D MICROCANONICAL ENSEMBLE

Any system left unperturbed tends towards thermal equilibrium with a defined temperature T, which, along with the other state variables that characterise the system. The internal thermal energy is reflected by the temperature and the heat capacity and is stored in the system in its microscopic degrees of freedom. In thermal equilibrium a continuous reshuffling of energy between all degrees of freedom is taking place, and these microscopical degrees of freedom follow energy-conserving Hamiltonian equations of motion.

Boltzmann was the first to realise that in thermal equilibrium all states compatible with the same energy are equally likely, as stated by the fundamental postulate of statistical mechanics, and that states of higher energy are less likely to be assumed, where the ratio of probabilities is given by the Boltzmann factor: A justification of these two properties of systems in thermodynamic equilibrium would be axiomatic in the sense of a fundamental postulate, or, it would in fact follow from the notion that the probability distributions have to maximise the Shannon-information entropy as a measure of randomness. Effectively, we will assume that in thermal equilibrium the randomness of how the states are occupied, is as random as possible, with a maximised Shannon-entropy. It is important to realise that statistical mechanics as the microscopic theory behind thermodynamics, is perfectly energy-conserving, despite the fact that it deals ultimately with thermal energy which is often a byproduct of dissipative processes: We can reconcile these two ideas: A mechanical system can be dissipative on large scales with e.g. kinetic energy being lost due to friction, but that the true microscopic degrees of freedom follow energy-conserving Hamiltonian dynamics. The second law of thermodynamics is an expression of the fact that it is incredibly improbable (but not) for all microscopic degrees of freedom to conspire and generate macroscopic motion out of a freak thermal fluctuation. In some sense, thermodynamics becomes then an effective theory for Hamiltonian systems with many (coupled) degrees of freedom.[1]

Hamiltonian systems with conserved energies are a consequence of Lagrange-functions that do not explicitly depend on time. In classical mechanics, the Lagrange-function $\mathcal{L}$ defines the action S

$$S = \int_{t_i}^{t_f} \mathrm{d}t\ \mathcal{L}(q, \dot{q}) \tag{D.112}$$

which for the mechanics of a particle is actually the non-relativistic limit of the arc-length of the trajectory through spacetime, measured in terms of proper time:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - dx^2 = c^2\mathrm{d}\tau^2 \quad \rightarrow \quad \mathrm{d}\tau = \sqrt{1 - \left(\frac{1}{c}\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2}\mathrm{d}t \quad \rightarrow \quad \mathrm{d}\tau = \frac{1}{\gamma}\mathrm{d}t \tag{D.113}$$

such that the arc length $S = \int \mathrm{d}s = c\int \mathrm{d}\tau = c\int \mathrm{d}t\,/\gamma$ with the Lorentz-factor γ.

---

[1] L. Boltzmann really saw the solution of equally probable states as an entropy-maximising distribution in his kinetic theory of gases in thermal equilibrium, but the entropy functional was only axiomatically defined by C. Shannon much later. Later still, the notion that thermodynamics is in reality a theory of information of physical systems was formulated by E.T. Jaynes. But at the time of the French Revolution, P.-S. Laplace formulated already for pure, non-physical probabilities the principle of indifference which states that without any specific knowledge, all outcomes of a random experiment must be considered equally probable.

As $1/\gamma \rightarrow 1-(\upsilon/c)/2$ for small velocities $\upsilon = \mathrm{d}x/\mathrm{d}t \ll c$, one recovers the classical Lagrange-function, where overall pre-factors and signs do not matter in the Euler-Lagrange equation.

⚠ *The variational principle is invariant under affine transform $\mathcal{L} \rightarrow a\mathcal{L}+b$ for the Lagrange-function.*

Hamilton's principle now stipulates that the physical trajectory taken by the system corresponds to an extremum of the action, i.e. $\delta S = 0$

$$\delta S = \int_{t_i}^{t_f} \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\right) = \int_{t_i}^{t_f} \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right)\delta q \tag{D.114}$$

with an integration by parts on $\delta\dot{q} = \mathrm{d}\delta q/\mathrm{d}t$ and a successive integration by parts, where the boundary terms do not contribute as the Lagrange-function is kept fixed at the boundary: This rearrangement gives the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} + \frac{\partial \mathcal{L}}{\partial q} = 0 \tag{D.115}$$

as a tool serving the purpose of isolating the action-extremising trajectory $q(t)$ as a differential equation.

The canonical momentum of the coordinate q is given by the relation

$$p = \frac{\partial \mathcal{L}}{\partial \dot{q}} \tag{D.116}$$

and helps to identify conserved quantities: If a coordinate is cyclic, $\mathcal{L}$ does not depend on it directly (and only on its derivative), so that $\partial\mathcal{L}/\partial q = 0$. Then, the Euler-Lagrange equation has the canonical momentum conserved in time evolution,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\mathrm{d}p}{\mathrm{d}t} = 0 \tag{D.117}$$

The conservation of energy in classical mechanics works in a different way, though, because the time t plays the role of a parameter describing motion, not that of a coordinate.

Typical Lagrange-functions in classical mechanics have the form

$$\mathcal{L} = \frac{m}{2}\dot{q}^2 - \Phi(q) \tag{D.118}$$

for a particle moving in a potential Φ, such that $\delta S = 0$ implies the Newtonian equation of motion $m\ddot{q} = -\mathrm{d}\Phi/\mathrm{d}q$. But there might be other systems that are perfectly amenable to a Lagrangian description, for instance an LC-circuit in electrical engineering:

$$\mathcal{L} = \frac{L}{2}\dot{Q}^2 - \frac{1}{2C}Q^2 \quad\rightarrow\quad \ddot{Q} + \frac{1}{LC}Q = 0 \tag{D.119}$$

with the charge Q and the current $\dot{Q}$: Recognising the Lagrange-function of a harmonic oscillator immediately suggests the angular frequency $\omega = 1/\sqrt{LC}$ and the corresponding equation of motion. Very often the kinetic and potential terms of the Lagrange-function are separate but that is not a necessity.

If $\partial\mathcal{L}/\partial\dot{q}$ still depends on q the differentiation $\mathrm{d}/\mathrm{d}t$ in the Euler-Lagrange equation will read:

$$\ddot{q}\frac{\partial^2\mathcal{L}}{\partial\dot{q}^2} + \dot{q}\frac{\partial^2\mathcal{L}}{\partial q\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 \tag{D.120}$$

While in one dimension the only requirement would now be that $\partial^2\mathcal{L}/\partial\dot{q}^2$ is non-zero (and it usually is, as it corresponds to the mass of the system), in more than one dimension $\partial^2\mathcal{L}/\partial\dot{q}^i/\partial\dot{q}^j$ would need to be an invertible matrix for isolating $\ddot{q}$.

D.1 *Virial theorem*

Is kinetic or potential energy the preferred form of energy of a system? Clearly, a physical system follow its dynamical equation and continuously reshuffles energy from one energy form to another, but taking averages over time, the system might preferentially be in a state where one energy form dominates over the other. To show this might be the case, let's start at the Euler-Lagrange-equation for any system,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 \tag{D.121}$$

and it with the coordinate q. Then, using the Leibnitz-rule,

$$q\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - q\frac{\partial\mathcal{L}}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t}\left(q\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) - \dot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} - q\frac{\partial\mathcal{L}}{\partial q} = 0 \tag{D.122}$$

We can break into this relation by individual inspection of the three terms:

1. time everage of the Lagrange-function $\mathcal{L}$ over $\Delta t \gg$ many dynamical time scales

$$\frac{1}{\Delta t}\int_0^{\Delta t}\mathrm{d}t\,\frac{\mathrm{d}}{\mathrm{d}t}\left(q\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = \frac{1}{\Delta t}\left(q\frac{\partial\mathcal{L}}{\partial\dot{q}}\right)\Bigg|_0^{\Delta t} = \frac{1}{\Delta t}\left(qp\right)\Big|_0^{\Delta t} \leq \frac{1}{\Delta t}(q_{\max}\cdot p_{\max}) \tag{D.123}$$

   where the result of the integration is estimated to be less than the product of the largest values for p and q that the integrand assumes between 0 and Δt. Clearly, this requires that the integrand and therefore the range of motion is bounded in p and q. Taking the limit $\Delta t \to \infty$, makes the term disappear, as Δt grows without bounds and $q_{\max}\cdot p_{\max}$ is finite.

2. $\mathcal{L}_{\mathrm{kin}}$ is homogeneous of degree 2 in $\dot{q}$ for non-relativistic mechanics:

$$\mathcal{L}_{\mathrm{kin}} = \mathrm{T} = \frac{m}{2}\dot{q}^2 \quad\to\quad \dot{q}\frac{\partial\mathcal{L}_{\mathrm{kin}}}{\partial\dot{q}} = 2\mathcal{L}_{\mathrm{kin}} = 2\mathrm{T} \tag{D.124}$$

3. $\mathcal{L}$ is homogeneous of degree k in q for power-law potentials $\Phi \propto q^k$

$$\mathcal{L}_{\mathrm{pot}} = \Phi = q^k \quad\to\quad q\frac{\partial\mathcal{L}_{\mathrm{pot}}}{\partial q} = k\mathcal{L}_{\mathrm{pot}} = k\Phi \tag{D.125}$$

Therefore, we obtain the relation

$$2\langle \mathrm{T} \rangle = k\langle \Phi \rangle \tag{D.126}$$

between the time-averages of the kinetic and potential energies, with the definition of the average kinetic energy $\langle \mathrm{T} \rangle$,

$$\lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_0^{\Delta t} \mathrm{d}t\ \mathrm{T}(t) = \langle \mathrm{T} \rangle \tag{D.127}$$

with an analogous average for $\langle \Phi \rangle$.

The harmonic oscillator has a potential $\Phi \propto q^2$ with the potential being a harmonic function of degree $k = 2$, and consequently, $\langle \mathrm{T} \rangle = \langle \Phi \rangle$, and both forms of energy are on average equal, which is a rather obvious result. After all, $q \propto \exp(\mathrm{i}\omega t)$ and $\dot{q} \propto \mathrm{i}\omega \exp(\mathrm{i}\omega t)$, such that the averages $\omega^2\langle q^2 \rangle$ and $\langle \dot{q}^2 \rangle$ are necessarily equal. For a Coulomb-potential we have $\Phi \propto 1/q$, and therefore $k = -1$ and $2\langle \mathrm{T} \rangle = -\langle \Phi \rangle$, which is fine, because Φ is negative. It is illustrative to increase k to high positive values: Then, the kinetic energy becomes on average dominant over the potential energy, $\langle \mathrm{T} \rangle = k/2\ \langle \Phi \rangle$. For high positive k one obtains a flat potential with very steep walls, where a particle zooms around in a state of high kinetic energy most of the time and spends only little time being deflected at the walls where the potential energy is high. A fun idea is the impossibility of a gravitationally bound ball of photons: There, the kinetic energy for ultra-relativsitic particles is a homogeneous function of order $k = 1$ such that the virial theorem would become $\langle \mathrm{T} \rangle = -\langle \Phi \rangle$, implying that the total energy is zero - but it would have to be negative for a bound system!

In summary, the virial theorem makes a statement about the average kinetic and potential energies in the course of the time evolution and is perfectly valid for a system with a single degree of freedom, as long as the motion is bounded in position and momentum. Often it is the case, thought, that systems consist of many particles and one might wonder if the average of kinetic and potential energy could be determined instead at a fixed time over the many particles, whether these ensemble averages are identical to the temporal averages, and whether the virial theorem applies to ensemble averages as well. Ergodic systems have in fact this property, although ergodicity is in particular systems difficult to demonstrate.

Fig. 8 shows the kinetic and potential energies for the motion of a particle inside a potential of the type $\Phi \propto -1/r^\alpha$ with $\alpha = 1$ and $\alpha = 0.9$. The total energy is negative, indicating that the particle is bound. The two energies clearly are conserved and average to some fixed value, whose ratio corresponds to the power law exponent of the potential: $-\alpha/2$.

Similarly, Fig. 8 illustrates the same virial relation for anharmonic potentials $\Phi \propto x^{2\alpha}$ with $\alpha = 3$ and $\alpha = 6$. With increasing α, the phases of high kinetic energy become longer and phases of high potential energy, when the particle changes its direction of motion, shorter. This is directly enforced by the flatter potential wells. In consequence, $\langle \mathrm{T} \rangle$ increases while $\langle \Phi \rangle$ decreases, such that $2\langle \mathrm{T} \rangle = \alpha\langle \Phi \rangle$ is fulfilled.

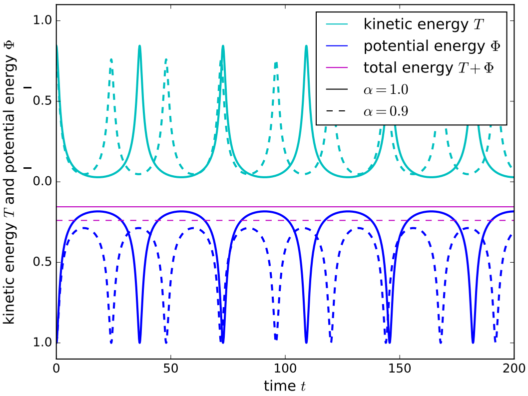

Figure 7: Kinetic energy T and potential energy Φ in Kepler-type potentials $\Phi \propto r^{-\alpha}$.

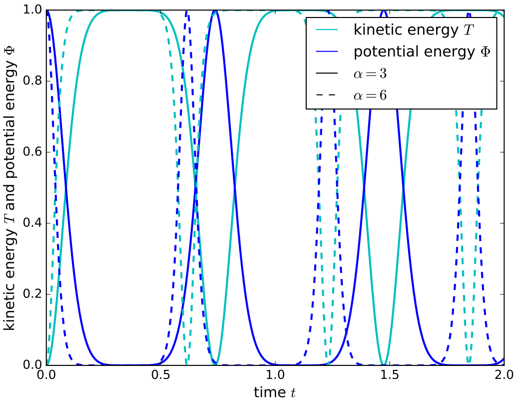

Figure 8: Kinetic energy T and potential energy Φ in anharmonic potentials of the type $\Phi \propto q^{2\alpha}$.

D.2 *Energy conservation*

Energy conservation is, from an arithmetic point of view, very similar: Instead of q we multiply the Euler-Lagrange-equation with $\dot{q}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 \tag{D.128}$$

to arrive at

$$\dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \dot{q}\frac{\partial\mathcal{L}}{\partial q} = m\dot{q}\ddot{q} + \dot{q}\frac{\mathrm{d}\Phi}{\mathrm{d}q} = \frac{m}{2}\frac{\mathrm{d}}{\mathrm{d}t}\dot{q}^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Phi(q) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m}{2}\dot{q}^2 + \Phi(q)\right) = 0 \tag{D.129}$$

with the replacements $\partial\mathcal{L}/\partial\dot{q} = m$ and $\partial\mathcal{L}/\partial q = -\Phi$ in the first step, followed by the time derivative $\dot{q}\mathrm{d}\Phi/\mathrm{d}q = \mathrm{d}\Phi/\mathrm{d}t$, suggested by the chain rule. The quantity $\mathcal{H}$

$$\mathcal{H}(q,\dot{q}) = \frac{m}{2}\dot{q}^2 + \Phi(q) \tag{D.130}$$

is the Hamilton-function of the system, which comes out as conserved, $\mathrm{d}\mathcal{H}/\mathrm{d}t = 0$.

D.3 *Legendre transforms and the Hamilton-function*

To be quite exact, $\mathcal{H}(q,\dot{q})$ is not yet the energy, because it is assumed to depend on $\dot{q}$ but not yet on the canonical momentum,

$$p = \frac{\partial\mathcal{L}}{\partial\dot{q}} \tag{D.131}$$

The Legendre-transform of the Lagrange-function $\mathcal{L}$

$$\mathcal{H}(q,p) = p\dot{q}(p) - \mathcal{L}(q,\dot{q}(p)) \tag{D.132}$$

which replaces the variable $\dot{q}$ by the variable p, can be shown to be conserved using this argument:

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \dot{p}\dot{q} + p\ddot{q} - \dot{q}\frac{\partial\mathcal{L}}{\partial q} - \ddot{q}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \frac{\partial\mathcal{L}}{\partial q}\dot{q} - \dot{q}\frac{\partial\mathcal{L}}{\partial q} = 0 \tag{D.133}$$

which is known as the Beltrami-identity. In the first step, we used the derivative of the canonical momentum,

$$\dot{p} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \frac{\partial\mathcal{L}}{\partial q} \tag{D.134}$$

together with the Euler-Lagrange equation. With the definition of the canonical momentum as the derivative $p = \partial\mathcal{L}/\partial\dot{q}$ one needs an invertible expression to be able to write down $\dot{q}(p)$ from $p(\dot{q})$, which is made sure by the convexity of the functional $\mathcal{L}$. Then, the Legendre-transform $\mathcal{H}$ is likewise convex, ensuring the existence of the inverse Legendre transform.

I am a bit picky to call only $p^2/(2m)$ kinetic energy and not $m\dot{q}^2/2$, for a very specific reason, even though the Legendre transform of a parabola $\dot{q}^2$ is of course a parabola p^2: The classical Lagrange-function is the non-relativistic limit of the arc-length of a spacetime-trajectory and therefore a geometric quantity, and acquires the interpretation of energy only after Legendre-transformation.

D.4 *Hamilton-equations of motion*

Instead of the Euler-Lagrange equations, which naturally lead to second-order equations of motions, one can write down equivalent coupled first-order equations of motions operating on the Hamilton-function $\mathcal{H}$ instead of the Lagrange-function $\mathcal{L}$.

$$\frac{\partial\mathcal{H}}{\partial p} = \dot{q} + p\frac{\partial\dot{q}}{\partial p} - \frac{\partial\dot{q}}{\partial p}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \dot{q} \quad\text{and}\quad \frac{\partial\mathcal{H}}{\partial q} = p\frac{\partial\dot{q}}{\partial q} - \frac{\partial\mathcal{L}}{\partial q} - \frac{\partial\dot{q}}{\partial q}\frac{\partial\mathcal{L}}{\partial\dot{q}} = -\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = -\dot{p} \tag{D.135}$$

The form of Hamilton's equation of motion allow an incredibly interesting notation, in particular for the harmonic oscillator, where $\mathcal{H} = p^2/2 + \omega^2 q^2/2$: Then, $\partial\mathcal{H}/\partial p = p = \dot{q}$ and $\partial\mathcal{H}/\partial q = \omega^2 q = -\dot{p}$, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & -\omega^2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} \tag{D.136}$$

which is the archetypical form of a symplectic differential equation. The solution to the system can be written in terms of a matrix exponential,

$$\begin{pmatrix} p \\ q \end{pmatrix} = \exp\left(\begin{pmatrix} 0 & -\omega^2 \\ 1 & 0 \end{pmatrix} t\right)\begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \tag{D.137}$$

applied to the initial conditions. Summing up the exponential series with this so-called symplectic matrix yields the familiar sine and cosine functions, effectively constituting a time evolution operator.

The question on energy conservation is answered in view of Hamilton's equation of motion in a completely different way: Forming the derivative of $\mathcal{H}$ which only depends on p and q (and not on t directly) yields

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial q}\dot{q} + \frac{\partial\mathcal{H}}{\partial p}\dot{p} = \frac{\partial\mathcal{H}}{\partial q}\frac{\partial\mathcal{H}}{\partial p} - \frac{\partial\mathcal{H}}{\partial p}\frac{\partial\mathcal{H}}{\partial q} = 0 = \{\mathcal{H}, \mathcal{H}\} \tag{D.138}$$

with the definition of Poisson-brackets,

$$\frac{\partial\mathcal{A}}{\partial q}\frac{\partial\mathcal{B}}{\partial p} - \frac{\partial\mathcal{B}}{\partial q}\frac{\partial\mathcal{A}}{\partial p} = \{\mathcal{A}, \mathcal{B}\} \tag{D.139}$$

for any two functions A and B dependent on the canonical momenta and coordinates. Then, the time evolution of any of such functions is determined by the Poisson equation of motion,

$$\frac{\mathrm{d}\mathrm{A}}{\mathrm{d}t} = \frac{\partial\mathrm{A}}{\partial t} + \dot{q}\frac{\partial\mathrm{A}}{\partial q} + \dot{p}\frac{\partial\mathrm{A}}{\partial p} = \frac{\partial\mathrm{A}}{\partial t} + \{\mathcal{H}, \mathcal{A}\} \tag{D.140}$$

Alternatively, the Lagrange-function can be thought to depend directly on (q, p) instead of $(q, \dot{q})$. From the definition of the Legendre-transformation

$$\mathcal{L}(p, q) = p\dot{q}(p) - \mathcal{H}(p, q) \tag{D.141}$$

one obtains from the two independent Euler-Lagrange equations for the coordinates p and q directly

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 = \dot{p} + \frac{\partial\mathcal{H}}{\partial q} \tag{D.142}$$

as well as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{p}} - \frac{\partial\mathcal{L}}{\partial p} = 0 = -\dot{q} + \frac{\partial\mathcal{H}}{\partial p} \tag{D.143}$$

i.e. again the two Hamilton equations of motion.

D.5 *Canonical transforms*

Within Hamiltonian mechanics there exists a class of coordinate transitions with covariant transformation of the equations of motion: canonical transforms. Effectively, the states of a mechanical system form a manifold that is parameterised by coordinates p and q, and transitions to new coordinates $\mathrm{Q} = \mathrm{Q}(p,q)$, $\mathrm{P} = \mathrm{P}(p,q)$ lead to new coordinates for which the same type of equation of motion is valid:

$$\dot{\mathrm{Q}} = \frac{\partial\bar{\mathcal{H}}}{\partial\mathrm{P}}, \quad \dot{\mathrm{P}} = -\frac{\partial\bar{\mathcal{H}}}{\partial\dot{\mathrm{Q}}}, \quad \text{with} \quad \bar{\mathcal{H}} = \mathcal{H}(p(\mathrm{P},\mathrm{Q}), q(\mathrm{P},\mathrm{Q})) \tag{D.144}$$

the equation of motion in (Q,P) needs to follow from the same variation of the Lagrange-function $\mathcal{L} = \mathrm{Q}\dot{\mathrm{P}} - \bar{\mathcal{H}}$ as the one in (q,p), so that

$$\mathrm{S} = \int_{t_i}^{t_f} \mathrm{d}t\, \bar{\mathcal{H}}(\mathrm{P},\mathrm{Q}) - \mathrm{Q}\dot{\mathrm{P}} = \int_{t_i}^{t_f} \mathrm{d}t\, \mathcal{H}(p,q) - \dot{q}p + \int_{t_i}^{t_f} \mathrm{d}t\, \frac{\mathrm{d}\bar{\mathrm{W}}}{\mathrm{d}t} \tag{D.145}$$

where the two actions can at most be different by a total time derivative, as

$$\int_{t_i}^{t_f} \mathrm{d}t\, \frac{\mathrm{d}\bar{\mathrm{W}}}{\mathrm{d}t} = \bar{\mathrm{W}}(t_f) - \bar{\mathrm{W}}(t_i) \tag{D.146}$$

does not contribute to the variation, which is fixed at t_i and t_f. But this new function can be used constructively to generate a canonical transform: For instance, the specific choice

$$\bar{\mathrm{W}}(q,\mathrm{P}) = \mathrm{W}(q,\mathrm{P}) - \mathrm{QP} \tag{D.147}$$

for $\bar{\mathrm{W}}$ with variables q and P

$$\frac{\mathrm{dW}}{\mathrm{d}t} = \dot{q}\underbrace{\frac{\partial\bar{\mathrm{W}}}{\partial q}}_{=p} + \dot{\mathrm{P}}\underbrace{\frac{\partial\mathrm{W}}{\partial\mathrm{P}}}_{=\mathrm{Q}} \tag{D.148}$$

is equivalent, if $\mathcal{H}(p,q) = \bar{\mathcal{H}}(\mathrm{P},\mathrm{Q})$ and if $p = \frac{\partial\mathrm{W}}{\partial q}$ and $\mathrm{Q} = \frac{\partial\mathrm{W}}{\partial\mathrm{P}}$. Then, in addition, the functional determinant is $\det\left(\partial(p,q)/\partial(\mathrm{P},\mathrm{Q})\right) = 1$.

A truly novel concept are infinitesimal canonical transformations: The identity transform $W = q \cdot P$ does not change anything

$$p = \frac{\partial W}{\partial q} = P\frac{\partial q}{\partial q} = P \quad \text{and} \quad Q = \frac{\partial W}{\partial P} = q\frac{\partial P}{\partial P} = q \tag{D.149}$$

such that we can define an infinitesimal transform $W(\epsilon) = qP + \epsilon w(q, P)$ controlled by a small $\epsilon > 0$

$$p = \frac{\partial W}{\partial q} = P + \epsilon\frac{\partial w}{\partial q} \quad \rightarrow \quad \Delta p = p - P = \epsilon\frac{\partial w}{\partial q} \tag{D.150}$$

as well as

$$Q = \frac{\partial W}{\partial P} = q + \epsilon\frac{\partial w}{\partial P} \quad \rightarrow \quad \Delta q = q - Q = -\epsilon\frac{\partial w}{\partial P} \tag{D.151}$$

In the spirit of a Lie-generator we take the limit $\epsilon \rightarrow 0$

$$\lim_{\epsilon\rightarrow 0}\frac{\Delta p}{\epsilon} = \frac{\partial W}{\partial q} \quad \text{and} \quad \lim_{\epsilon\rightarrow 0}\frac{\Delta q}{\epsilon} = \frac{\partial W}{\partial p} \tag{D.152}$$

Setting $\epsilon = \Delta t$ and $W = \mathcal{H}(p, q)$

$$\lim_{\Delta t\rightarrow 0}\frac{\Delta q}{\Delta t} = \dot{q} = -\frac{\partial \mathcal{H}}{\partial p} \quad \text{and} \quad \lim_{\Delta t\rightarrow 0}\frac{\Delta p}{\Delta t} = \dot{p} = \frac{\partial \mathcal{H}}{\partial q} \tag{D.153}$$

In summary, the time evolution of a Hamiltonian system is itself a canonical transformation, and the Hamilton equations of motion are the Lie-generators of the transform: Therefore, the entire dynamical evolution of the system can be mapped onto evolution equations for the coordinates and the evolution becomes just a coordinate transform. The phase space volume is invariant under canonical transformations, and therefore the phase space volume needs to be conserved in time evolution too: This is exactly the statement of Liouville's theorem. A counter-example are dissipative systems, for instance the harmonic oscillator with velocity-dependent friction, where the phase space volume decreases with time, as shown by Fig. 9.

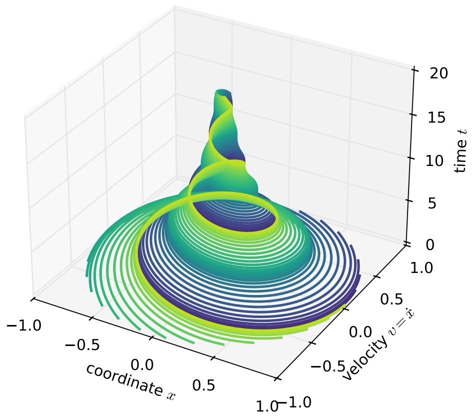

Figure 9: Time evolution of the enclosed phase space volume for an ensemble of damped harmonic oscillators.