In classical mechanics, for a system with Lagrangian $L$, one is interested in the extremals of
\[ \int_{t_1}^{t_2} L(t, q(t), \dot{q}(t)) \, dt \]
along curves $t \mapsto q(t)$ with $q(t_1) = q_1$ and $q(t_2) = q_2$ fixed.

Extremals satisfy E-L equations. Given $(t_0, q_0)$, and a choice of $\dot{q}_0$, one can integrate E-L to produce an extremal curve (not necessarily passing through specified end-points), generating a field of extremals. A given extremal is contained in a central field of extremals if the map $\dot{q}_0 \mapsto q$ is nonregular.
In that case, one can define correctly

\[ S(t, q) = \int_{t_0}^{t} L(q, \dot{q}(s), \ddot{q}(s))\,ds \]

where \( q \to q(s) \) is an \underline{extremal} connecting \((t_0, q_0)\) to \((t, q)\).

We call \( S \) the action function (with parameter \((t_0, q_0)\)).

**Lemma**

\[ ds = p\,dq - H\,dt \]

where \( p = \frac{\partial L}{\partial \dot{q}} \), \( H = p\dot{q} - L \)

**Remark** Proof is an application of Stokes theorem.

Now \[ ds = \frac{\partial S}{\partial t}\,dt + \frac{\partial S}{\partial q}\,dq \,.

The lemma says

\[ p = \frac{\partial S}{\partial q} \]

\[ \frac{\partial S}{\partial t} + H(t, q, \frac{\partial S}{\partial q}) = 0 \]
This last equation is the Hamilton–Jacobi equation of mechanics. It has a very nice generalization to control theory called the Hamilton–Jacobi–Bellman equation (or simply Bellman equation for short).

We approach this by a discretization of the problem of continuous-time optimal control

\[
\begin{align*}
\dot{x}(t) &= f(t, x(t), u(t)) \\
x(t_0) &= x_0 \\
\min_{u(t)} & \quad \phi(x(T)) + \int_{0}^{T} L(t, x(t), u(t)) \, dt \\
\text{subject to} & \quad (1) 
\end{align*}
\]

This is a free-end point problem with a terminal cost \( \phi(x(T)) \).

Approximate (1) by the discrete-time system

\[
x_{k+1} = x_k + s \cdot f(k, x_k, u_k)
\]

\( k = 0, 1, 2, \ldots, N-1 \),  \( s = T/N \)

where \( x_k \approx x(ks) \), \( u_k \approx u(ks) \).
The cost is approximated by

\[ \phi(x_N) + \sum_{k=0}^{N-1} L(k, x_k, u_k) \delta \]

Let

\[ J^*(k, x) \]

be the optimal cost-to-go at time \( k \) and state \( x \).

We also call this the value function. Clearly

\[ J^*(N, x) = \phi(x) \quad \text{(end state, no decision)} \]

\[ J^*(k, x) = \min_v \left\{ L(k, x, v)\delta + J^*(k+1, x+f(k, x, v)) \right\} \]

\[ k = 0, 1, 2, \ldots, N-1 \]

These are the equations of dynamic programming (follow from the principle of optimality). We have a functional equation for unknown \( J^* \). Solve it backwards from \( k = N \).

Suppose \( J^* \) has sufficient differentiability (in the continuous time and space setting).
Then one can write,
\[ J^* ((k+1), x + f(k, x, \nu) s) \]
\[ = J^* (k, x) + \frac{\partial}{\partial t} J^* (k, x) \cdot s \]
\[ + \frac{\partial}{\partial x} J^* (k, x) \cdot f(k, x, \nu) s \]
\[ + o(s) \]
\[ \Rightarrow J^*(k, x) = \min_{\nu} \left\{ L(k, x, \nu) s + J^*(k, x) \right\} \]
\[ + \frac{\partial}{\partial t} J^* (k, x) \cdot s \]
\[ + \frac{\partial}{\partial x} J^* (k, x) \cdot f(k, x, \nu) s \]
\[ + o(s) \]

Cancelling \( J^*(k, x) \) from both sides, dividing by \( s \) and sending \( s \to 0 \) we get
\[ 0 = \min_{\nu} \left\{ L(t, x, \nu) + \frac{\partial}{\partial t} J^* (t, x) \right\} \]
\[ + \frac{\partial}{\partial x} J^* (t, x) \cdot f(t, x, \nu) \]
\[ \Leftrightarrow \frac{\partial}{\partial t} J^* (t, x) = - \min_{\nu} \left\{ L(t, x, \nu) + \frac{\partial}{\partial x} J^* (t, x) \cdot f(t, x, \nu) \right\} \]

This is the Hamilton–Jacobi–Bellman (HJB) equation or simply the Bellman equation in continuous time.
Sufficiency Theorem.

Suppose \( V(t,x) \) is a solution to

\[
\Theta = \min_{\nu} \left\{ L(t,x,v) + \nabla V(t,x) \cdot \nabla V(t,x) \right\}
\]

\[
+ \nabla V(t,x) \cdot f(t,x,v) \quad \forall t,x
\]

\[
V(T,x) = \phi(x) \quad \forall x
\]

Suppose that \( \mu^*(t,x) \) attains the minimum in HJB, \( \forall t \) and \( x \). Let \( \{ x^*(t) \mid t \in [0,T] \} \) be the state trajectory satisfying

\[
x^*(t) = f (t, x^*(t), \mu^*(t, x^*(t)))
\]

\[
x^*(0) = x_0
\]

Suppose that this closed loop dynamical has a unique solution starting from at any pair \((t,x)\) and that the control trajectory \( \mu^*(t, x^*(t)) \) is piecewise continuous as a function \( t \). Then \( V \) is the unique solution to HJB, equal to the optimal cost-to-go (value) function

\[
V(t,x) = J^*(t,x) \quad \forall t,x
\]

where

\[
J^*(t,x) = \min_{u(\cdot)} \left[ \int_{t}^{T} L (s, x(s), u(s)) \, ds + \phi(x(T)) \right]
\]
subject to \[ \dot{x}(t) = f(\sigma, x(t), u(t)) \quad \sigma \geq t \]
\[ x(t) = x_0 \]

Furthermore \[ u^*(t) = \mu^*(t, x^*(t)) \] is optimal \[ 0 \leq t \leq T. \]

Proof: let \( \bar{u}(\cdot) \) be any admissible control trajectory and let \( \bar{x}(t) \) \( 0 \leq t \leq T \) be state correspond to \( \bar{u}(\cdot) \). From HJB

\[ 0 \leq L(t, \bar{x}(t), \bar{u}(t)) \]
\[ + \nabla_t V(t, \bar{x}(t)) \]
\[ + \nabla_x V(t, \bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) \]

\[ = L(t, \bar{x}(t), \bar{u}(t)) \]
\[ + \frac{d}{dt} V(t, \bar{x}(t)) \cdot \frac{dx}{dt} \]

Integrate both sides to get

\[ 0 \leq \int_0^T L(\sigma, \bar{x}(\sigma), \bar{u}(\sigma)) d\sigma \]
\[ + V(T, \bar{x}(T)) - V(0, \bar{x}(0)) \]

\[ = \int_0^T + \psi(\bar{x}(T)) - V(0, x_0) \]

- \nabla -
\[ V(0, x_0) = \phi(\bar{x}(T)) + \int_0^T L(\sigma, \bar{x}(\sigma), \bar{u}(\sigma)) d\sigma \]

If we \( u^*(t) \) and \( x^*(t) \) instead of \( \bar{u}(t) \)

and \( \bar{x}(t) \), the preceding inequalities become

equalities (starting from HJB).

\[ V(0, x_0) = h(x^*(T)) + \int_0^T L(\sigma, x^*(\sigma), u^*(\sigma)) d\sigma \]

Thus the cost corresponding to \((u^*, x^*)\)

is \( \leq \) cost corresponding \((\bar{u}, \bar{x})\) for any admissible \( \bar{u} \).

\[ u^*(t) \text{ is optimal} \]

and \[ V(0, x_0) = J^*(0, x_0) \]

In the preceding arguments, we can replace \((0, x_0)\) by \((t, x)\) with no change.

\[ V(t, x) = J^*(t, x) \]