Derivation of Newton's Law from Lagrangian and general gradient.

Peeter Joot

August 9, 2008

1

In the classical limit the Lagrangian action for a point particle in a general position dependent field is:

$$S = \frac{1}{2} m \mathbf{v}^2 - \varphi \tag{1}$$

Given the Lagrange equations that minimize the action, it is fairly simple to derive the Newtonian force law.

$$\begin{aligned} 0 &= \frac{\partial S}{\partial x^i} - \frac{d}{dt} \frac{\partial S}{\partial \dot{x}^i} \\ &= -\frac{\partial \varphi}{\partial x^i} - \frac{d}{dt} \left(m \dot{x}^i \right) \end{aligned}$$

Multiplication of this result with the unit vector $\mathbf{e}_i$, and summing over all unit vectors we have:

$$\sum \mathbf{e}_i \frac{d}{dt} \left(m \dot{x}^i \right) = -\sum \mathbf{e}_i \frac{\partial \varphi}{\partial x^i}$$

Or, using the gradient operator, and writing $\mathbf{v} = \sum \mathbf{e}_i \dot{x}^i$, we have:

$$\mathbf{F} = \frac{d(m\mathbf{v})}{dt} = -\nabla \varphi \tag{2}$$

1.1 The mistake hiding above.

Now, despite the use of upper and lower pairs of indexes for the basis vectors and coordinates, this result is not valid for a general set of basis vectors. This initially confused the author, since the RHS sum $\mathbf{v} = \sum \mathbf{e}_i v^i$ is valid for any set of basis vectors independent of the orthonormality of that set of basis vectors. This is assuming that these coordinate pairs follow the usual reciprocal relationships:

$$\mathbf{x} = \sum \mathbf{e}_i x^i$$

$$x^i = \mathbf{x} \cdot \mathbf{e}^i$$

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta^i{}_j$$

However, the LHS that implicitly defines the gradient as:

$$\nabla = \sum \mathbf{e}_i \frac{\partial}{\partial x^i}$$

is a result that is only valid when the set of basis vectors $\mathbf{e}_i$ is orthonormal. The general result is expected instead to be:

$$\nabla = \sum \mathbf{e}^i \frac{\partial}{\partial x^i}$$

This is how the gradient is defined (without motivation) in Doran/Lasenby. One can however demonstrate that this definition, and not $\nabla = \sum \mathbf{e}_i \frac{\partial}{\partial x^i}$, is required by doing a computation of something like $\nabla \|\mathbf{x}\|^\alpha$ with $\mathbf{x} = \sum x^i \mathbf{e}_i$ for a general basis $\mathbf{e}_i$ to demonstrate this. An example of this can be found in the appendix below.

So where did things go wrong? It was in one of the "obvious" skipped steps: $\mathbf{v} = \sum \dot{x}^i \dot{x}^i$. It is in that spot where there is a hidden orthonormal frame vector requirement since a general basis will have mixed product terms too (ie: non-diagonal metric tensor).

Expressed in full for general frame vectors the action to minimize is the following:

$$S = \frac{1}{2} m \sum \dot{x}^i \dot{x}^j \mathbf{e}_i \cdot \mathbf{e}_j - \varphi \tag{3}$$

Or, expressed using a metric tensor $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$, this is:

$$S = \frac{1}{2} m \sum \dot{x}^i \dot{x}^j g_{ij} - \varphi \tag{4}$$

1.2 Equations of motion for vectors in a general frame.

Now we are in shape to properly calculate the equations of motion from the Lagrangian action minimization equations.

$$
\begin{aligned}
0 &= \frac{\partial S}{\partial x^k} - \frac{d}{dt}\frac{\partial S}{\partial \dot{x}^k} \\
&= -\frac{\partial \varphi}{\partial x^k} - \frac{d}{dt}\left(\frac{1}{2}m\sum g_{ij}\frac{\partial \dot{x}^i}{\partial \dot{x}^k}\dot{x}^j + \dot{x}^i\frac{\partial \dot{x}^j}{\partial \dot{x}^k}\right) \\
&= -\frac{\partial \varphi}{\partial x^k} - \frac{d}{dt}\left(\frac{1}{2}m\sum g_{ij}({\delta^i}_k\dot{x}^j + \dot{x}^i{\delta^j}_k)\right) \\
&= -\frac{\partial \varphi}{\partial x^k} - \frac{d}{dt}\left(\frac{1}{2}m\sum(g_{kj}\dot{x}^j + g_{ik}\dot{x}^i)\right) \\
&= -\frac{\partial \varphi}{\partial x^k} - \frac{d}{dt}\left(m\sum g_{kj}\dot{x}^j\right) \\
\frac{d}{dt}\left(m\sum g_{kj}\dot{x}^j\right) &= -\frac{\partial \varphi}{\partial x^k} \\
\sum \mathbf{e}^k\frac{d}{dt}\left(m\sum g_{kj}\dot{x}^j\right) &= -\sum \mathbf{e}^k\frac{\partial \varphi}{\partial x^k} \\
\frac{d}{dt}\left(m\sum_j \dot{x}^j \underbrace{\sum_k \mathbf{e}^k\mathbf{e}_k \cdot \mathbf{e}_j}_{=\mathbf{e}_j}\right) &= \\
\frac{d}{dt}\left(m\sum_j \dot{x}^j\mathbf{e}_j\right) &=
\end{aligned}
$$

The requirement for reciprocal pairs of coordinates and basis frame vectors is due to the summation $\mathbf{v} = \sum \mathbf{e}_i\dot{x}^i$, and it allows us to write all of the Lagrangian equations in vector form for an arbitrary frame basis as:

$$
\mathbf{F} = \frac{d(m\mathbf{v})}{dt} = -\sum \mathbf{e}^k\frac{\partial \varphi}{\partial x^k} \tag{5}
$$

If we are calling this RHS a gradient relationship in an orthonormal frame, we therefore must define the following as the gradient for the general frame:

$$
\nabla = \sum \mathbf{e}^k\frac{\partial}{\partial x^k} \tag{6}
$$

The Lagrange equations that minimize the action still generate equations of motion that hold when the coordinate and basis vectors cannot be summed in this fashion. In such a case, however, the ability to merge the generalized coordinate equations of motion into a single vector relationship will not be possible.

2 Appendix. Scratch calculations.

2.1 frame vector in terms of metric tensor, and reciprocal pairs.

$$
\begin{aligned}
e_j &= \sum a_k e^k \\
e_j \cdot e_k &= \sum a_i e^i \cdot e_k \\
e_j \cdot e_k &= a_k \\
&\implies \\
e_j &= \sum e_j \cdot e_k e^k \\
e_j &= \sum g_{jk} e^k
\end{aligned}
$$

2.2 Gradient calculation for an absolute vector magnitude function.

As a verification that the gradient as defined in equation 6 works as expected, lets do a calculation that we know the answer as computed with an orthonormal basis.

$$
f(\mathbf{r}) = \|\mathbf{r}\|^{\alpha}
$$

$$
\begin{aligned}
\nabla f(\mathbf{r}) &= \nabla \|\mathbf{r}\|^{\alpha} \\
&= \sum \mathbf{e}^k \frac{\partial}{\partial x^k} \left(\sum x^i x^j g_{ij} \right)^{\alpha/2} \\
&= \frac{\alpha}{2} \sum \mathbf{e}^k \left(\sum x^i x^j g_{ij} \right)^{\alpha/2-1} \quad \frac{\partial}{\partial x^k} \left(\sum x^i x^j g_{ij} \right) \\
&= \alpha \|\mathbf{r}\|^{\alpha-2} \sum \mathbf{e}^k x^i g_{ki} \\
&= \alpha \|\mathbf{r}\|^{\alpha-2} \sum_i x^i \underbrace{\sum_k \mathbf{e}^k \mathbf{e}_k \cdot \mathbf{e}_i}_{=\mathbf{e}_i} \\
&= \alpha \|\mathbf{r}\|^{\alpha-2} \mathbf{r}
\end{aligned}
$$