Newton's Method Convergence

We're solving:

$$F(v) = v - \delta t I^{-1} R^t\left(\tfrac{v}{2}\right) L = 0$$

The Theorem of Kantorovich tells us when the Newton's iteration converges. We assume that out first iteration starts at $v = 0$. Then we need to find constants β, η and κ such that

$$\begin{aligned}
\left|F'(0)^{-1}\right| &\leq \beta \\
\left|F'(0)^{-1}F(0)\right| &\leq \eta \\
\left|F'(u) - F'(v)\right| &\leq \lambda|u - v| \text{ for any } u \text{ and } v
\end{aligned}$$

and if these constants respect:

$$\beta\eta\lambda \leq \tfrac{1}{2}$$

then the Newton's method applied to $F(v) = 0$ starting at $v = 0$ converges and the iterates are contained inside a ball centered at $v = 0$ and of radius:

$$r = \frac{2\eta}{1 + \sqrt{1 - 2\lambda\beta\eta}}$$

First we compute the derivative of F:

$$F'(v) = \mathrm{Id} - \tfrac{\delta t}{2} I^{-1}\left[R^t\left(\tfrac{v}{2}\right)L\right]_\times T^t\left(\tfrac{v}{2}\right)$$

and we get the first bound:

$$\begin{aligned}
\left|F'(0)^{-1}\right| &= \left|\left(\mathrm{Id} - \tfrac{\delta t}{2} I^{-1} L_\times\right)^{-1}\right| \\
&= \left|\sum_{i=0}^{\infty}\left(\tfrac{\delta t}{2} I^{-1} L_\times\right)^i\right| \\
&\leq \sum_{i=0}^{\infty}\left|\tfrac{\delta t}{2} I^{-1} L_\times\right|^i \\
&= \frac{1}{1 - \tfrac{\delta t}{2}\left|I^{-1} L_\times\right|} \\
&\leq \frac{1}{1 - \tfrac{\delta t}{2}\sigma(I^{-1})|L|} =: \beta
\end{aligned}$$

We need to assume that $\frac{\delta t}{2}\sigma(I^{-1})|L| < 1$, or the geometric series might not converge.

For the second bound, we notice that:

$$\left|F(0)\right| = \delta t\left|I^{-1}L\right| \leq \delta t \cdot \sigma(I^{-1}) \cdot |L|$$

and we can set

$$\eta := \frac{\delta t \cdot \sigma(I^{-1}) \cdot |L|}{1 - \frac{\delta t}{2}\sigma(I^{-1})|L|}$$

This can be made much tighter by using some properties of the product $F'(0)^{-1}F(0)$, but making the math much more complex.

Finally the last bound is:

$$\begin{aligned}
&\tfrac{\delta t}{2}\left|I^{-1}\left[R^t\left(\tfrac{v}{2}\right)L\right]_\times T^t\left(\tfrac{v}{2}\right) - I^{-1}\left[R^t\left(\tfrac{u}{2}\right)L\right]_\times T^t\left(\tfrac{u}{2}\right)\right| \\
&\leq \frac{\delta t \cdot \sigma(I^{-1})}{2}\left|\left[R^t\left(\tfrac{v}{2}\right)L\right]_\times T^t\left(\tfrac{v}{2}\right) - \left[R^t\left(\tfrac{u}{2}\right)L\right]_\times T^t\left(\tfrac{u}{2}\right)\right| \\
&\leq \frac{\delta t \cdot \sigma(I^{-1})}{2} \cdot \frac{|v - u|}{2} \cdot \left|L\right|
\end{aligned}$$

where the last inequality follows from the fact that the derivative of the exponential is 1-Lipschitz (see next section), and we can define:

$$\lambda := \frac{\delta t \cdot \sigma(I^{-1}) \cdot \left|L\right|}{4}$$

To simplify the notation, let's set:

$$\Gamma := \sigma(I^{-1})|L|$$

We have our 3 constants:

$$\begin{aligned}
\beta &= \frac{1}{1 - \frac{\delta t}{2}\Gamma} \\
\eta &= \frac{\delta t \cdot \Gamma}{1 - \frac{\delta t}{2}\Gamma} \\
\lambda &= \frac{\delta t \cdot \Gamma}{4}
\end{aligned}$$

The condition

$$\beta\eta\lambda \leq \tfrac{1}{2}$$

becomes:

$$\frac{\delta t^2 \Gamma^2}{2\left(1 - \frac{\delta t}{2}\Gamma\right)^2} \leq 1$$

But we already assumed that $\frac{\delta t}{2}\Gamma < 1$, implying the term under the square in the denominator is positive, and we have:

$$\delta t \leq \frac{2}{\sqrt{2} + 1}\Gamma^{-1}$$

The radius of the ball containing the iterates is:

$$\begin{aligned}
r &= \frac{2\eta}{1 + \sqrt{1 - 2\lambda\beta\eta}} \\
&\leq 2\eta = \frac{2\delta t \Gamma}{1 - \frac{\delta t}{2}\Gamma} \\
&\leq 2\sqrt{2}
\end{aligned}$$

Derivative is 1-Lipschitz

Let's prove that the derivative of R is 1-Lipschitz. Which is to say that for any vectors X and Y, we have:

$$\left|\left[\left[\mathrm{D}_u\left(R(u)X\right)\right]Y - \left[\mathrm{D}_v\left(R(v)X\right)\right]Y\right]\right| \leq |u - v||X||Y|$$

for any u and v. This is done using the integral definition of T:

$$T(v) = \int_0^1 R(su)ds = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} R(\tfrac{iv}{n})$$

So

$$
\begin{aligned}
\left|\left[\mathrm{D}_u\left(R(u)X\right)\right]Y-\left[\mathrm{D}_v\left(R(v)X\right)\right]Y\right| &= \left|[R(u)X]\times[T(u)Y]-[R(v)X]\times[T(v)Y]\right| \\
&= \lim_{n\to\infty}\frac{1}{n}\left|\sum_{i=0}^{n}[R(u)X]\times\left[R(\tfrac{iu}{n})Y\right]-[R(v)X]\times\left[R(\tfrac{iv}{n})Y\right]\right| \\
&\leq \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\left|[R(u)X]\times\left[R(\tfrac{iu}{n})Y\right]-[R(v)X]\times\left[R(\tfrac{iv}{n})Y\right]\right|
\end{aligned}
$$

and we only need to prove that each summand is less than $|u-v||X||Y|$:

$$
\begin{aligned}
&\left|[R(u)X]\times\left[R(\tfrac{iu}{n})Y\right]-[R(v)X]\times\left[R(\tfrac{iv}{n})Y\right]\right| \\
&=\left|R\left(\tfrac{i}{n}u\right)\left[\left(R\left(\tfrac{n-i}{n}u\right)X\right)\times Y\right]-R\left(\tfrac{i}{n}v\right)\left[\left(R\left(\tfrac{n-i}{n}v\right)X\right)\times Y\right]\right| \\
&=\left|\left[R\left(\tfrac{i}{n}u\right)-R\left(\tfrac{i}{n}v\right)\right]\left[\left(R\left(\tfrac{n-i}{n}u\right)X\right)\times Y\right]-R\left(\tfrac{i}{n}v\right)\left[\left(\left(R\left(\tfrac{n-i}{n}v\right)-R\left(\tfrac{n-i}{n}u\right)\right)X\right)\times Y\right]\right| \\
&\leq\left|\left[R\left(\tfrac{i}{n}u\right)-R\left(\tfrac{i}{n}v\right)\right]\left[\left(R\left(\tfrac{n-i}{n}u\right)X\right)\times Y\right]\right|+\left|R\left(\tfrac{i}{n}v\right)\left[\left(\left(R\left(\tfrac{n-i}{n}v\right)-R\left(\tfrac{n-i}{n}u\right)\right)X\right)\times Y\right]\right| \\
&\leq\left|\tfrac{i}{n}u-\tfrac{i}{n}v\right|\left|X\right|\left|Y\right|+\left|\tfrac{n-i}{n}u-\tfrac{n-i}{n}v\right|\left|X\right|\left|Y\right| \\
&=|u-v||X||Y|
\end{aligned}
$$

where we used the invariance of the cross product under rotation, the triangle inequality and that R is 1-Lipschitz.