Supplement —Phys410—Fall 2011 Prof. Ted Jacobson
Room 4115, (301)405-6020
www.physics.umd.edu/grt/taj/410a/ jacobson@umd.edu

Lagrange multipliers and constraints

Lagrange multipliers

To explain this let me begin with a simple example from multivariable calculus: suppose $f(x, y, z)$ is constant on the $z = 0$ surface. Then although we can't say that $\nabla f = 0$ when $z = 0$, we can say $\nabla f = w\hat{\mathbf{z}}$ when $z = 0$. For notational simplicity (as well as a hidden agenda), I'll use df for the gradient of f, so the last equation would be written $df = w\,\hat{\mathbf{z}}$.

Now let's generalize this example to consider a function that is constant on a surface defined by a constraint equation $C(x, y, z) = 0$. Then df must be parallel to dC when $C = 0$. Put differently,

$$df = w\,dC \quad \text{when} \quad C = 0, \text{ for some function } w. \tag{1}$$

In the example $C = z$, so $dC = \hat{\mathbf{z}}$, and this agrees with what we just said above. Another (and better) way to see that this is the right condition is to take the dot product with an arbitrary vector $\mathbf{v}$, which gives $\mathbf{v} \cdot df = w\mathbf{v} \cdot dC$. If $\mathbf{v}$ is tangent to the constraint surface then $\mathbf{v} \cdot dC = 0$, so in that case the equation implies $\mathbf{v} \cdot df = 0$, i.e. the rate of change of f along the directions that lie in the surface is zero. For $\mathbf{v}$ that is not tangent to the surface, f can change.

Now what if there are two constraints? For example $C_1 = z$ and $C_2 = x^2 + y^2 + z^2 - R^2$. So $C_1 = C_2 = 0$ implies the point is both on the $z = 0$ plane and on the sphere of radius R. That is, it lies on the circle of radius R in the xy plane, centered on the origin. The gradient of a function that is constant on this circle must satisfy $df = w_1 dC_1 + w_2 dC_2$ when $C_1 = C_2 = 0$. This is equivalent to saying that the derivative of f in any direction tangent to both constraint surfaces is zero.

One more formal point before applying this to Lagrangians: Instead of writing $df = wdC$ we can equally well write $df = d(wC)$, because $d(wC) = wdC + Cdw$, and when $C = 0$ the second term vanishes. So the condition on f can also be written as $d(f - wC) = 0$ when $C = 0$. Since w is undetermined at this stage anyway, we can also flip the sign (to agree with other conventions) and write this condition as

$$d(f + wC) = 0 \quad \text{when} \quad C = 0. \tag{2}$$

If we have two constraints, the condition can be written as $d(f + w_1C_1 + w_2C_2) = 0$ when $C_1 = C_2 = 0$. The same holds for any number of constraints.

Lagrange multipliers and mechanics

Let's illustrate how this applies to constrained mechanics by an example. Consider a simple pendulum of length R. We've seen we can just impose from the beginning the constraint $r = R$, using the angle θ as our sole generalized coordinate. This is equivalent to just demanding that the action be stationary with respect to variations of the path $(r(t), \theta(t))$ that respect, for each time t, the constraint $C(t) = r(t) - R = 0$.

This constrained variational principle on the action functional $S[r(t), \theta(t)]$ is just like what was discussed above for functions. Instead of one or two constraints however we have an infinite number of constraints $C(t) = 0$, one for each t. If we add them all to S, multiplied by a Lagrange multiplier function $w(t)$ and integrating over t, we arrive at an equivalent, but unconstrained variational principle: the variation of $S + \int w(t)C(t)\,dt$ should be zero for *any* variation, when $C(t) = 0$ holds. Or, in terms of the Lagrangian, the variation of $\int (L + wC)\,dt$ must vanish (the t-dependence of w and C is not explicitly indicated but it's there). So in the end it's quite simple: we just add to the Lagrangian an arbitrary multiple of the constraint(s).

For the pendulum, the θ equation is unchanged, but now that r is not fixed a priori we get an r equation of motion. The action is the integral of

$$L + wC = \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}mr^2\dot{\theta}^2 + mgr\cos\theta + w(r - R), \tag{3}$$

so the r equation is

$$m\ddot{r} = mr\dot{\theta}^2 + mg\cos\theta + w, \tag{4}$$

where the w term comes from $\partial(wC)/\partial r$ when $C = 0$. Recall that this is supposed to hold only when the constraint $C = r - R = 0$ holds, so it is really the condition

$$0 = mR\dot{\theta}^2 + mg\cos\theta + w. \tag{5}$$

Since w is so far an arbitrary function this doesn't impose any condition on anything else, of course. In fact, we can solve this equation for w,

$$w = -mR\dot{\theta}^2 - mg\cos\theta. \tag{6}$$

Forces of constraint

What is the meaning of w? It's whatever it must be for the r equation of motion to be satisfied when r is fixed at $r = R$. So w must be closely related to the force of tension of the string. In fact, in this case, it is exactly the tension force, as we can see with a Newtonian calculation: the force in the radial direction is the radial component of the gravitational force minus the tension, $mg\cos\theta - T$, where T is the magnitude of the tension force. The radial acceleration is the centripetal acceleration $-R\dot{\theta}^2$. The radial component of $\mathbf{F} = m\mathbf{a}$ then yields $T = mg\cos\theta + mR\dot{\theta}^2$. Hence $w = -T$. The minus sign is because the tension force is in the negative $\hat{\mathbf{r}}$ direction.

What is the general relation between the Lagrange multiplier $w(t)$ and the force of constraint? The answer is simple: whatever the wC term produces in the equation of motion, that is the generalized force for the corresponding generalized coordinate. That is, $w\,\partial C/\partial q$ is the generalized force. In the pendulum example, $C = r - R$, and the coordinate is r, so $\partial C/\partial r = 1$, and w is just the constraint force in the r direction. If the constrained coordinate q had been an angle, $w\,\partial C/\partial q$ would be the torque of constraint. If q is some more unusual generalized coordinate, then we'd just have some unusual generalized force of constraint.