APPENDIX B

COMPARISON OF NEWTON DESCENT WITH VARIATIONAL PRINCIPLE IN HAMILTONIAN SYSTEM

Variational principle in Lagrangian mechanics is explained. Its usage in calculating periodic orbits is discussed and compared with the Newton descent method in ref. [158].

B.1 Variational principle in Hamiltonian systems

Variational principle is a powerful tool in theoretical physics. In Hamiltonian mechanics, it plays such a central role that much of the recent development of the subject is based on it, analytically or numerically. The variational formulation of mechanics is also a convenient bridge between classical mechanics and other branches of physics like statistical physics, quantum physics.

Although the following formulation is valid for the most general case, for brevity, we just consider a particle with unit mass ($m = 1$) moving in the 3-d space under the influence of the external potential $V(q,t)$. Newton's second law tells us that the acceleration $a(t) = \frac{d^2q}{dt^2}$ is proportional to the external force $f = -\frac{\partial V}{\partial q}$, that is

$$a(t) = -\frac{\partial V}{\partial q} \,. \tag{152}$$

Eq.(152) describes the motion locally while variational principle gives a global view of particle orbits. We can describe the dynamics in either the configuration (Lagrangian formulation) or the phase space (Hamiltonian formulation). In configuration space, the important physical quantity is the Lagrangian

$$L(q,\dot{q},t) = \dot{q}^2/2 - V(q,t) \,. \tag{153}$$

The equation of motion is determined by requiring the variation of the action integral

$$S = \int_{t_1}^{t_2} dt\, L(q,\dot{q},t) \tag{154}$$

to be dependent only on the surface terms [65].

In the most general case, the path varies in the configuration space with

$$q(t) \to q'(t') = q(t) + \delta q(t) \,, \quad t \to t'(t) = t + \delta t(t) \,. \tag{155}$$

These equations define the variation δq and δt in terms of the original variables: q and t. The time change induces a change in the integration measure in (154) by

$$\delta(dt) = d(t + \delta t) - dt = dt\frac{d}{dt}\delta t(t)\,.$$

In many applications, the time is not varied and we use δ_0 instead of δ to indicate this variation, *i.e.* $\delta_0 t = 0$. The variation of $q(t)$ is then to the highest order given by

$$\delta q(t) = q'(t + \delta t) - q(t) = q'(t) - q(t) + \delta t\frac{dq'(t)}{dt} := \delta_0 q(t) + \delta t\frac{dq(t)}{dt}\,.$$

Similarly,

$$\delta\dot{q}(t) = \delta_0\dot{q}(t) + \delta t\frac{d}{dt}\dot{q} = \frac{d}{dt}(\delta q) - \dot{q}\frac{d}{dt}\delta t\,.$$

So, the variation of the action S is

$$\begin{aligned}
\delta S &= \int_{t_1}^{t_2} L\,\delta(dt) + \delta L\,dt \\
&= \int_{t_1}^{t_2} dt\,L\frac{d}{dt}\delta t + \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} + \frac{\partial L}{\partial t}\delta t\right) \\
&= \int_{t_1}^{t_2} dt\,\frac{d}{dt}(L\delta t + \frac{\partial L}{\partial \dot{q}}\delta q - \dot{q}\frac{\partial L}{\partial \dot{q}}\delta t) + \\
&\quad \int_{t_1}^{t_2} dt\,\left[\frac{\partial L}{\partial q} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}})\right]\delta q + \left[\frac{\partial L}{\partial t} + \frac{d}{dt}(\dot{q}\frac{\partial L}{\partial \dot{q}} - L)\right]\delta t\,.
\end{aligned} \tag{156}$$

As usual, we may define the conjugate momentum $p := \frac{\partial L}{\partial \dot{q}}$; and the Hamiltonian $H = \dot{q}\frac{\partial L}{\partial \dot{q}} - L$. According to our variational principle, the variation of the action with respect to a classical orbit only comes from the surface term,

$$\delta S = (p\,\delta q - H\,\delta t)|_{t_1}^{t_2}\,,$$

so the rest term in (156) should be identically zero. The variations δq give the equation

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0\,. \tag{157}$$

This is the famous Euler-Lagrangian equation and is able to produce the equation of motion in terms of the generalized coordinates. It is easy to see that if q is a cyclic coordinate (absent from L), the conjugate momentum p is a conserved quantity. The variations δt give the equation

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}\,. \tag{158}$$

If the potential V is independent of time, then $\frac{dH}{dt} = 0$, *i.e.*, $H = E$ is a constant. In other words, if time t is treated as a coordinate and is cyclic, Hamiltonian H would be the corresponding conjugate momentum and is conserved. In fact, (158) can be directly derived from (157) since we have the identity

$$\frac{\partial L}{\partial t} + \frac{d}{dt}\left(\dot{q}\frac{\partial L}{\partial \dot{q}} - L\right) = -\left[\frac{\partial L}{\partial q} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}})\right]\dot{q}\,.$$

With this observation, up to the surface term, δS depends only on $\delta_0 q$. The variation δt in time can well be contained in the position variation δq, which leaves some arbitrariness in the time parametrization of the loop as long as numerical calculation is concerned. For the same reson, the only equation that the particle trajectory has to satisfy is (157), the right hand side of which is the gradient of S with respect to δq (or $\delta_0 q$).

B.2 Newton descent and variational principle

If we are seeking a periodic orbit of known period T, we may set $\delta t = 0$, and the variational principle in this case reads

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta_0 q = 0 \,. \tag{159}$$

So, the periodic orbit is a critical point (minimum, maximum or saddle) of the action functional. If the periodic orbit is an extremum (minimum or maximum), gradient-descent method [208, 51, 81] could be used. The more general way is to solve directly the zero gradient condition (157). Go back to the simple case and assume that v does not contain time explicitly, *i.e.*, $L(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - v(q)$. The q gradient of S is then

$$g(q) := f(q) - \ddot{q} = \; a - \tilde{a} \,, \tag{160}$$

where $a := f(q)$ and $\tilde{a} := \ddot{q}$ make the notations consistent with those in [158]. We would like to find zeros of (160) in the loop space $\{q(t) : q(t) = q(t{+}T), t \in [0, T]\}$. Different approaches can be applied at this point. In view of the periodic boundary condition, Helleman and Bountis [120] derived the gradient $g(\hat{q})$ in Fourier space and use Newton's method to locate its zeros.

In the more complicated case, Newton descent is a efficient method for finding zeros of $g(q)$. Let's review the Newton descent method in the Hamiltonian case. Our starting equation is

$$\frac{\partial}{\partial \tau}(\tilde{a} - \lambda^2 a) = -(\tilde{a} - \lambda^2 a) \,. \tag{161}$$

As the period T is known by assumption it is convenient to set $\lambda = 1$. Then (161) is equivalent to

$$\frac{\partial}{\partial \tau}\, g(q) = -g(q) \,, \tag{162}$$

or more explicitly,

$$\frac{\delta g(q)}{\delta q} \cdot \frac{\partial q}{\partial \tau} = -g(q) \,. \tag{163}$$

This is just the infinitesimal version of Newton's method. As we make the loop deformation continuous, we call it *Newton descent*. Therefore basically, Newton descent is a variant of Newton's method for locating zeros of the stationary variation condition $g(q) = 0$. In practice, we can choose to work in the configuration q-space or in the Fourier $\hat{q}$-space. In q-space numerical calculation, the large matrix derived from the left hand side of (163) is sparse, which is easier for inversion, especially in high-dimensional space. When we work in $\hat{q}$-space, much fewer modes are sufficient to capture the essential shape of the periodic orbit if it is 'simple' (high frequency components are highly supressed).

Solve (163) for $\delta_0 q$ and the corresponding variation in S is given by

$$\delta S = -\int_{t_1}^{t_2} dt\, g(q) \left(\frac{\delta g}{\delta q}\right)^{-1} (q) g(q)\, \delta\tau \,. \tag{164}$$

If the loop $q(t)$ is close to a periodic orbit, we can check where δS heads. Near a minimum (maximum), all the eigenvalues of $\frac{\delta g}{\delta q}$ being positive (negative), δS is therefore decreasing (increasing) to make S approach its minimum (maximum) value. Near a saddle, situation becomes more complicated: δS could be positive or negative. But (162) ensures that $g(q)$ always approaches zero and thus $q(t)$ evolves toward a periodic orbit.

In chapter 3, we introduced the following cost functional

$$F^2[\tilde{x}] = \frac{1}{2\pi} \oint_{L(\tau)} ds\, \left(\tilde{v}(\tilde{x}) - \lambda v(\tilde{x})\right)^2 \,. \tag{165}$$

The Newton descent flow always decreases F^2 and every solution of (157) minimizes F^2. However, F^2 is not equivalent to the action integral S in the Lagrangian mechanics, because there may be local minima (maxima, saddle) of (165) that are not solutions to (157). So, stationary variation condition $\delta F^2 = 0$ may not yield solutions that are physical while $\delta S = 0$ always does. At a critical point S assumes any value while F^2 can only take value 0 in order to identify a true solution of the original equation. The Newton descent (162) directly searches zeros of $g(q)$, which is better than ordinary minimization of F^2 (see chapter 9 of [208]).

B.3 Conclusion

From the above comparison, we see that the basic equation of Lagrangian mechanics can be derived from the variational principle. The principle is also applicable analytically or numerically to determining periodic orbits, when the desired orbit is an extremum of the action integral. If the periodic orbit appears as a saddle which is the case in many instances, we cannot apply the principle directly. In practice, various versions of Newton's method are used to locate zeros of the gradient of the action. The Newton descent is one variant of Newton's method among others.