A MOTION

A.1 *Scales in physical laws: Poisson vs. Yukawa*

A good example of a scale-free physical law is the $1/r$-potential in electrostatics or in Newton-gravity in 3d dimensions: It follows as a vacuum solution of the Poisson-equation

$$\Delta\Phi = -4\pi\rho \tag{A.1}$$

in the Gauß-system of units. Assuming spherical symmetry for the field away from a point charge one can verify that $\Phi \propto 1/r$ is in fact a solution to

$$\Delta\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) = 0 \tag{A.2}$$

The solution $\Phi \propto 1/r$ is perfectly scale free as a power law; increasing the charge can be absorbed in an increased distance. This can be seen directly by the scale transformation $r \to \alpha r$, under which $\partial_r \to \alpha^{-1}\partial_r$ and consequently $\Delta \to \alpha^{-2}\Delta$. Then, $\Phi \to \alpha^{-1}\Phi$ because $\rho \to \alpha^{-3}\rho$, and two powers of α cancel, making Φ consistent with the scaling of r.

This scale-invariance expressed by the power law is broken in the Yukawa-equation

$$\left[\Delta - \lambda^2\right]\Phi = -4\pi\rho \tag{A.3}$$

with a parameter λ: It has units of inverse length and allows to distinguish between the regimes $\lambda r \ll 1$ and $\lambda r \gg 1$, because despite the fact that the field equation is still linear, scale-invariance is violated. As a solution one finds $\Phi \propto \exp(-\lambda r)/r$ in 3 dimensions, which behaves $\Phi \propto 1/r$ for small distances, where $\exp(-\lambda r) \simeq 1 - \lambda r \pm \ldots$, but at large distances the solution drops faster to zero than $1/r$. Therefore, one has constructed a scale-dependent modification of the Poisson-equation. From a physical point of view, Yukawa aimed at a short-range force for explaining the binding of nucleons, and almost at exactly the same time, Debye considered electric fields in electrolytes, where the shielding of ions led to a fast decrease of electric fields around charges.

Please note that much of the arguments are only applicable in 3 dimensions or more. In two dimensions the Poisson-equation reads

$$\Delta\Phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Phi}{\partial r}\right) = \frac{1}{r}\frac{\partial\Phi}{\partial r} + \frac{\partial^2\Phi}{\partial r^2} = 0 \tag{A.4}$$

which is solved by $\Phi \propto \ln r$: While this is mathematically perfect, there are a couple of issues concerning the physical application. The potential does not vanish for $r \to \infty$ and there is no scale-free behaviour of the solution despite the fact that the Poisson-equation is scale free. Adding a Yukawa-type term

$$\left[\Delta - \lambda^2\right]\Phi = \frac{1}{r}\frac{\partial\Phi}{\partial r} + \frac{\partial^2\Phi}{\partial r^2} - \lambda^2\Phi = 0 \tag{A.5}$$

gives rise to a differential equation that is known as Emden-Fowler-type and has a (very complicated) solution in terms of Bessel-functions J_0 and Y_0, where λ appears as the wavenlength of the oscillation in the Bessel-functions: Surely it plays the role of a scale, but not as clearly as in 3 dimensions.

A.2 *Buckingham's* Π*-theorem and the Navier-Stokes-equation*

The example about the Poisson- and Yukawa-equation showed how scales can be introduced in a linear equation, and we should investigate if there can be scale-free behaviour in a nonlinear equation. This is in fact the case, as the example of the dimensionless Navier-Stokes-equation in fluid mechanics shows:

$$\partial_t \upsilon^i + (\upsilon_j \partial^j)\upsilon^i = -\frac{1}{\rho}\partial^i p - \partial^i \Phi + \mu \partial_j \partial^j \upsilon^i \tag{A.6}$$

The Navier-Stokes equation describes the acceleration of a fluid with velocity υ^i under the action of forces, for instance gradients in pressure p, in the gravitational potential Φ and viscous forces with the shear viscosity μ, all under the condition of incompressible fluids with $\partial_i \upsilon^i = 0$. Multiplying with ρ to make things a bit more transparent gives an equation where every term has units of mass/length^2/time^2.

If we introduce typical scales, we could reach a form of the Navier-Stokes equation where it would become scale free: It would become an dimensionless equation, and flow patterns of different physical dimension, if they fall back onto the same dimensionless equation, would be scaled versions of each other: Introducing a length scale L for $x \to x^* = x/\mathrm{L}$, a time scale T for $t \to t^* = t/\mathrm{T}$, a velocity scale V for $\upsilon \to \upsilon^* = \upsilon/\mathrm{V}$, a pressure scale P for $p \to p^* = p/\mathrm{P}$ and finally a scale for the gravitational acceleration G for $g \to g^* = g/\mathrm{G}$ yields

$$\frac{\rho \mathrm{V}}{\mathrm{T}}\partial_t^* \upsilon^{*i} + \frac{\rho \mathrm{V}^2}{\mathrm{L}}(\upsilon^*{}_j \partial^{*j})\upsilon^{*i} = -\frac{\mathrm{P}}{\mathrm{L}}\partial^{*i} p^* - \frac{\rho \mathrm{G}}{\mathrm{L}}\partial^{*i}\Phi^* + \frac{\mu\rho \mathrm{V}}{\mathrm{L}^2}\partial^*{}_j \partial^{*j}\upsilon^{*i} \tag{A.7}$$

with the dimensionless derivatives

$$\frac{\partial}{\partial x} = \frac{\partial x^*}{\partial x}\frac{\partial}{\partial x^*} = \frac{1}{\mathrm{L}}\partial_x^* \quad \text{and} \quad \frac{\partial}{\partial t} = \frac{\partial t^*}{\partial t}\frac{\partial}{\partial t^*} = \frac{1}{\mathrm{T}}\partial_t^* \tag{A.8}$$

It should be noted that L and T are typical scales on which the flow changes, and that the scale V is independent of L/T: You can have a slowly-varying high velocity flow or, vice versa, a rapidly changing low-velocity flow.

In eqn. A.7 one has reached a curious ordering of all terms: The units are concentrated in the prefactors, while all terms involving quantities with a superscript-$*$ are dimensionless. Dividing the entire formula by the prefactor of the second, nonlinear term then gives rise to:

$$\underbrace{\frac{\mathrm{L}}{\mathrm{TV}}}_{\text{Strouhal}}\partial_t^* \upsilon^{*i} + (\upsilon^*{}_j \partial^{*j})\upsilon^{*i} = -\underbrace{\frac{\mathrm{P}}{\rho \mathrm{V}^2}}_{\text{Euler}}\partial^{*i} p^* - \underbrace{\frac{\mathrm{G}}{\mathrm{V}^2}}_{\text{Froude}^{-2}}\partial^{*i}\Phi^* + \underbrace{\frac{\mu}{\mathrm{VL}}}_{\text{Reynolds}^{-1}}\partial^*{}_j \partial^{*j}\upsilon^{*i} \tag{A.9}$$

Flows with identical scaling numbers can be mapped onto each other, and the primary application is indeed technical: When designing airplanes, it might be difficult to construct a full-size airplane model and to test it in a wind tunnel at actual velocities. Instead, one can try out a much smaller model at lower air speeds; if the scaling numbers are identical between the two situations, the flow patterns are scaled versions of each other. In summary, scales might be present in linear laws and there might be scale-free behaviour in nonlinear laws.

There is a fantastic way of memorising the Reynolds number, which is associated with turbulence: VL/μ means, that stirring a coffee fast with a big spoon is making the flow turbulent, but it would not work in honey!

A.3 *Constants of Nature and Planck's system of units*

There is a clear distinction between classical physics and modern physics: In classical physics, the purpose of constants is to sort out the units and to relate quantities in a phenomenological way: From this point of view there really is not much of a difference between the spring constant k in Hooke's law

$$F = -kr \tag{A.10}$$

and the gravitational constant G in Newton's law of gravity

$$F = -G\frac{mM}{r^2} \tag{A.11}$$

Modern physics on the other hand distinguishes between different regimes where Nature behaves classical or shows a markedly new behaviour, for instance at high velocities close to c, motion at low action close to $\hbar$, at low energies comparable to the thermal energy $k_B T$ and finally at distances close to GM/c^2 at massive objects. In these cases, classical physics gets replaced by special relativity, by quantum mechanics, by statistical physics and finally by general relativity, respectively.

As first noticed by Planck, the four constants c, $\hbar$, G and k_B can be used to define a *natural* system of units which is universally valid and does not depend on any human concept for length, time, mass or temperature. For instance, a fundamental mass could be constructed by setting

$$m_P = c^\alpha \hbar^\beta G^\gamma = \text{length}^{\alpha+3\beta+2\gamma}\text{time}^{-\alpha-2\beta-\gamma}\text{mass}^{-\beta+\gamma} \tag{A.12}$$

which is solved by $\alpha = -\beta = \gamma = 1/2$, defining the Planck-mass m_P,

$$m_P = \sqrt{\frac{c\hbar}{G}} \simeq 10^{-8}\text{kg} \simeq 10^{16}\text{GeV}/c^2 \tag{A.13}$$

Similarly, one can define a length-scale l_P, a time scale t_P and a temperature scale T_P,

$$l_P = \sqrt{\frac{G\hbar}{c^2}} \simeq 10^{-35}\text{m}, \quad t_P = \frac{l_P}{c} \simeq 10^{-43}\text{s}, \quad T_P = \frac{1}{k_B}\sqrt{\frac{c^3\hbar}{G}} \simeq 10^{30}\text{K} \tag{A.14}$$

This beautiful idea is somewhat tainted by the realisation that there are in fact two constants in gravity, G and the cosmological constant Λ. This second constant makes the construction of a fundamental system of units ambiguous, and what's even more puzzling, starting from c, G and Λ defines a system which very well characterises the Universe today, with a length scale $1/\sqrt{\Lambda} \simeq 3\text{Gpc}/h$ and an age of $1/(\sqrt{\Lambda}c) \simeq 10^{17}$s, while even derived quantities like the density scale come out correctly.

A.4 *Classical Lagrange-functions*

Classical mechanics describes motion axiomatically with a Lagrange-function $\mathcal{L}(q^i, \dot{q}^i)$ as a function of the (generalised) coordinates q^i and the velocities $\dot{q}^i$, defined as the rate of change of the coordinates as the time-parameter evolves. An integration over t then defines the action S

$$S = \int_{t_i}^{t_f} \mathrm{d}t \, \mathcal{L}(q^i, \dot{q}^i) \tag{A.15}$$

as a functional over the trajectory $q^i(t)$. Hamilton's principle

$$\delta S = 0 \tag{A.16}$$

then asserts that the physical motion is the one that extremises the action functional, and incidentally we realise that the linearity of the variation δS induces that the action is affinely invariant. $S \to aS + b$ would not change anything in Hamilton's principle, as $\delta(aS + b) = a\delta S = 0$ shows the irrelevance of a and b.

Carrying out the variation is done by writing

$$\delta S = \int_{t_i}^{t_f} \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial q^i}\delta q^i + \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\delta \dot{q}^i\right) = \int_{t_i}^{t_f} \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right)\delta q^i = 0 \tag{A.17}$$

after setting $\delta\dot{q}^i = \mathrm{d}/\mathrm{d}t \; \delta q^i$, followed by an integration by parts. The boundary term vanishes if the variation on the boundary vanishes, $\delta q^i(t_i) = \delta q^i(t_f) = 0$, or at least if their difference is constant. From the last expression we can isolate the Euler-Lagrange-equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}^i} = \frac{\partial \mathcal{L}}{\partial q^i} \tag{A.18}$$

If one now chooses the Lagrange function to be

$$\mathcal{L} = \frac{m}{2}\gamma_{ab}\dot{q}^a\dot{q}^b - \Phi(q^i) \tag{A.19}$$

with the Euclidean metric γ_{ab} and a potential Φ, the Euler-Lagrange-function becomes equivalent to Newton's equation of motion: The gradient of the Lagrange-function with respect to the coordinate yields

$$\frac{\partial \mathcal{L}}{\partial q^i} = -\frac{\partial \Phi}{\partial q^i} \tag{A.20}$$

Please always rename the indices in the kinetic term of the Lagrange-function before substituting it into the Euler-Lagrange equation!

and the derivative of the kinetic term becomes

$$\frac{\partial \mathcal{L}}{\partial \dot{q}^i} = \frac{m}{2}\gamma_{ab}\Big(\underbrace{\frac{\partial \dot{q}^a}{\partial \dot{q}^i}}_{\delta^a_i}\dot{q}^b + \dot{q}^a\underbrace{\frac{\partial \dot{q}^b}{\partial \dot{q}^i}}_{\delta^b_i}\Big) = m\dot{x}_i \tag{A.21}$$

Finally, we arrive at Newton's equation of motion $m\ddot{q}_i = -\partial_i\Phi$ by differentiation with respect to t. One might not always have such a convenient separation into a term involving only $\dot{q}^i$ and only q^i, for instance, the harmonic oscillator $\mathcal{L} = \dot{q}^2/2 - \omega^2 q^2/2$ could be rewritten as $\mathcal{L} = (\dot{q} + \omega q)(\dot{q} - \omega q)/2$. In these cases, the time-derivative might act on a function $\partial\mathcal{L}/\partial\dot{q}$ which is still a function of q, so one needs to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}^i} = \ddot{q}^j\frac{\partial^2\mathcal{L}}{\partial\dot{q}^i\partial\dot{q}^j} + \dot{q}^j\frac{\partial^2\mathcal{L}}{\partial q^j\partial\dot{q}^i} = \frac{\partial\mathcal{L}}{\partial q^i} \tag{A.22}$$

and a solution for $\ddot{q}^j$ depends on the invertibility of the matrix $\partial^2\mathcal{L}/\partial\dot{q}^i\partial\dot{q}^j$:

$$\ddot{q}^j = \left(\frac{\partial^2\mathcal{L}}{\partial\dot{q}^i\partial\dot{q}^j}\right)^{-1}\left(\frac{\partial\mathcal{L}}{\partial q^i} - \dot{q}^j\frac{\partial^2\mathcal{L}}{\partial q^j\partial\dot{q}^i}\right) \tag{A.23}$$

and of course for 1-dimensional motion, it would be enough for $\partial^2\mathcal{L}/\partial\dot{q}^2$ to be nonzero. Typically, $\partial^2\mathcal{L}/\partial\dot{q}^2$ is just the mass or inertia of the system, which is strictly positive such that the $\ddot{q}$-term can be isolated.

While the Lagrange-formalism seems straightforward as an axiomatic foundation of classical mechanics, there seem to be many issues: There is no fundamental justification for $\mathcal{L}$ or S, as they are both not measurably quantities. S is only determined up to an affine transform, and so must be $\mathcal{L}$. At least for motion in a vector space, there is no advantage of using Lagrangian mechanics over the Newton equation of motion, and one might wonder what the relation between Hamilton's principle for the motion of objects and Fermat's principle for the propagation of light might be.

A.5 *Classical universality and mechanical similarity*

The Lagrange-function $\mathcal{L}$ is invariant under affine transformations,

$$\mathcal{L} \to a\mathcal{L} + b \tag{A.24}$$

with two constants a and b, which is no more than a novelty: Clearly, both constants drop out of the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 \tag{A.25}$$

b, because it gets lots in the differentiation and a because both differentiations are linear, so it appears as an irrelevant overall prefactor. But there is a way in which this affine invariance of the Lagrange-function can be used in a sensible way: If one rescales the coordinates $q \to \alpha q$ and the time parameter $t \to \beta t$, the kinetic energy T scales $\mathrm{T} \to (\alpha/\beta)^2\mathrm{T}$ and the potential energy $\Phi \to \alpha^n\Phi$ for a scale-free power-law potential $\Phi \propto q^n$. Because the scaling of T and Φ are inherently different, one needs to assume a relation between them, such that the Lagrange-function $\mathcal{L} = \mathrm{T} - \Phi$ just changes by an (irrelevant) overall factor:

Please keep in mind that in classical mechanics the time is just a parameter to describe motion!

$$\frac{\alpha^2}{\beta^2} \propto \alpha^n \quad \text{or, equivalently,} \quad \beta^2 \propto \alpha^{2-n} \tag{A.26}$$

This scaling can be read off from Newton's equation of motion as well (surely it is consistent with the Lagrange-function $\mathcal{L} = \mathrm{T} - \Phi$):

$$\ddot{q} = -\frac{\partial\Phi}{\partial q} \to \frac{\alpha}{\beta^2}\ddot{q} = -\frac{\alpha^n}{\alpha}\frac{\partial\Phi}{\partial q} \quad \text{implying} \quad \beta^2 \propto \alpha^{2-n} \tag{A.27}$$

for the specific form $\Phi \propto q^n$. Therefore, the length and time scales need to be in that particular relation given by the similarity condition $\beta^2 = \alpha^{2-n}$, which we can specifically try out for the most common scale-free potentials:

1. $\Phi \propto q^2$, $n = 2$: harmonic oscillator

   In the case of the harmonic oscillator, similarity implies $t^2 = \text{const}$, which indicates that the time scale of e.g. a pendulum is independent of amplitude.

2. $\Phi \propto q$, $n = 1$: inclined plane with a constant slope

   Here, time and length scale are related by $t^2 \propto q$, typical for uniformly accelerated motion.

3. $\Phi \simeq \text{const}$, $n = 0$: flat potential

   A flat potential is characterised by $t^2 \propto q^2$, or equivalently, inertial motion at constant velocity, as no acceleration takes place

4. $\Phi \propto 1/q$, $n = -1$: Coulomb-potential

   In a Coulomb-potential, Kepler's third law is valid, as $t^2 \propto q^3$.

These four examples illustrate the *principle of mechanical similarity* where we can say something profound about motion without performing the variation or solving the actual equation of motion. For instance, we found out that all planetary orbits are scaled version of each other as every orbit needs to fulfil Kepler's law. To formulate this in a very extreme way, for determining the distances of the planets to the Sun one just needs a calendar.

A cute example of mechanical similarity is the motion of astronauts on the surface of the Moon, at a fraction of Earth's gravity: There, everything seems to be happening in slow motion, because accelerations are much lower. Speeding up a movie of astronauts would make everything appear normal again. You might as well have the association that the motion of the astronauts looks as if they were under water: That's sensible, too, because buoyancy reduces the effective gravitational acceleration, leading to the same effect of longer time constants.

A.6 *Total derivatives in the Lagrange-function*

The Lagrange-function is only determined up to a total derivative $\mathrm{d}M(q^i,t)/\mathrm{d}t$ of a function $M(q^i,t)$ which may depend on the coordinates q^i and on the time parameter t, but *not* on the velocities $\dot{q}^i$. In fact, transforming the Lagrange-function

$$\mathcal{L}(q^i,\dot{q}^i) \rightarrow \mathcal{L}(q^i,\dot{q}^i) + \frac{\mathrm{d}}{\mathrm{d}t}M(q^i,t) \tag{A.28}$$

implies a transformation of the action

$$S = \int \mathrm{d}t\; \mathcal{L}(q^i,\dot{q}^i) \rightarrow S + \int \mathrm{d}t\; \frac{\mathrm{d}}{\mathrm{d}t}M(q^i,t) \tag{A.29}$$

but Hamilton's principle $\delta S = 0$ invalidates the new term: Writing the variation with a Euler-Lagrange-operator acting on M

$$\delta S = \delta \int dt\, \mathcal{L} + \frac{d}{dt}M = \delta \int dt\, \mathcal{L} + \int dt \left[\frac{d}{dt}\frac{\partial}{\partial \dot{q}} - \frac{\partial}{\partial q}\right]\frac{dM}{dt} \tag{A.30}$$

lets us treat each term separately. For the second term, there is

$$\frac{\partial}{\partial q}\frac{dM}{dt} = \frac{\partial}{\partial q}\left(\dot{q}\frac{\partial M}{\partial q} + \frac{\partial M}{\partial t}\right) = \frac{\partial \dot{q}}{\partial q}\frac{\partial M}{\partial q} + \dot{q}\frac{\partial^2 M}{\partial q^2} + \frac{\partial^2 M}{\partial q \partial t} \tag{A.31}$$

because M depends on q and t, but not on $\dot{q}$. For the first term, we get

$$\frac{\partial}{\partial \dot{q}}\frac{dM}{dt} = \frac{\partial}{\partial \dot{q}}\left(\dot{q}\frac{\partial M}{\partial q} + \frac{\partial M}{\partial t}\right) = \frac{\partial \dot{q}}{\partial \dot{q}}\frac{\partial M}{\partial q} = \frac{\partial M}{\partial q} \tag{A.32}$$

because $\partial\dot{q}/\partial\dot{q} = 1$. A successive time derivative yields then

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}}\frac{dM}{dt} = \dot{q}\frac{\partial^2 M}{\partial q^2} + \frac{\partial^2 M}{\partial q \partial t} \tag{A.33}$$

so that all additional terms cancel, because

$$\frac{\partial \dot{q}}{\partial q} = \frac{\partial}{\partial q}\frac{\partial q}{\partial t} = \frac{\partial}{\partial t}\frac{\partial q}{\partial q} = \frac{\partial}{\partial t}1 = 0 \tag{A.34}$$

where we've use the interchangeability of the second partial derivatives.

Alternatively, one can argue that adding the total derivative changes the action according to

$$S \to S + \int_{t_i}^{t_f} dt\, \frac{dM}{dt} = S + M(q(t_f), t_f) - M(q(t_i), t_i) \tag{A.35}$$

The variation δq vanishes at the endpoints t_i and t_f by construction, this however does not constrain the value of $\delta\dot{q}(t)$ at the endpoints. Because $M(q,t)$ is only a function of q and not of $\dot{q}$ we can be sure that δM vanishes for both $\delta q(t_i)$ and $\delta q(t_f)$, cancelling the boundary term.

A.7 *Virial theorem*

Lagrangian systems are energy-conserving if $\mathcal{L}$ does not depend direclty on time t. This can be seen explicitly in Newton's equation of motion

$$m\ddot{q} = -\frac{\partial}{\partial q}\Phi \tag{A.36}$$

if multiplied with $\dot{q}$:

$$m\dot{q}\ddot{q} = m\frac{d}{dt}\frac{\dot{q}^2}{2} = -\dot{q}\frac{\partial}{\partial q}\Phi = -\frac{d}{dt}\Phi \quad \to \quad \frac{d}{dt}\left(\frac{m}{2}\dot{q}^2 + \Phi\right) = 0 \tag{A.37}$$

with the energy $E = m\dot{q}^2/2 + \Phi$, because Φ depends on t only through the trajectory $q(t)$, and obviously not explicitly. While the total energy is conserved and while the equation of motion constantly changes kinetic into potential energy and back, one might ask the rather sensible question if the system likes to spend more time in a state of high kinetic energy or in a state of high potential energy.

The answer to this question is the virial theorem: Multiplying the equation of motion with q instead of $\dot{q}$ and averaging over a time interval Δt gives

$$0 = \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t \left(mq\ddot{q} + q\frac{\partial\Phi}{\partial q}\right) = \frac{1}{\Delta t} mq\dot{q}|_0^{\Delta t} - \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t \left(m\dot{q}^2 - q\frac{\partial\Phi}{\partial q}\right) \tag{A.38}$$

after an integration by parts of the first term. The term $mq\dot{q}$ gets evaluated at 0 and Δt and can be estimated to be less than the maximum coordinate q_{max} times the maximum velocity $\dot{q}_{\mathrm{max}}$ over the time interval from 0 to Δt, *if* the motion is bounded:

$$\frac{1}{\Delta t} mq\dot{q}|_0^{\Delta t} \leq \frac{1}{\Delta t} mq_{\mathrm{max}}\dot{q}_{\mathrm{max}} \to 0 \quad \text{as} \quad \Delta t \to \infty \tag{A.39}$$

and vanishes then if the average is taken over arbitrarily large time intervals. The term

$$\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, m\dot{q}^2 = 2\langle \mathrm{T}\rangle \tag{A.40}$$

becomes twice the average kinetic energy, and for proceeding with the potential term, we need to make an assumption about its functional shape: *If* it is a homogeneous function of order k, $\Phi \propto q^k$, we get

$$\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, q\frac{\partial\Phi}{\partial q} = \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, k\Phi = k\langle\Phi\rangle \tag{A.41}$$

because $q\partial_q\Phi = q\partial_q q^k = kqq^{k-1} = kq^k = k\Phi$. Therefore, the average energies are related to each other by the virial law

$$2\langle \mathrm{T}\rangle = k\langle\Phi\rangle \tag{A.42}$$

A prime example for this is the harmonic oscillator, where both T and Φ are homogeneous functions of order $k = 2$ in $\dot{q}$ and q, respectively, resulting in equal average kinetic and potential energies.

It is perhaps a bit more transparent to derive the virial law from the Euler-Lagrange-equation as the equivalent equation of motion directly. Multiplying with the coordinate q and averaging gives

$$0 = \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, q\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q}\right) = \frac{1}{\Delta t} q\frac{\partial\mathcal{L}}{\partial\dot{q}}|_0^{\Delta t} - \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t \left(\dot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} + q\frac{\partial\mathcal{L}}{\partial q}\right) \tag{A.43}$$

which we solve by an integration by parts in the first term. The derivative $\partial\mathcal{L}/\partial\dot{q}$ is the canonical momentum p and we can invoke the same argument about bounded systems,

$$\frac{1}{\Delta t} mqp\Big|_0^{\Delta t} \leq \frac{1}{\Delta t} mq_{\max}p_{\max} \to 0 \quad \text{as} \quad \Delta t \to 0 \tag{A.44}$$

now in phase space, so that the two remaining averages determine the virial law: Typically, the Lagrange-function is a homogeneous function of order 2 in $\dot{q}$,

$$\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, \dot{q}\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, \dot{q}\frac{\partial \mathrm{T}}{\partial \dot{q}} = 2\langle \mathrm{T}\rangle \tag{A.45}$$

and in the case of power laws a homogeneous function of order k in q with an additional minus-sign.

$$\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, q\frac{\partial \mathcal{L}}{\partial q} = -\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t\, q\frac{\partial \Phi}{\partial q} = -k\langle \Phi\rangle \tag{A.46}$$

and the virial law is established:

$$2\langle \mathrm{T}\rangle = k\langle \Phi\rangle \tag{A.47}$$

An illustrative example might be to choose a rather high value of k: Then, the potential is essentially a box with a flat bottom and high walls, in which the particle zooms from left to right and back in a state of high kinetic energy essentially all the time, and spends little time climbing up the walls and changing its direction of motion. For high k, $\langle \mathrm{T}\rangle$ is much higher than $\langle \Phi\rangle$. The second example is the impossibility of a gravitationally bound ball of photons: There, the kinetic energy is a homogeneous function of order $k = 1$ as energy depends linearly on momentum, and for the gravitational potential $\Phi \propto 1/q$ we have $k = -1$ as the degree, so the virial law becomes: $\langle \mathrm{T}\rangle = -\langle \Phi\rangle$, and the total energy $\mathrm{E} = \langle \mathrm{T}\rangle + \langle \Phi\rangle = 0$, but it would need to be negative for a bound system. Lastly, a peculiar case is a harmonic oscillator with $k = 2$: Then, the average kinetic and potential energies are exactly equal, $\langle \dot{x}^2\rangle = \omega^2\langle x^2\rangle$.

A.8 *Galilei-invariance of classical systems*

Classical mechanics uses Galilean relativity, meaning that the equation of motions are identical in every Galilei-frame, which in turn is defined as the class of frames moving at constant relative velocities where inertial forces are absent. Mathematically they are defined as the coordinate transformations $q \to q + \upsilon t$ with a constant velocity υ, such that $\dot{q} \to \dot{q} + \upsilon$ and $\ddot{q} \to \ddot{q}$, leaving the Newtonian equation of motion unchained.

On the level of the Lagrange-function there is a change,

$$\mathcal{L} = \frac{m}{2}\dot{q}^2 \to \frac{m}{2}(\dot{q}+\upsilon)^2 = \frac{m}{2}\left(\dot{q}^2 + 2\dot{q}\upsilon + \upsilon^2\right) = \frac{m}{2}\dot{q}^2 + \frac{\mathrm{d}}{\mathrm{d}t}\left(mq\upsilon + \frac{m}{2}\upsilon^2 t\right) \tag{A.48}$$

where the additional terms can be absorbed into a the time derivative of a function $\mathrm{M}(q,t)$ which depends on the coordinate q and t (please keep in mind that υ is constant!), but not on $\dot{q}$ directly, so the action $\mathrm{S} = \int \mathrm{d}t\, \mathcal{L}$ is effectively unchanged.

While this looks very convincing there is something fundamental that is being overlooked in Galilean, non-relativistic mechanics. In the process of varying the action, one transitions from an *invariant*, scalar Lagrange-function to a *covariant* vectorial or tensorial equation of motion with consistent transformation properties. For instance, the Lagrange-function

$$\mathcal{L} = \frac{1}{2}\gamma_{ij}\dot{x}^i\dot{x}^j - \Phi(x^i) \tag{A.49}$$

is rotationally invariant, clearly because of the scalar product $\gamma_{ij}\dot{x}^i\dot{x}^j$ involving the Euclidean metric γ_{ij}, but also because of the scalar potential Φ, which doesn't have any internal degrees of freedom that would be affected by a rotation. After variation, the equation of motion

$$m\ddot{x}_i = -\partial_i\Phi \tag{A.50}$$

puts a linear form $\ddot{x}_i$ into relation with the gradient $\partial_i\Phi$, again written as a linear form, so that the entire formula transforms consistently. Clearly, one could use the inverse Euclidean metric γ^{ij} to write it in vector form, $m\ddot{x}^i = -\partial^i\Phi$, with $\ddot{x}^i = \gamma^{ij}\ddot{x}_j$ and $\partial^i\Phi = \gamma^{ij}\partial_j\Phi$. This property of the variational principles is known as the invariance-covariance principle: You always obtain a covariant equation of motion (or field equation) from an invariant Lagrange-function (or density).

The curiosity is now that actually boosts and rotations form a common group, the proper Lorentz-group, so classical mechanics based on Galilean relativity instead of Lorentzian relativity needs to realise the invariance-covariance principle differently: Time is universal and identical in all frames, and excluded from coordinate transforms. This enables the invariance of the accelerations $\ddot{q}$ in all frames instead of dealing with a construction a covariant equation of motion.

A.9 *Alternatives to the Lagrange-function*

The Lagrange-function $\mathcal{L} = T(\dot{q}^i) - \Phi(q^i)$ is defined axiomatically in classical mechanics in order to make it consistent with the Newtonian equation of motion. You might want to ask if one could have other terms in the Lagrange-function that would be compatible with a linear, second order equation of motion. As the order of the powers of q^i and $\dot{q}^i$ decreases by one through the differentiation in the Euler-Lagrange-equation, there should be at most squares in the Lagrange-function. Higher-order derivatives like $\ddot{q}^i$ are excluded by the Ostrogradsky-instability (we will come to that!). Therefore, one could imagine a Lagrange-function

$$\mathcal{L} = \gamma_{ij}\dot{q}^i\dot{q}^j - \gamma_{ij}q^i\ddot{q}^j - \Phi + \lambda_i q^i + \mu_i\dot{q}^i + \alpha_{ij}q^iq^j + \beta_{ij}q^i\dot{q}^j + \epsilon + \ldots \tag{A.51}$$

and possibly many more terms. But actually, one is quite restricted: $-\gamma_{ij}q^i\ddot{q}^j$ is just $\gamma_{ij}\dot{q}^i\dot{q}^j$ after an integration by parts, $\lambda_i q^i$, $\alpha_{ij}q^iq^j$ and ϵ are particular potentials, and $\beta_{ij}q^i\dot{q}^j$ as well as $\mu_i\dot{q}^i$ would vanish: After all, they are just total time derivatives of the functions $\beta_{ij}q^iq^j$ and μ_iq^i which just depend on time and position.

It is very interesting to see that any reformulation of the Lagrange-function that can be achieved by integration by parts gives rise to exactly the same equation of motion: That is the case because $\mathcal{L}$ only ever appears in the action integral $S = \int dt\, \mathcal{L}$ with a fixed boundary. But for dealing with a term like $\gamma_{ij}q^i\ddot{q}^j$ of second order we need a generalisation of the Euler-Lagrange-equation: Performing a variation to second order yields:

$$\delta S = \int \mathrm{d}t\,\left(\frac{\partial\mathcal{L}}{\partial q}\delta q + \frac{\partial\mathcal{L}}{\partial\dot{q}}\delta\dot{q} + \frac{\partial\mathcal{L}}{\partial\ddot{q}}\delta\ddot{q}\right) = \int \mathrm{d}t\,\left(\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\mathcal{L}}{\partial\ddot{q}}\right)\delta q = 0 \quad (A.52)$$

with a single integration by parts for the second, and a double integration by parts in the third term. Then, Hamilton's principle defines the generalisation of the Euler-Lagrange-equation to higher orders:

$$\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\mathcal{L}}{\partial\ddot{q}} = 0 \quad (A.53)$$

A.10 *Beltrami-identity and the conservation of energy*

The conservation of energy in classical mechanics is realised very differently compared to other conservation laws: In those, one can identify cyclic variables q defined by the condition $\partial\mathcal{L}/\partial q = 0$, so that the Euler-Lagrange-equation makes sure that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = 0 \quad \text{and consequently, the canoncial momentum} \quad p = \frac{\partial\mathcal{L}}{\partial\dot{q}} \quad (A.54)$$

is conserved, $\mathrm{d}p/\mathrm{d}t = 0$. Time, however, is *not* a coordinate in classical mechanics, so the definition of energy as the canoncial momentum $\partial\mathcal{L}/\partial\dot{t}$ is *impossible*, it is completely unclear what $\dot{t}$ should actually be if not 1.

Instead, one needs the Beltrami-identity: By constructing

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{L} - \dot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = \dot{q}\frac{\partial\mathcal{L}}{\partial q} + \ddot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \ddot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \dot{q}\left(\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = 0 \quad (A.55)$$

because the Euler-Lagrange-equation sets the brackets to zero. Hence, there is a conserved quantity $\mathcal{H} = \dot{q}p - \mathcal{L}$, referred to as the Hamilton-function $\mathcal{H}$, which depends on the canonical momentum p and the coordinate q. This definition already suggests that $\mathcal{H}(p,q)$ is the Legendre-transform of $\mathcal{L}(q,\dot{q})$.

Let's investigate Ostrogradsky's idea that things become unstable if higher derivatives of q are included and write $\mathcal{L} = \mathcal{L}(q,\dot{q},\ddot{q},\dddot{q})$, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{L} - \dot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = \dot{q}\frac{\partial\mathcal{L}}{\partial q} + \ddot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} + \dddot{q}\frac{\partial\mathcal{L}}{\partial\ddot{q}} + \ldots - \ddot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \dot{q}\left(\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) + \dddot{q}\frac{\partial\mathcal{L}}{\partial\ddot{q}} + \ldots = \quad (A.56)$$

and subsituting the general Euler-Lagrange-equation one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{L} - \dot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = \dot{q}\left(-\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\mathcal{L}}{\partial\ddot{q}} + \frac{\mathrm{d}^3}{\mathrm{d}t^3}\frac{\partial\mathcal{L}}{\partial\dddot{q}} + \ldots\right) + \dddot{q}\frac{\partial\mathcal{L}}{\partial\ddot{q}} + \ldots \neq 0 \quad (A.57)$$

which would never work out to be zero.

The Hamilton-function $\mathcal{H}$ takes on a nice, directly interpretable form for the standard Lagrange-function $\mathcal{L} = m/2\,\dot{q}^2 - \Phi(q)$: The canonical momentum is $p = \partial\mathcal{L}/\partial\dot{q} = m\dot{q}$ and therefore, $\dot{q} = p/m$, yielding

$$\mathcal{H}(p,q) = p\dot{q} - \mathcal{L}(q,\dot{q}(p)) = \frac{p^2}{m} - \frac{p^2}{2m} + \Phi(q) = \frac{p^2}{2m} + \Phi(q) \quad (A.58)$$

Then, $\mathrm{d}\mathcal{H}/\mathrm{d}t = 0$ and $\mathcal{H}$ is conserved.

A.11 *Convexity of the Lagrange-function*

The variational principle relies heavily on the fact that the Lagrange-function $\mathcal{L}$ is a convex function in $\dot{q}$, and that the action S is a convex functional. Only then, there is a uniquely defined extremum and $\delta S = 0$ defines the actual equation of motion. Imagine if Hamilton's principle had multiple solutions for $\delta S = 0$! One would clearly end up in an impossible situation where multiple equations of motion would try to determine the evolution of a system.

Furthermore, the Hamilton-function is determined as the Legendre-transform of the Lagrange-function. For that to be feasible one needs the Lagrange function to be convex in $\dot{q}$: The canonical momentum $p = \partial\mathcal{L}/\partial\dot{q}$ is needed for replacing $\dot{q}$ with p, and therefore the relation $p(\dot{q})$ needs to be invertible to give $\dot{q}(p)$. Invertibility is given if $p = \partial\mathcal{L}/\partial\dot{q}$ is monotonically increasing and this is true if $\mathcal{L}$ is a convex function in $\dot{q}$. Then, $\mathcal{H}$ becomes a convex function in p, too, and the inverse Legendre-transform back to $\mathcal{L}$ is well-defined. Convexity (or concavity, any overall sign in $\mathcal{L}$ is undetermined because of affine invariance $\mathcal{L} \to a\mathcal{L} + b$ with $a = -1$ and $b = 0$) is surely given for a standard form $\mathcal{L} \propto \dot{q}^2$, but is there a more fundamental reason for it? The answer to this profound question is relativity:

A.12 *Lorentz- and Galilei-relativity*

Lagrange-mechanics is really a bit of physics that was discovered 100 years too early, as many aspect don't make much sense without relativity: $\mathcal{L} = T - V$ is not measurable and ad-hoc to result in Newton's equation of motion, and the covariance under Galilei-transforms and rotations is realised in very different ways. So, let's approach Lagrange-mechanics through relativity!

In the absence of potentials or curvature, spacetime should be homogeneous as no point or instance in time should play a particular role, and this homogeneity should be reflected in the transformation between different coordinate frames. An observer looking at two coordinate choices could measure the rate at which the coordinates x^μ and x'^μ containing the collection of spatial coordinates x^i and time t are drifting by as a function of her or his proper time τ, defining the velocity as the rate of change of the coordinates

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \text{const.}\ \text{, and identically in S}' : \frac{\mathrm{d}x'^\mu}{\mathrm{d}\tau} = \text{const.} \tag{A.59}$$

which is constant for inertial motion and suitably chosen coordinates, and the corresponding acceleration

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = 0,\ \text{and identically in S}' : \frac{\mathrm{d}^2x'^\mu}{\mathrm{d}\tau^2} = 0 \tag{A.60}$$

then vanishes in both systems. The relation between the two velocities and accelerations is given by

$$\frac{\mathrm{d}x'^\mu}{\mathrm{d}\tau} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \tag{A.61}$$

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\tau^2} + \frac{\partial^2 x'^\mu}{\partial x^\nu\partial x^\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}. \tag{A.62}$$

If the gradient of the Jacobian of the coordinate change vanishes,

$$\frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\rho} = 0 \tag{A.63}$$

vanishing accelerations in one frame imply vanishing accelerations in the other. Then, the rate of change of the coordinate passage $dx^\mu/d\tau$ is constant, and the transitions respects homogeneity. The solution for $x'^\mu(x^\nu)$ is consequently given by a linear, affine relation,

$$x'^\mu = A^\mu{}_\nu\, x^\nu + a^\mu. \tag{A.64}$$

with two sets of integration constants $A^\mu{}_\nu$ and a^μ.

Let's construct this transform from the most general transition between two frames, where we align for simplicity the coordinate axes with the direction of relative motion, taken to be the x-axis. There is an event with coordinates x^μ in S and x'^μ in S′, and the two frames move with a relative (constant) velocity υ. A linear, affine transform would then be the only one to respect the homogeneity of spacetime (nonlinear transforms would always single out certain spacetime points), so we make the ansatz:

$$x' = ax + bt, \text{ but } x = \upsilon t \text{ must imply } x' = 0 \tag{A.65}$$

$$x' = 0 = a\upsilon t + bt = (a\upsilon + b)\, t \;\Rightarrow\; b = -a\upsilon, \text{ and:} \tag{A.66}$$

$$x' = a(x - \upsilon t) \tag{A.67}$$

Reversing the roles of S and S′ implies that

$$x = ax' + bt' \text{ but } x' = -\upsilon t \text{ must imply } x = 0 \tag{A.68}$$

$$x = 0 = -a\upsilon t' + bt' = (-a\upsilon + b)\, t' \;\Rightarrow\; b = +a\upsilon, \text{ and:} \tag{A.69}$$

$$x = a(x' + \upsilon t') \tag{A.70}$$

But this relation between x and x' is not yet fixed without an additional assumption that determines the value of a. Here, Nature would have in fact a choice! Either, Nature could work with a universal time coordinate (or rather, a parameter, as it does not participate in transforms unlike the other coordinates). A universal time parameter would require that $t = t'$, which is the defining property of Galilei-transforms. Then,

$$x = a(x' + \upsilon t) = a(a(x - \upsilon t) + \upsilon t) = a^2 x + (1 - a)\, \upsilon t = x \tag{A.71}$$

which can only be realised if $a = 1$. Nature chose instead, for very good reasons, the speed of light to be equal in all frames, $c = c'$, which requires Lorentz- instead of Galilei-transforms between frames. In this choice,

$$x' = ct' = a(ct - \upsilon t) \tag{A.72}$$

$$x = ct = a(ct' - \upsilon t') \tag{A.73}$$

and consequently

$$c^2 tt' = a^2 (c - \upsilon)(c + \upsilon) \cdot tt', \tag{A.74}$$

where the third equation was obtained by multiplying the first two. Dividing by tt' and solving for a yields the Lorentz-factor γ,

$$a = \gamma = \frac{1}{\sqrt{1-\beta^2}}, \quad \text{with} \quad \beta = \frac{\upsilon}{c} \tag{A.75}$$

We should note that Lorentz-transformations, due to their linearity, do not 'mix' the spatial coordinates. The Lorentz-factor γ diverges at $\beta = 1$ and would indeed become imaginary for values $\beta > 1$. Taylor-expanding γ for small velocities β gives the result that

$$\gamma \sim 1 + \left.\frac{\partial^2\gamma}{\partial\beta^2}\right|_{\beta=0} \cdot \frac{\beta^2}{2} = 1 + \frac{\beta^2}{2}, \text{ with } \left.\frac{\partial\gamma}{\partial\beta}\right|_{\beta=0} = 1 \tag{A.76}$$

which is perfectly consistent with the fact that for low velocities $\beta \ll 1$ and $\gamma \simeq 1$, Lorentz- and Galilei-transforms are indistinguishable. Writing ct and arranging the temporal and spatial coordinates into a vector $x^\mu = \begin{pmatrix} ct \\ x \end{pmatrix}$ allows to use the standard matrix-form of the Lorentz-transformation:

$$x' = \gamma\,(x - \upsilon t) = \gamma\,(x - \beta c t) \tag{A.77}$$

$$ct' = \gamma\,(ct - \beta x), \tag{A.78}$$

so that one arrives at

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{A.79}$$

encapsulating the Lorentz-transform in a matrix $\Lambda^\mu{}_\nu$, with $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. Therefore, there are only two possible linear transformations, where Nature chooses to conserve the speed c, and we will see how this is related to the causal structure of spacetime and the hyperbolic evolution of field equations.

Just as rotations leave Euclidean scalar products $r^2 = \gamma_{ij} x^i x^j$ invariant, quadratic forms $s^2 = \eta_{\mu\nu} x^\mu x^\nu$ formed with the Minkowski-metric $\eta_{\mu\nu}$ are untouched by Lorentz-transforms, as one can see by direct computation:

$$s'^2 = (ct')^2 - x'^2 = \gamma^2 \left[(ct)^2 - 2ct\beta x + \beta^2 x^2 - x^2 + 2x\beta ct - \beta^2 (ct)^2\right] = \underbrace{\gamma^2(1-\beta^2)}_{1}\left((ct)^2 - x^2\right) = s^2 \tag{A.80}$$

With this realisation, one can define an orthogonality relation:

$$s'^2 = \eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta x^\alpha x^\beta = \eta_{\alpha\beta} x^\alpha x^\beta = s^2 \tag{A.81}$$

and therefore

$$\eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta} \tag{A.82}$$

as an orthogonality relation for $\Lambda^\mu{}_\alpha$. The physical interpretation of the invariant s is the proper time τ displayed on a comoving clock, i.e. a clock inside the local rest frame, where $x = 0$ constantly,

$$s^2 = \eta_{\mu\nu}x^\mu x^\nu = (ct)^2 - x^2 = c^2\tau^2 \tag{A.83}$$

Differentially, this implies

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = c^2\mathrm{d}\tau^2 = c^2\mathrm{d}t^2 - \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{A.84}$$

and therefore

$$\mathrm{d}\tau = \sqrt{1 - \gamma_{ij}\frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}}\mathrm{d}t = \sqrt{1-\beta^2}\mathrm{d}t = \frac{\mathrm{d}t}{\gamma} \tag{A.85}$$

with $\beta^2 = \gamma_{ij}\upsilon^i\upsilon^j/c^2$. As $\gamma \geq 1$ always, $\mathrm{d}\tau < \mathrm{d}t$ and one observes a relativistic dilation of proper time relative to the coordinate time. $\pm 1/\gamma = \pm\sqrt{1-\beta^2}$ is in fact a semi-circle, so it's a perfectly convex (concave for the other sign choice) function, and would make an excellent candidate for the Lagrange-function. Additionally, its Taylor-expansion

$$\frac{1}{\gamma} = 1 - \frac{\beta^2}{2} \pm \ldots \tag{A.86}$$

at low velocities $\beta \ll 1$ gives a term that is reminiscent of the classical kinetic energy!

A.13 *Rapidity*

Lorentz-boosts form a group as their combination always gives a boost, but clearly the velocity is not additive, which can be verified by direct combination of two boosts: The sum of the velocities β always needs to stay below 1. One might wonder then whether there is an additive parameter for the Lorentz-transforms which replaces the velocity β. From the range of values of the terms in the Lorentz-transform $\Lambda^\mu{}_\alpha$, where $0 \leq \beta\gamma < \infty$ (for positive β with a sign change for negative β!) and $1 \leq \gamma < \infty$ on could think of a parameterisation $\cosh\psi = \gamma$ and $\sinh\psi = \beta\gamma$, such that

$$\tanh\psi = \beta \quad \rightarrow \quad \psi = \mathrm{artanh}\beta \tag{A.87}$$

with the rapidity ψ replacing the physical velocity β. Then, the Lorentz-transform can be written as a hyperbolic rotation,

$$\Lambda^\mu{}_\alpha = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \tag{A.88}$$

where the Lorentz-invariance of $s^2 = \eta_{\mu\nu}x^\mu x^\nu$ is then re-expressed as the property $\cosh^2\psi - \sinh^2\psi = \gamma^2 - (\beta\gamma)^2 = \gamma^2(1-\beta^2) = 1$ of the hyperbolic functions. Two successive boosts in this representation then shows that rapidity is additive as a parameter, which we'll elaborate in the chapter about Lie-groups. For the time being, we use the addition theorem of the hyperbolic tangent to get

$$\beta_{\phi+\psi} = \tanh(\phi+\psi) = \frac{\tanh\phi + \tanh\psi}{1 + \tanh\phi\tanh\psi} = \frac{\beta_\phi + \beta_\psi}{1 + \beta_\phi\beta_\psi} \tag{A.89}$$

which falls back onto $\beta_\phi + \beta_\psi$ for small $\beta \ll 1$, where the hyperbolic tangent, at the same time, is approximated well by its argument, $\beta = \tanh\psi \simeq \psi$. As $\tanh\psi$ is bounded by 1 even as $\psi \to \infty$, successive boosts do not exceed $\beta = 1$ or $\upsilon = c$.

A.14 *Proper time and the relativistic Lagrange-function*

The rate of change of coordinates $\mathrm{d}x^\mu$ can be measured in units of proper time $\mathrm{d}\tau$, leading to the definition of 4-velocity,

$$x^\mu = \begin{pmatrix} ct \\ x^i \end{pmatrix} \quad \text{implies that} \quad u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma \begin{pmatrix} c \\ \upsilon^i \end{pmatrix} \tag{A.90}$$

with $\gamma = \mathrm{d}t/\mathrm{d}\tau$ and $\upsilon^i = \mathrm{d}x^i/\mathrm{d}t$. If the proper time τ is used to parameterise the trajectory $x^\mu(\tau)$ in this way, the tangent vector $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$ is normalised to c,

$$\eta_{\mu\nu}u^\mu u^\nu = \gamma^2(c^2 - \upsilon^2) = c^2\gamma^2(1-\beta^2) = c^2 \tag{A.91}$$

which is true even for a particle at rest: There, only $u^t = c$ is nonzero, $\upsilon^i = 0$ and $\eta_{\mu\nu}u^\mu u^\nu = c^2$, effectively, the particle drifts along the ct-axis at velocity c. Motivated by the idea that $1/\gamma$ could be a good candidate for the relativistic Lagrange-function, we could imagine that the arc-length $S = \int \mathrm{d}s = c\tau$ through spacetime of a trajectory $x^\mu(\tau)$ could be extremised, as a very intuitive concept:

$$S = -mc\int \mathrm{d}s = -mc^2\int \mathrm{d}\tau = -mc^2\int \mathrm{d}t\,\sqrt{1-\beta^2} = -mc^2\int \frac{\mathrm{d}t}{\gamma}, \tag{A.92}$$

therefore $\mathcal{L} = -mc^2/\gamma$, with the nonrelativistic limit $-mc^2(1-\beta^2)$, i.e. up to an affine transform the actual kinetic energy $m\upsilon^2/2$.

A.15 *Geometric view on motion*

If the relativistic arc-length through spacetime as a candidate for the Lagrange-function were true, force-free motion should proceed along a straight line, as a reflection of the law of inertia: Technically, we replace the variation of the abstract classical action by the much more concrete variation of an arc through spacetime and monitor how the length $S = \int \mathrm{d}s$ would change under a variation,

$$\delta S = -mc^2\,\delta\int \mathrm{d}\tau = -mc^2\int \frac{\eta_{\mu\nu}}{2\mathrm{d}\tau}\left[\mathrm{d}x^\mu\delta\mathrm{d}x^\nu + \delta\mathrm{d}x^\mu\mathrm{d}x^\nu\right] =$$

$$-mc^2\int \eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\underbrace{\delta\mathrm{d}x^\nu}_{=\mathrm{d}\delta x^\mu} = mc^2\int \mathrm{d}\tau\,\eta_{\mu\nu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}\delta x^\nu \tag{A.93}$$

starting from

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \quad \to \quad c\mathrm{d}\tau = \sqrt{\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} \tag{A.94}$$

using that the Minkowski-metric is symmetric, $\eta_{\mu\nu} = \eta_{\nu\mu}$ and finally that the differential can be expanded as

$$\mathrm{d}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\mathrm{d}\tau = \frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}\mathrm{d}\tau \tag{A.95}$$

after reshuffling the differentiations by an integration by parts. Hamilton's principle $\delta S = 0$ therefore implies that

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = 0 \quad \rightarrow \quad x^\mu(\tau) = a^\mu\tau + b^\mu \tag{A.96}$$

i.e. a straight line through spacetime, with two integration constants a^μ and b^μ. These trajectories result from minimising the arc-length, which is a convex functional and bounded by $S = 0$ on the light cone. Affine invariance of the arc-length just means that you're free to measure it as proper time with any unit and from any zero-point.

With the law of inertia explained, could the formalism be adopted to (gravitational) potentials? There, we would indeed expect accelerations $\mathrm{d}^2x^\mu/\mathrm{d}\tau^2$ as a consequence of gradients in Φ. General relativity is really much beyond the scope of the tooltips-lecture but let's try this idea out with a weakly perturbed Minkowskian spacetime. There, the line element is given by

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \left(1+2\frac{\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1-2\frac{\Phi}{c^2}\right)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{A.97}$$

with the metric $g_{\mu\nu}$ instead of the Minkowski-metric $\eta_{\mu\nu}$, and a weak gravitational potential Φ with $|\Phi| \ll c^2$ (Already at this point you can see that without a speed of light we could not say whether the potential is weak or strong!) on top of a Minkowski-vector space (I'm a bit adamant here, because just from $g_{\mu\nu} \neq \eta_{\mu\nu}$ you can *not* infer the existence of gravitational potentials if there is full freedom in choosing the coordinates.) The first thing we should check if there is an influence of the gravitational potential on the passage of time: After all, proper time should differ from coordinate time which is first of all caused by special relativistic time dilation due to motion. For a stationary object $\mathrm{d}x^i = 0$ because there is no change in coordinate, and we get

$$\mathrm{d}\tau = \sqrt{1+2\frac{\Phi}{c^2}}\mathrm{d}t \simeq \left(1+\frac{\Phi}{c^2}\right)\mathrm{d}t = \left(1-\frac{GM}{c^2r}\right)\mathrm{d}t \tag{A.98}$$

by substitution of a Newtonian potential $\Phi = -GM/r$, where $2GM/c^2 = r_S$ defines the Schwarzschild-radius: It seems to be the case that $G/c^2 \simeq 10^{-28}$m/kg assigns a length-scale to a mass, and for an object like the Sun with $M = 10^{30}$kg the Schwarzschild-radius comes out with a few hundred meters. As $\mathrm{d}\tau \leq \mathrm{d}t$, we observe a gravitational time dilation of proper time relative to coordinate time, and Φ seems to have an influence on the relativistic arc-length just as velocity would. Repeating the above derivation with $\mathrm{d}s = c\mathrm{d}\tau = \sqrt{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu}$ shows

$$S = -mc^2\int\mathrm{d}\tau = -mc\int\mathrm{d}t\,\sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}} = -mc\int\mathrm{d}t\,\sqrt{\left(1+2\frac{\Phi}{c^2}\right)c^2 - \left(1-2\frac{\Phi}{c^2}\right)\upsilon^2} \tag{A.99}$$

Ignoring $\Phi\upsilon$ as a higher-order term then yields after approximating $\sqrt{1+x} = 1 + x/2$

$$S = -mc^2 \int \mathrm{d}t \sqrt{1 + 2\frac{\Phi}{c^2} - \frac{\upsilon^2}{c^2}} \simeq \int \mathrm{d}t \left(\frac{m\upsilon^2}{2} - m\Phi - mc^2\right) \tag{A.100}$$

i.e. the classical Lagrange-function $\mathcal{L} = T - \Phi$ with the rest mass as an additional term, which is irrelevant due to affine invariance of $\mathcal{L}$. Weirdly enough, motion in classical mechanics proceeds, in the aim to minimise S, in a way that time dilation is extremised, by being far down in gravitational potentials or by being fast.

A.16 *Relativistic energy and momentum*

We have seen in the last chapter that the Lagrange-function is much more a statement of causal motion in spacetime and has little to do with energies: Those appear after Legendre-transform, which is always well defined because the Lagrange function is a convex functional in $\dot{x}$ - this is, incidentally, the same reason why the variation yields a unique result and finds a unique extremum. It is important to realise that the conservation of the various canonical momenta are ensured by cyclic coordinates, but that energy conservation is a consequence of the Beltrami-identity.

The Legendre-transform of the relativistic Lagrange-function should provide a relativistic dispersion relation, i.e. a relation between energy and momentum. The canonical momentum p is derived from the Lagrange-function $\mathcal{L} = -1/\gamma = -\sqrt{c^2 - \dot{x}^2}$,

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{c^2 - \dot{x}^2}} \tag{A.101}$$

where this relation can be inverted from $p(\dot{x})$ to $\dot{x}(p)$, as a consequence of the convexity of $\mathcal{L}$, and which is needed for computing the Legendre-transform

$$\mathcal{H}(p, x) = p\dot{x}(p) - \mathcal{L}(x, \dot{x}(p)) \tag{A.102}$$

With the inversion

$$p^2 = \frac{\upsilon^2}{c^2 - \upsilon^2} \quad \rightarrow \quad p^2(c^2 - \upsilon^2) = \upsilon^2 \quad \rightarrow \quad p^2c^2 = \upsilon^2(1 + p^2) \quad \rightarrow \quad \upsilon = \frac{cp}{\sqrt{1 + p^2}} \tag{A.103}$$

with $\upsilon = \dot{x}$ and consequently

$$\mathcal{H}(p, x) = \upsilon \underbrace{\frac{\upsilon}{\sqrt{c^2 - \upsilon^2}}}_{p} + \underbrace{\sqrt{c^2 - \upsilon^2}}_{\frac{\upsilon}{p}} = \upsilon p + \frac{\upsilon}{p} = \upsilon\left(p + \frac{1}{p}\right) = \frac{cp}{\sqrt{1 + p^2}}\frac{1 + p^2}{p} = c\sqrt{1 + p^2} \tag{A.104}$$

Including the rest mass m would yield the relativistic dispersion relation

$$\mathcal{H} = \sqrt{(mc^2)^2 + (cp)^2} \tag{A.105}$$

which is approximated by $\mathcal{H} = mc^2 + p^2/(2m)$ for $p \ll mc$ and by $\mathcal{H} = cp$ for $p \gg mc$. The Hamilton-function $\mathcal{H}$ in turn is again convex and allows an inverse Legendre-transform to recover $\mathcal{L}$.

A.17 *Causality and light cones*

The Lorentz-invariant $\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ makes it possible to differentiate between time-like $\mathrm{d}s^2 > 0$, space-like $\mathrm{d}s^2 < 0$ and light-like $\mathrm{d}s^2 = 0$ separations in spacetime. Weirdly enough, the causal ordering of events, i.e. a statement of $\mathrm{d}t > 0$ or $\mathrm{d}t < 0$ depends on the chosen frame, and it is always possible to change the sign of $\mathrm{d}t$ in space-like separated events from the point of view of a fast enough moving Lorentz-frame.

Only inside the light cone, i.e. for all time-like separated events one observes the same causal order from all Lorentz-frames, which necessarily move at speeds $\beta < 1$. Being located inside the light-cone is certainly true for all massive particles: Their 4-velocities u^μ are normalised to $\eta_{\mu\nu}u^\mu u^\nu = c^2 > 0$, and in fact the light cone is the convex hull of all possible trajectories of massive particles. Approaching $\beta = 1$ has the Lorentz-factor γ diverge, which is often rewritten as a relativistic mass increase,

$$p^\mu = m u^\mu = \gamma m \begin{pmatrix} c \\ \upsilon^i \end{pmatrix} \quad \text{from} \quad u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma \begin{pmatrix} c \\ \upsilon^i \end{pmatrix} \tag{A.106}$$

and therefore, relativistic mass increase is purely a consequence of the dilation of proper time. Asking whether it would be possible to accelerate a (charged) particle with electromagnetic forces will be answered negatively: Starting with the Lorentz-equation

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{q}{m}\mathrm{F}^{\mu\nu}u_\nu \tag{A.107}$$

with the field-tensor $\mathrm{F}^{\mu\nu}$ that contains the electric and magnetic fields. Multiplying the Lorentz-equation with u_μ gives:

$$u_\mu\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}(u_\mu u^\mu) = \frac{q}{m}\mathrm{F}^{\mu\nu}u_\mu u_\nu = 0 \tag{A.108}$$

such that the normalisation of u^μ is conserved to be $c^2 > 0$, and u^μ remains time-like and inside the light cone despite being accelerated: The reason is purely geometrical, as the contraction of the antisymmetric tensor $\mathrm{F}^{\mu\nu}$ with the symmetric tensor $u_\mu u_\nu$ is necessarily zero, making it impossible for $u_\mu u^\mu$ to change.

At this point, we should start to be careful not to link the Lorentz-geometry to any particular coordinate choice. When considering light cone coordinates, $\mathrm{d}u = \mathrm{c}\mathrm{d}t + \mathrm{d}x$ and $\mathrm{d}v = \mathrm{c}\mathrm{d}t - \mathrm{d}x$ the line element is given by

$$\mathrm{d}s^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\mathrm{d}x^\nu = \mathrm{c}^2\mathrm{d}t^2 - \mathrm{d}x^2 = (\mathrm{c}\mathrm{d}t + \mathrm{d}x)(\mathrm{c}\mathrm{d}t - \mathrm{d}x) = \mathrm{d}u \cdot \mathrm{d}v, \tag{A.109}$$

and the corresponding Lorentzian metric is represented by the matrix

$$\eta_{\mu\nu} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{A.110}$$

in these coordinates. Surely, the geometry is identical and has not been changed by the new definition of coordinates, and the spectrum of eigenvalues of the new metric is exactly +1 and −1.