebraic context.

Definition 4.16 (Complete integrability (Poisson case)). Suppose that the Poisson tensor is of constant rank $2n$ on a dense open subset of a Poisson manifold M of dimension d, and that the algebra of Casimir functions is maximal, i.e. it contains $d - 2n$ independent functions. A Hamiltonian system on M is said to be completely integrable if it admits $d - n$ independent functions (including the Hamiltonian H) which are in involution.

Example 4.17 (Euler top). For a rigid body rotating freely about a fixed point, the angular momentum is a point in the phase space $M = \mathbb{R}^3$ with the Poisson bracket (4.13), and the Hamiltonian is

$$H = -\frac{1}{2}\left(\frac{(L_1)^2}{I_1} + \frac{(L_2)^2}{I_2} + \frac{(L_3)^2}{I_3}\right),$$

where $I = \operatorname{diag}(I_1, I_2, I_3)$ is the diagonalized inertia tensor. Hamilton's equations $\dot{L}_j = \{L_j, H\}$ for $j = 1, 2, 3$ can be written in vector form as

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{L} \times \boldsymbol{\omega}, \tag{4.16}$$

where $\boldsymbol{\omega} = I^{-1}\mathbf{L}$ is the angular momentum. The Casimir (4.15) and H are two independent functions in involution, and the other conditions of Definition (4.16) are satisfied, so the system (4.16) is completely integrable.

Remark 4.18. The full description of the rigid body involves a rotation $R(t) \in \mathrm{SO}(3)$ which satisfies a second-order equation, corresponding to geodesic motion on the group SO(3) with respect to a suitable metric [14].

In the theory of integrable systems, and especially in infinite dimensions, there are numerous examples of systems with more than one Hamiltonian structure.

Definition 4.19 (Bi-Hamiltonian system). Two Poisson brackets $\{,\}_{1,2}$ are said to be compatible if any linear combination

$$\lambda_1 \{,\}_1 + \lambda_2 \{,\}_2$$

is also a Poisson bracket. A system is said to be bi-Hamiltonian if the flow can be written as a Hamiltonian vector field with respect to two independent Poisson brackets that are compatible with each other, i.e.

$$\{\cdot, H_1\}_1 = \{\cdot, H_2\}_2$$

for two different Hamiltonian functions $H_{1,2}$.

Note that, for two Poisson brackets to be compatible, it is enough to check that their sum satisfies the Jacobi identity. It turns out that for the Euler top there is another Poisson bracket that is compatible with (4.13).

Example 4.20 (Bi-Hamiltonian structure for Euler top). Denote the bracket (4.13) by $\{,\}_1$, set $H_1 = H$, and define

$$\{L_1, L_2\}_2 = \frac{L_3}{I_3}, \quad \{L_2, L_3\}_2 = \frac{L_1}{I_1}, \quad \{L_3, L_1\}_2 = \frac{L_2}{I_2}.$$

Then $\{,\}_{1,2}$ are compatible, and equations (4.16) can be written as

$$\dot{L}_j = \{L_j, H_2\}_2, \quad j = 1, 2, 3, \text{ with } H_2 = \frac{1}{2}|\mathbf{L}|^2,$$

so the Euler top is bi-Hamiltonian.

Remark 4.21. One way to verify the preceding result is by observing that, under the flow (4.16), any function F evolves according to

$$\frac{dF}{dt} = \det \frac{\partial(F, H_1, H_2)}{\partial(L_1, L_2, L_3)}.$$

In the above, the Jacobian determinant on the right-hand-side defines the Nambu bracket of three functions on $\mathbb{R}^3$ [19], denoted $\{F, H_1, H_2\}$.

5. Classical Field Theory

There are two obvious ways to generalize the variational approach to classical mechanics: firstly, one can consider Lagrangians with higher order derivatives; and secondly, one can take derivatives with respect to additional independent variables (space as well as time). The canonical Hamiltonian framework for Lagrangians with higher derivatives was derived by Ostrogradsky (see [3] for applications), but higher order Lagrangians have several features which make them undesirable for a physical theory [24]. In this section we describe first-order Lagrangian densities for field theories and the canonical Hamiltonian formalism, as well as more general Hamiltonian formulations of evolutionary partial differential equations (PDEs).

5.1. *Lagrangians for scalar fields*

For a single scalar field ϕ in Minkowski spacetime M of dimension $n = d+1$, the action takes the form

$$\mathcal{S} = \int_M \mathcal{L}\, d^{d+1}x, \tag{5.1}$$

where the Lagrangian density is

$$\mathcal{L} = \tfrac{1}{2}\partial^\mu\phi\partial_\mu\phi - \mathcal{V}(\phi), \tag{5.2}$$

with the Lorentz metric $g = \mathrm{diag}(1, -1, \ldots, -1)$, and

$$\partial_\mu\phi = \frac{\partial\phi}{\partial x^\mu}, \quad \partial^\mu\phi = g^{\mu\nu}\partial_\nu\phi;$$

and all indices are lowered and raised with $g_{\mu\nu}$ and $g^{\mu\nu}$, the components of g and its inverse (co-metric). Typically $n = 4$, and the coordinate indices are labelled from $\mu = 0$, so $(x^\mu) = (x^0, x^1, x^2, x^3) = (t, \mathbf{x})$ (setting the speed of light $c = 1$), and the Lagrangian is Lorentz invariant. With different coordinates, or in a curved spacetime, the appropriate metric should be used instead, and the volume element $d^n x = d^{d+1}x$ in (5.1) should be replaced by the invariant volume form (3.3). Different field theories result from the choice of function $\mathcal{V}$, which specifies the potential energy of the field.

Taking the action (5.1) with a first-order Lagrangian density, and applying the Principle of Least Action, $\delta\mathcal{S} = 0$, with vanishing boundary conditions at infinity yields the Euler–Lagrange equations

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0. \tag{5.3}$$

For $\mathcal{L}$ given by (5.2), this becomes

$$\partial^\mu\partial_\mu\phi + \mathcal{V}'(\phi) = 0. \tag{5.4}$$

Example 5.1 (Klein–Gordon field). This corresponds to the choice $\mathcal{V} = \frac{1}{2}m^2\phi^2$, where m is the mass. This is called a free field theory: the resulting field equation (see (5.4) below) is linear.

Example 5.2 (ϕ^4 theory). This is a nonlinear theory with the discrete symmetry $\phi \to -\phi$, which arises by setting $\mathcal{V} = \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4$ (with λ being a coupling constant).

Example 5.3 (Sine-Gordon theory). The choice $\mathcal{V} = m^2(1 - \cos\phi)$ results in an integrable field theory in dimension $n = 2$, both at the classical and quantum level. It has exact multiple soliton solutions called *kinks*.

For the case of a first-order Lagrangian with fields taking values in a target space of dimension > 1 (e.g. sigma models, where ϕ takes values in a Lie group, or Yang–Mills–Higgs theories like (2.3), with gauge fields), each component of each field satisfies an equation of the form (5.3).

5.2. *Hamiltonian field theory: canonical case*

In order to formulate a field theory as a Hamiltonian system, it is necessary to separate out space and time, which breaks Lorentz covariance. After rewriting the action (5.1) as

$$\mathcal{S} = \int L\, dt, \quad L = \int \mathcal{L}\, d^d x,$$

the total energy (Hamiltonian) is defined by the Legendre transformation

$$H = \int_{\mathbb{R}^d} \pi\, \partial_t\phi\, d^d x \,-\, L, \quad \pi = \frac{\partial \mathcal{L}}{\partial(\partial_t\phi)},$$

where π is called the momentum density. For the Lagrangian density (5.2) the Hamiltonian becomes

$$H = \int_{\mathbb{R}^d} \left(\frac{1}{2}\pi^2 + \frac{1}{2}|\nabla\phi|^2 + \mathcal{V}(\phi) \right)\, d^d x.$$

The canonical form of Hamilton's equations is

$$\begin{pmatrix} \partial_t\phi \\ \partial_t\pi \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \dfrac{\delta H}{\delta \phi} \\ \dfrac{\delta H}{\partial \pi} \end{pmatrix} \tag{5.5}$$

(cf. equation (4.9) in finite dimensions), where the *Fréchet derivative* of a functional H with respect to a field u is defined by

$$\left\langle \frac{\delta H}{\delta u}, v \right\rangle = \frac{d}{d\epsilon} H[u + \epsilon v]|_{\epsilon=0},$$

with $\langle , \rangle$ denoting the L^2 pairing on $\mathbb{R}^d$: $\langle f, g\rangle = \int_{\mathbb{R}^d} fg\, d^d x$.

5.3. *Generalized Hamiltonian structures for PDEs*

In the $1+1$-dimensional case ($d = 1$), we now consider PDEs involving higher spatial derivatives, which can be written as Hamiltonian systems (first order in time).[f] For a single field u, Hamilton's equations take the form

$$u_t = \mathcal{J}\frac{\delta H}{\delta u}, \tag{5.6}$$

where $\mathcal{J}$ is a skew-symmetric operator defining a Poisson bracket between pairs of functionals F, G according to

$$\{F, G\} = \left\langle \frac{\delta F}{\delta u}, \mathcal{J}\frac{\delta G}{\delta u} \right\rangle = \int \frac{\delta F}{\delta u}\, \mathcal{J}\frac{\delta G}{\delta u}\, dx.$$

(Integrals are taken over the whole of $\mathbb{R}$, and it is assumed that all fields and their derivatives vanish at infinity, so $\int F_x\, dx = 0$ for any quantity F.)

Example 5.4 (First Hamiltonian structure for KdV). The Korteweg–de Vries (KdV) equation, describing long waves moving on a shallow canal, is

$$u_t = u_{3x} + 6uu_x. \tag{5.7}$$

Letting D_x denote total derivative with respect to x, KdV can be written in Hamiltonian form by taking $\mathcal{J} = \mathcal{J}_1$ and $H = H_1$, where

$$\mathcal{J}_1 = D_x, \quad H_1 = \int \left(-\frac{1}{2}u_x^2 + u^3\right) dx. \tag{5.8}$$

For PDEs, the notion of integrability can be defined by the existence of an infinite set of commuting symmetries, and in the Hamiltonian setting this is linked to infinitely many conservation laws, which can be derived recursively from a bi-Hamiltonian structure.

Example 5.5 (Second Hamiltonian structure for KdV). Let

$$\mathcal{J}_2 = D_x^3 + 4uD_x + 2u_x, \quad H_2 = \int \frac{1}{2}u^2\, dx. \tag{5.9}$$

[f]For convenience, here we use subscripts to denote partial derivatives, so $u_t = \frac{\partial u}{\partial t}$, $u_{xx} = u_{2x} = \frac{\partial^2 u}{\partial x^2}$, $u_{xxx} = u_{3x} = \frac{\partial^3 u}{\partial x^3}$, and so on.

The Poisson brackets $\{,\}_k$ defined by $\mathcal{J}_k$ for $k = 1, 2$ are compatible with each other, and (5.7) can be written in bi-Hamiltonian form:

$$u_t = \mathcal{J}_1 \frac{\delta H_1}{\delta u} = \mathcal{J}_2 \frac{\delta H_2}{\delta u}.$$

Furthermore, KdV has an infinite sequence of commuting symmetries given in terms of the *recursion operator* $\mathfrak{R} = \mathcal{J}_2 \mathcal{J}_1^{-1}$ by

$$u_{t_j} = \mathfrak{R}\, u_x, \quad j = 0, 1, 2, \ldots,$$

with a corresponding sequence of conserved quantities in involution with respect to both brackets.

If a conserved quantity is a local functional, given in the form $H[u] = \int \mathcal{H}\, dx$ for a density $\mathcal{H} = \mathcal{H}(u, u_x, u_{xx}, \ldots)$ given in terms of u and its derivatives, then its Fréchet derivative is given in terms of the Euler operator $\mathfrak{E}$:

$$\frac{\delta H}{\delta u} = \mathfrak{E} \cdot \mathcal{H} := \sum_{j=0}^{\infty} (-D_x)^j \frac{\partial \mathcal{H}}{\partial u_{jx}}.$$

Example 5.6 (Camassa–Holm equation). The equation

$$u_t - u_{xxt} = u u_{xxx} + 2 u_x u_{xx} - 3 u u_x \tag{5.10}$$

has three simple conserved quantities given in terms of u:

$$H_0 = \int u\, dx, \quad H_1 = \int \frac{1}{2}(u_x^2 + u^2)\, dx, \quad H_2 = \int \frac{1}{2}(u u_x^2 + u^3)\, dx.$$

To write it in bi-Hamiltonian form [4], introduce the compatible pair

$$\mathcal{J}_1 = -(m D_x + D_x m), \quad \mathcal{J}_2 = D_x^3 - D_x, \text{ with } m = u - u_{xx}.$$

Then, since $u(x,t) = (1 - D_x^2)^{-1} m(x,t) = \frac{1}{2} \int_{\mathbb{R}} e^{-|x-y|} m(y,t)\, dy$, (5.10) is

$$m_t = \mathcal{J}_1 \frac{\delta H_1}{\delta m} = \mathcal{J}_2 \frac{\delta H_2}{\delta m}.$$

Remark 5.7 (Soliton solutions). Integrable Hamiltonian PDEs typically have exact solutions in the form of localized nonlinear waves called *solitons*, that preserve their speed and amplitude after collisions. The construction of solitons can be achieved using the inverse scattering problem

for an associated linear system known as a *Lax pair* (see [2, 13]). For the KdV equation (5.7) the 1-soliton solution is

$$u(x,t) = 2k^2 \mathrm{sech}^2(k(x - x_0) + 4k^3 t),$$

where $k > 0$ and x_0 are arbitrary parameters, and an N-soliton solution is a nonlinear superposition of N of these. The Camassa–Holm equation (5.10) has peaked solitons called *peakons*, which are weak solutions with

$$u(x,t) = \sum_{j=1}^{N} p_j(t)\, e^{-|x-q^j(t)|}, \quad m(x,t) = 2\sum_{j=1}^{N} p_j(t)\, \delta(x - q^j(t)),$$

where the peak positions q^j and amplitudes p_j evolve according to a geodesic flow with Hamiltonian (4.7) and co-metric

$$g^{jk}(\mathbf{q}) = e^{-|q^j - q^k|}.$$

6. Geometry and Soliton Dynamics

In this section, we discuss the application of differential geometry to the dynamics of topological solitons. We also introduce some concepts in topology.

6.1. *Homotopy theory*

Given a manifold M and an interval $I = [0,1]$ we can define *paths*

$$\alpha : I \to M : t \mapsto \alpha(t), \quad \text{where } \alpha(0) = p_0,\ \alpha(1) = p_1.$$

A *loop* is a path with $p_0 = p_1$. Paths can be multiplied via

$$\alpha * \beta(s) = \begin{cases} \alpha(2s) & 0 \le s \le \frac{1}{2}, \\ \beta(2s-1) & \frac{1}{2} \le s \le 1. \end{cases}$$

The constant path is $c(s) = p_0$ for all $s \in I$. The inverse of a paths is $\alpha^{-1}(s) = \alpha(1-s)$. *This is not a group, yet!*

Definition 6.1 (Homotopy). Let $\alpha, \beta : I \to M$ be loops at p_0. The loops α and β are *homotopic*, denoted by $\alpha \sim \beta$, if there exists a continuous map $F : I \times I \to M$ such that $F(s,0) = \alpha(s)$ and $F(s,1) = \beta(s)$ for all $s \in I$. Furthermore, $F(0,t) = F(1,t) = p_0$ for all $t \in I$.

It can be shown that $\alpha \sim \beta$ is an equivalence relation. Let $[\alpha]$ be the equivalence class which contains α. Define a product on equivalence classes by $[\alpha] * [\beta] = [\alpha * \beta]$. This gives the *fundamental group* $\pi_1(M, p_0)$.[g]

Example 6.2. The winding number $\pi_1(S^1) = \mathbb{Z}$ counts how many times a path goes around the circle. Similarly, for the punctured plane $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$ counts how many times a path encircles the origin. The fundamental group of the torus is $\pi_1(T^2) = \mathbb{Z} \oplus \mathbb{Z}$; more generally, $\pi_1(M \times N) = \pi_1(M) \oplus \pi_1(N)$.

This generalizes naturally to higher homotopy groups: Consider maps from the cube $I^n = I \times \cdots \times I$ to a manifold M such that all the points on the boundary ∂I^n of the cube are mapped to $p_0 \in M$:

$$\alpha : (I^n, \partial I^n) \to (M, p_0).$$

Again we can form the product $\alpha * \beta$ and define the equivalence classes $[\alpha]$ (also known as homotopy classes), giving the nth homotopy group $\pi_n(M)$.

Remark 6.3 (Summary of important results).

- Homotopy groups are Abelian for $n > 1$, i.e. $[\alpha] * [\beta] = [\beta] * [\alpha]$.
- Manifolds M with $\pi_1(M) = 1$ are called *simply-connected*.
- The *degree* of a map is $\pi_n(S^n) = \mathbb{Z}$ related to the number of pre-images.
- $\pi_n(S^d) = 1$ for $1 \leq n < d$: the map is not onto, therefore contractible.
- $\pi_{n+1}(S^n) = \mathbb{Z}_2$, for $n \geq 3$, but $\pi_3(S^2) = \mathbb{Z}$ which is related to the Hopf bundle in Example 2.3.
- Homotopy groups of spheres are really complicated, e.g. $\pi_{n+2}(S^2) = \mathbb{Z}_2$ for $n \geq 2$.
- *Spectral sequences* are an important tool: Let G be a Lie group with subgroup H; then

$$\cdots \to \pi_n(H) \to \pi_n(G) \to \pi_n(G/H) \to \pi_{n-1}(H) \to \pi_{n-1}(G) \to \pi_{n-1}(G/H) \to \cdots$$

  is a long exact sequence. Example: $G = S^3$, $H = S^1$ and $G/H = S^2$.

6.2. *Homotopy groups and field theory*

Field configurations at a fixed time are maps $\phi : \mathbb{R}^d \to M$, from flat space to a target space. Homotopies of maps occur naturally (e.g. time evolution is continuous and connects different field configurations in the same homotopy

[g] If M is arcwise connected then $\pi_1(M, p_0)$ is isomorphic to $\pi_1(M, p_1)$.

class). Two scenarios naturally give rise to homotopy groups, both arising from boundary conditions (due to finite energy):

(1) *One-point compactification:* There is a unique vacuum $v_0 \in M$, namely, $\phi(\mathbf{x}) = v_0$ for $\mathbf{x} \to \infty$. We can identify all these points, so that topologically $\mathbb{R}^d \cup \{\infty\} = S^d$. So we need $\pi_d(M)$.
(2) *Non-trivial maps at infinity:* The vacuum is degenerate and forms a submanifold N of M. Then in the limit $|\mathbf{x}| \to \infty$ there is a continuous map $\phi|_\infty : S^{d-1}_\infty \to N$. So we need $\pi_{d-1}(N)$.

This leads to the following classification of solitons.

$\pi_n(S^k)$	ungauged	gauged
$\pi_1(S^1)$	Kinks	**Vortices**
$\pi_2(S^2)$	Baby-Skyrmions, Lumps	Monopoles
$\pi_3(S^3)$	Skyrmions	Instantons
$\pi_3(S^2)$	Hopf Solitons	

6.3. *Ginzburg–Landau vortices*

The Ginzburg–Landau energy is given by (3.5). In coordinates $\mathbf{x} = (x^1, x^2) = (x, y)$ it is invariant under gauge transformations

$$\phi(\mathbf{x}) \mapsto e^{\mathrm{i}\alpha(\mathbf{x})}\phi(\mathbf{x}), \quad a_i(\mathbf{x}) \mapsto a_i(\mathbf{x}) + \partial_i\alpha(\mathbf{x}),$$

where $e^{\mathrm{i}\alpha(\mathbf{x})}$ is a spatially varying phase. The quantity

$$B \equiv f_{12} = \partial_1 a_2 - \partial_2 a_1$$

is the magnetic field. The vacuum is $\phi = 1$, $a_i = 0$ and gauge transformations of this, and we require $|\phi| \to 1$ as $|\mathbf{x}| \to \infty$.

Transforming to polar coordinates $(x, y) = (\rho\cos\theta, \rho\sin\theta)$, the energy is

$$V = \frac{1}{2}\int_0^\infty \int_0^{2\pi} \left(B^2 + \overline{D_\rho\phi} D_\rho\phi + \frac{1}{\rho^2}\overline{D_\theta\phi} D_\theta\phi + \frac{\lambda}{4}\left(1 - \overline{\phi}\phi\right)^2 \right) \rho \; d\rho \; d\theta.$$

For finite energy fields we can fix the gauge asymptotically, so that

$$\lim_{\rho\to\infty} \phi(\rho, \theta) = e^{\mathrm{i}\alpha(\theta)},$$

where α is a continuous function of θ. As θ increases from 0 to 2π, $\alpha(\theta)$ increases by $2\pi N$ (ϕ is single valued). The *winding number* N is an arbitrary integer that cannot change under smooth deformations of the field, so remains constant in time, and is also invariant under smooth gauge transformations.

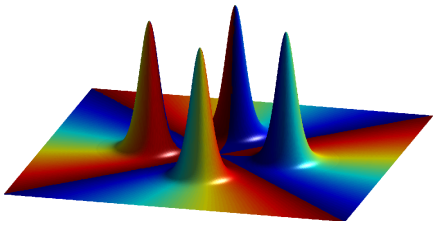
For finite energy, the covariant derivative $D_\theta\phi = (\partial_\theta - \mathrm{i}a_\theta)\phi$ has to vanish as $\rho \to \infty$, so $\phi \sim e^{\mathrm{i}\alpha(\theta)}$ implies $a_\theta = \frac{d\alpha}{d\theta}$. Hence, by Stokes' theorem,

$$\int_{\mathbb{R}^2} B \; d^2x = \int_0^{2\pi} a_\theta \; d\theta|_{\rho\to\infty} = \alpha(2\pi) - \alpha(0) = 2\pi N, \tag{6.1}$$

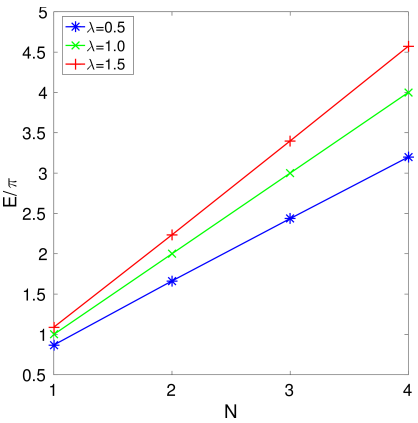
so N measures the units of magnetic flux in the plane. If ϕ has only isolated zeros, then the number of these (counted with multiplicity) is N — see Fig. 3(a). A zero of ϕ is said to have multiplicity k, if on a small circle enclosing it, $-\arg\phi$ increases by $2\pi k$. For simple zeros $k = \pm 1$.

Let E_N be the minimal energy V of N vortices. There are three different regimes: **(i) type I:** $\lambda < 1$, $E_N < NE_1$ — the vortices attract; **(ii) type II:** $\lambda > 1$, $E_N > NE_1$ — the vortices repel; **(iii) critical coupling:** $\lambda = 1$, $E_N = NE_1$ — no forces between static vortices. (See Fig. 3(b).) At critical coupling, by "completing the square" V can be written as

$$V = \frac{1}{2}\int \left(\left(B - \frac{1}{2}\left(1 - \overline{\phi}\phi\right)\right)^2 + \left(\overline{D_1\phi + \mathrm{i}D_2\phi}\right)\left(D_1\phi + \mathrm{i}D_2\phi\right) + B \right) d^2x$$
$$\geq \pi N,$$



(a)



(b)

Fig. 3. Figure (a) shows the energy density of a vortex configuration of charge $N = 4$. The four peaks correspond to four single vortices. The colour indicates the phase of the Higgs field ϕ. Clearly, the colour circle is completed four times as one goes around the "circle at infinity" once. Figure (b) shows the energy of Ginzburg–Landau vortices for charges $N = 1, \ldots, 4$ for coupling constant $\lambda = 0.5, 1$ and 1.5. For $\lambda = 1$ we observe $E_N = NE_1$.

using (6.1). Demanding that the squares vanish yields the Bogomolny equations

$$D_1\phi + \mathrm{i}D_2\phi = 0, \quad B - \tfrac{1}{2}\left(1 - \overline{\phi}\phi\right) = 0. \tag{6.2}$$

These equations cannot be solved analytically. However, a lot is known about the solutions. For given topological charge N, the Bogomolny equations have a $2N$-dimensional manifold of static solutions, known as the *moduli space* M_N. (Gauge equivalent solutions are identified.) All zeros of ϕ have positive multiplicity (generically there are only simple zeros). A solution is completely determined by the locations of these zeros, which can be anywhere, so M_N is parametrized by N unordered points in $\mathbb{R}^2$, requiring $2N$ coordinates. There are no *static* forces between vortices for $\lambda = 1$, however, there will be *velocity-dependent* forces.

6.4. *Relativistic vortex dynamics*

In $2+1$ dimensions with $(x^\mu) = (t, \mathbf{x})$, the standard relativistic Lagrangian is

$$\mathcal{L} = \frac{1}{2}\overline{D_\mu\phi}D^\mu\phi - \frac{1}{4}f_{\mu\nu}f^{\mu\nu} - \frac{\lambda}{8}\left(1 - \overline{\phi}\phi\right)^2.$$

In the following, we will often use complex coordinates $z = x + \mathrm{i}y$.

Parametrizing the moduli space for $\lambda = 1$ by vortex positions $z = Z_i$, assumed to be time dependent, gives a reduced Lagrangian for geodesics on M_N,

$$L_{\text{red.}} = \tfrac{1}{2}\big(g_{rs}\dot{Z}_r\dot{Z}_s + g_{r\overline{s}}\dot{Z}_r\dot{\bar{Z}}_s + g_{\overline{r}\overline{s}}\dot{\bar{Z}}_r\dot{\bar{Z}}_s\big) - \pi N, \tag{6.3}$$

summed for $1 \le r, s \le N$. Setting $h = \log|\phi|^2$ in (6.2) implies

$$\nabla^2 h + 1 - e^h = 4\pi\sum_{r=1}^{N}\delta^2(z - Z_r).$$

The δ functions arise because h has logarithmic singularities at the zeros Z_r of ϕ. Expanding h around the point Z_r gives

$$h(z,\bar{z}) = 2\log|z - Z_r| + a_r + \tfrac{1}{2}\bar{b}_r(z - Z_r) + \tfrac{1}{2}b_r(\bar{z} - \bar{Z}_r) + \cdots.$$

After a long calculation, we find $g_{rs} = 0 = g_{\overline{r}\overline{s}}$ in (6.3), so this purely kinetic Lagrangian gives rise to the moduli space metric

$$g = \pi\left(\delta_{rs} + 2\frac{\partial b_s}{\partial Z_r}\right)dZ_r d\bar{Z}_s,$$

which can be shown to be Kähler. This structure provides a lot of information about the metric, although it is only know implicitly. The moduli space approximation captures the dynamics of vortices, in particular right-angle scattering.

We can consider physical spaces X with a different metric, e.g.

$$ds^2 = dt^2 - \Omega(dx^2 + dy^2),$$

where $\Omega = \Omega(x, y)$ is a conformal factor defining the Riemannian metric on X. Again we can "complete the square" and obtain the Bogomolny equations

$$D_1\phi + \mathrm{i}D_2\phi = 0, \quad B - \frac{\Omega}{2}\left(1 - \overline{\phi}\phi\right) = 0. \tag{6.4}$$

For a surface X with metric defined by Ω the integral

$$c_1 = \frac{1}{2\pi}\int_X f = \frac{1}{2\pi}\int_X B \; d^2x$$

is an integer. This topological invariant is known as the first *Chern number.*

If X has a finite area A then we can integrate the second Bogomolny equation over X to obtain

$$2\int_X B\, d^2x + \int_X |\phi|^2\, \Omega\, d^2x = \int_X \Omega\, d^2x \Longrightarrow A$$
$$= 4\pi N + \int_X |\phi|^2\, \Omega\, d^2x \geq 4\pi N,$$

which is called the Bradlow limit. In other words, each vortex needs an area of at least 4π. At the Bradlow bound $A = 4\pi N$ both equations (6.4) can trivially be solved by $\phi = 0$ and $B = \frac{\Omega}{2}$. For the torus T^2 the moduli space metric has been calculated as an expansion around the Bradlow limit. For the sphere S^2 moduli space metric is essentially the Fubini–Study metric close to the Bradlow limit.

For the hyperbolic plane, the background metric on the Poincaré disc $|z| < 1$ is $ds^2 = \Omega\, dz\, d\overline{z}$, with the conformal factor $\Omega = 8(1 - |z|^2)^{-2}$. Setting $h = \log|\phi|^2$ we can again derive an equation for h, namely

$$\nabla^2 h + \Omega - \Omega e^h = 4\pi \sum_{r=1}^{N} \delta^2(z - Z_r).$$

and this can be transformed to Liouville's equation, which is integrable. In this case, the moduli space is known explicitly, and

$$\phi = \frac{1 - |z|^2}{1 - |f|^2}\frac{df}{dz}, \quad \text{where } f(z) = \prod_{i=1}^{N+1}\left(\frac{z - c_i}{1 - \overline{c}_i z}\right),$$

with $|c_i| < 1$. The positions of the vortices are the zeros of $\frac{df}{dz}$, and the moduli space metric is

$$g = \pi \left(\Omega(Z_r)\delta_{rs} + 2\frac{\partial b_s}{\partial Z_r} \right) dZ_r d\bar{Z}_s$$

but now we can calculate b_s for special cases. The metric for n vortices on a regular polygon with m vortices fixed at the origin is given by

$$ds^2 = \frac{4\pi n^3 |\alpha|^{2n-2} d\alpha \; d\bar{\alpha}}{\left(1 - |\alpha|^{2n}\right)^2} \left(1 + \frac{2n\left(1 + |\alpha|^{2n}\right)}{\sqrt{(m+1)^2\left(1 - |\alpha|^{2n}\right)^2 + 4n^2|\alpha|^{2n}}} \right)$$

for $n \neq m + 1$, and by

$$ds^2 = \frac{12\pi n^3 |\alpha|^{2n-2} d\alpha \; d\bar{\alpha}}{\left(1 - |\alpha|^{2n}\right)^2}$$

for $n = m+1$. The non-trivial zeros are at $z = \alpha \, \mathrm{e}^{2\pi \mathrm{i} k/n}$ for $k = 0, \ldots, n-1$.

7. Summary

In these notes we have tried to give a flavour of the most basic differential geometric ideas that are important in mathematical physics, and provide a dictionary between mathematical and physical terminology. The interested reader is encouraged to delve into the bibliography to gain a deeper understanding.

Acknowledgements

A. Hone acknowledges the support of EPSRC Fellowship EP/M004333/1. S. Krusch was supported by the EPSRC First Grant EP/I034491/1. S. Krusch wants to thank J. Ashcroft for creating Figures 3(a) and 3(b).

References

[1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics, Vol. 60. Springer-Verlag (1978).

[2] O. Babelon, D. Bernard and M. Talon, *Introduction to Classical Integrable Systems*. Cambridge Monographs on Mathematical Physics. Cambridge University Press (2007).

[3] M. Błaszak, *Multi-Hamiltonian Theory of Dynamical Systems*. Springer (1998).
[4] R. Camassa and D. D. Holm, An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* **71**, 1661–1664 (1993).
[5] Y. Choquet-Bruhat, *Géometrie Differentielle et Systèmes Extérieurs*. Dunod (1968).
[6] W. A. Fedak and J. J. Prentis, The 1925 Born and Jordan paper "On quantum mechanics". *Amer. J. Phys.* **77**, 128–139 (2009).
[7] T. Frankel, *The Geometry of Physics*, 3rd edn. Cambridge University Press (2012).
[8] D. S. Freed and K. K. Uhlenbeck (eds.), *Geometry and Quantum Field Theory*. IAS/Park City Mathematics Series, Vol. 1. American Mathematical Society (1995).
[9] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*. Dover (2000).
[10] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics*. Cambridge University Press (1984).
[11] V. Guillemin and S. Sternberg, *Variations on a Theme by Kepler*. AMS Colloquium Publications, Vol. 42. American Mathematical Society (1990).
[12] K. C. Hannabuss, *An Introduction to Quantum Theory*. Oxford University Press (1997).
[13] N. J. Hitchin, G. B. Segal and R. S. Ward, *Integrable Systems: Twistors, Loop Groups, and Riemann Surfaces*. Oxford University Press (1999).
[14] D. D. Holm, *Geometric mechanics. Part II: Rotating, Translating and Rolling*, 2nd edn. Imperial College Press (2011).
[15] L. P. Hughston and K. P. Tod, *An Introduction to General Relativity*. London Mathematical Society Student Texts, Vol. 5. Cambridge University Press (1991).
[16] F. Mandl, *Statistical Physics*. Wiley (1971).
[17] N. J. Manton and P. M. Sutcliffe, *Topological Solitons*. Cambridge Monographs on Mathematical Physics. Cambridge University Press (2004).
[18] M. Nakahara, *Geometry, Topology and Physics*. Graduate Student Series in Physics, 2nd edn. Institute of Physics (2003).
[19] Y. Nambu, Generalized Hamiltonian dynamics. *Phys. Rev. D* **7**, 2405–2412 (1973).
[20] P. J. Olver, *Applications of Lie Groups to Differential Equations*. Graduate Texts in Mathematics, Vol. 107. Springer (1986).
[21] A. Sudbery, *Quantum Mechanics and the Particles of Nature: An Outline for Mathematicians*. Cambridge University Press (1986).
[22] P. Vanhaecke, *Integrable Systems in the Realm of Algebraic Geometry*. Lecture Notes in Mathematics, Vol. 1638, 2nd edn. Springer (2001).
[23] A. Weinstein, The local structure of Poisson manifolds. *J. Differential Geom.* **18**, 523–557 (1983).
[24] R. P. Woodard, The theorem of Ostrogradsky, preprint (2015), `arXiv:1506.02210v2`.