1   Derivation of kinetic energy in terms of quaternions

A representation of the quaternion algebra are the 2×2 Pauli matrices, or equivalently, the 3×3 rotation generator matrices. Pauli matrices are a bit nicer than the generator matrices because they have very useful trace and determinant properties. Furthermore, they, plus the identity matrix, form an orthonormal basis for all 2×2 matrices. We can express every *unit* quaternion with three indepenent parameters (restricting them to be of unit magnitude removes one of our degrees of freedom) which can be represented as

$$\mathbf{q} = e^{\frac{i}{2}\theta_\mu \boldsymbol{\sigma}_\mu}.$$

I'm going to try to stick to the convention that Greek indices run from 1 to 3 and label either components of a vector (here a vector of rotation parameters θ_μ) or a tensor (later we'll have the moment of inertia tensor). Later, I'll use Latin indices to label components of matrices. Furthermore, when I say that an object is a quaternion, I am writing it as a matrix. Namely, *any* quaternion can be represented as:

$$\mathbf{z} = z_0 \mathbb{1} + z_\mu \boldsymbol{\sigma}_\mu.$$

And as well, Einstein summation notation is presumed. Also, bold objects represent matrices, unless I write them out with their indices explicitly.

The goal is to get from $\mathbf{q}$ to the angular velocity $\vec{\omega}$, then to the kinetic energy T, and then finally to the conjugate variable to $\mathbf{q}$, which we'll call $\mathbf{M}$. This will help us write down a Hamiltonian for a rotating particle, and then ultimately, we will construct the partition function for a group of particles who interact by way of their orientations. So first thing's first, we need $\vec{\omega}$.

1.1   Angular velocity vector

I won't prove this, but if we represent the orientation of some object by a quaternion $\mathbf{e}$ which is fixed, we can rotate it into a different position by *conjugating* by $\mathbf{q}$:

$$\mathbf{e}' = \mathbf{q}^\dagger \mathbf{e} \mathbf{q}.$$

Here, the dagger is a complex transpose, so we have

$$\mathbf{q}^\dagger = e^{-\frac{i}{2}\theta_\mu \boldsymbol{\sigma}_\mu}.$$

So say now we make $\mathbf{q}$ time dependent and move it forward in time by dt. The new orientation of our object is now:

$$\mathbf{e}'' = (d\mathbf{q})^\dagger \mathbf{e}'(d\mathbf{q}), \qquad d\mathbf{q} = e^{\frac{i}{2}dt\omega_\mu \boldsymbol{\sigma}_\mu} \approx 1 + \frac{i}{2}dt\omega_\mu \boldsymbol{\sigma}_\mu.$$

So we can also instead write

$$\mathbf{e}'' = \mathbf{q}'^\dagger \mathbf{e} \mathbf{q}', \qquad \mathbf{q}' = \mathbf{q}(d\mathbf{q}).$$

So we have

$$\dot{\mathbf{q}}dt = \mathbf{q}(d\mathbf{q}) - \mathbf{q} = \mathbf{q}\left(\frac{i}{2}\omega_\mu\boldsymbol{\sigma}_\mu\right)dt.$$

Because $\mathbf{q}$ is a unit quaternion, we have

$$\mathbf{q}^\dagger\mathbf{q} = \mathbf{q}\mathbf{q}^\dagger = \mathbb{1}.$$

So we can multiply from the left by $\mathbf{q}^\dagger$ and get

$$\omega_\mu\boldsymbol{\sigma}_\mu = -2i\mathbf{q}^\dagger\dot{\mathbf{q}}.$$

The left-hand side is a Hermitian matrix, so we can also write

$$\omega_\mu\boldsymbol{\sigma}_\mu = 2i\dot{\mathbf{q}}^\dagger\mathbf{q}.$$

So this is all good, but we'd like the individual components of the vector ω_μ. To do this, we can use the algebraic properties of the Pauli matrices:

$$[\boldsymbol{\sigma}_\mu, \boldsymbol{\sigma}_\nu] = 2i\epsilon_{\mu\nu\gamma}\boldsymbol{\sigma}_\gamma, \qquad \{\boldsymbol{\sigma}_\mu, \boldsymbol{\sigma}_\nu\} = 2\mathbb{1}\delta_{\mu\nu}.$$

$$\boldsymbol{\sigma}_\mu\boldsymbol{\sigma}_\nu = \frac{1}{2}\{\boldsymbol{\sigma}_\mu, \boldsymbol{\sigma}_\nu\} + \frac{1}{2}[\boldsymbol{\sigma}_\mu, \boldsymbol{\sigma}_\nu] = \mathbb{1}\delta_{\mu\nu} + i\epsilon_{\mu\nu\gamma}\boldsymbol{\sigma}_\gamma.$$

$$\text{Tr}(\boldsymbol{\sigma}_\mu) = 0, \qquad \text{Tr}(\boldsymbol{\sigma}_\mu\boldsymbol{\sigma}_\nu) = \text{Tr}(\mathbb{1}\delta_{\mu\nu}) = 2\delta_{\mu\nu}.$$

So we can multiply by another Pauli matrix and take the trace of both sides of our equation to isolate the components of the vector $\vec{\omega}$.

$$\omega_\mu\text{Tr}(\boldsymbol{\sigma}_\mu\boldsymbol{\sigma}_\nu) = -2i\text{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu).$$

$$\boxed{\omega_\nu = -i\text{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu).}$$

We can also take the Hermitian conjugate of both sides, and since ω_ν is real, we also obtain

$$\boxed{\omega_\nu = i\text{Tr}(\dot{\mathbf{q}}^\dagger\mathbf{q}\boldsymbol{\sigma}_\nu).}$$

1.2 Conjugate momentum and matrix derivatives

The kinetic energy of our rotating object is simply expressed with the moment of inertia tensor:

$$T = \frac{1}{2}\omega_\mu I_{\mu\nu}\omega_\nu.$$

So we can simply plug in our work from above:

$$T = -\frac{1}{2}I_{\mu\nu}\text{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu)\text{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu) = -\frac{1}{2}I_{\mu\nu}\text{Tr}(\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu\mathbf{q}^\dagger)\text{Tr}(\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu\mathbf{q}^\dagger).$$

On the far right hand side, I used the cyclic property of the trace to move the time derivative to the front, which will be useful now when we take derivatives with respect to it.

In order to construct the Hamiltonian, we need to first find the conjugate momentum to our dynamical variable **q**. This is done with a derivative with respect to its time derivative:

$$\mathbf{M} = \frac{\partial T}{\partial \dot{\mathbf{q}}}.$$

But what does the derivative with respect to a matrix mean? Well it just means a derivative with respect to all of its components. For the sake of illustration, we can calculate the derivative of a trace of a matrix $f(\mathbf{X}) = \mathrm{Tr}(XY)$ in the following way:

$$f(\mathbf{X}) = X_{ij}Y_{ji}.$$

$$\frac{\partial f}{\partial X_{k\ell}} = \delta_{ik}\delta_{j\ell}Y_{ji} = Y_{\ell k}.$$

$$\frac{\partial f}{\partial \mathbf{X}} = \mathbf{Y}^T.$$

This will really be all we need. For more details, I recommend *The Matrix Cookbook* (`http://www.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf`) written by some smart folks at MIT. As for now, we can go back to finding the conjugate momentum:

$$\mathbf{M}^T = \left(\frac{\partial T}{\partial \dot{\mathbf{q}}}\right)^T = -\frac{1}{2}I_{\mu\nu}\left(\boldsymbol{\sigma}_\mu\mathbf{q}^\dagger\mathrm{Tr}(\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu\mathbf{q}^\dagger) + \boldsymbol{\sigma}_\nu\mathbf{q}^\dagger\mathrm{Tr}(\dot{\mathbf{q}}^\dagger\boldsymbol{\sigma}_\mu\mathbf{q})\right).$$

The two objects in parentheses are the same tensor, but transposed in the indices μ and ν. But we're summing over these indices with the moment of inertia tensor $I_{\mu\nu}$, which is symmetric. So we really have two of the same object, and we can combine them together. Thus we have

$$\mathbf{M}^T = -I_{\mu\nu}\boldsymbol{\sigma}_\mu\mathbf{q}^\dagger\mathrm{Tr}(\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu\mathbf{q}^\dagger).$$

In index notation:

$$M_{ij} = -I_{\mu\nu}\left(\boldsymbol{\sigma}_\mu\mathbf{q}^\dagger\right)_{ji}\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu).$$

A useful quantity is the Hermitian conjugate:

$$M_{ij}^\dagger = -I_{\mu\nu}\left(\mathbf{q}\boldsymbol{\sigma}_\mu\right)_{ji}\mathrm{Tr}(\dot{\mathbf{q}}^\dagger\mathbf{q}\boldsymbol{\sigma}_\nu).$$

1.3   Kinetic energy in terms of conjugate momentum

In order to use the power of the Hamiltonian formalism, we need to express the kinetic energy in terms of the conjugate momentum instead of the time derivative of our generalized coordinate

$\mathbf{q}$. However, I can't think of a way to invert the above two relations to get $\dot{\mathbf{q}}$ in terms of $\mathbf{M}$. However, we can totally sub out $\dot{\mathbf{q}}$ for $\mathbf{M}$ in a particular case when the moment of inertia tensor is proportional to the identity matrix. First we begin with a particularly useful quantity, the trace over $\mathbf{M}^\dagger\mathbf{M}$:

$$\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right) = M_{ij}^\dagger M_{ji} = I_{\mu\nu}I_{\gamma\eta}(\mathbf{q}\boldsymbol{\sigma}_\mu)_{ji}(\boldsymbol{\sigma}_\gamma\mathbf{q}^\dagger)_{ij}\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu)\mathrm{Tr}(\dot{\mathbf{q}}^\dagger\mathbf{q}\boldsymbol{\sigma}_\eta).$$

$$\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right) = I_{\mu\nu}I_{\gamma\eta}\mathrm{Tr}(\mathbf{q}\boldsymbol{\sigma}_\mu\boldsymbol{\sigma}_\gamma\mathbf{q}^\dagger)\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu)\mathrm{Tr}(\dot{\mathbf{q}}^\dagger\mathbf{q}\boldsymbol{\sigma}_\eta).$$

In the first trace, because it's cyclic, we can shuttle the $\mathbf{q}^\dagger$ to the front which cancels with the $\mathbf{q}$. And we have a very simple expression for the trace over a product of Pauli matrices:

$$\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right) = 2I_{\mu\nu}I_{\gamma\eta}\delta_{\mu\gamma}\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu)\mathrm{Tr}(\dot{\mathbf{q}}^\dagger\mathbf{q}\boldsymbol{\sigma}_\eta).$$

$$\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right) = -2I_{\nu\eta}^2\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu)\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\eta).$$

Here, I've used $\mathbf{q}^\dagger\dot{\mathbf{q}} = -\dot{\mathbf{q}}^\dagger\mathbf{q}$. So now this looks extremely similar to our result for the kinetic energy.

$$T = -\frac{1}{2}I_{\mu\nu}\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu)\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\nu).$$

The only difference is a factor of 4 and an extra power of the tensor I. As far as I can see, there doesn't seem to be enough information to solve for T in terms of $\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right)$ for a general I. However, if $I_{\mu\nu} = I\delta_{\mu\nu}$, then this simplifies things tremendously.

$$T = -\frac{I}{2}\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu)\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu).$$

$$\mathrm{Tr}(\mathbf{M}^\dagger\mathbf{M}) = -2I^2\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu)\mathrm{Tr}(\mathbf{q}^\dagger\dot{\mathbf{q}}\boldsymbol{\sigma}_\mu).$$

$$\boxed{T = \frac{1}{4I}\mathrm{Tr}\left(\mathbf{M}^\dagger\mathbf{M}\right).}$$