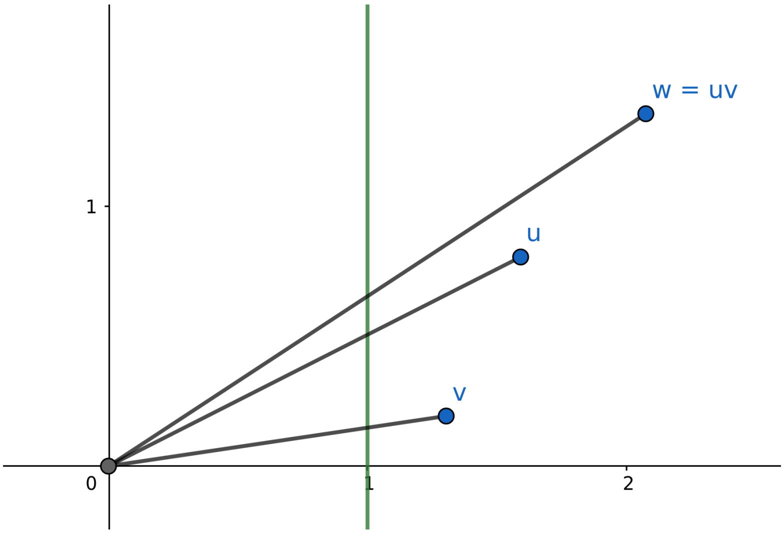
Dual Numbers

The dual numbers are defined analogously to complex numbers by introducing an element ε such that $\varepsilon^2 = 0$.

We then have for multiplication $(a+\varepsilon b)(c+\varepsilon d) = ac+\varepsilon(ad+bc)$ and for the (squared) magnitude $||a+\varepsilon b||^2 := (a+\varepsilon b)(a-\varepsilon b) = a^2$ and $||a+\varepsilon b|| = a$ resp. Thus unit dual numbers have the form $1+\varepsilon x$ or $-1+\varepsilon x$. Pure dual numbers εx have zero magnitude and therefore cannot be normalized.

Just as with complex numbers, we can plot dual numbers on a plane to better understand their properties. Any normalizable number can be written as $r(1+\varepsilon y)$ where y is where a line extending from the origin intersects the $x = 1$ line. Thus multiplication becomes $r_1(1+\varepsilon y_1)r_2(1+\varepsilon y_2) = r_1 r_2(1+\varepsilon(y_1+y_2))$. If we only consider the points on the $x = 1$ line, this is simply a 1D translation, and the magnitudes multiply.



Note also that if we define a function $T(x) := 1+\varepsilon x$, we see that $T(x)T(y) = T(x+y)$. These results are very reminiscent of complex number multiplication and so we should immediately consider whether $T(x) = e^{\varepsilon x}$.

We start with the definition $e^x := \lim_{n\to\infty}(1+\frac{x}{n})^n$, so in our case $e^{\varepsilon x} = \lim_{n\to\infty}(1+\frac{\varepsilon x}{n})^n$. But this is easy: we saw that multiplication of unit dual numbers is simply addition in the dual part, so exponentiation becomes multiplication and the n's cancel:

$$\begin{aligned} e^{\varepsilon x} &= \lim_{n\to\infty}\left(1+\frac{\varepsilon x}{n}\right)^n \\ &= \lim_{n\to\infty} 1+n\frac{\varepsilon x}{n} \\ &= 1+\varepsilon x = T(x) \end{aligned}$$

So just like e^{ix} moves around the unit circle $e^{\varepsilon x}$ moves along the unit line.

Automatic Differentiation

Dual numbers have another interesting property: $f(x+\varepsilon b) = f(x)+\varepsilon b f'(x)$, i.e. one automatically gets the derivative in the dual part. Let's look at some examples:

(NB: $x = 0$ has to be considered separately here)

$$
\begin{aligned}
(x+\varepsilon)^n &= \left(x\left(1+\frac{\varepsilon}{x}\right)\right)^n \\
&= x^n\left(1+n\frac{\varepsilon}{x}\right) \\
&= x^n + \varepsilon n x^{n-1}
\end{aligned}
$$

$$
\begin{aligned}
e^{x+\varepsilon} &= e^x e^\varepsilon \\
&= e^x(1+\varepsilon) \\
&= e^x + \varepsilon e^x
\end{aligned}
$$

$$
\begin{aligned}
\log(x+\varepsilon) &= \log\left(x\left(1+\varepsilon\frac{1}{x}\right)\right) \\
&= \log(x) + \log\left(1+\varepsilon\frac{1}{x}\right) \\
&= \log(x) + \varepsilon\frac{1}{x}
\end{aligned}
$$

That this works in general (or at least for functions that can be expressed with a Taylor series) is easy to show:

$$
\begin{aligned}
f(x+\varepsilon b) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x+\varepsilon b-a)^n \\
&= \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}(\varepsilon b)^n \text{ (with } a = x\text{)} \\
&= f(x) + \varepsilon b f'(x) + \varepsilon^2(...) \\
&= f(x) + \varepsilon b f'(x)
\end{aligned}
$$

We also get the product- and chain-rule (the sum-rule is trivial):

$$
\begin{aligned}
f(x+\varepsilon)g(x+\varepsilon) &= (f(x)+\varepsilon f'(x))(g(x)+\varepsilon g'(x)) \\
&= f(x)g(x) + \varepsilon(f'(x)g(x) + f(x)g'(x))
\end{aligned}
$$

$$
\begin{aligned}
f(g(x+\varepsilon)) &= f(g(x)+\varepsilon g'(x)) \\
&= f(g(x)) + \varepsilon f'(g(x))g'(x)
\end{aligned}
$$

Dual Quaternions

For geometric purposes dual numbers only are of limited use because projecting onto the unit line only leaves us with one dimension. Fortunately we can add as many dimensions as we like if we demand $\varepsilon_i\varepsilon_j = 0$ (the same does not work for complex numbers because what is ij if there is no k?). This gives us translation in any dimension but that is also of limited use if we cannot rotate: we want quaternions. Interestingly we can combine both ideas to give us dual quaternions.

A general dual quaternion looks like $q = q_r + \varepsilon q_d$ where q_r and q_d are regular quaternions and ε commutes with qvectors. Multiplication then gives $qp = q_r p_r + \varepsilon(q_r p_d + q_d p_r)$.

Now consider translation. We've seen that $e^{\varepsilon\mathbf{t}} = 1 + \varepsilon\mathbf{t}$ and two dual qvectors $\mathbf{t}$ and $\mathbf{s}$ are added as we expect:

$$\begin{aligned} e^{\varepsilon\mathbf{t}}e^{\varepsilon\mathbf{s}} &= (1+\varepsilon\mathbf{t})(1+\varepsilon\mathbf{s}) \\ &= 1+\varepsilon(\mathbf{t}+\mathbf{s}) \\ &= e^{\varepsilon(\mathbf{t}+\mathbf{s})} \end{aligned}$$

To combine rotation and translation however we need a general dual quaternion and to use the more general sandwich product, this however is easy to construct.

Note that a pure dual qvector is unaffected by a sandwich product like this:

$$\begin{aligned} e^{\varepsilon\frac{1}{2}\mathbf{t}}(\varepsilon\mathbf{v})e^{\varepsilon\frac{1}{2}\mathbf{t}} &= (1+\varepsilon\frac{1}{2}\mathbf{t})\varepsilon\mathbf{v}(1+\varepsilon\frac{1}{2}\mathbf{t}) \\ &= \varepsilon\mathbf{v} \end{aligned}$$

If, however, we add 1 to this dual qvector, it is translated by $\mathbf{t}$:

$$\begin{aligned} e^{\varepsilon\frac{1}{2}\mathbf{t}}(1+\varepsilon\mathbf{v})e^{\varepsilon\frac{1}{2}\mathbf{t}} &= e^{\varepsilon\frac{1}{2}\mathbf{t}}e^{\varepsilon\frac{1}{2}\mathbf{t}} + e^{\varepsilon\frac{1}{2}\mathbf{t}}\varepsilon\mathbf{v}e^{\varepsilon\frac{1}{2}\mathbf{t}} \\ &= e^{\varepsilon\mathbf{t}} + \varepsilon\mathbf{v} \\ &= 1+\varepsilon\mathbf{t}+\varepsilon\mathbf{v} \\ &= 1+\varepsilon(\mathbf{t}+\mathbf{v}) \end{aligned}$$

Both types of objects are affected by rotation (with $q = e^{\mathbf{r}\frac{\theta}{2}}$):

$$q(\varepsilon\mathbf{v})q^* = \varepsilon(q\mathbf{v}q^*)$$

$$\begin{aligned} q(1+\varepsilon\mathbf{p})q^* &= q1q^* + \varepsilon(q\mathbf{p}q^*) \\ &= 1+\varepsilon(q\mathbf{p}q^*) \end{aligned}$$

Thus we can identify points with $e^{\varepsilon\mathbf{p}}$ and have them be affected by rotation and translation, and vectors with $\varepsilon\mathbf{v}$ and have them only be affected by rotation. This is exactly what we want.

Now we only have to consider the composition of these transformations, but for this we first need a new type of conjugate. Under conjugation quaternion multiplication reverses its order: $(qp)^* = p^*q^*$. This is an important property when dealing with sandwich products because it lets us calculate the object on the one side from the one on the other side. If translation and rotation are to combine nicely, we need a conjugate on the right side of translation as well, but regular quaternion conjugation does not give us the correct result. We need a new conjugate with the property $(1+\varepsilon\mathbf{t})^\circ = 1+\varepsilon\mathbf{t}$, which at the same time includes the quaternion conjugate $\mathbf{t}^*$. $(1+\varepsilon\mathbf{t})^\circ = 1-\varepsilon\mathbf{t}^*$ has this property, so now let us investigate if $(qp)^\circ = p^\circ q^\circ$ holds for any dual quaternions with $q^\circ := q_r^* - \varepsilon q_d^*$:

$$\begin{aligned} (qp)^\circ &= (q_r p_r + \varepsilon(q_r p_d + q_d p_r))^\circ \\ &= (q_r p_r)^* - \varepsilon((q_r p_d)^* + (q_d p_r)^*) \\ &= p_r^* q_r^* - \varepsilon(p_d^* q_r^* + p_r^* q_d^*) \\ &= (p_r^* - \varepsilon p_d^*)(q_r^* - \varepsilon q_d^*) \\ &= p^\circ q^\circ \end{aligned}$$

Using this definition we can now express rotation around the origin (with quaternion r) followed by a translation by $\mathbf{t}$:

$$
\begin{aligned}
e^{\varepsilon\frac{1}{2}\mathbf{t}} r(1+\varepsilon\mathbf{v})r^* e^{\varepsilon\frac{1}{2}\mathbf{t}} &= (1+\varepsilon\frac{1}{2}\mathbf{t})r(1+\varepsilon\mathbf{v})r^*(1+\varepsilon\frac{1}{2}\mathbf{t}) \\
&= (r+\varepsilon\frac{1}{2}\mathbf{t}r)(1+\varepsilon\mathbf{v})(r^*+\varepsilon\frac{1}{2}r^*\mathbf{t}) \\
&= (r+\varepsilon\frac{1}{2}\mathbf{t}r)(1+\varepsilon\mathbf{v})(r^*-\varepsilon\frac{1}{2}r^*\mathbf{t}^*) \\
&= (r+\varepsilon\frac{1}{2}\mathbf{t}r)(1+\varepsilon\mathbf{v})(r^*-\varepsilon\frac{1}{2}(\mathbf{t}r)^*) \\
&= (r+\varepsilon\frac{1}{2}\mathbf{t}r)(1+\varepsilon\mathbf{v})(r+\varepsilon\frac{1}{2}\mathbf{t}r)^\circ
\end{aligned}
$$

Thus we have arrived at a system that allows us to express and compose translations and rotations around arbitrary axes.