APPENDIX N

Lagrange Multipliers Method

Imagine a Cartesian coordinate system of $n + m$ dimension with the axes labeled $x_1, x_2, \ldots, x_{n+m}$ and a function[1] $E(\boldsymbol{x})$, where $\boldsymbol{x} = (x_1, x_2, \ldots, x_{n+m})$. Suppose that we are interested in finding the lowest value of E, but only among such $\boldsymbol{x}$ that satisfy m conditions (*conditional extremum*):

$$W_i(\boldsymbol{x}) = 0 \tag{N.1}$$

for $i = 1, 2, \ldots, m$. The constraints cause the number of the *independent* variables to be n.

If we calculated the differential $\mathrm{d}E$ in point $\boldsymbol{x}_0$ that corresponds to an extremum of E, then we obtain 0:

$$0 = \sum_{j=1}^{n+m} \left(\frac{\partial E}{\partial x_j} \right)_0 \mathrm{d}x_j, \tag{N.2}$$

where the derivatives are computed at the point of the extremum. The quantities $\mathrm{d}x_j$ stand for the infinitesimally small increments. From Eq. (N.2), we cannot draw the conclusion that $\left(\frac{\partial E}{\partial x_j}\right)_0$ equals 0. This would be true if the increments $\mathrm{d}x_j$ were independent, but they are not. Indeed, one finds the relations between them by making the differentials of the conditions W_i:

$$\sum_{j=1}^{n+m} \left(\frac{\partial W_i}{\partial x_j} \right)_0 \mathrm{d}x_j = 0 \tag{N.3}$$

for $i = 1, 2, \ldots, m$ (the derivatives are computed for the extremum).

This means that the number of the truly independent increments is only n. Let us try to exploit that. To this end, let us multiply Eq. (N.3) by a number ϵ_i (*Lagrange multiplier*),

Joseph Louis de Lagrange (1736–1813), French mathematician of Italian origin, self-taught, and professor at the Artillery School of Torino, then at École Normale Supérieure in Paris. His main achievements are in variational calculus, mechanics, number theory, algebra, and mathematical analysis.

[1] Symbol E is chosen to suggest that in our applications, the quantity will have the meaning of energy.

Ideas of Quantum Chemistry, Second Edition. http://dx.doi.org/10.1016/B978-0-444-59436-5.00034-9

which will be fixed in a moment. Then, let us add together all the conditions [Eq. (N.3)], and subtract the result from Eq. (N.2). One gets

$$\sum_{j=1}^{n+m}\left[\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0\right]\mathrm{d}x_j=0,$$

where the summation extends over $n+m$ terms. The summation may be carried out in two steps. First, let us sum up the first n terms, and afterward add together the other terms:

$$\sum_{j=1}^{n}\left[\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0\right]\mathrm{d}x_j+\sum_{j=n+1}^{n+m}\left[\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0\right]\mathrm{d}x_j=0.$$

The multipliers ϵ_i had been treated until now as undetermined. Well, we may force them to make each of the terms in the second summation to equal zero:[2]

$$\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0=0 \tag{N.4}$$

for $j=n+1,\ldots,n+m$.

Hence, the first summation alone is 0:

$$\sum_{j=1}^{n}\left[\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0\right]\mathrm{d}x_j=0,$$

which means that now we have only n increments $\mathrm{d}x_j$ and therefore they are *independent*. Since for any (small) $\mathrm{d}x_j$ the sum is always 0, then the only reason for that could be that each parenthesis [] equals zero individually:

$$\left(\frac{\partial E}{\partial x_j}\right)_0-\sum_i \epsilon_i\left(\frac{\partial W_i}{\partial x_j}\right)_0=0 \qquad \text{for } j=1,\ldots,n.$$

This set of n equations (the so-called *Euler equation*), together with the m conditions Eq. (N.1) and m Eq. (N.4) gives a set of $n+2m$ equations with $n+2m$ unknowns (m epsilons and $n+m$ components x_i of the vector $\boldsymbol{x}_0$).

[2] This is possible if the determinant build of coefficients $\left(\frac{\partial W_i}{\partial x_j}\right)_0$ is nonzero (this is what we have to assume). For example, if several conditions were identical, then the determinant would be zero.

For a conditional extremum, the constraint $W_i(\boldsymbol{x}) = 0$ has to be satisfied for $i = 1, 2, \ldots, m$ and $\left(\frac{\partial E}{\partial x_j}\right)_0 - \sum_i \epsilon_i \left(\frac{\partial W_i}{\partial x_j}\right)_0 = 0$ for $j = 1, \ldots, n+m$. The x_i found from these equations determine the position $\boldsymbol{x}_0$ of the conditional extremum E.

Whether it is a minimum, a maximum, or a saddle point determines the analysis of the matrix of the second derivatives (Hessian). If its eigenvalues computed at $\boldsymbol{x}_0$ are all positive (negative), then it is a minimum[3] (maximum); otherwise, it is a saddle point.

Example 1. ***Minimizing a Paraboloid Going Along a Straight Line Off Center***

Let us take a paraboloid

$$E(x, y) = x^2 + y^2.$$

This function has, of course, a minimum at (0, 0), but the minimum is of no interest to us. What we want to find is a minimum of E, but only *when* x and y satisfy some conditions. In our case, there will be only one of them:

$$W = \frac{1}{2}x - \frac{3}{2} - y = 0. \tag{N.5}$$

This means that we are interested in a minimum of E *when going along a straight line* $y = \frac{1}{2}x - \frac{3}{2}$.

The Lagrange multipliers method works as follows:

- We differentiate W and multiply by an unknown (Lagrange) multiplier ϵ, thus getting: $\epsilon(\frac{1}{2}dx - dy) = 0$.
- This result (i.e., 0) is subtracted[4] from $dE = 2xdx + 2ydy = 0$, and we obtain $dE = 2xdx + 2ydy - \frac{1}{2}\epsilon dx + \epsilon dy = 0$.
- In the last expression, the coefficients at dx and at dy have to equal zero.[5] In this way, we obtain two equations: $2x - \frac{1}{2}\epsilon = 0$ and $2y + \epsilon = 0$.
- The third equation needed is the constraint $y = \frac{1}{2}x - \frac{3}{2}$.
- The solution to this three equations gives a set of x, y, ϵ that corresponds to an extremum. We obtain $x = \frac{3}{5}, y = -\frac{6}{5}, \epsilon = \frac{12}{5}$. Thus, we have obtained not only the position of the minimum: $x = \frac{3}{5}, y = -\frac{6}{5}$, but also the Lagrange multiplier ϵ. The minimum value of E, which has been encountered when going along the straight line $y = \frac{1}{2}x - \frac{3}{2}$, is equal to $E\left(\frac{3}{5}, -\frac{6}{5}\right) = \left(\frac{3}{5}\right)^2 + \left(-\frac{6}{5}\right)^2 = \frac{9+36}{25} = \frac{9}{5}$.

[3] In this way, we find a minimum; no information is available, whether it is global or local.

[4] It also can be added–that does not matter (in such a case, we get another value of ϵ).

[5] Only now is this possible.

Example 2. ***Minimizing a Paraboloid Going Along a Circle (Off Center)***

Let us take the same paraboloid [Eq. (N.5)], but put in another constraint:

$$W = (x - 1)^2 + y^2 - 1 = 0. \tag{N.6}$$

This condition means that we want to go around a circle of radius 1 centered at $(1, 0)$ and watch for which point (x, y) we will have the lowest value[6] of E. The example is chosen in such a way as to answer the question first *without any calculations*. Indeed, the circle goes through $(0, 0)$, so this point has to be found as a minimum. Besides that, we should find a maximum at $(2, 0)$ because this is the point of the circle that is most distant from $(0, 0)$.

Well, let us see whether the Lagrange multipliers method will give the same result.

After differentiation of W, multiplying it by the multiplier ϵ, subtracting the result from dE and rearranging the terms, we obtain

$$dE = [2x - \epsilon(2x - 2)]dx + 2y(1 - \epsilon)dy = 0,$$

which (after forcing the coefficients at dx and dy to be zero) gives the set of three equations:

$$\begin{aligned} 2x - \epsilon(2x - 2) &= 0, \\ 2y(1 - \epsilon) &= 0, \\ (x - 1)^2 + y^2 &= 1. \end{aligned}$$

Check that this set has the following solutions: $(x, y, \epsilon) = (0, 0, 0)$ and $(x, y, \epsilon) = (2, 0, 2)$. The solution $(x, y) = (0, 0)$ corresponds to the minimum, while the solution $(x, y) = (2, 0)$ corresponds to the maximum.[7] This is what we expected to get.

Example 3. ***Minimizing the Mean Value of the Harmonic Oscillator Hamiltonian***

This example is different: it will pertain to the extremum of a *functional*.[8] *This is what we are often going to encounter in the methods of quantum chemistry.* Let us take the energy functional

$$E[\phi] = \int_{-\infty}^{\infty} dx \phi^* \hat{H} \phi \equiv \langle \phi | \hat{H} \phi \rangle,$$

where $\hat{H}$ stands for the harmonic oscillator Hamiltonian: $\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2$. If we were asked about what function ϕ ensures the minimum value of $E[\phi]$, then such a function could be found right away–it is $\phi = 0$. This happens because the kinetic energy integral, as well as the potential energy integral, are positive numbers, except the situation when $\phi = 0$, where the result is zero. But this is not what we have thought. We want that ϕ has a probabilistic interpretation, as any wave function, and therefore $\langle \phi | \phi \rangle = 1$, and not zero. Well, in such a case, we want

[6] Or, in other words, we intersect the paraboloid by the cylinder surface of radius 1 and the cylinder axis (parallel to the symmetry axis of the paraboloid) shifted to $(1, 0)$.

[7] The method does not give us information about the kind of extremum found.

[8] The argument of a functional is a function that produces the value of the functional (a number).

to minimize $E\left[\phi\right]$, *but forcing the normalization condition is always satisfied.* Therefore, we search for the extremum of $E\left[\phi\right]$ *with the condition* $W = \langle\phi|\phi\rangle - 1 = 0$. It is easy to foresee that what the method has to produce (if it is of any value) is the normalized ground-state wave function for the harmonic oscillator. How will the Lagrange multipliers method get this result?

The answer is on p. 234.