

Sammanfattning av SI2360 Analytisk mekanik och klassisk fältteori

Yashar Honarmandi
yasharh@kth.se

27 mars 2019

Sammanfattning

Innehåll

1	Coordinates	1
2	Tensors	1
3	Geometry	3
4	Classical mechanics	5

1 Coordinates

Coordinates A general set of coordinates on \mathbb{R}^n is n numbers $x^a, a = 1, \dots, n$ that uniquely define a point in the space.

Example: Cartesian coordinates In cartesian coordinates we introduce an orthonormal basis \mathbf{e}_i . Vi kan då skriva $\mathbf{x} = x^i \mathbf{e}_i$.

Basis vectors There are two different choices of coordinate bases.

The first is the tangent basis of vectors

$$\mathbf{E}_a = \partial_{x_a} \mathbf{x} = \partial_a \mathbf{x}.$$

The second is the dual basis

$$\mathbf{E}^a = \vec{\nabla} x^a.$$

Vector coordinates Any vector can now be written as

$$\mathbf{v} = v^a \mathbf{E}_a = v_a \mathbf{E}^a.$$

The v^a are called contravariant components and the v_a are called covariant components.

We can now compute the scalar product

$$\mathbf{E}_a \cdot \mathbf{E}^b = \partial_a \mathbf{x} \cdot \vec{\nabla} x^b = \delta_a^b.$$

Coordinate transformations Suppose that a vector can be written as

$$\mathbf{v} = v^a \mathbf{E}_a = v^{a'} \mathbf{E}'_{a'}.$$

How do we transform between these? A single component is given by

$$v^{a'} = \mathbf{E}'_{a'} \cdot v^a \mathbf{E}_a = v^a (\vec{\nabla} x^{a'} \cdot \partial_a \mathbf{x}) = v^a \partial_a x^{a'}.$$

Tangents to curves The tangent to a curve is given by

$$\dot{\gamma} = \frac{d\mathbf{x}}{dt} = \partial_a \mathbf{x} \frac{dx^a}{dt} = \dot{x}^a \mathbf{E}_a.$$

Gradients The gradient of a curve is given by

$$\vec{\nabla} f = \partial_a f \vec{\nabla} x^a = \mathbf{E}^a \partial_a f.$$

Rates of change along a curve The rate of change of a quantity along a path is given by

$$\frac{df}{dt} = \partial_a f \frac{dx^a}{dt} = \vec{\nabla} f \cdot \dot{\gamma}.$$

2 Tensors

Definition A tensor of rank N is a multilinear map from N vectors to a scalar.

Components of a tensor The components of a tensor are defined by

$$T(\mathbf{E}^{a_1}, \dots, \mathbf{E}^{a_N}) = T^{a_1, \dots, a_N}.$$

These are called the contravariant components of the tensor, and the covariant components are defined similarly. Mixed components can also be defined.

Rules for tensors Tensors obey the following rules:

$$(T_1 + T_2)(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}) = T_1(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}) + T_2(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}),$$

$$(kT)(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}) = kT(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}).$$

In component form:

$$(T_1 + T_2)^{a_1 \dots a_n} = T_1^{a_1 \dots a_n} + T_2^{a_1 \dots a_n},$$

$$(kT)^{a_1 \dots a_n} = kT^{a_1 \dots a_n}.$$

The metric tensor The metric tensor g is a rank 2 tensor defined by $g(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$. Its components satisfy

$$v_a = \mathbf{E}_a \cdot v^b \mathbf{E}_b = g(\mathbf{E}_a, \mathbf{E}_b) v_b = g_{ab} v^b,$$

and likewise

$$v^a = g^{ab} v_b.$$

Tensor product Given two tensors T_1 and T_2 of ranks n_1 and n_2 , we can define the rank $n_1 + n_2$ tensor $T_1 \otimes T_2$ as

$$(T_1 \otimes T_2)(\mathbf{v}_1, \dots, \mathbf{v}_{n_1}, \mathbf{w}_1, \dots, \mathbf{w}_{n_2}) = T_1(\mathbf{v}_1, \dots, \mathbf{v}_{n_1}) T_2(\mathbf{w}_1, \dots, \mathbf{w}_{n_2}).$$

In component form:

$$(T_1 \otimes T_2)^{a_1 \dots a_{n_1+n_2}} = T_1^{a_1 \dots a_{n_1}} T_2^{a_{n_1+1} \dots a_{n_1+n_2}}.$$

Tensors as linear combinations Using the tensor product, all tensors can be written as linear combinations of certain basis elements due to their bilinearity. Define

$$e_{a_1 \dots a_n} = \mathbf{E}_{a_1} \otimes \dots \otimes \mathbf{E}_{a_n}$$

to be the tensor that satisfies

$$e_{a_1 \dots a_n}(\mathbf{E}^{b_1}, \dots, \mathbf{E}^{b_n}) = (\mathbf{E}_{a_1} \cdot \mathbf{E}^{b_1}) \dots (\mathbf{E}_{a_n} \cdot \mathbf{E}^{b_n}) = \delta_{a_1}^{b_1} \dots \delta_{a_n}^{b_n}.$$

Then any tensor can be written as

$$T = T^{a_1 \dots a_n} e_{a_1 \dots a_n}$$

where the $T^{a_1 \dots a_n}$ are exactly the contravariant components of T .

Tensors as linear transforms on tensors A rank n tensor can also be viewed as a linear map from rank m tensors to rank $n - m$ tensors. To do this, we first define, given T , the rank $n - m$ tensor $\tilde{T}(\mathbf{w}_1 \otimes \dots \otimes \mathbf{w}_m)$ such that

$$(\tilde{T}(\mathbf{w}_1 \otimes \dots \otimes \mathbf{w}_m))(\mathbf{v}_1, \dots, \mathbf{v}_{n-m}) = T(\mathbf{w}_1, \dots, \mathbf{w}_m, \mathbf{v}_1, \dots, \mathbf{v}_{n-m}).$$

This map is also linear in all the \mathbf{w}_i . Next, given a rank $n - m$ tensor \tilde{T} , one can define the rank $n - m$ tensor $T(\mathbf{w}_1, \dots, \mathbf{w}_m)$ such that

$$T(\mathbf{w}_1, \dots, \mathbf{w}_m, \mathbf{v}_1, \dots, \mathbf{v}_{n-m}) = (\tilde{T}(\mathbf{w}_1 \otimes \dots \otimes \mathbf{w}_m))(\mathbf{v}_1, \dots, \mathbf{v}_{n-m}).$$

This is a linear rank n tensor.

Tensor contraction Given a complete set of vectors \mathbf{v}_i and their dual \mathbf{v}^i such that $\mathbf{v}_i \cdot \mathbf{v}^i = \delta_i^j$, the contraction $e_{12}T$ of two arguments of a rank n tensor is the tensor of rank $n - 2$ satisfying

$$(e_{12}T)(\mathbf{w}_1, \dots, \mathbf{w}_{n-2}) = T(\mathbf{v}_i, \mathbf{v}^i, \mathbf{w}_1, \dots, \mathbf{w}_{n-2}).$$

In component form:

$$(e_{12}T)^{a_1 \dots a_{n-2}} = T_c^{a_1 \dots a_{n-2}}.$$

The definition is similar (I assume) for the contraction of other arguments.

3 Geometry

Covariant derivatives and Christoffel symbols When computing a derivative, one must account both for the change in the quantity itself and the change of basis. We have

$$\partial_b \mathbf{E}_a = \Gamma_{ba}^c \mathbf{E}_c$$

where the Γ_{ba}^c are called Christoffel symbols. These satisfy

$$\mathbf{E}^c \cdot \partial_b \mathbf{E}_a = \mathbf{E}^c \cdot \Gamma_{ba}^d \mathbf{E}_d = \delta_d^c \Gamma_{ba}^d = \Gamma_{ba}^c.$$

Note that

$$\partial_a \mathbf{E}_b = \partial_a \partial_b \mathbf{x} = \partial_b \partial_a \mathbf{x} = \partial_b \mathbf{E}_a,$$

which implies

$$\Gamma_{ba}^c = \Gamma_{ab}^c.$$

Using this, we can compute the partial derivate of $\mathbf{v} = v^a \mathbf{E}_a$ with respect to χ^a as

$$\partial_a \mathbf{v} = \mathbf{E}_b \partial_a v^b + v^b \partial_a \mathbf{E}_b = \mathbf{E}_b \partial_a v^b + v^b \Gamma_{ab}^c \mathbf{E}_c.$$

Renaming the summation indices yields

$$\partial_a \mathbf{v} = \mathbf{E}_b (\partial_a v^b + v^c \Gamma_{ac}^b),$$

which contains one term from the change in the coordinates and one term from the change in basis. We now define the covariant derivative of the contravariant components of \mathbf{v} as

$$\vec{\nabla}_a v^b = \partial_a v^b + v^c \Gamma_{ac}^b.$$

We would also like to define the covariant derivative of the covariant components of a vector field. To do this, we use the fact that

$$\partial_a \mathbf{E}_b \cdot \mathbf{E}^c = \partial_a \delta_b^c = 0.$$

The product rule yields

$$\mathbf{E}_b \cdot \partial_a \mathbf{E}^c + \mathbf{E}^c \cdot \partial_a \mathbf{E}_b = \mathbf{E}_b \cdot \partial_a \mathbf{E}^c + \mathbf{E}^c \cdot \Gamma_{ab}^d \mathbf{E}_d = \mathbf{E}_b \cdot \partial_a \mathbf{E}^c + \delta_d^c \cdot \Gamma_{ab}^d = \mathbf{E}_b \cdot \partial_a \mathbf{E}^c + \Gamma_{ab}^c,$$

which implies

$$\partial_a \mathbf{E}^c = -\Gamma_{ab}^c \mathbf{E}^b.$$

Repeating the steps above now yields

$$\vec{\nabla}_a v_b = \partial_a v_b - \Gamma_{ab}^c v_c.$$

Curve length Consider some curve parametrized by t , and let $\dot{\gamma}$ denote its tangent. The curve length is given by

$$ds^2 = d\mathbf{x} \cdot d\mathbf{x} = g(\dot{\gamma}, \dot{\gamma}) dt^2 = g_{ab} \dot{\chi}^a \dot{\chi}^b dt^2.$$

The curve length is now given by

$$L = \int dt \sqrt{g_{ab} \dot{\chi}^a \dot{\chi}^b}.$$

Geodesics A geodesic is a curve that extremises the curve length between two points. From variational calculus, it is known that such curves satisfy the Euler-Lagrange equations, and we would like a differential equation that describes such a curve. By defining $\mathcal{L} = \sqrt{g_{ab}\dot{\chi}^a\dot{\chi}^b}$, the Euler-Lagrange equations for the curve length becomes

$$\partial_{\chi^a}\mathcal{L} - \frac{d}{dt}\partial_{\dot{\chi}^a}\mathcal{L} = 0.$$

Computing the derivatives yields

$$\frac{1}{2\mathcal{L}} \left(\dot{\chi}^a\dot{\chi}^b\partial_c g_{ab} - \frac{d}{dt}(2\dot{\chi}^a g_{ac}) \right) = 0.$$

We note that the expression in the paranthesis is the Euler-Lagrange equation for the integral of \mathcal{L}^2 . Expanding the derivative yields

$$\frac{1}{\mathcal{L}} \left(\frac{1}{2}\dot{\chi}^a\dot{\chi}^b\partial_c g_{ab} - g_{ac}\ddot{\chi}^a - \dot{\chi}^a\dot{\chi}^b\partial_b g_{ac} \right) = 0.$$

Multiplying this by $-g^{cd}\mathcal{L}$ yields

$$g_{ac}g^{cd}\ddot{\chi}^a + \frac{1}{2}\dot{\chi}^a\dot{\chi}^b g^{cd}(2\partial_b g_{ac} - \partial_c g_{ab}) = g_{ac}g^{cd}\ddot{\chi}^a + \frac{1}{2}\dot{\chi}^a\dot{\chi}^b g^{cd}(\partial_b g_{ac} + \partial_b g_{ac} - \partial_c g_{ab}) = 0.$$

The a and b indices are summed over, and may thus be swapped. Combined with the symmetry of the metric tensor, this yields

$$g_{ac}g^{cd}\ddot{\chi}^a + \frac{1}{2}\dot{\chi}^a\dot{\chi}^b g^{cd}(\partial_b g_{ac} + \partial_a g_{cb} - \partial_c g_{ab}) = 0.$$

Summation of the first term over a gives $g_{ac}g^{cd}\ddot{\chi}^a = g_{ca}g^{cd}\ddot{\chi}^a = g^{cd}\ddot{\chi}_c$, and summation over c gives $g^{cd}\ddot{\chi}_c = g^{cd}\ddot{\chi}^d$. This thus yields

$$\ddot{\chi}^d + \frac{1}{2}\dot{\chi}^a\dot{\chi}^b g^{cd}(\partial_b g_{ac} + \partial_a g_{cb} - \partial_c g_{ab}) = 0.$$

Christoffel symbols and the geodesic equation Consider a straight line with a tangent vector of constant magnitude. In euclidean space, this is a geodesic. This curve satisfies

$$\frac{d\dot{\gamma}}{dt} = (\dot{\gamma} \cdot \vec{\nabla})\dot{\gamma} = \dot{\chi}^a\partial_a\dot{\gamma} = \dot{\chi}^a(\vec{\nabla}_a\dot{\chi}^d)\mathbf{E}_d = (\dot{\chi}^a\partial_a\dot{\chi}^d + \dot{\chi}^a\dot{\chi}^c\Gamma_{ac}^d)\mathbf{E}_d.$$

Comparison to the chain rule yields

$$\frac{d\dot{\gamma}}{dt} = (\ddot{\chi}^a + \dot{\chi}^a\dot{\chi}^c\Gamma_{ac}^d)\mathbf{E}_d.$$

Comparing this to the geodesic equation yields

$$\Gamma_{ab}^d = \frac{1}{2}g^{dc}(\partial_b g_{ac} + \partial_a g_{cb} - \partial_c g_{ab}).$$

A better approach would have been to go through the derivation of the geodesic equation again, identifying the Christoffel symbols as you go, but I have no idea how to do that.

The geometry of curved space In curved space, we face the restriction that there is no position vector. All vectors in curved space are instead restricted to the tangent space. It turns out that tangent vectors at a point have coordinates $\dot{\chi}^a$ and that the tangent vectors consist of the tangent vectors to the coordinate lines, i.e. partial derivatives.

We can also impose a metric tensor such that $\mathbf{v} \cdot \mathbf{w} = g_{ab}v^aw^b$, where the metric tensor is symmetric and positive definite.

Dual vectors can be defined as linear maps from tangent vectors to scalars, i. e. on the form

$$V(\mathbf{w}) = V_a w^a.$$

In particular, the dual vector df can be defined as

$$df(\mathbf{v}) = v^a \partial_a f = \frac{df}{dt}$$

along a curve with \mathbf{v} as a tangent. A basis for the space of dual vectors is $e^a = d\chi^a$. The tangent and dual spaces, if a metric exists, are related by $v_a = g_{ab}v^b$.

Curve lengths are defined and computed as before. By defining geodesics as curves that extremize path length, this gives a set of Christoffel symbols and therefore a covariant derivative and a sense of what it means for a vector to change along a curve.

4 Classical mechanics

In classical mechanics, configuration space is the space of all possible configurations of a system. We can impose coordinates χ^a on this space in order to use what we know.

Kinetic energy Kinetic energy is defined by a rank 2 tensor as

$$E_k = \frac{1}{2} T_{ab} \dot{\chi}^a \dot{\chi}^b,$$

where the dot now really represents the time derivative.

Hamilton's principle We define the Lagrangian of a system as $\mathcal{L} = E_k - V$, where V is the potential energy and taken to be a function on coordinate space. The action of a system over time is defined as

$$S = \int dt \mathcal{L}.$$

Hamilton's principle states that for the motion of the system in configuration space, $\delta S = 0$. This can be expressed as

$$\delta S = \int dt \delta \mathcal{L} = \int dt \left(\partial_{\chi^a} \mathcal{L} - \frac{d}{dt} \partial_{\dot{\chi}^a} \mathcal{L} \right) \delta \chi^a = 0.$$

The kinetic metric Consider a system with no potential energy. The Lagrangian simply becomes $\mathcal{L} = \frac{1}{2} T_{ab} \dot{\chi}^a \dot{\chi}^b$. This is very similar to the integral of curve length (or, rather its square, the extremum of which was noted to be the same), except g_{ab} has been replaced by T_{ab} . This inspires us to define T_{ab} as the kinetic metric, with corresponding Christoffel symbols.

Motion of a classical system By defining $a^b = \dot{\chi}^a \vec{\nabla}_a \dot{\chi}^b$, the previous work leads us to a system with no potential satisfying $a^b = \ddot{\chi}^b + \Gamma_{ac}^b \dot{\chi}^a \dot{\chi}^c = 0$. In other words, a system with no potential moves along the geodesics of the kinetic metric.

For a system with a potential, only the $\partial_{\chi^a} \mathcal{L}$ term is affected, and

$$a^b = -T^{ba} \partial_a V = T^{ba} F,$$

which is a generalization of Newton's second law.

Noether's theorem Noether's theorem relates symmetries of physical systems to conservation laws.

What is a symmetry, then? Consider a one-parameter transformation $t \rightarrow \tau(t, s)$, $q^a \rightarrow Q^a(q, s)$, where s is the parameter with respect to which the system is transformed, such that $\tau(t, 0) = t$, $Q^a(q, s) = q^a$ and for small $s = \varepsilon$ that $t \rightarrow t + \varepsilon \delta t$, $q^a \rightarrow q^a + \varepsilon \delta q^a$. This is assumed to be normalized such that δt is either 0 or 1. How? Don't ask. A quasi-symmetry of a system with Lagrangian \mathcal{L} is a transformation such that

$$\varepsilon \delta \mathcal{L} = \mathcal{L}(Q, \dot{Q}, \tau) - \mathcal{L}(q, \dot{q}, t) = \varepsilon \frac{dF}{dt}$$

for some F . The variation of the Lagrangian can be written as

$$\delta\mathcal{L} = \partial_{q^a}\mathcal{L}\delta q^a + \partial_{\dot{q}^a}\mathcal{L}\delta\dot{q}^a + \partial_t\mathcal{L}\delta t.$$

The total time derivative of the Lagrangian is given by

$$\frac{d\mathcal{L}}{dt} = \partial_t\mathcal{L} + \partial_{q^a}\mathcal{L}\dot{q}^a + \partial_{\dot{q}^a}\mathcal{L}\ddot{q}^a,$$

which yields

$$\delta\mathcal{L} = \partial_{q^a}\mathcal{L}(\delta q^a - \dot{q}^a\delta t) + \partial_{\dot{q}^a}\mathcal{L}(\delta\dot{q}^a - \ddot{q}^a\delta t) + \frac{d\mathcal{L}}{dt}\delta t.$$

The equations of motion are $\partial_{q^a}\mathcal{L} = \frac{d}{dt}\partial_{\dot{q}^a}\mathcal{L}$. For a set of coordinates that satisfy this - a so-called on-shell solution - we have

$$\begin{aligned}\delta\mathcal{L} &= \frac{d}{dt}\partial_{\dot{q}^a}\mathcal{L}(\delta q^a - \dot{q}^a\delta t) + \partial_{\dot{q}^a}\mathcal{L}(\delta\dot{q}^a - \ddot{q}^a\delta t) + \frac{d\mathcal{L}}{dt}\delta t \\ &= \frac{d}{dt}(\partial_{\dot{q}^a}\mathcal{L}(\delta q^a - \dot{q}^a\delta t) + \mathcal{L}\delta t).\end{aligned}$$

If the transformation is a quasi-symmetry of the system, then this is equal to a total time derivative of F , and the quantity

$$J = F - \partial_{\dot{q}^a}\mathcal{L}\delta q^a + (\dot{q}^a\partial_{\dot{q}^a}\mathcal{L} - \mathcal{L})\delta t$$

thus satisfies $\frac{dJ}{dt} = 0$. We can introduce the general momenta $p_a = \partial_{\dot{q}^a}\mathcal{L}$ and the Hamiltonian $\mathcal{H} = p_a\dot{q}^a - \mathcal{L}$ to write

$$J = F - p_a\delta q^a + \mathcal{H}\delta t.$$

We arrive at the conclusion that J is a conserved quantity under a quasi-symmetry of the system. Identifying the conservation laws of a system is thus a matter of identifying the quasi-symmetries of a system and computing J under that transformation.

Example: A free particle in space Consider a free particle in space. Its Lagrangian is given by $\mathcal{L} = \frac{1}{2}m\dot{\mathbf{x}}^2$, and the variation of this is

$$\delta\mathcal{L} = m\dot{\mathbf{x}} \cdot \delta\dot{\mathbf{x}}.$$

Its general momentum is

$$\mathbf{p} = \partial_{\dot{\mathbf{x}}}\mathcal{L} = m\dot{\mathbf{x}}.$$

The Hamiltonian is

$$\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{x}} - \mathcal{L} = \frac{1}{2}m\dot{\mathbf{x}}^2.$$

We now want to identify quasi-symmetries of the system that make the variation of the Lagrangian either zero or the time derivative of some quantity. A key idea here is that we are only allowed to change the variations (or so I think).

A first attempt is keeping $\delta\mathbf{x}$ constant and not varying time (a spatial translation), which implies $\delta\dot{\mathbf{x}} = \mathbf{0}$ and $\delta\mathcal{L} = 0$. This implies that F is constant. The conserved quantity is thus

$$J = F - \mathbf{p} \cdot \delta\mathbf{x} = F - \mathbf{p} \cdot \mathbf{c},$$

i.e. the momentum of the system is conserved. We also note that the constant F in this case is arbitrary, and we might as well have set it to 0. This will be the case at least sometimes.

A second attempt is varying time, i.e. $\delta t = 1$, but keeping the coordinates fixed, i.e. $\delta\mathbf{x} = \mathbf{0}$ (a time translation). This yields $\delta\dot{\mathbf{x}} = \mathbf{0}$ and $\delta\mathcal{L} = 0$. Once again F is constant and taken to be zero, and the conserved quantity is thus $J = H$, i.e. the Hamiltonian of the system is conserved.

A third attempt is to somehow make the scalar product in the variation of the Lagrangian zero, without varying time. An option is $\delta\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}$, where $\boldsymbol{\omega}$ is a constant vector. This yields $\delta\dot{\mathbf{x}} = \boldsymbol{\omega} \times \dot{\mathbf{x}}$ and $\delta\mathcal{L} = 0$. The conserved quantity is thus

$$\begin{aligned}J &= -\mathbf{p} \cdot (\boldsymbol{\omega} \times \mathbf{x}) \\ &= -\boldsymbol{\omega} \cdot (\mathbf{x} \times \mathbf{p}).\end{aligned}$$

Since $\boldsymbol{\omega}$ is constant, that means that $\mathbf{x} \times \mathbf{p}$, i.e. the angular momentum, is conserved.