1 Complex numbers and rotations in two dimensions

One of the easiest ways to rotate of a point (x, y) around the origin, counterclockwise, by an angle θ, is to represent it as a COMPLEX NUMBER, thus

$$x + i\,y$$

and multiply the latter by $\exp(i\ \theta) = \cos\theta + i\ \sin\theta$. The new coordinates are then given as the REAL (x coordinate) and PURELY IMAGINARY (y coordinate) part of

$$(x + i\,y)(\cos\theta + i\ \sin\theta) = (x\cos\theta - y\sin\theta) + i\,(y\cos\theta + x\sin\theta)$$

as can be easily verified graphically.

It should be obvious that $\omega \equiv \exp(i\ t)$, where $t \in [0, 2\pi)$ is a parametric representation of the UNIT CIRCLE, centered at the origin - note that this corresponds to the point $(1, 0)$ rotated by angle t.

Similarly, $\omega + \frac{\varepsilon}{\omega}$, where $\varepsilon < 1$, is an ELLIPSE.

Proof:

$$\begin{aligned} x &= (1+\varepsilon)\cos t \\ y &= (1-\varepsilon)\sin t \end{aligned}$$

implies

$$\frac{x^2}{(1+\varepsilon)^2} + \frac{y^2}{(1-\varepsilon)^2} = 1$$

The corresponding ECCENTRICITY is

$$\frac{\sqrt{(1+\varepsilon)^2 - (1-\varepsilon)^2}}{1+\varepsilon} = \frac{2\sqrt{\varepsilon}}{1+\varepsilon}$$

□

2 Quaternions and rotations in three dimensions

To extend the previous idea (of a rotation facilitated by complex numbers) to three dimensions, we need to understand the ALGEBRA OF QUATERNIONS. Its elements have the form of

$$A + a_x\mathfrak{i} + a_y\mathfrak{j} + a_z\mathfrak{k} \equiv A + \mathbf{a}$$

where A and $\mathbf{a}$ are referred to as the SCALAR and VECTOR parts, respectively. Note that now we have *three* imaginary units $\mathfrak{i}$, $\mathfrak{j}$ and $\mathfrak{k}$, each squaring to -1, and anticommuting with each other (i.e. $\mathfrak{i} \odot \mathfrak{j} = -\mathfrak{j} \odot \mathfrak{i}$, etc.); furthermore, $\mathfrak{i} \odot \mathfrak{j} = \mathfrak{k}$.

One can easily verify that multiplying any two quaternions is equivalent to

$$(A + \mathbf{a}) \odot (B + \mathbf{b}) = AB - \mathbf{a}\cdot\mathbf{b} + A\mathbf{b} + B\mathbf{a} + \mathbf{a}\times\mathbf{b} \tag{1}$$

Addition (subtraction) of two quaternions is a trivial, element-wise operation. So is multiplication of a quaternion by a scalar.

Note that (by utilizing Taylor series), we can evaluate simple functions of quaternions, for example

$$\exp_{\odot}(A+\mathbf{a}) = (\cos a + \hat{\mathbf{a}}\sin a)\exp A \tag{2}$$

where $a \equiv \sqrt{a_x^2 + a_y^2 + a_z^2}$ and $\hat{\mathbf{a}} \equiv \frac{\mathbf{a}}{a}$.

Consider a general quaternion $\mathfrak{A} \equiv A + \mathbf{a}$. Its QUATERNIONIC CONJUGATE is denoted by $\bar{\mathfrak{A}}$ and defined as $A - \mathbf{a}$ (each imaginary unit changing sign). It can be easily verified that

$$\overline{\mathfrak{A} \odot \mathfrak{B}} = \bar{\mathfrak{B}} \odot \bar{\mathfrak{A}} \tag{3}$$

Rotating a vector $\mathbf{r}$, right-hand-screw-wise, around an axis (passing through the origin) of direction $\boldsymbol{\theta}$, by an angle equal to the magnitude of $\boldsymbol{\theta}$, is achieved by

$$\exp_{\odot}(\frac{\boldsymbol{\theta}}{2}) \odot \mathbf{r} \odot \exp_{\odot}(-\frac{\boldsymbol{\theta}}{2})$$

Proof:

$$\exp_{\odot}(\frac{\boldsymbol{\theta}}{2}) \odot (\mathbf{r}_{\|} + \mathbf{r}_{\perp}) \odot \exp_{\odot}(-\frac{\boldsymbol{\theta}}{2}) = \mathbf{r}_{\|} + \exp_{\odot}(\boldsymbol{\theta}) \odot \mathbf{r}_{\perp} =$$
$$\mathbf{r}_{\|} + (\cos\theta + \hat{\boldsymbol{\theta}}\sin\theta) \odot \mathbf{r}_{\perp} = \mathbf{r}_{\|} + \mathbf{r}_{\perp}\cos\theta + \hat{\boldsymbol{\theta}} \times \mathbf{r}_{\perp}\sin\theta$$

where $\mathbf{r}_{\|} + \mathbf{r}_{\perp}$ is decomposing $\mathbf{r}$ into parts parallel and perpendicular to $\boldsymbol{\theta}$. The rest is rather routine (once you get used to it) quaternion algebra. Note that when two vectors (say $\mathbf{q}$ and $\mathbf{w}$) are perpendicular, they anticommute (when multiplied as two quaternions). This implies $\mathbf{q} \odot \exp_{\odot}\mathbf{w} = \exp_{\odot}(-\mathbf{w}) \odot \mathbf{q}$. Similarly, parallel vectors commute, implying $\mathbf{q} \odot \exp_{\odot}\mathbf{w} = \exp_{\odot}(\mathbf{w}) \odot \mathbf{q}$. $\square$

Let us point out that any quaternion $\mathfrak{R}$ which meets

$$\mathfrak{R} \odot \bar{\mathfrak{R}} \equiv \bar{\mathfrak{R}} \odot \mathfrak{R} = 1$$

can be written in the

$$\mathfrak{R} = \exp_{\odot}(\frac{\boldsymbol{\theta}}{2})$$

form (proof quite trivial).

Alternately, each such $\mathfrak{R}$ (and therefore any rotation) can be written as

$$\exp_{\odot}(\mathfrak{k}\frac{\phi}{2}) \odot \exp_{\odot}(\mathfrak{i}\frac{\theta}{2}) \odot \exp_{\odot}(\mathfrak{k}\frac{\psi}{2}) \equiv \cos\frac{\theta}{2}\exp_{\odot}(\mathfrak{k}\frac{\psi+\phi}{2}) + \mathfrak{i}\sin\frac{\theta}{2} \odot \exp_{\odot}(\mathfrak{k}\frac{\psi-\phi}{2}) \tag{4}$$

where ϕ, θ and ψ are the usual EULER ANGLES.

One can also show that when $\mathfrak{R}$, rather than being a *fixed* rotation, becomes a function of time t,

$$2\dot{\mathfrak{R}} \odot \bar{\mathfrak{R}}$$

is a vector whose direction agrees with the instantaneous axis of rotation, and whose length yields the corresponding angular speed. Note that the dot implies element-wise differentiation with respect to t.

For (4), this yields

$$\mathfrak{i}\,(\dot{\theta}\cos\phi+\dot{\psi}\sin\phi\sin\theta)+\,\mathfrak{j}\,(\dot{\theta}\sin\phi-\dot{\psi}\cos\phi\sin\theta)+\,\mathfrak{k}\,(\dot{\phi}+\dot{\psi}\cos\theta)\equiv\mathfrak{k}\,\dot{\phi}+\mathbf{n}\dot{\theta}+\mathbf{e}\dot{\psi}$$

where $\mathbf{n}$ is the so called NODAL direction, and $\mathbf{e}$ is the new (rotated) z direction. This is quite obvious geometrically.

3 Bi-quaternions and special relativity

To deal with four dimensions of Special Relativity, we need to extend Quaternion Algebra of the previous section by allowing each of its components to be a complex number (whose imaginary unit, say I, is different from the quaternionic imaginary units $\mathfrak{i}$, $\mathfrak{j}$ and $\mathfrak{k}$, commuting with each of them). An element of the new (so called PAULI or BI-QUATERNION) ALGEBRA can thus be written as

$$A+I\,B+(a_x+I\,b_x)\mathfrak{i}+(a_y+I\,b_y)\mathfrak{j}+(a_z+I\,b_z)\mathfrak{k}=A+I\,B+\mathbf{a}+I\,\mathbf{b}\equiv\Gamma+\boldsymbol{\alpha}$$

where Γ and the components of $\boldsymbol{\alpha}$ are *complex*.

In addition to quaternion conjugation, which is now called REFLECTION and denote $\mathfrak{A}^-$, we can also define COMPLEX CONJUGATION $(I\to-I)$, denoting it by $\mathfrak{A}^*$. The combination of both (we will not give it a name) will be denoted $\mathfrak{A}^+\equiv(\mathfrak{A}^-)^*=(\mathfrak{A}^*)^-$.

Multiplication of two elements of this algebra can still be done using (1), but for our purpose it is more convenient to define a new product by

$$(\Gamma+\boldsymbol{\alpha})\circ(\Lambda+\boldsymbol{\beta})\equiv\Gamma\Lambda+\boldsymbol{\alpha}\cdot\boldsymbol{\beta}+\Gamma\boldsymbol{\beta}+\Lambda\boldsymbol{\alpha}+I\,\boldsymbol{\alpha}\times\boldsymbol{\beta}$$

Note that this multiplication is isomorphic to (1) by

$$\begin{array}{ccc}\Gamma+\boldsymbol{\alpha} & \leftrightarrow & \Gamma+I\,\boldsymbol{\alpha}\\ \circ & \leftrightarrow & \odot\end{array}$$

Also note that, in addition to (3), which now reads

$$(\mathfrak{A}\circ\mathfrak{B})^-=\mathfrak{B}^-\circ\mathfrak{A}^-$$

we also have

$$(\mathfrak{A}\circ\mathfrak{B})^*=\mathfrak{B}^*\circ\mathfrak{A}^*$$

implying that

$$(\mathfrak{A}\circ\mathfrak{B})^+=\mathfrak{A}^+\circ\mathfrak{B}^+$$

Finally, (2) changes to

$$\exp_\circ(\Gamma+\boldsymbol{\alpha})=(\cosh\alpha+\hat{\boldsymbol{\alpha}}\sinh\alpha)\exp\Gamma \tag{5}$$

where all quantities are now, in general, *complex*, including $\alpha = \sqrt{\alpha_x^2 + \alpha_y^2 + \alpha_z^2}$. Note that (5), being an *even* function of α, remains single-valued. Also note that there are important special cases of (5), e.g.

$$\exp_\circ(I\,\mathbf{b}) = (\cos b + I\,\hat{\mathbf{b}}\sin b)$$

In Special Relativity (taking $c = 1$, by a choice of units), points of space-time (*real* elements of Dirac algebra, denoted $t + \mathbf{r}$ and called 4-VECTORS) can be transformed (defining the so called LORENTZ TRANSFORMATION) by

$$t' + \mathbf{r}' = \mathfrak{R} \circ (t + \mathbf{r}) \circ \mathfrak{R}^*$$

where

$$\mathfrak{R} \circ \mathfrak{R}^- = \mathfrak{R}^- \circ \mathfrak{R} = 1$$

This transormation clearly *preserves* the following measure of space-time 'distance' (can be of either sign)

$$(t + \mathbf{r}) \circ (t + \mathbf{r})^-$$

and keeps the new coordinates *real*.

The laws of Physics retain the same form in the new (primed) coordinate system.

It is quite easy to see that $\mathfrak{R}$ must have the form of $\exp_\circ(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ is a *complex* vector. It is only slightly more difficult to prove that this can also be written as

$$\mathfrak{R} = \exp_\circ\left(\tfrac{\mathbf{a}}{2}\right) \circ \exp_\circ\left(\tfrac{\mathbf{b}}{2I}\right)$$

where $\mathbf{a}$ and $\mathbf{b}$ are *real* vectors. The first of these factors corresponds to a so called BOOST in the direction of $\mathbf{a}$, with the speed of $\tanh a$, the second factor is the usual 3-dimensional rotation of $\mathbf{r}$, around the $\mathbf{b}$-direction, by angle b.

In addition to 4-vectors, we need to be able to transform 4-COVECTORS, such as for example $(t + \mathbf{r})^-$ and $\partial_t + \boldsymbol{\nabla}$, thus:

$$\mathfrak{R}^+ \circ (\partial_t + \boldsymbol{\nabla}) \circ \mathfrak{R}^-$$

MIXED 4-VECTORS (4-vector ∘ 4-covector), by

$$\mathfrak{R} \circ \mathfrak{M} \circ \mathfrak{R}^-$$

and MIXED 4-COVECTORS (4-covector ∘ 4-vector):

$$\mathfrak{R}^+ \circ \mathfrak{M}^* \circ \mathfrak{R}^*$$

These are the only physically meaningful quantities (unlike, for example, a product of two 4-vectors, which would be impossible to transform). Note that a mixed 4-vector (mixed 4-covector) can be raised to any integer power, remaining a mixed 4-vector (mixed 4-covector). It is thus physically meaningful to compute simple functions, such as (5), of mixed 4-(co)vectors.

Also note that the operation of reflection converts a mixed 4-vector into another mixed 4-vector (same with mixed 4-covertors), which implies the scalar part of any mixed 4-vector must be an invariant (have the same value, after any Lorentz transformation).

A typical mixed 4-covector is constructed by

$$(\partial_t + \boldsymbol{\nabla}) \circ (\varphi + \mathbf{A}) \equiv -\mathbf{E} + I\,\mathbf{B}$$

imposing the Lorentz condition of $\partial_t \varphi + \boldsymbol{\nabla} \cdot \mathbf{A} \equiv 0$. This converts the 4-potential $\varphi + \mathbf{A}$ into the corresponding electromagnetic field.

Maxwell equations then read:

$$(\partial_t - \boldsymbol{\nabla}) \circ (-\mathbf{E} + I\,\mathbf{B}) = \rho + \mathbf{J}$$

where the right hand side is the charge and current density. Note that a 4-vector multiplied by a mixed 4-covector (the left hand side) returns a 4-vector.

If $\mathfrak{X} = t + \mathbf{r}$ are coordinates of a moving, unit-mass (can be always arranged, by a choice of units), point-like particle, the corresponding 4-momentum is $\mathfrak{P} \equiv \overset{\bullet}{\mathfrak{X}}$, where the dot indicates differentiation with respect to PROPER TIME τ (a scalar invariant), defined by

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{d\mathbf{r}}{dt} \cdot \frac{d\mathbf{r}}{dt}}$$

More explicitely,

$$\mathfrak{P} = \gamma(1 + \mathbf{v})$$

where $\mathbf{v} \equiv \frac{d\mathbf{r}}{dt}$ and $\gamma \equiv \frac{1}{\sqrt{1 - \mathbf{v} \cdot \mathbf{v}}}$.

Furthermore, when the particle is of unit charge, the Lorentz-force law reads:

$$\overset{\bullet}{\mathfrak{P}} = \frac{(\mathbf{E} + I\,\mathbf{B}) \circ \mathfrak{P} + \mathfrak{P} \circ (\mathbf{E} - I\,\mathbf{B})}{2}$$

This equation simplifies by introducing

$$\mathfrak{P} = \mathfrak{U} \circ \mathfrak{P}_0 \circ \mathfrak{U}^*$$

where $\mathfrak{P}_0$ is the initial 4-momentum and $\mathfrak{U}$ solves

$$\overset{\bullet}{\mathfrak{U}} = \frac{(\mathbf{E} + I\,\mathbf{B}) \circ \mathfrak{U}}{2}$$

with $\mathfrak{U}_0 = 1$.

When $\mathbf{E}$ and $\mathbf{B}$ are uniform (i.e. space and time independent) fields, this yields

$$\mathfrak{U} = \exp_\circ\left(\tfrac{\mathbf{E} + I\,\mathbf{B}}{2}\tau\right)$$

and, consequently,

$$\mathfrak{P} = \exp_\circ\left(\tfrac{\mathbf{E} + I\,\mathbf{B}}{2}\tau\right) \circ \mathfrak{P}_0 \circ \exp_\circ\left(\tfrac{\mathbf{E} - I\,\mathbf{B}}{2}\tau\right) \tag{6}$$

To express the previous solution in a form which can be easily integrated (we still have to solve $\overset{\bullet}{\mathfrak{X}} = \mathfrak{P}$) is a bit tricky in general, but the tasks simplifies in a few special cases, e.g. when $\mathbf{E}$ and $\mathbf{B}$ are orthogonal, having the same magnitude, say h. We can then write (6) as follows:

$$(1+\tfrac{\tau}{2}\mathbf{E}+\tfrac{\tau}{2}i\mathbf{B})\circ\mathfrak{P}_0\circ(1+\tfrac{\tau}{2}\mathbf{E}-\tfrac{\tau}{2}i\mathbf{B}) = \gamma_0+\gamma_0(\hat{\mathbf{E}}\cdot\mathbf{v}_0)h\tau+\gamma_0[1-(\hat{\mathbf{E}}\times\hat{\mathbf{B}})\cdot\mathbf{v}_0]\tfrac{h^2\tau^2}{2}+ \\ \gamma_0\mathbf{v}_0+\gamma_0[1-(\hat{\mathbf{B}}\times\mathbf{v}_0)]h\tau+\gamma_0[1-(\hat{\mathbf{E}}\times\hat{\mathbf{B}})\cdot\mathbf{v}_0](\hat{\mathbf{E}}\times\hat{\mathbf{B}})\tfrac{h^2\tau^2}{2}$$

Integrating the last expression yields

$$\begin{aligned} t &= t_0+\gamma_0[\tau+(\mathbf{E}\cdot\mathbf{v}_0)\frac{\tau^2}{2}+\gamma_0[\mathbf{E}^2-(\mathbf{E}\times\mathbf{B})\cdot\mathbf{v}_0]\tfrac{\tau^3}{6} \\ \mathbf{r} &= \mathbf{r}_0+\gamma_0\mathbf{v}_0\tau+\gamma_0[\mathbf{E}-(\mathbf{B}\times\mathbf{v}_0)]\frac{\tau^2}{2}+\gamma_0[(\mathbf{E}\cdot\mathbf{v}_0)\mathbf{E}+(\mathbf{B}\cdot\mathbf{v}_0)\mathbf{B}+\mathbf{E}\times\mathbf{B}]\tfrac{\tau^3}{6} \end{aligned}$$

[1] D. Hestenes: *New Foundations for Classical Mechanics, 2^{nd} Edition* (1999) Kluwer Academic Publishers