

Euler among the Pedestrians

W. Smith

Computational Science and Engineering Department,
C.C.L.R.C. Daresbury Laboratory,
Daresbury, Warrington WA4 4AD, U.K.

1 Introduction

All molecular dynamicists worth their salt will be familiar with Euler's celebrated equations (1) describing the torque induced motion of a rigid body:

$$\begin{aligned}\tau_x^p &= I_{xx}^p \dot{\omega}_x^p - \omega_y^p \omega_z^p (I_{yy}^p - I_{zz}^p) \\ \tau_y^p &= I_{yy}^p \dot{\omega}_y^p - \omega_z^p \omega_x^p (I_{zz}^p - I_{xx}^p) \\ \tau_z^p &= I_{zz}^p \dot{\omega}_z^p - \omega_x^p \omega_y^p (I_{xx}^p - I_{yy}^p)\end{aligned}\tag{1}$$

where $\vec{\tau}^p$ is the torque, $\vec{\omega}^p$ the angular velocity, I_{xx}^p , I_{yy}^p , I_{zz}^p are the moments of inertia and $\dot{\vec{\omega}}^p$ the is angular velocity - all defined in the so-called principal frame of reference. These equations are augmented by equations of motion for the Euler angles that define the orientation (here expressed in the *x-convention* [1]):

$$\begin{aligned}\dot{\phi} &= \frac{1}{\sin\theta} (\sin\psi \omega_x^p + \cos\psi \omega_y^p) \\ \dot{\theta} &= \cos\psi \omega_x^p - \sin\psi \omega_y^p \\ \dot{\psi} &= -\frac{1}{\tan\theta} (\sin\psi \omega_x^p + \cos\psi \omega_y^p) + \omega_z^p\end{aligned}\tag{2}$$

These equations present a number of difficulties to a molecular dynamicist, some of which are conceptual and some practical¹. The purpose of this note is to help make newcomers aware of them, and hopefully clarify some issues, which are rarely, if ever, discussed in mathematical texts.

¹There is also an infamous and common mistake to be aware of. When implementing Euler's equations in software, and also when manipulating formulae, it is a common error to forget what frame of reference is being used, and thereby leave out the centrifugal terms altogether. What makes this error so pernicious is the fact that the resulting code will run very nicely, conserving energy and generating seemingly meaningful results. This blunder is therefore very hard to detect from an inspection of the results.

2 Spinning into trouble

Perhaps the first conceptual difficulty with Euler's equations (1) is that they are presented in the principal frame. This is, of course, the frame of reference in which the moment of inertia tensor \mathbf{I} is diagonal. (The moment of inertia tensor is a 3x3 symmetric matrix, which can *always* be diagonalised, and so we may assume a suitable body fixed frame of reference always exists.) The body fixed frame of reference rotates with the rigid body and so the tensor \mathbf{I} remains diagonal in that frame. This is in marked contrast to the behaviour of \mathbf{I} in the laboratory frame, where the six independent components vary as the body rotates.

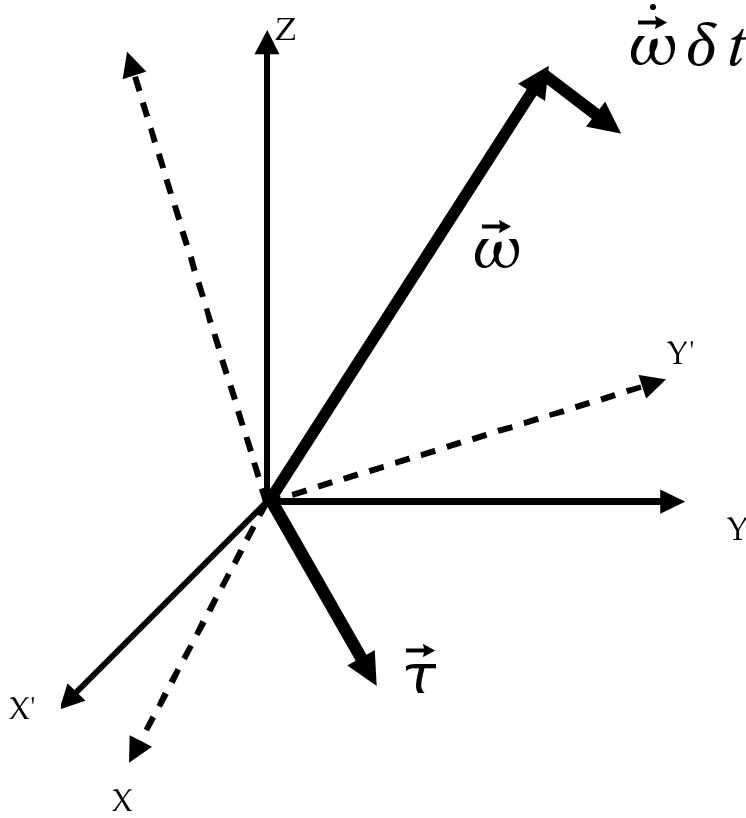
So where's the conceptual difficulty? This, I'm sure, can be expressed in various ways by different people, but for me the difficulty lies in the fact that the motion of the body is being described in a frame of reference which moves with the body. This seems absurd, surely if the frame of reference is fixed in the body, how can the body move with respect to it? If one tries to think along such lines, one can, at a pinch, imagine some situations where this does make sense, such as a body that grows or shrinks in the rotating frame, a strange form of motion perhaps, but one that doesn't interfere with one's grasp of what the frame of reference means. However, in general there is little profit in thinking in this way. It is much better to think in terms of vectors. (In so doing it is best always to think of them as real physical objects, which retain their identity in any frame of reference. This is particularly helpful when dealing with angular velocity, where it is easy to slip into thinking that the angular velocity vector of a rotating body vanishes in the body frame, because the body is no longer 'rotating' in that frame.)

How does a vector representation clarify matters? Imagine the rotating rigid body 'frozen' at an instant in time. At this instant a simple relationship exists between the components of a vector \vec{v} in the body (principal) frame and the laboratory frame:

$$\vec{v}^p = \mathbf{R} \vec{v} \tag{3}$$

where \mathbf{R} is the familiar rotation matrix. Despite having different components, vectors \vec{v} and \vec{v}^p are, of course, the same physical object. This immutability is universal, and applies equally to the angular velocity vector $\vec{\omega}$ and the torque vector $\vec{\tau}$. This situation is shown in Figure 1, where the laboratory and rotating frames (X,Y,Z) and (X',Y',Z') respectively, appear with the angular velocity vector $\vec{\omega}$ and the torque vector $\vec{\tau}$. Now, as a result of the presence of the torque vector, in a very short time interval δt , the angular momentum vector will change by the addition of a small vector $\vec{\omega} \delta t$. Meanwhile, the rigid body, which was initially rotating about the vector $\vec{\omega}$, after the interval δt , ends up rotating about the vector $\vec{\omega} + \vec{\omega} \delta t$. As this happens, of course, the body fixed axes must follow the change. This is the essential physics. What Euler's equations provide is the description (i.e. the components) of the important vectors in the frame (X',Y',Z') at the instant it was 'frozen'. This gives us some physical insight into what is actually going on in Euler's formulation.

So much for a pictorial description, let's see what's going on mathematically.



3 Doing the maths

If we go back to basic texts on vectors [1] to find out how they behave in a rotating system, then we soon encounter the famous equation:

$$\frac{d}{dt} \vec{v} = \left\{ \frac{d}{dt} \vec{v} \right\}_r + \vec{\omega} \times \vec{v} \quad (4)$$

This equation, which is central to the derivation of Euler's equation, relates the rate of change of a vector \vec{v} in the laboratory frame, to its rate of change in a rotating frame. It is clear that the time derivative on the left of this equation refers to an inertial frame (i.e. one not accelerating with respect to what Newton called absolute space). However the derivative on the right (with the subscript r , for *rotation*) refers to a rotating, and therefore a non-inertial frame. It is therefore no surprise to see the appearance of the

vector product $\vec{\omega} \times \vec{v}$ on the right, which provides the adjustment necessary to translate the observed change in the rotating frame, back to the laboratory frame. This additional term is sometimes called the centrifugal term.

One of the easiest ways to visualise equation (3) is to imagine a constant (i.e. fixed) vector in the rotating frame, set at some angle to the axis of rotation, and then think how this might appear to a laboratory observer. (The rotating vector traces out a ‘coolie’ hat!) The tip of the vector traces a circle (the rim of the coolie hat) in the laboratory frame, and the velocity of the tip along this circle, at any instant, is the time derivative of the vector in inertial space. If, in addition, the vector in the rotating frame is changing with time, the ‘velocity’ (i.e. rate-of-change) vector of that must be added vectorially to the circular velocity.

All this make good physical and intuitive sense, but there is a potential source of confusion. This arises when one asks what are the *components* of the vectors in these equations. Clearly, for this equation to make sense, all the vectors must be defined with respect to the same frame of reference, otherwise corresponding components would not have the same meaning. But on the left of (4) the time derivative must be in an inertial frame, while on the right it must be in the rotating frame. (And speaking frankly, at this point, I *don't know* what frame the cross product is meant to be in, except that $\vec{\omega}$ and \vec{v} should together be in the same frame to provide a meaningful vector product.)

There is a smart trick to deal with this. If we construct an inertial frame which at one instant in time just so happens to coincide with the body fixed frame, then at that one instant, the components of the time derivative on both sides of (3) are in the same frame of reference. If we also evaluate the cross product in that same frame, we can sum up *all* the components on the right in a way that is physically meaningful. Note that this instantaneous frame is otherwise unrelated to the rotating frame, and it has no rotational motion of its own. It follows that when constructing this frame, we can use a simple rotational transformation of our original laboratory frame, without adding extra centrifugal terms. The disadvantage of this ‘trick’ is that it makes one think that Euler’s equation is expressed in a frame of reference that is co-rotating with the body. This is simply not true, such a frame would be non-inertial. It is expressed in a frame of reference which is only instantaneously coincident with the body frame, and which is fundamentally inertial in character.

So, after all this discussion, we can write (4) in the form

$$\frac{d}{dt}\vec{v}^p = \left\{ \frac{d}{dt}\vec{v}^p \right\}_r + \vec{\omega}^p \times \vec{v}^p \quad (5)$$

to indicate that everthing (including $\vec{\omega}$ and \vec{v}) are expressed in the principal frame.

Now, going back to Euler’s equations, let us derive them with what we know.

4 Deriving Euler's equations - 1

Any standard textbook [1] that introduces the reader to rigid body motion will show the derivation of the rotational equation of motion of a rigid body about its own centre of mass in an inertial frame:

$$\vec{\tau} = \frac{d}{dt} \vec{J} \quad (6)$$

where \vec{J} is the angular momentum vector. Following the discussion leading to equation (5) we can immediately transform equation (6) and obtain

$$\vec{\tau}^p = \left\{ \frac{d}{dt} \vec{J}^p \right\}_r + \vec{\omega}^p \times \vec{J}^p \quad (7)$$

The angular momentum is

$$\vec{J}^p = \mathbf{I}^p \vec{\omega}^p \quad (8)$$

and using the chain rule for vector differentiation the time derivative can be written as

$$\left\{ \frac{d}{dt} \vec{J}^p \right\}_r = \dot{\mathbf{I}}^p \vec{\omega}^p + \mathbf{I}^p \dot{\vec{\omega}}^p \quad (9)$$

In the principal frame of reference, the moment of inertia tensor is constant and its time derivative vanishes. Hence the time derivative of the angular momentum in the body frame is

$$\frac{d}{dt} \vec{J}^p = \mathbf{I}^p \dot{\vec{\omega}}^p \quad (10)$$

So we may write (7) as

$$\vec{\tau}^p = \mathbf{I}^p \dot{\vec{\omega}}^p + \vec{\omega}^p \times \mathbf{I}^p \vec{\omega}^p \quad (11)$$

Which is mathematically identical to Euler's equation (1).

5 Deriving Euler's equations - 2

As is apparent from the above that equation (4) is the key to deriving Euler's equation. However it is equally apparent that one has to deal with some slippery concepts on the way. Can one take a more pedestrian route, one that doesn't require so much physical intuition? The answer is yes, but the derivation is much longer, which probably accounts for its absence from the text books.

We begin by noting that equation (8) is not restricted to the principal frame, and can be generally written as

$$\vec{J} = \mathbf{I} \vec{\omega}. \quad (12)$$

Inserting this into equation (6) and performing the time differentiation, we obtain

$$\vec{\tau} = \mathbf{I} \dot{\vec{\omega}} + \dot{\mathbf{I}} \vec{\omega} \quad (13)$$

which is an equation set entirely in the laboratory frame. In consequence the tensor \mathbf{I} is generally not diagonal, and its components vary with time.

We can now consider an instant in time and ask what must be done to diagonalise \mathbf{I} . The answer is that we must find a matrix \mathbf{R} (and its transpose \mathbf{R}^\dagger) such that

$$\mathbf{I}^p = \mathbf{R} \mathbf{I} \mathbf{R}^\dagger \quad (14)$$

The matrix \mathbf{R} is the matrix of eigenvectors of \mathbf{I} , which, of course, is a rotation matrix. It may be conveniently obtained by standard methods such Jacobi diagonalisation. We may now multiply equation (13) by the rotation matrix and obtain

$$\mathbf{R} \vec{\tau} = \mathbf{R} \mathbf{I} \mathbf{R}^\dagger \mathbf{R} \dot{\vec{\omega}} + \mathbf{R} \dot{\mathbf{I}} \mathbf{R}^\dagger \mathbf{R} \vec{\omega}, \quad (15)$$

In which we have also used the well known property of the rotation matrix that

$$\mathbf{R}^\dagger \mathbf{R} = \mathbf{1} \quad (16)$$

where $\mathbf{1}$ is the unit matrix. Completing the transformation, our equation now becomes

$$\vec{\tau}^p = \mathbf{I}^p \dot{\vec{\omega}}^p + \mathbf{R} \dot{\mathbf{I}} \mathbf{R}^\dagger \vec{\omega}^p, \quad (17)$$

It is useful to pause at this point and examine what we have obtained. The resemblance of this equation to Euler's is apparent, provided we can convert the final term on the right into the centrifugal form given by Euler. If, for the moment, we assume this can be done, we already have learned an important lesson. It is obvious that the equation remains in the inertial frame, since all we have done is apply a rotational operation \mathbf{R} to the equation, which carries no dynamical implications; we have just reorientated our inertial axes. This is a valuable observation.

The reader now has three choices. The first is to accept that equations (1) and (17), are one and the same (which is where the physics stops) and skip to the next section. The second is to follow the rest of the demonstration and appreciate the brevity of Euler's approach. The third, strictly for young Turks, is to go away and complete it independently.

To proceed further, we need to look closely at the moment of inertia tensor. We shall consider the case where it is derived for a rigid molecule consisting of point masses m_i at displacements \vec{d}_i from the molecular centre of mass. Thus

$$\mathbf{I} = \sum_i m_i \{ d_i^2 \mathbf{1} - [\vec{d}_i \vec{d}_i] \} \quad (18)$$

where the term in square brackets [. .] is an outer product of the contained vectors. Equivalently we have for each component of \mathbf{I}

$$I_{\alpha\beta} = \sum_i m_i \{ d_i^2 \delta_{\alpha\beta} - d_{i\alpha} d_{i\beta} \}. \quad (19)$$

The time derivative of this is easily obtained:

$$\dot{I}_{\alpha\beta} = - \sum_i m_i \{ \dot{d}_{i\alpha} d_{i\beta} + d_{i\alpha} \dot{d}_{i\beta} \}. \quad (20)$$

where we have used the fact the d_i^2 is a constant for a rigid molecule. We now need the product $\mathbf{R} \dot{\mathbf{I}} \mathbf{R}^\dagger$, which is given by

$$R_{\mu\alpha} \dot{I}_{\alpha\beta} R_{\beta\lambda}^\dagger = \sum_i m_i \{ -R_{\mu\alpha} \dot{d}_{i\alpha} d_{i\beta} R_{\beta\lambda}^\dagger - R_{\mu\alpha} d_{i\alpha} \dot{d}_{i\beta} R_{\beta\lambda}^\dagger \} \quad (21)$$

in which we have used the Einstein convention regarding repeated indices α, β etc. This equation readily collapses into the following form.

$$\dot{I}_{\mu\lambda}^p = \sum_i m_i \{ -\dot{d}_{i\mu}^p d_{i\lambda}^p - d_{i\mu}^p \dot{d}_{i\lambda}^p \} \quad (22)$$

Now we must recognise that

$$\dot{d}_i^p = \vec{\omega}^p \times \vec{d}_i^p \quad (23)$$

or in expanded form

$$\dot{d}_{i\alpha}^p = \omega_{|\alpha+1|}^p d_{i|\alpha+2|}^p - \omega_{|\alpha+2|}^p d_{i|\alpha+1|}^p. \quad (24)$$

This clumsy notation is meant to indicate the cyclic nature of the index α . Thus when $\alpha = 2$ (i.e. the y component), $|\alpha + 1| = 3$ (the z component) and therefore $|\alpha + 3| = 1$ (the x component), and so on. Substituting this into (22) gives

$$\dot{I}_{\mu\lambda}^p = - \sum_i m_i \{ \omega_{|\mu+1|}^p d_{i|\mu+2|}^p d_{i\lambda}^p - \omega_{|\mu+2|}^p d_{i|\mu+1|}^p d_{i\lambda}^p + \omega_{|\lambda+1|}^p d_{i|\lambda+2|}^p d_{i\mu}^p - \omega_{|\lambda+2|}^p d_{i|\lambda+1|}^p d_{i\mu}^p \} \quad (25)$$

(where fortunately, the Einstein convention is not needed!) We can now inspect individual components of the matrix $\dot{\mathbf{I}}$, while remembering the definition of the matrix \mathbf{I} in equation (19) and that off-diagonal components of the matrix \mathbf{I}^p are zero. Thus we find that

$$\dot{I}_{xx}^p = \dot{I}_{yy}^p = \dot{I}_{zz}^p = 0 \quad (26)$$

and

$$\begin{aligned} \dot{I}_{xy}^p = \dot{I}_{yx}^p &= \omega_z^p (I_{xx}^p - I_{yy}^p) \\ \dot{I}_{xz}^p = \dot{I}_{zx}^p &= \omega_y^p (I_{zz}^p - I_{xx}^p) \\ \dot{I}_{yz}^p = \dot{I}_{zy}^p &= \omega_x^p (I_{yy}^p - I_{zz}^p). \end{aligned} \quad (27)$$

Finally, we need the product $\dot{\mathbf{I}}^p \vec{\omega}^p$, which is easily obtained. The result is

$$\begin{aligned} (\dot{\mathbf{I}}^p \vec{\omega}^p)_x &= \omega_y^p \omega_z^p (I_{zz}^p - I_{yy}^p) \\ (\dot{\mathbf{I}}^p \vec{\omega}^p)_y &= \omega_z^p \omega_x^p (I_{xx}^p - I_{zz}^p) \\ (\dot{\mathbf{I}}^p \vec{\omega}^p)_z &= \omega_x^p \omega_y^p (I_{yy}^p - I_{xx}^p) \end{aligned} \quad (28)$$

These are the terms which must replace the final term in equation (17), thus establishing its equivalence to Eulers equations (1).

6 Turning the screw

So far we have said little about the auxiliary equations for the motion of the Euler angles (2). These are obtained in a straightforward manner by defining angular velocity vectors $\vec{\omega}_\phi, \vec{\omega}_\theta, \vec{\omega}_\psi$, associated with rotation through the respective Euler angles, and transforming these vectors into the principal frame.

These equations can, in principle, be solved directly by some numerical algorithm, but nobody does this any more. This is because the equations (2) generate infinities when the angle θ becomes zero, as it is perfectly free to do! For this reason Denis Evans and Sohail Murad famously introduced quaternions into molecular dynamics [2], which did away with this difficulty forever. I don't propose to elaborate on the topic, given the fact that the method is now universal and is described in detail in the standard texts [3]. I would however like to suggest an alternative approach.

First we should note that equation (25) can also be written in the laboratory frame:

$$\dot{I}_{\alpha\beta} = \omega_{|\alpha+1|} I_{|\alpha+2|\beta} - \omega_{|\alpha+2|} I_{|\alpha+1|\beta} + \omega_{|\beta+1|} I_{|\beta+2|\alpha} - \omega_{|\beta+2|} I_{|\beta+1|\alpha} \quad (29)$$

This is derived in a manner completely analogous to (25) and we have gathered the contributions to the components of the moment of inertia tensor. (Notice this formula works equally well for sums involving the diagonal components on the right). We now have an expression for $\dot{\mathbf{I}}$ which we shall represent briefly as:

$$\dot{\mathbf{I}} = F(\vec{\omega}, \mathbf{I}). \quad (30)$$

We now write equation (13) as

$$\dot{\omega} = \mathbf{I}^{-1}(\vec{\tau} - \dot{\mathbf{I}} \vec{\omega}) \quad (31)$$

In which the inverse matrix is trivial, since it is only a 3×3 matrix. Equations (30) and (31) constitute a set of coupled equations which can be solved by a leapfrog algorithm as follows.

1. At each time step n we have $\vec{\tau}^n$, \mathbf{I}^n , $\vec{\omega}^{n-1/2}$ and the angular momentum $\vec{J}^{n-1/2}$.

2. Integrate (6):

$$\vec{\mathbf{J}}^{n+1/2} = \vec{\mathbf{J}}^{n-1/2} + \Delta t \vec{\tau}^n$$

3. Estimate $\vec{\omega}^n$:

$$\vec{\omega}^n = \mathbf{I}_n^{-1}(\vec{\mathbf{J}}^{n-1/2} + \vec{\mathbf{J}}^{n+1/2})/2$$

4. Calculate $\dot{\mathbf{I}}^n$ using (29):

$$\dot{\mathbf{I}}^n = F(\vec{\omega}^n, \mathbf{I}^n)$$

5. Integrate (31):

$$\vec{\omega}^{n+1/2} = \vec{\omega}^{n-1/2} + \Delta t \mathbf{I}_n^{-1}(\vec{\tau}^n - \dot{\mathbf{I}}^n \vec{\omega}^n)$$

6. Calculate interim $\mathbf{I}^{n+1/2}$:

$$\mathbf{I}^{n+1/2} = \mathbf{I}^n + (\Delta t/2) \dot{\mathbf{I}}^n$$

7. Calculate $\dot{\mathbf{I}}^{n+1/2}$ using (29):

$$\dot{\mathbf{I}}^{n+1/2} = F(\vec{\omega}^{n+1/2}, \mathbf{I}^{n+1/2})$$

8. Calculate \mathbf{I}^{n+1} :

$$\mathbf{I}^{n+1} = \mathbf{I}^n + \Delta t \dot{\mathbf{I}}^{n+1/2}$$

The experienced practitioner will recognise opportunities to iterate on both $\vec{\omega}^n$ and $\mathbf{I}^{n+1/2}$ in this scheme, and so improve the accuracy of the interim values. This, admittedly long-winded, procedure has the advantage that it solves the rotational equations of motion entirely in the the laboratory frame, without any of the inherent instabilities of equation (2).

There is an important omission however. Nowhere do we calculate the rotation matrix \mathbf{R} , and without it we cannot find the positions of the atoms of the rigid molecule in space. Thus the calculation of the torque becomes problematical! But this is not an insurmountable problem as we can diagonalise the inertia tensor using Jacobi's method and so obtain \mathbf{R} as the matrix of eigenvectors. This need not be as expensive an operation as it first seems, since the rotation matrix from the proceeding time step will be an excellent approximation to it, and can be used to partially diagonalise \mathbf{I} . The number of iterations in the Jacobi method will therefore be much reduced. I offer this algorithm as an alternative for those who find quaternions too esoteric!

7 Summary

The equations of Euler have been derived and the conceptual difficulties inherent in the standard derivation examined. An alternative derivation has been presented. Finally an altogether different approach to the solution of the rotational equations of motion has been presented.

References

- [1] H. Goldstein, *Classical Mechanics*, Addison-Wesley 1980.
- [2] D.J. Evans and S. Murad, *Molecular Physics* **34** (1977) 327.
- [3] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press (1980).