Lagrangeの未定乗数法 Lagrange Undertermined Multipliers

拘束のない系では，Lagrangian の中に現われたすべての $q_i(i=1,2,\cdots,n)$ を全部独立として変分を行なう。つまり q_i を $q_i+\delta q_i$ と変化させて，δq_i を全部独立の無限小として扱う。もし q_i がすべて独立なものでなく，ある束縛条件

$$g(q_1,q_2,\cdots,q_f)=0 \tag{1}$$

が存在している場合には，むろんすべての q_i を独立とみなすことはできない. 変分をとる場合も，$q_i+\delta q_i$ が束縛条件を満たすようにとられなければならない. そのときはいうまでもなく自由度は $f-1$ に減っている。δq_i の許される範囲は

$$g(q_1+\delta q_1,\cdots,q_f+\delta q_f)=g(q_1,\cdots,q_f)+\sum_i \frac{\partial g}{\partial q_i}\delta q_i=0$$

より

$$\sum_i \frac{\partial g}{\partial q_i}\delta q_i=0 \tag{2}$$

のものに限られる. (1) のもとに, たとえば関数

$$F(q_1,\cdots,q_f) \tag{3}$$

が極値をとるように q_i を定めるには, Lagrange によって考案された次の方法が有効である. これを Lagrarange の未定乗数法 (undetermined multip1iers) という. 今，(3) の変分をとりそれを 0 とおく. すなわち

$$\delta F=\sum_{i=1}^{f}\frac{\partial F}{\partial q_i}\delta q_i=0 \tag{4}$$

ところが，δq_i は (2) を満たすものに限られるから，(4) の δq_i の係数を直ちに 0 とおくわけにはいかない.(2) を用いて, たとえば δq_f を消去してから $\delta q_i(i=1,2,\cdots,f-1)$ の係数を 0 とすればよいが, それでは $i=1,2,\cdots,f$ の変数の間の対称性が失われて気持ちが悪いばかりでなく, 1 個以上の束縛条件がある場合, 独立でない δq_i を消去するのがたいへんである. そこで, このような不便をさけるために,(4) と (2) をいっしょにして

$$\delta F=\sum_{i=1}^{f}\left(\frac{\partial F}{\partial q_i}+\lambda\frac{\partial g}{\partial q_i}\right)\delta q_i=0 \tag{5}$$

を考えよう.(2)(4) が成り立つかぎり (5) も成り立つ. ここに λ は $q_1,\cdots,q_f$ の任意の関数でよいが, この任意性をうまく利用して, たとえば δq_f が消えるようにすることができる. それには,λ を

$$\frac{\partial F}{\partial q_f}+\lambda\frac{\partial g}{\partial q_f}=0 \tag{6}$$

と選べばよい. すると (5) から δq_f は消えてしまい

$$\delta F=\sum_{i=1}^{f-1}\left(\frac{\partial F}{\partial q_i}+\lambda\frac{\partial g}{\partial q_i}\right)\delta q_i=0 \tag{7}$$

となる.$\delta q_i(i=1,\cdots,f-1)$ は独立だから,(7) を満たすものは

$$\frac{\partial F}{\partial q_i}+\lambda\frac{\partial g}{\partial q_i}=0,\quad i=1,2,\cdots,f-1 \tag{8}$$

に限られる. このとき λ は (6) を満たすようなものである. したがって (6) と (8) をいっしょにして

$$\frac{\partial F}{\partial q_i}+\lambda\frac{\partial g}{\partial q_i}=0,\quad i=1,2,\cdots,f \tag{9}$$

となる. すなわち (6) を満たすように λ を決めて, その λ を用いて (8) とおくと, はじめの関数 F は条件 (1) のもとで極値をとる. 以上の手続きをさらに手ぎわよく行なうには,F の代わりに

$$\bar{F} \equiv F + \lambda g \tag{10}$$

を定義し,$\bar{F}$ を独立な $q_1,,,q_f,\lambda$ に対して極値をとるように定めればよい。この場合,(1) のあったことを忘れて変分をとると,(1) が結果として出てくる。$f+1$ 個の独立変数 $q_1,\cdots,q_f,\lambda$ をそれぞれ $q_1+\delta q_1,\cdots,q_f+\delta q_f,\lambda+\delta\lambda$ と変えると

$$\delta\bar{F} = \sum_{i=1}^{f}\left(\frac{\partial F}{\partial q_i} + \lambda\frac{\partial g}{\partial q_i}\right)\delta q_i + g\delta\lambda = 0 \tag{11}$$

これを 0 とおくと, $\delta q_i (i=1,2,\cdots,f)$ と $\delta\lambda$ は全部独立だから

$$\frac{\partial F}{\partial q_i} + \lambda\frac{\partial g}{\partial q_i} = 0, \quad i=1,2,\cdots,f \tag{12}$$

$$g = 0 \tag{13}$$

が得られる.(12,13) は, 条件 (1) のもとに,f が極値をとる条件と全く同じである.λ のことを,Lagrange の未定乗数という。条件つきの極値問題では, しばしば利用される便利な量である.

もし, 束縛条件が n 個 $(n<f)$ あれば,n 個の未定乗数を用いて

$$\bar{F} \equiv F + \sum_{\mu=1}^{n}\lambda_\mu g_\mu \tag{14}$$

を作り,$q_1,\cdots,q_f,\lambda_1,\cdots,\lambda_n$ を全部独立変数として扱えば,F の極値は,

$$\frac{\partial F}{\partial q_i} + \sum_{\mu=1}^{n}\lambda_\mu\frac{\partial g_\mu}{\partial q_i} = 0, \quad i=1,2,\cdots,f \tag{15}$$

$$g_\mu = 0, \quad \mu=1,2,\cdots,n \tag{16}$$

で与えられる.

例 1: 簡単な微分の問題として,

$$f(x,y) = (x^2+y^2)/2 \tag{17}$$

が, 条件

$$x+y=a \tag{18}$$

のもとに極値をとる点を求めてみよう. Lagrange の未定乗数によらないで行なうには, まず y を消去する。すると

$$f(x,a-x) = x^2 - ax + a^2/2 \tag{19}$$

が得られるから, これを x で微分して 0 とおくと

$$f'(x,a-x) = 2x - a = 0 \tag{20}$$

$$x = a/2 \tag{21}$$

一方,Lagrange の未定乗数法によって同じ問題を扱うには,

$$\bar{f}(x,y,\lambda) \equiv f(x,y) + \lambda(x+y-a) \tag{22}$$

を定義し, これを,x,y,λ を全部独立として極値を求める. それには,

$$\frac{\partial\bar{f}}{\partial x} = x+\lambda = 0 \tag{23}$$

$$\frac{\partial\bar{f}}{\partial y} = y+\lambda = 0 \tag{24}$$

$$\frac{\partial\bar{f}}{\partial\lambda} = x+y-a = 0 \tag{25}$$

を解けばぽよい. はじめの 2 式から λ を消去して, 第 3 式と組み合わせると,

$$x = a/2, \quad y = a/2 \tag{26}$$

得られて,(21) と完全に一致することがわかる.

1 Euler-Lagrange equtaion of motion with constraints

Now we can write the Lagrangian of the system

$$L = L(q_1, q_2, \cdots, q_N, \dot{q}_1, \dot{2}_2, \cdots, \dot{q}_N) \tag{27}$$

The action integral S is written by

$$S = \int_{t_1}^{t_2} dt L(q_1, q_2, \cdots, q_N, \dot{q}_1, \dot{2}_2, \cdots, \dot{q}_N) \tag{28}$$

The Hamilton principle says that the motion of which the action integral has extremal value is the actual motion. This pinciple gives us the Eular-Langrange equation.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \cdots, f \tag{29}$$

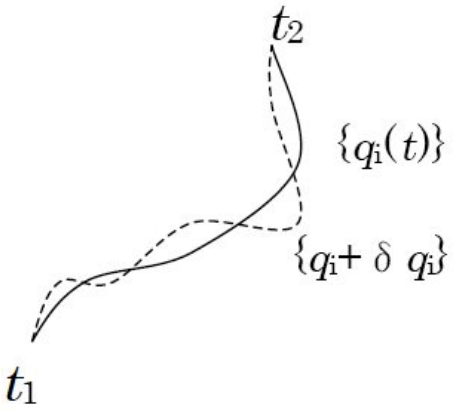

Figure 1: Action integral

Now we consider the holonomic constaraint

$$g_k(q_1, q_2, \cdots, q_f) = 0, \qquad k = 1, 2, \cdots, K \tag{30}$$

Using the Lagrnage undetermined multipliers method

$$\begin{aligned}
\delta \int_{t_1}^{t^2} dt(L + \sum_k^K \lambda_k g_k) &= \int_{t_1}^{t_2} dt \left[L(\{q_i + \delta q_i\}, \{\dot{q}_i + \delta \dot{q}_i\}) + \sum_k^K \lambda_k g_k(\{q_i + \delta q_i\}) \right] \\
&\quad - \int_{t_1}^{t_2} dt \left[L(\{q_i\}, \{\dot{q}_i\}) + \sum_k^K \lambda_k g_k(\{q_i\}) \right] && (31) \\
&= \int_{t_1}^{t^2} dt \sum_i^f \left[\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \sum_k^K \lambda_k \frac{\partial g_k}{\partial q_i}\delta q_i \right] && (32) \\
&= \int_{t_1}^{t^2} dt \sum_i^f \left[\frac{\partial L}{\partial q_i}\delta q_i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i + \sum_k^K \lambda_k \frac{\partial g_k}{\partial q_i}\delta q_i \right] && (33) \\
&= \sum_i^f \int_{t_1}^{t^2} dt \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \sum_k^K \lambda_k \frac{\partial g_k}{\partial q_i} \right] \delta q_i && (34)
\end{aligned}$$

Here we used

$$\int_{t_1}^{t_2} dt \frac{\partial L}{\partial \dot{q}_i}\frac{d\delta q_i}{dt} = \left[\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i \tag{35}$$

and $\delta q(t_1) = \delta q(t_2) = 0$. For arbitary δq_i we have $\delta(L + \sum \lambda_k g_k) = 0$ if

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \sum_k^K \lambda_k \frac{\partial g_k}{\partial q_i} \tag{36}$$

2 Constraint dynamics: SHAKE to RATTLE

To reduce the computaion, one sometimes want to handle the dynamics of a molecular system in which certain arbitrarily selected degrees of freedom (such as bond lengths) are constrained, while others remain free to evolve under the influence of intermolecular and intramolecular forces.

SHAKE is based on the Verlet algorithm. The Verlet algorithm has some drawbacks. The velocities of the atoms are not among the variables used in integrating the equations of motion, and they can be obtained only with extra effort or strorage. The velocity version of Verlet algorithm eliminates these problems. The velocity version of SHAKE algorithm is called RATTLE, and it has two advantages over SHAKE. It has higher precision and deals velocities directly. The latter advandatge is vely useful to carry out constant temerature and pressure simulation and nonequilibrium sumulation.

The constraint conditions $\sigma_{ij}(\{\mathbf{r}\})$ may be given by [1]

$$\sigma_{ij}(\{\mathbf{r}\}) \equiv (\mathbf{r}_i - \mathbf{r}_j)^2 - d_{ij}^2 = 0 \tag{39}$$

The Lagrange equation of motion under these constraints becomes

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}_i}\right) - \frac{\partial L}{\partial \mathbf{r}_i} = -\sum_j{}' \lambda_{ij}(t)\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i} \tag{40}$$

Here $\{\lambda_{ij}\}$ are the time-dependent Lagrange undetermined multipliers and the prime denotes a summation over only those atoms j that are connected with atom i by a constraint. Note that $\sigma_{ij} = \sigma_{ji}, \lambda_{ij} = \lambda_{ji}$

$$\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i(t)} = 2[\mathbf{r}_i(t) - \mathbf{r}_j(t)] \tag{41}$$

As the Lagrangean is given by

$$L = \sum_i \frac{1}{2} m_i \mathbf{r}_i^2 - V(\{\mathbf{r}_i(t)\}), \tag{42}$$

then the equation of motion becomes

$$m_i \ddot{\mathbf{r}}_i = \vec{\mathbf{F}}_i + \sum_j{}' \lambda_{ij}(t)\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i}. \tag{43}$$

$$\frac{d\sigma_{ij}}{dt} = 2[\dot{\mathbf{r}}_i(t) - \dot{\mathbf{r}}_j(t)] \cdot [\mathbf{r}_i(t) - \mathbf{r}_j(t)] = 0 \tag{44}$$

$$\begin{aligned}
\vec{\mathbf{r}}_i(t+\delta t) &= \vec{\mathbf{r}}_i(t) + \delta t \vec{\mathbf{v}}_i(t) + \\
&\quad + \frac{(\delta t)^2}{2m_i}\left[\vec{\mathbf{F}}_i(t) - \sum_j{}' \lambda_{ij}(t)\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i}\right] \\
&= \vec{\mathbf{r}}_i(t) + \delta t \vec{\mathbf{v}}_i(t) + \frac{(\delta t)^2}{2m_i}\Big[\vec{\mathbf{F}}_i(t) \\
&\quad - \sum_j{}' 2\lambda_{ij}^{RR}(t)\{\mathbf{r}_i(t) - \mathbf{r}_j(t)\}\Big]
\end{aligned} \tag{45}$$

and

$$\begin{aligned}
\vec{\mathbf{v}}_i(t+\delta t) &= \vec{\mathbf{v}}_i(t) + \frac{\delta t}{2m_i}[\vec{\mathbf{f}}_i(t+\delta t) + \vec{\mathbf{f}}_i(t)] \\
&\quad - \frac{\delta t}{m_i}\left[\sum_j{}' \lambda_{ij}^{RR}(t)\{\vec{\mathbf{r}}_i(t) - \vec{\mathbf{r}}_j(t)\} + \right. \\
&\quad \left. \sum_j{}' \lambda_{ij}^{RV}(t+\delta t)\{\vec{\mathbf{r}}_i(t+\delta t) - \vec{\mathbf{r}}_j(t+\delta t)\}\right]
\end{aligned} \tag{46}$$

[1] Question: How about

$$g = z - z_0 = 0 \tag{37}$$

In this case the total momentum of the system is not conserved then new constraint

$$g_1 = \sum_i m_i \dot{z}_i = 0 \tag{38}$$

should be also considered. (ref: Okazaki MD book)

2.1 Iterative procedure for RATTLE calculations

Suppose that the positions, velocities, and intermolecular forces are known at time t, and we wish to calculate the corresponding quantities for the time $t + \delta t$. This can be done by a straightforward modification of the iterative procedure of Ryckaert et al. [4] for SHAKE. Let us define

$$g_{ij} = \delta t \lambda_{ij}^{RR}(t), \tag{47}$$

$$k_{ij} = \delta t \lambda_{ij}^{RV}(t + \delta t), \tag{48}$$

$$\vec{\mathbf{q}}_i = \dot{\mathbf{r}}_i(t) + (\delta t/2m_i)\vec{\mathbf{F}}_i(t) - (1/m_i)\sum_i{}' g_{ij}\vec{\mathbf{r}}_{ij}(t) \tag{49}$$

Then the above Eqs. can be expressed as

$$\vec{\mathbf{r}}_i(t + \delta t) = \vec{\mathbf{r}}_i(t) + \delta t \vec{\mathbf{q}}_i, \tag{50}$$

$$\dot{\mathbf{r}}_i(t + \delta t) = \vec{\mathbf{q}}_i + (\delta t/2m_i)\vec{\mathbf{F}}_i(t + \delta t) - (1/m_i)\sum_i{}' k_{ij}\vec{\mathbf{r}}_{ij}(t + \delta t) \tag{51}$$

First we solve for the $\vec{\mathbf{q}}_i$ by interation. To start, we let

$$\vec{\mathbf{q}}_i = \dot{\mathbf{r}}_i(t) + (\delta t/2m_i)\vec{\mathbf{F}}_i(t) \qquad i = 1, ...N \tag{52}$$

At this point the iterative loop begins. Pick a constraint. Suppose it involves atoms i and j. Let

$$\mathbf{s} = \mathbf{r}_i(t) + \delta t \mathbf{q}_i(t) - \mathbf{r}_j(t) - \delta t \mathbf{q}_j(t) \tag{53}$$

Then $\mathbf{s}$ is the current approximation for the vector displacement of atoms i and j. If $|\mathbf{s}|^2 - d_{ij}^2$ differs from zero by an amount less than an acceptable tolerance, go to the beginning of the iterative loop and pick a new constraint. If not, then we want to find corrections for $\mathbf{q}_i$ and $\mathbf{q}_j$ to make the constraint be satisfied more closely. Let

$$\mathbf{r}_i^T = \mathbf{r}_i(t) + \delta t[\mathbf{q}_i - g\mathbf{r}_{ij}(t)/m_i] \tag{54}$$

and

$$\mathbf{r}_j^T = \mathbf{r}_j(t) + \delta t[\mathbf{q}_j + g\mathbf{r}_{ij}(t)/m_j] \tag{55}$$

These are the new values for $\mathbf{r}_i(t+\delta t)$ and $\mathbf{r}_j(t+\delta t)$, when the corrections proportional to g are made to $\mathbf{q}_i$ and $\mathbf{q}_j$. We want to choose g so that

$$|\mathbf{r}_i^T - \mathbf{r}_j^T|^2 = d_{ij}^2 \tag{56}$$

Solving for g we find

$$g = (s^2 - d_{ij}^2)/\{2\delta t[\mathbf{s} \cdot \mathbf{r}_{ij}](m_i^{-1} + m_j^{-1})\} \tag{57}$$

which we have neglected quantities of order g^2. Then we replace $\mathbf{q}_i$ by the old value of $\mathbf{q}_i$ minus $g\mathbf{r}_{ij}(t)/m_i$ and $\mathbf{q}_j$ by the old value of $\mathbf{q}_j$ plus $g\mathbf{r}_{ij}(t)/m_j$, go to the beginning of the iterative loop, and choose a new constraint. This iterative procedure is continued until all the constraints are satisfied to within the acceptable tolerance.

The procedure converges to the correct result, At each stage of the iterative procedure, the $\mathbf{q}$'s are corrected by an amount of the proper form, and the procedure terminates when all the constraints on the interatomic distances are satisfied to within the desired accuracy.

Now that $\mathbf{q}_i$. and $\mathbf{r}_i(t + \delta t)$ are known fof all i, th forces at time $t + \delta t$ can be calculated. Just before doing this, the positions at time $t+\delta t$ should be placed in the memory locations that previously held the positions at time t, and the $\mathbf{q}_i, i = 1, , ..N$, should be placed in the memory locations that previously held the velocities at time t. This allows the algorithm to be implemented using just $3N$ memory locations for N degrees of freedom.

Next we solve for the $\dot{\mathbf{r}}_i(t + \delta t)$ by iteration. To start, we let

$$\dot{\mathbf{r}}_i(t + \delta t) = \mathbf{q}_i + \delta \mathbf{F}_i(t + \delta t)/2m_i, \qquad i = 1, ..., N. \tag{58}$$

At this point the iterative loop begins. Pick a constraint. Suppose it involves atom i and j. Calculate the dot product of $\mathbf{r}_{ij}(t+\delta t)$ and $\dot{\mathbf{r}}_{ij}(t+\delta t)$. If it differs from zero by less than an acceptable

tolerance, then go to the beginning of the iterative loop and pick another constraint. If it differs from zero by more than the acceptable tolerance, then we want to correct the two velocities, $\dot{\mathbf{r}}_i$ and $\dot{\mathbf{r}}_j$. Let

$$\dot{\mathbf{r}}_i^T = \dot{\mathbf{r}}_i(t+\delta t) - k\mathbf{r}_{ij}(t+\delta t)/m_i \tag{59}$$

and

$$\dot{\mathbf{r}}_j^T = \dot{\mathbf{r}}_j(t+\delta t) + k\mathbf{r}_{ij}(t+\delta t)/m_j \tag{60}$$

These are the new values of $\dot{\mathbf{r}}_i(t+\delta t)$ and $\dot{\mathbf{r}}_j(t+\delta t)$ when corrections proportional to k are made. We want to choose k so that $\dot{\mathbf{r}}_i^T - \dot{\mathbf{r}}_j^T$ is perpendicular to $\mathbf{r}_{ij}(t+\delta t)$. This leads to the following choice:

$$k = \mathbf{r}_{ij}(t+\delta t)\cdot[\dot{\mathbf{r}}_i(t+\delta t) - \dot{\mathbf{r}}_j(t+\delta t)]/\{d_{ij}^2(m_i^{-1}+m_j^{-1})\}. \tag{61}$$

Then we replace $\dot{\mathbf{r}}_i(t+\delta t)$ by $\dot{\mathbf{r}}_i^T$, and $\dot{\mathbf{r}}_j(t+\delta t)$ by $\dot{\mathbf{r}}_j^T$, go to the beginning of the iterative loop, and pick another constraint.

This procedure converges to the correct result. At each stage of the iterative procedure, the $\dot{\mathbf{r}}_i(t+\delta t)$ are corrected by an amount of the proper form. and the procedure terminates when all the constraints on the velocities are satisrred to within the desired accuracy.

[REFERENCES:Hans C. Andersen, J. Comput. Phys. 52, 24-34 (1983).]