CLASSICAL FIELD THEORY

Action and Lagrangian in Classical Mechanics

Before we study quantum field theory we need to learn about the classical fields and the equations governing their dynamics. This dynamics is usually encoded in the action functional, which generalizes a rather formal approach to classical mechanics. So let us start with a quick review of the Lagrangian, the action, and the least action principle in classical mechanics.

For simplicity, let's start with a single dynamical variable $q(t)$, for example a position of a particle moving in one dimension, straight or curved. Consider any smooth "trajectory" $q(t)$ of this variable between times t_1 and t_2 — regardless of whether that trajectory obeys some equations of motion or is completely random like a drunkard's walk — and define the action integral

$$S[q(t)] \;=\; \int_{t_1}^{t_2} L(q,\dot{q})\,dt. \tag{1}$$

In this formula, $L(q,\dot{q})$ is the Lagrangian, for example

$$L(q,\dot{q}) \;=\; E_{\rm kin}(\dot{q}) \;-\; E_{\rm pot}(q) \;=\; \frac{m}{2}\,\dot{q}^2 \;-\; V(q). \tag{2}$$

In the context of the integral (1), the Lagrangian is an ordinary function L of 2 variables q and $\dot{q}$ evaluated for $q(t)$ and $dq(t)/dt$ for the specific trajectory $q(t)$ for each time t. On the other hand, the action S is a *functional* of the entire trajectory $q(t)$ — *i.e.*, the entire history of its motion — rather than its value of any particular time.

Now consider the set of all possible trajectories $q(t)$ which begin at the same value q_1 at time t_1 and end at the same value q_2 at time t_2. The *least action principle* states that trajectory which has the lowest action integral $S[q(t)]$ among this set is the classical trajectory — it happens to obey the classical equations of motion such as Newton Laws.

To see how this works, take any test trajectory $\bar{q}(t)$ and consider its infinitesimal variations $q(t) = \bar{q}(t) + \delta q(t)$. If the trajectory $\bar{q}(t)$ happens to minimize the action functional —

or even if it is any other kind of a stationary point — then the first variation of the action must vanish, $\delta S = 0$. So let's calculate this first variation

$$\delta S \;=\; \int_{t_1}^{t_2} \delta L \, dt. \tag{3}$$

Obviously

$$\delta L @ (\text{time } t) \;=\; \frac{\partial L}{\partial q} \times \delta q(t) \;+\; \frac{\partial L}{\partial \dot{q}} \times \frac{d}{dt} \delta q(t) \tag{4}$$

where the derivatives of L are evaluated for $q = \bar{q}(t)$ and $\dot{q} = d\bar{q}(t)/dt$. Consequently

$$\begin{aligned} \delta S \;&=\; \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} \times \delta q(t) \;+\; \frac{\partial L}{\partial \dot{q}} \times \frac{d}{dt} \delta q(t) \right) dt \\ &\qquad \langle\!\langle \, \text{integrating the second term by parts} \, \rangle\!\rangle \\ &=\; \int_{t_1}^{t_2} \delta q(t) \times \left(\frac{\partial L}{\partial q} \;-\; \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \;+\; \delta q \times \frac{\partial L}{\partial \dot{q}} \Bigg|_{t_1}^{t_2} . \end{aligned} \tag{5}$$

Moreover, due to the initial and final conditions — fixed $q(t_1) = q_1$ and $q(t_2) = q_2$ — we must have $\delta q(t_1) = \delta q(t_2) = 0$, which kills the boundary term in eq. (5). Thus, we are left with

$$\delta S[q] \;=\; \int \delta q \times \left(\frac{\partial L}{\partial q} \;-\; \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) dt. \tag{6}$$

Now, we want this variation to vanish for *any* $\delta q(t)$, and the only way this can happen if the expression inside $(\cdots)$ vanishes for all times t. In this way, the least action principle leads to the *Euler–Lagrange equation of motion*

$$\frac{\partial L}{\partial q} \;-\; \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \;=\; 0 \tag{7}$$

where the derivatives of L are evaluated for the would-be classical trajectory $\bar{q}(t)$. For

example, for the Lagrangian $L = \frac{1}{2}m\dot{q}^2 - V(q)$, we have

$$\frac{\partial L}{\partial q} \;=\; -\frac{dV}{dq} \;=\; +F(q) \quad \text{while} \quad \frac{\partial L}{\partial \dot{q}} \;=\; m\dot{q}, \tag{8}$$

so the Euler–Lagrange equation (7) becomes

$$F(q) \;-\; \frac{d}{dt}\big(m\dot{q}\big) \;=\; 0, \tag{9}$$

or equivalently the Second Law of Newton

$$m\ddot{q} \;=\; F(q) \;=\; -\frac{dV}{dq}\,. \tag{10}$$

$\star \quad \star \quad \star$

For the multiple dynamical variables $q_1(t), \ldots, q_N(t)$ — for example, the 3D positions of $N/3$ particles — the Lagrangian $L(q_1, \ldots, q_N; \dot{q}_1, \ldots \dot{q}_N)$ is a function of $2N$ arguments — N "positions" q_i and N "velocities" $\dot{q}_i$. For $N/3$ particles with purely potential forces between them we have

$$L \;=\; \sum_i \frac{m_i}{2} \times \dot{q}_i^{\,2} \;-\; V(q_1, \ldots, q_n) \tag{11}$$

while more generally L may entangle all $2N$ of its arguments in a complicated fashion. In any case, the action

$$S[\text{all } q_i(t)] \;=\; \int L(q_1, \ldots, q_N; \dot{q}_1, \ldots \dot{q}_N)\, dt \tag{12}$$

is a functional of the combined history of all N dynamical variables at all times (in some interval from t_1 to t_2). Similar to what we had for a single variable, the action functional is defined for all kinds of variables' trajectories regardless of any equations they might obey or disobey. But the least action principle says that the classical trajectories — which obey the Euler–Lagrange equations of motion — yield the lowest action among all other trajectories which begin and end at the same values of the q_i.

The Euler–Lagrange equations for the multiple variables obtain from demanding zero first variation of the action with respect to the independent infinitesimal variations $\delta q_i(t)$ of all N variables. Since this exercise is very similar to what we have done for a single variable, let me skip the intermediate steps and write down the end result: For each variable q_i and for every moment t, we must have

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \;-\; \frac{\partial L}{\partial q_i} \;=\; 0 \quad (\forall i = 1, \ldots N). \tag{13}$$

For example, for the Lagrangian (11), the Euler Lagrange equations of motion become

$$\frac{d}{dt}(m_i \dot{q}_i) \;+\; \frac{\partial V}{\partial q_i} \;=\; 0, \tag{14}$$

or equivalently Newton equations for multiple bodies

$$m_i \ddot{q}_i \;=\; -\frac{\partial V}{\partial q_i} \;=\; +F_i(q_1, \ldots, q_N). \tag{15}$$

$\star \quad \star \quad \star$

Normally, the Lagrangian L involves the dynamical variables q_i themselves and their *first* time derivatives $\dot{q}_i$ but no higher derivatives; consequently, the Euler–Lagrange equations of motion are second-order differential equations. If we include higher time derivatives into the Lagrangian, the equation of motion would also be higher-order in d/dt. For example, let $L(q, \dot{q}, \ddot{q})$ involve the second time derivative; then the Euler–Lagrange equation becomes

$$\frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}} \;-\; \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \;+\; \frac{\partial L}{\partial q} \;=\; 0, \tag{16}$$

which is of the 4^{th} order in d/dt (assuming $\partial L/\partial \ddot{q}$ depends on the $\ddot{q}$). Mathematically, this equation is equivalent to two second-order equations for two interacting variables $q_1(t)$ and $q_2(t)$ (for example, $q_1 = q$ and $q_2 = \ddot{q}$), and to pick a specific solution we need to fix 4 initial conditions $q(0)$, $\dot{q}(0)$, $\ddot{q}(0)$, and $\dddot{q}(0)$, or equivalently two variables $q_1(0)$, $q_2(0)$ and their first derivatives $\dot{q}_1(0)$, $\dot{q}_2(0)$. Physically, this corresponds to two independent (but interacting) degrees of freedom $q_1(t)$ and $q_2(t)$ rather than a single $q(t)$.

More generally, a dynamical degree of freedom is an independent variable obeying a second-order (in d/dt) equation of motion which takes two initial conditions (the variable itself and its velocity) to fix a specific solution. Thus, a variable obeying an equation of motion of degree $2n$ describes n dynamical degrees of freedom rather than one, and it would be more transparent to replace it with n variables obeying second-order equations. On the other hand, a pair of first order equations for a pair of variables — such the coordinate and the momentum — describes a single degree of freedom rather than two.

In the Lagrangian terms, if L depends on higher derivatives of some variable q, that variable involves multiple degrees of freedom. If the highest derivative entering L is $q^{(n)}$ — and if L is a non-linear function of $q^{(n)}$ — then $q(t)$ describes n degrees of freedom. For transparency's sake, it is best to replace such $q(t)$ with n variables which enter L with their first derivatives only. On the other hand, if a variable enters the Lagrangian without any time derivatives at all — *i.e.*, L depends on a q_i but not on $\dot{q}_i$, *etc.*, — then the q_i is not a dynamical variable at all but rather a Lagrange multiplier enforcing some time-independent constraint on the other variables. Finally, if L is linear on a time derivative $\dot{q}_i$ — *i.e.*, L depends on the $\dot{q}_i$ but $\partial L/\partial\dot{q}_i$ does not depend on the $\dot{q}_i$ itself or other velocities, — then the $q_i(t)$ obeys a first-order equation and describes $1/2$ of a degree of freedom; obviously, there should be an even number of such variables.

Classical Fields and Lagrangian Densities

A classical field is a *continuous family of dynamical variables* $\Phi_{\mathbf{x}}(t)$, one variable for each space point $\mathbf{x} = (x, y, z)$; mathematically, this family is a function $\Phi(\mathbf{x}, t)$ of all 4 spacetime coordinates $(x, y, z; t)$. Note that the space coordinates (x, y, z) here are not dynamical variables as they do not describe the position of any physical body. Instead, it's the field which comprises dynamical variables $\Phi_{\mathbf{x}}(t)$ at every point in space, and the coordinates $(x, y, z) = \mathbf{x}$ merely label these variables.

Most generally, the Lagrangian $L[\Phi(\mathbf{x}), \dot{\Phi}(\mathbf{x})]$ for a classical field Φ is a functional of the entire space configuration of the field $\Phi(\mathbf{x})$ and its time derivative $\dot{\Phi}(\mathbf{x}) = \partial\Phi(\mathbf{x}, t)/\partial t$ at any particular time. However, in field theories which do not allow interactions at a distance

— such as all relativistic theories, — the Lagrangian L is a space integral

$$L \;=\; \int d^3\mathbf{x}\, \mathcal{L}(\Phi, \nabla\Phi, \dot{\Phi}) \tag{17}$$

of the *Lagrangian density*, $\mathcal{L}$ which is an ordinary local function of the field Φ and its space and time derivatives $\nabla\Phi$ and $\dot{\Phi}$. Consequently, the action is a 4D spacetime integral of the Lagrangian density

$$S[\Phi(\mathbf{x},t)] \;=\; \int dt \int d^3\mathbf{x}\, \mathcal{L}(\Phi, \nabla\Phi, \dot{\Phi}). \tag{18}$$

Similarly to mechanics, the equations of motion for a classical field obtain from the least action principle: Among all the time evolutions of a field with given initial and final values (for all $\mathbf{x}$), the evolution which obeys the field equations has the least action (18). Thus, the classical field equations themselves obtain from infinitesimal variations of some field configuration $\overline{\Phi}(\mathbf{x},t) \mapsto \overline{\Phi}(\mathbf{x},t) + \delta\Phi(\mathbf{x},t)$ and demand $\delta S = 0$ for all $\delta\Phi(\mathbf{x},t)$. Under such infinitesimal variations

$$\delta\mathcal{L}@(\mathbf{x},t) \;=\; \frac{\partial\mathcal{L}}{\partial\Phi} \times \delta\Phi \;+\; \frac{\partial\mathcal{L}}{\partial\dot{\Phi}} \times \frac{\partial\delta\Phi}{\partial t} \;+\; \frac{\partial\mathcal{L}}{\partial(\nabla\Phi)} \cdot \nabla\delta\Phi \tag{19}$$

where the derivatives of L are taken for the $\Phi = \overline{\Phi}(\mathbf{x},t)$, $\dot{\Phi} = \dot{\overline{\Phi}}(\mathbf{x},t)$, and $\nabla\Phi = \nabla\overline{\Phi}(\mathbf{x},t)$, hence

$$\begin{aligned}
\delta S \;&=\; \int dt \int d^3\mathbf{x} \left(\frac{\partial\mathcal{L}}{\partial\Phi}\,\delta\Phi \;+\; \frac{\partial\mathcal{L}}{\partial\dot{\Phi}}\,\frac{\partial\delta\Phi}{\partial t} \;+\; \frac{\partial\mathcal{L}}{\partial(\nabla\Phi)} \cdot \nabla\delta\Phi \right) \\
&\qquad \langle\!\langle\, \text{integrating by parts} \,\rangle\!\rangle \\
&= \int dt \int d^3\mathbf{x}\, \delta\Phi(\mathbf{x},t) \left(\frac{\partial\mathcal{L}}{\partial\Phi} \;-\; \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\Phi}} \right) \;-\; \nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\Phi)} \right) \right) @(\mathbf{x},t) \\
&\qquad + \int d^3\mathbf{x} \left(\frac{\partial\mathcal{L}}{\partial\dot{\Phi}}\,\delta\Phi \right)\Bigg|_{t_1}^{t_2} \;+\; \int dt \iint\limits_{\substack{\text{boundary} \\ \text{of 3D space}}} d^2\mathbf{Area} \cdot \frac{\partial\mathcal{L}}{\partial(\nabla\Phi)}\,\delta\Phi\,.
\end{aligned} \tag{20}$$

On the last line here, the time boundary term vanishes because the $\delta\Phi(\mathbf{x},t)$ variation must vanish at the initial and the final times $t = t_1$ and $t = t_2$. As to the space boundary term, we

may eliminate it by assuming that for $|\mathbf{x}| \to \infty$, $\delta\Phi(\mathbf{x},t) \to 0$ fast enough that the surface integral asymptotes to zero. Consequently, we are left with the

$$\delta S \;=\; \int dt \int d^3\mathbf{x}\, \delta\Phi(\mathbf{x},t) \left(\frac{\partial \mathcal{L}}{\partial \Phi} \;-\; \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\Phi}} \right) \;-\; \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial (\nabla \Phi)} \right) \right) @(\mathbf{x},t), \tag{21}$$

which must vanish for any $\delta\Phi(\mathbf{x},t)$ which obeys the boundary conditions. Clearly, this requires the expression inside the big $(\cdots)$ to vanish at all $\mathbf{x}$ and t, hence the Euler–Lagrange field equations

$$\frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\Phi}} \right) \;+\; \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial (\nabla \Phi)} \right) \;-\; \frac{\partial \mathcal{L}}{\partial \Phi} \;=\; 0. \tag{22}$$

For example, consider the free relativistic scalar field Φ with the Lagrangian

$$\mathcal{L} \;=\; \frac{1}{2c^2}\, \dot{\Phi}^2 \;-\; \frac{1}{2}\, (\nabla \Phi)^2 \;-\; \frac{\kappa^2}{2}\, \Phi^2. \tag{23}$$

For this Lagrangian

$$\frac{\partial \mathcal{L}}{\partial \dot{\Phi}} \;=\; \frac{1}{c^2}\, \dot{\Phi}, \qquad \frac{\partial \mathcal{L}}{\partial (\nabla \Phi)} \;=\; -\nabla \Phi, \qquad \frac{\partial \mathcal{L}}{\partial \Phi} \;=\; -\kappa^2 \Phi, \tag{24}$$

so the Euler–Lagrange equation (22) becomes

$$\frac{\partial}{\partial t} \left(\frac{1}{c^2}\, \frac{\partial \Phi}{\partial t} \right) + \;\nabla \cdot \big(-\nabla \Phi \big) \;-\; \big(-\kappa^2 \Phi \big) \;=\; 0. \tag{25}$$

In a more compact form, this is the Klein–Gordon equation

$$\left(\frac{1}{c^2}\, \frac{\partial^2}{\partial t^2} \;-\; \nabla^2 \;+\; \kappa^2 \right) \Phi(\mathbf{x},t) \;=\; 0 \tag{26}$$

for a free relativistic field of mass $m = (\hbar/c)\kappa$.

Most field theories have multiple fields — *i.e.*, multiple dynamical variables $\phi_1(\mathbf{x},t)$, …, $\phi_N(\mathbf{x},t)$ at each space point $\mathbf{x}$. Such multiple fields can be different species of scalar fields, or different components of a vector, tensor, *etc.*, field, or any combination thereof: for example, 3 components of one vector field, 3 more components of another vector field, plus 5 scalar fields for $N = 11$ fields altogether. For the moment, let me ignore the difference between the components and the species and simply focus on having N independent $\phi_i(\mathbf{x},t)$ dynamical variables at each space point.

Assuming local interactions between the fields, the action has form

$$S[\phi_i(\mathbf{x},t) \text{ for all } i \text{ at all } \mathbf{x} \text{ and all } t] \;=\; \int\!dt \int\!d^3\mathbf{x}\, \mathcal{L}(\phi_1,\ldots,\phi_N;\nabla\phi_1,\ldots,\nabla\phi_N;\dot{\phi}_1,\ldots,\dot{\phi}_N), \tag{27}$$

where the Lagrangian density $\mathcal{L}$ is a function of $5N$ variables: N fields ϕ_i themselves, N time derivatives $\dot{\phi}_i$, and $3N$ components of N space gradients $\nabla\phi_i$. Minimizing the action (27) gives rise to N coupled Euler–Lagrange equations, one for each field:

$$\forall i = 1,\ldots,N: \quad \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}_i}\right) \;+\; \nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\phi_i)}\right) \;-\; \frac{\partial\mathcal{L}}{\partial\phi_i} \;=\; 0. \tag{28}$$

For example, for N relativistic scalar fields with Lagrangian density

$$\mathcal{L} \;=\; \sum_{i=1}^{N}\left(\frac{1}{2c^2}(\dot{\phi}_i)^2 \;-\; \frac{1}{2}(\nabla\phi_i)^2\right) \;-\; V(\phi_1,\ldots,\phi_N) \tag{29}$$

we have

$$\frac{\partial\mathcal{L}}{\partial\dot{\phi}_i} \;=\; \frac{1}{c^2}\,\dot{\phi}_i\,, \quad \frac{\partial\mathcal{L}}{\partial(\nabla\phi_i)} \;=\; -\nabla\phi_i\,, \quad \frac{\partial\mathcal{L}}{\partial\phi_i} \;=\; -\frac{\partial V}{\partial\phi_i}\,, \tag{30}$$

hence the Euler–Lagrange field equations (28) become

$$\forall i = 1,\ldots,N: \quad \frac{1}{c^2}\,\frac{\partial^2}{\partial t^2}\phi_i \;-\; \nabla^2\phi_i \;+\; \frac{\partial V}{\partial\phi_i} \;=\; 0. \tag{31}$$

In general $\partial V/\partial\phi_i$ depends on both the ϕ_i itself and on the other fields $\phi_{j\neq i}$, so eqs. (31) are not independent but coupled to each other.

For another example, consider a non-relativistic complex (*i.e.*, complex number valued) scalar field $\psi(\mathbf{x},t)$ describing the superfluid component of liquid helium II. Specifically, its magnitude describes the superfluid density while the gradient of its phase describes the superfluid's velocity according to

$$\rho(\mathbf{x},t) \;=\; M\,|\psi(\mathbf{x},t)|^2\,, \quad \mathbf{v}(\mathbf{x},t) \;=\; \frac{\hbar}{M}\,\nabla\,\mathrm{phase}\big(\psi(\mathbf{x},t)\big) \tag{32}$$

where M is the mass of a helium atom. In the Landau–Ginzburg theory of superfluidity, the field equations for $\psi(\mathbf{x},t)$ obtain from a local Lagrangian density

$$\mathcal{L} \;=\; -\hbar\,\mathrm{Im}(\psi^*\dot{\psi}) \;-\; \frac{\hbar^2}{2M}\,\nabla\psi^*\cdot\nabla\psi \;+\; \mu|\psi|^2 \;-\; \frac{\lambda}{2}\,|\psi|^4. \tag{33}$$

A complex field $\psi(\mathbf{x},t)$ is equivalent to two real fields $\mathrm{Re}\,\psi(\mathbf{x},t)$ and $\mathrm{Im}\,\psi(\mathbf{x},t)$, so we should be getting two coupled field equations by independently varying $\delta\,\mathrm{Re}\,\psi(\mathbf{x},t)$ and $\delta\,\mathrm{Im}\,\psi(\mathbf{x},t)$. However, it is easier to derive the field equation(s) in a manifestly complex form by treating the $\delta\psi(\mathbf{x},t)$ and $\delta\psi^*(\mathbf{x},t)$ as independent variations, hence Euler–Lagrange equations

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\psi}}\right) \;+\; \nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\psi)}\right) \;-\; \frac{\partial\mathcal{L}}{\partial\psi} \;=\; 0 \\ \text{and}\qquad &\\ &\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\psi}^*}\right) \;+\; \nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\psi^*)}\right) \;-\; \frac{\partial\mathcal{L}}{\partial\psi^*} \;=\; 0. \end{aligned} \tag{34}$$

For the Landau–Ginzburg Lagrangian density (33)

$$-\,\mathrm{Im}(\psi^*\dot{\psi}) \;=\; \tfrac{i}{2}\psi^*\times\dot{\psi} \;-\; \tfrac{i}{2}\psi\times\dot{\psi}^*, \tag{35}$$

hence

$$\frac{\partial\mathcal{L}}{\partial\dot{\psi}} \;=\; \frac{i\hbar}{2}\,\psi^*, \quad \frac{\partial\mathcal{L}}{\partial\dot{\psi}^*} \;=\; -\frac{i\hbar}{2}\,\psi, \tag{36}$$

$$\frac{\partial\mathcal{L}}{\partial(\nabla\psi)} \;=\; -\frac{\hbar^2}{2M}\,\nabla\psi^*, \quad \frac{\partial\mathcal{L}}{\partial(\nabla\psi^*)} \;=\; -\frac{\hbar^2}{2M}\,\nabla\psi, \tag{37}$$

$$\frac{\partial \mathcal{L}}{\partial \psi} \;=\; -\frac{i\hbar}{2}\,\dot{\psi}^{*} \;+\; (\mu - \lambda|\psi|^2)\psi^{*}, \tag{38}$$

$$\frac{\partial \mathcal{L}}{\partial \psi^{*}} \;=\; +\frac{i\hbar}{2}\,\dot{\psi} \;+\; (\mu - \lambda|\psi|^2)\psi, \tag{39}$$

so eqs. (34) become

$$\frac{i\hbar}{2}\,\dot{\psi}^{*} \;-\; \frac{\hbar^2}{2M}\,\nabla^2\psi^{*} \;+\; \frac{i\hbar}{2}\,\dot{\psi}^{*} \;-\; (\mu - \lambda|\psi|^2)\psi^{*} \;=\; 0, \tag{40}$$

$$-\frac{i\hbar}{2}\,\dot{\psi} \;-\; \frac{\hbar^2}{2M}\,\nabla^2\psi \;-\; \frac{i\hbar}{2}\,\dot{\psi} \;-\; (\mu - \lambda|\psi|^2)\psi \;=\; 0, \tag{41}$$

which in a more compact form becomes a Schrödinger-like equation

$$i\hbar\frac{\partial \psi}{\partial t} \;=\; -\,\frac{\hbar^2}{2M}\,\nabla^2\psi \;+\; (\lambda|\psi|^2 - \mu)\psi \tag{42}$$

for the ψ field,[⋆] and a complex-conjugate equation

$$-i\hbar\frac{\partial \psi^{*}}{\partial t} \;=\; -\,\frac{\hbar^2}{2M}\,\nabla^2\psi^{*} \;+\; (\lambda|\psi|^2 - \mu)\psi^{*} \tag{43}$$

for the conjugate field ψ^{*}. Note that differential eqs. (42) and (43) are first-order in the time derivative (albeit second-order in space derivatives), so together $\psi(\mathbf{x},t)$ and $\psi^{*}(\mathbf{x},t)$ — or equivalently the two real fields $\mathrm{Re}\,\psi(\mathbf{x},t)$ and $\mathrm{Im}\,\psi(\mathbf{x},t)$ — describe $2 \times \frac{1}{2} = 1$ degree of freedom (at each $\mathbf{x}$) rather than 2 degrees of freedom one naively expects from a complex field.

In non-relativistic theories, the fields' equations of motion often have more space derivatives then time derivatives. Consequently, the Lagrangian density function may depend on higher space derivatives of the fields, thus $\mathcal{L}(\phi, \nabla\phi, \nabla\nabla\phi, \ldots; \dot{\phi})$ or even

⋆ Despite similarity, eq. (42) is not a true Schrödinger equation for a wave function. In particular, eq. (42) has a non-linear term $\lambda|\psi|^2 \times \psi$, while a true Schrödinger equation is always linear, come hell or high water. Also, the $\psi(\mathbf{x},t)$ here is a classical field rather than a wave function of anything; indeed, a true wave function of the superfluid would depend on a position of every helium atom in a superfluid rather than on a single position $\mathbf{x}$. Nevertheless, some old textbooks miscall $\psi(\mathbf{x},t)$ a "macroscopic wave function"; this is a very misleading terminology, so please be careful.

$\mathcal{L}(\phi, \nabla\phi, \nabla\nabla\phi, \ldots; \dot{\phi}, \nabla\dot{\phi}, \ldots)$; all such Lagrangian density functions are OK as long as they do not involve higher *time* derivatives. The general Euler–Lagrange equation for such a higher-derivative Lagrangian density is

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \nabla \cdot \frac{\partial \mathcal{L}}{\partial(\nabla\dot{\phi})} + \cdots\right) - \left(\frac{\partial \mathcal{L}}{\partial \phi} - \nabla \cdot \frac{\partial \mathcal{L}}{\partial(\nabla\phi)} + \nabla_i\nabla_j \frac{\partial \mathcal{L}}{\partial(\nabla_i\nabla_j\phi)} - \cdots\right) = 0. \tag{44}$$

Moreover, the Lagrangian L maybe non-local in space and involve multiple space integrals rather than a single integral of the Lagrangian density $\mathcal{L}$. For example, an improvement of the Landau–Ginzburg theory which accounts for a finite-range two-body forces between helium atoms has

$$\begin{aligned} L \;=\; & \int d^3\mathbf{x} \left(-\hbar\, \mathrm{Im}(\psi^* \dot{\psi}) \;-\; \frac{\hbar^2}{2M}\, \nabla\psi^* \cdot \nabla\psi \;+\; \mu|\psi|^2 \right) \\ & - \frac{1}{2} \int d^3\mathbf{x} \int d^3\mathbf{y}\, V_2(\mathbf{x}-\mathbf{y}) \times |\psi(\mathbf{x})|^2 \times |\psi(\mathbf{y})|^2. \end{aligned} \tag{45}$$

where $V_2(\mathbf{x}-\mathbf{y})$ is the interaction potential for 2 helium atoms, one at $\mathbf{x}$ the other at $\mathbf{y}$. The field equations stemming from this Lagrangian are differential in time but integro-differential in space, namely

$$i\hbar \frac{\partial}{\partial t}\psi(\mathbf{x},t) \;=\; -\frac{\hbar^2}{2M}\, \nabla^2\psi(\mathbf{x},t) \;-\; \mu\psi(\mathbf{x},t) \;+\; \psi(\mathbf{x},t) \times \int d^3\mathbf{y}\, V_2(\mathbf{x}-\mathbf{y})\, |\psi(\mathbf{y},t)|^2, \tag{46}$$

and the complex-conjugate equation for the $\psi^*(\mathbf{x},t)$.

We shall return to such non-local-in-space classical (and quantum) field theories in 3–4 weeks. But for the next few lectures we shall focus on the relativistic field theories, which are always local.

Relativistic Fields: Conventions and Notations

Following the *Peskin & Schroeder* textbook, I shall use Greek letters — especially the mid-alphabet letters λ, μ, ν — for the 4D Lorentz indices and make them run from 0 to 3: 0 for time and 1,2,3 for space. I shall use the $(+,-,-,-)$ signature for the metric,

$$g_{\mu\nu} \;=\; \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{47}$$

thus spacetime coordinates $x^\mu = (x^0 = ct; x^1, x^2, x^3)$ but $x_\mu = (+ct, -x^1, -x^2, -x^3)$. Likewise, for most 4–vectors A^μ with an upper index, the space components (A^1, A^2, A^3) form a 3–vector with the usual sign, while for a 4–vector A_μ with a lower index, the $\mu = 1, 2, 3$ components have an opposite sign. Only the time component $A^0 = A_0$ has the usual sign regardless of whether the index $\mu = 0$ is upstairs or downstairs. For example, for the energy-momentum 4–vector, $P^\mu = (E/c; +\mathbf{p})$ while $P_\mu = (+E/c; -\mathbf{p})$. Likewise, for the electric current 4–vector $J^\mu = (c\rho; +\mathbf{J})$ while $J_\mu = (+c\rho; -\mathbf{J})$; for the electromagnetic 4–vector potential $A^\mu = (\Phi; +\mathbf{A})$ while $A_\mu = (+\Phi, -\mathbf{A})$, *etc.*, *etc.* The only exception from this rule is the derivative 4-vector

$$\partial_\mu \;=\; \frac{\partial}{\partial x^\mu} \;=\; \left(\frac{1}{c}\frac{\partial}{\partial t}; +\nabla\right) \quad \text{while} \quad \partial^\mu \;=\; \frac{\partial}{\partial x_\mu} \;=\; \left(\frac{1}{c}\frac{\partial}{\partial t}; -\nabla\right). \tag{48}$$

The reason for this exception is to have $\partial_\mu x^\nu = \delta_\mu^\nu$ and $\partial^\mu x_\nu = \delta_\nu^\mu$ without any extra signs. On the other hand, $\partial_\mu x_\nu = g_{\mu\nu}$ and $\partial^\mu x^\nu = g^{\mu\nu}$ involve the metric tensor.

Throughout this class, I shall use Einstein's summation convention: When the same Lorentz index appears twice within a product of 4–vectors or 4–tensors, once upstairs and once downstairs, then such index is implicitly summed over. For example,

$$\begin{aligned} A^\mu B_\mu \;=\; \sum_{\mu=0,1,2,3} A^\mu B_\mu \;&=\; A^0 B_0 \,+\, A^1 B_1 \,+\, A^2 B_2 \,+\, A^3 B_3 \\ &=\; A^0 B^0 \,-\, A^1 B^1 \,-\, A^2 B^2 \,-\, A^3 B^3 \\ &=\; A^0 B^0 \,-\, \mathbf{A}\cdot\mathbf{B}. \end{aligned} \tag{49}$$

On the other hand, if the same index appears twice upstairs, or twice downstairs, or more

then twice, then this is probably a mistake of some kind. In very rare cases when this is not a mistake, one should explicitly say whether such index should be summed over or not, but either way the expression would break the Lorentz symmetry.

In relativistic notations, a classical field is $\phi(x)$ where x denotes the 4-vector x^μ, its space and time derivatives $\nabla\phi$ and $\dot\phi$ combine into a 4–vector $\partial_\mu\phi(x)$, thus the Lagrangian density $\mathcal{L}(\phi, \partial_\mu\phi)$, for example for a free relativistic scalar field

$$\mathcal{L} \;=\; \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) \;-\; \frac{(mc/\hbar)^2}{2}\,\phi^2. \tag{50}$$

The action of a relativistic field theory is a 4D spacetime integral of the Lagrangian density,

$$S[\phi(x)] \;=\; \int dt \int d^3\mathbf{x}\,\mathcal{L} \;=\; \frac{1}{c}\int d^4x\,\mathcal{L}. \tag{51}$$

The 4D volume element d^4x is invariant under Lorentz transformations, so a Lorentz-invariant action calls for a Lorentz-invariant (or rather Lorentz scalar) Lagrangian density $\mathcal{L}$ such as (50). We shall see other examples of Lorentz-scalar Lagrangian densities next lecture and in the homework set#1.

The relativistic form of the Euler–Lagrange field equation is

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\right) \;-\; \frac{\partial\mathcal{L}}{\partial\Phi} \;=\; 0. \tag{52}$$

For example, for the Lagrangian density (50),

$$\frac{\partial\mathcal{L}}{\partial\Phi} \;=\; -(mc/\hbar)^2\phi \quad\text{while}\quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \;=\; +\partial^\mu\Phi, \tag{53}$$

— note how the index μ moves upstairs from downstairs, — hence eq. (52) becomes

$$\partial_\mu\big(\partial^\mu\Phi\big) \;+\; (mc/\hbar)^2\Phi \;=\; 0, \tag{54}$$

or in a more compact form

$$[\partial^2 + (mc/\hbar)^2]\Phi(x) \;=\; 0 \tag{55}$$

where $\partial^2 \;=\; \partial_\mu\partial^\mu$ is the Lorentz square of the derivative operator, also known as the

D'Alembert operator

$$\partial^2 \;=\; \Box \;=\; \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \;-\; \nabla^2. \tag{56}$$

For the multiple fields (or field components) $\phi_1(x), \ldots, \phi_N(x)$, the Lagrangian density $\mathcal{L}$ is a function of all the ϕ_i and their first derivative 4–vectors $\partial_\mu \phi_i$, thus

$$S \;=\; \frac{1}{c}\int d^4x\, \mathcal{L}(\phi_1, \ldots, \phi_N; \partial_\mu\phi_1, \ldots, \partial_\mu\phi_N), \tag{57}$$

and minimizing this action leads to N coupled Euler–Lagrange field equations

$$\forall i = 1, \ldots, N: \quad \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_i)} \right) \;-\; \frac{\partial \mathcal{L}}{\partial \Phi_i} \;=\; 0. \tag{58}$$

For example, for N scalar fields with Lagrangian density

$$\mathcal{L} \;=\; \sum_{i=1}^{N} \frac{1}{2}(\partial_\mu \Phi_i)(\partial^\mu \Phi_i) \;-\; V(\Phi_1, \ldots, \Phi_N), \tag{59}$$

the field equations are

$$\forall i = 1, \ldots, N: \quad \partial^2 \Phi_i \;+\; \frac{\partial V}{\partial \Phi_i} \;=\; 0. \tag{60}$$

More interesting examples of multiple components of a vector or tensor field will be discussed in the next set of my notes and in the homework set#1.