Underpinning theory of the generalized scaling and squaring approximated exponential integrators, and some other things.

Partecipants: Sebastiano Ferraris, Marco Lorenzi, Pankaj Daga, Marc Modat, Tom Vercauteren

1 Generalized scaling and squaring with approximated exponential integrator

Details of the proof of the generalized scaling and squaring with approximated exponential integrator formula.

Thesis: Let $\Omega \subset \mathbb{R}^d$ be an image's domain ($d = 2, 3$ for bi and tri-dimensional images), $\phi : \Omega \to \Omega$ a diffeomorphism and $\mathbf{v}$ its tangent vector field that defines the transformation's velocities and directions at each point of Ω.
The relationship between $\mathbf{v}$ and ϕ is given by the stationary ODE

$$\frac{\mathrm{d}\phi_t}{\mathrm{d}t} = \mathbf{v}(\phi_t), \qquad \phi_0 = \mathrm{Id} \ , \tag{1}$$

where ϕ_0 coincides with the identity function Id defined on Ω, and the solution at the time point $t = 1$ coincides with the diffeomorphism ϕ. The solution is given by

$$\phi_1(\mathbf{x}) \simeq \mathbf{x} + \mathbf{v}(\mathbf{x}) + \frac{1}{2}\mathrm{J}_{\mathbf{v}(\mathbf{x})}\mathbf{v}(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega \ . \tag{2}$$

Some introductory facts: At the core of the ODE (1) there is the concept of *flow of diffeomorphisms*. It is defined as the family of diffeomorphisms $\{\phi_t\}_{t\in\mathbf{R}}$ continuously parametrized by a time-parameter t, such that ϕ_0 equals the identity Id and that satisfy the one-parameter subgroup property $\phi_t \circ \phi_s = \phi_{t+s}$.

When the one-parameter subgroup is applied to a point $\mathbf{x}$ in Ω, the new point $\phi_t(\mathbf{x})$ can be *denoted* with $\mathbf{x}(t)$ and its time derivative with $\dot{\mathbf{x}}(t)$. Equation (1), when considered for one particular point, can be rewritten as

$$\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}(t)) \qquad \mathbf{x}(0) = \mathbf{x} \ . \tag{3}$$

Some more introductory facts: For any real $d \times d$ matrix A and any d-dimensional column vector $\mathbf{b}$, it holds

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{b} \qquad \Longrightarrow \qquad \mathbf{x}(t) = \varphi_0(tA)\mathbf{b} \quad \forall t \in \mathbb{R} \ . \tag{4}$$

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \mathbf{b}, \quad \mathbf{x}(0) = \mathbf{0} \quad \Longrightarrow \quad \mathbf{x}(t) = t\varphi_1(tA)\mathbf{b} \quad \forall t \in \mathbb{R} \ . \tag{5}$$

Where $\varphi_0(Z)$ is the matrix exponential of $e^Z = \sum_{j=0}^{\infty} \frac{Z^k}{j!}$, whose numerical computation is performed with expm, and $\varphi_1(Z)$ is the shifted Taylor expansion given by $(\varphi_0(Z) - I)Z^{-1} = \sum_{j=0}^{\infty} \frac{Z^k}{(j+1)!}$ [3]. For a positive integer k, $\varphi_k(Z) = \sum_{j=0}^{\infty} \frac{Z^k}{(j+k)!}$.

Explanation of the equation (2): Without any loss of generality it is always possible to *translate the coordinate frame* so that the initial position of $\mathbf{x}(0)$ coincides with the origin of the axis $\mathbf{0}$. The translation is given by $\mathbf{y}(t) := \mathbf{x}(t) - \mathbf{x}(0)$, and in this new frame we have $\mathbf{y}(0) = \mathbf{0}$, $\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{x}(0)$, and equation (3) can be written as:

$$\dot{\mathbf{y}}(t) = \mathbf{v}(\mathbf{y}(t) + \mathbf{x}(0)) \qquad \mathbf{y}(0) = \mathbf{0} \ . \tag{6}$$

The Taylor expansion of the SVF $\mathbf{v}$ around $\mathbf{x}(0)$ computed at $\mathbf{y}(t) + \mathbf{x}(0)$ for any real t provides[1]

$$\dot{\mathbf{y}}(t) = \mathbf{v}(\mathbf{y}(t) + \mathbf{x}(0)) = \mathbf{v}(\mathbf{x}(0)) + \mathrm{J}_{\mathbf{v}(\mathbf{x}(0))}\mathbf{y}(t) + \mathcal{O}(\mathbf{y}(t)^2) \ .$$

This last expansion, gives us the hint ot follow the *exponential integrators approach* [2]. It consists in the strategy of separating the linear part (whose integration is provided by the exponential map expm) and the non-linear part of the SVF. Equation (6) can be written as

$$\dot{\mathbf{y}}(t) = \mathbf{v}_0 + \mathrm{J}_{\mathbf{v}_0}\mathbf{y}(t) + \mathcal{N}_{\mathbf{v}}(\mathbf{y}(t)) \ , \qquad \mathbf{y}(0) = \mathbf{0} \ , \tag{7}$$

where $\mathbf{v}(\mathbf{x}(0))$ is indicated with $\mathbf{v}_0$ for notation convenience and $\mathrm{J}_{\mathbf{v}_0}$ is the $d \times d$ spatial Jacobian of $\mathbf{v}_0$. The non-linear part of the SVF, indicated with $\mathcal{N}_{\mathbf{v}}(\mathbf{y}(t))$, can be seen as an operator on the space of the SVF over Ω, that subtract the linear part of $\mathbf{v}$ computed with the Taylor expansion of $\mathbf{v}$ in $\mathbf{x}(t)$ around $\mathbf{x}(0)$. For a fixed $\mathbf{x} = \mathbf{x}(0) \in \Omega$, on which is acting a one-parameter subgroup of diffeomorphisms, it is defined by:

$$\begin{aligned} \mathcal{N}_{\mathbf{v}}(\mathbf{y}(t)) &= \mathbf{v}(\mathbf{y}(t) + \mathbf{x}(0)) - \big(\mathbf{v}_0 + \mathrm{J}_{\mathbf{v}_0}\mathbf{y}(t)\big) \\ &= \mathbf{v}(\mathbf{x}(t)) - \mathbf{v}(\mathbf{x}(0)) - \mathrm{J}_{\mathbf{v}(\mathbf{x}(0))}(\mathbf{x}(t) - \mathbf{x}(0)) \ . \end{aligned}$$

It follows easily that $\mathcal{N}_{\mathbf{v}}(\mathbf{y}(0)) = \mathbf{0}$ and $\mathcal{N}_{\mathbf{v}}(\mathbf{y}(t)) \in \mathcal{O}\big((\mathbf{x}(t) - \mathbf{x}(0))^2\big)$ when $\mathbf{x}(t) \to \mathbf{x}(0)$.

When the time-parameter t is in a small neighbour of the origin (as it happens when scaling the SVF by an appropriate factor in the generalized scaling and

[1] Given f, g and h, vector valued function in an Euclidean space, with the notation $f(x) = g(x) + \mathcal{O}(h(x))$ for $x \to x_0$ we mean that exists a real positive M and a δ such that $\|f(x) - g(x)\| < M\|h(x)\|$ when $\|x - x_0\| < \delta$.

squaring framework), the non linear part is small, and at the initial ODE problem we can associate a linearised version:

$$\dot{\mathbf{y}}(t) = \mathbf{v}_0 + \mathrm{J}_{\mathbf{v}_0}\mathbf{y}(t) \ , \qquad \mathbf{y}(0) = \mathbf{0} \ . \tag{8}$$

Since $\mathcal{N}_{\mathbf{v}}(\mathbf{y}(t)) \in \mathcal{O}\big(\mathbf{y}(t)^2\big)$, the solution of the linearised problem 8 approximates the solution of the initial problem (6).

Passing in *homogeneous coordinates*, equation (8) can be written as

$$\begin{bmatrix} \dot{\mathbf{y}} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathrm{J}_{\mathbf{v}_0} & \mathbf{v}_0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ 1 \end{bmatrix}$$

And using the implication (4), we have that the solution, written again in homogeneous coordinates, is

$$\begin{bmatrix} \mathbf{y}(t) \\ 1 \end{bmatrix} = \varphi_0\left(t \begin{bmatrix} \mathrm{J}_{\mathbf{v}_0} & \mathbf{v}_0 \\ 0 & 0 \end{bmatrix} \right) \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \tag{9}$$

which is the exact solution of the linearised ODE, that approximates the sought solution for any time parameter t close enough to 0.

In order to *avoid the computational cost of the exponential of a matrix*, and to have an approach that can be easily vectorized, we can apply (5) to the linearised problem (8) to obtain the following solution:

$$\mathbf{y}(t) = t\varphi_1(t\mathrm{J}_{\mathbf{v}_0})\mathbf{v}_0 \ .$$

and, by translating the coordinate frame to the initial one with $\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}(0)$, it follows:

$$\mathbf{x}(t) = \mathbf{x}(0) + t\varphi_1(t\mathrm{J}_{\mathbf{v}_0})\mathbf{v}_0 \ , \qquad \mathbf{x}(1) = \mathbf{x}(0) + \varphi_1(\mathrm{J}_{\mathbf{v}_0})\mathbf{v}_0 \ ,$$

that is the solution of the linearised ODE associated to (1) at the point $\mathbf{x}$.

For its *numerical computation, we can approximate* φ_1 truncating it at its second order:

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x} + t\varphi_1(t\mathrm{J}_{\mathbf{v}_0})\mathbf{v}_0 \\ &= \mathbf{x} + t\big(\mathrm{I} + \frac{t\mathrm{J}_{\mathbf{v}_0}}{2} + \frac{t^2\mathrm{J}_{\mathbf{v}_0}^2}{6} + \cdots\big)\mathbf{v}_0 \\ &= \mathbf{x} + t\mathbf{v}_0 + \frac{t^2}{2}\mathrm{J}_{\mathbf{v}_0}\mathbf{v}_0 + \mathcal{O}\big(t^3\mathrm{J}_{\mathbf{v}_0}^2\mathbf{v}_0\big) \ . \end{aligned}$$

for $t \to 0$. Therefore the solution to the ODE 1 can be written as

$$\phi_t(\mathbf{x}) = \mathbf{x} + t\mathbf{v}_0 + \frac{t^2}{2}\mathrm{J}_{\mathbf{v}_0}\mathbf{v}_0 + \mathcal{O}\big(t^3\mathrm{J}_{\mathbf{v}_0}^2\mathbf{v}_0\big) + \mathcal{O}\big((\mathbf{x}(t) - \mathbf{x}(0))^2\big) \ , \tag{10}$$

where the first asymptotic error limit is a consequence of having truncated φ_1, and the second as a consequence of having linearised the problem.

For $t \to 0$, the last equation can be rewritten as a function defined over Ω as:

$$\phi_t = \mathrm{Id} + t\mathbf{v} + \frac{t^2}{2}\mathrm{J}_{\mathbf{v}}\mathbf{v} + \mathcal{O}\big(t^3\mathrm{J}^2_{\mathbf{v}}\mathbf{v} + (\mathbf{x}(t) - \mathbf{x}(0))^2\big) \ , \tag{11}$$

where $\mathrm{J}_{\mathbf{v}}$ is the Jacobian function that at each $\mathbf{x}$ provides the vector valued operators $\mathrm{J}_{\mathbf{v}(\mathbf{x})}$.

When $t = 1$ we have that the initial ODE system can have approximation solution as:

$$\phi_1 \simeq \mathrm{Id} + \mathbf{v} + \frac{1}{2}\mathrm{J}_{\mathbf{v}}\mathbf{v} \ , \tag{12}$$

Using equation (12) in the generalized scaling and squaring: The equation just derived can be applied to the SVF $\mathbf{w}$, after having it reduced by a multiplicative factor of 2^N in the scaling and squaring framework. The final algorithm that improves the currently used scaling and squaring is given by the following steps:

1. Scaling of $\mathbf{w}$ by a factor of 2^N: $\mathbf{v} = \mathbf{w}/2^N$.
2. The approximation of the Lie exponential, indicated with $\widetilde{\exp}(\mathbf{v})$ si computed as
$$\widetilde{\exp}(\mathbf{v}) = \mathbf{x} + \mathbf{v}(\mathbf{x}) + \frac{1}{2}\mathrm{J}_{\mathbf{v}(\mathbf{x})}\mathbf{v}(\mathbf{x}) \ .$$
3. The result $\widetilde{\exp}(\mathbf{v})$ is pair-wise composed by itself 2^N-times.

2 Product of two vector fields: an attempt to compute the Taylor series

There are no natural way in defining the product between stationary velocity fields (or in general vector fields). As an attempt we exploit the concept of directional derivative; aim of this paragraph is to present how this is defined, how the Jacobian matrix can be involved in the computations and how the notation $\frac{\partial}{\partial x_i}$ to indicate the elements of the basis of the real vector space can be considered.

Let Ω be a subset of $\mathbf{R}^d$ domain of a continuous real valued function $f : \Omega \to \mathbf{R}$ and domain of a velocity field $\mathbf{u} : \Omega \to \mathbf{R}^d$. Considering $d = 2$ to simplify the notation, we can write $\mathbf{u}$ in components as

$$\mathbf{u}(x, y) = (u_x(x, y), u_y(x, y))$$

Where the components of $\mathbf{u}$ are real valued functions defined over the domain, $u_x : \Omega \to \mathbf{R}$, $u_y : \Omega \to \mathbf{R}$.

The directional derivative of f in the direction of $\mathbf{u}$ is, by definition, given by

$$\mathrm{D}_{\mathbf{u}} f := u_x \frac{\partial f}{\partial x} + u_y \frac{\partial f}{\partial y} \ , \tag{13}$$

and it is a function defined over Ω if the partial derivative of f are defined.

The directional derivative of f in the direction of $\mathbf{u}$ at the point (x_0, y_0) is defined as

$$\mathrm{D}_{\mathbf{u}} f(x_0, y_0) := u_x(x_0, y_0) \frac{\partial f(x_0, y_0)}{\partial x} + u_y(x_0, y_0) \frac{\partial f(x_0, y_0)}{\partial y} \ .$$

Given two SVF $\mathbf{u}$ and $\mathbf{v}$ (or more generally two vector fields)

$$\mathbf{u}(x, y) = (u_x(x, y), u_y(x, y)) \qquad \mathbf{v}(x, y) = (v_x(x, y), v_y(x, y)) \ ,$$

one of the possible definition of product between them is as directional derivative of the components of the vector field $\mathbf{v}$ in the direction of $\mathbf{u}$:

$$\mathrm{D}_{\mathbf{u}} \mathbf{v} := \big(\mathrm{D}_{\mathbf{u}} v_x, \mathrm{D}_{\mathbf{u}} v_y \big) \ . \tag{14}$$

It results that

$$\mathrm{D}_{\mathbf{u}} \mathbf{v} := \big(u_x \frac{\partial v_x}{\partial x} + u_y \frac{\partial v_x}{\partial y}, u_x \frac{\partial v_y}{\partial x} + u_y \frac{\partial v_y}{\partial y} \big)$$

We observe that $\mathrm{D}_{\mathbf{u}} \mathbf{v}$ is defined if v_x and v_y are differentiable and that it is a vector field as well.

First consequence is that the product introduced is equivalent to the product of the Jacobian of $\mathbf{v}$, indicated with $\mathrm{J}_{\mathbf{v}}$, times the vector field $\mathbf{u}$:

$$\mathrm{D}_{\mathbf{u}} \mathbf{v} = \big(u_x \frac{\partial v_x}{\partial x} + u_y \frac{\partial v_x}{\partial y}, u_x \frac{\partial v_y}{\partial x} + u_y \frac{\partial v_y}{\partial y} \big) = \begin{bmatrix} \frac{\partial v_x}{\partial x} & \frac{\partial v_x}{\partial x} \\ \frac{\partial v_y}{\partial x} & \frac{\partial v_y}{\partial y} \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \mathrm{J}_{\mathbf{v}} \mathbf{u} \ .$$

The Lie bracket of two SVF based on this definition, is provided by

$$[\mathbf{u}, \mathbf{v}] := \mathrm{D}_{\mathbf{u}} \mathbf{v} - \mathrm{D}_{\mathbf{v}} \mathbf{u} = \mathrm{J}_{\mathbf{v}} \mathbf{u} - \mathrm{J}_{\mathbf{u}} \mathbf{v} \ ,$$

that is defined in analogy with the Lie bracket for matrix Lie algebra: let A and B square matrices, their Lie bracket is given by

$$[A, B] := AB - BA \ .$$

The sequential application of directional derivative of SVF is not free of deceptions. When three SVF, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$, are given it is relatively easy to see that we can have two possible chain derivative:

$$\mathrm{D}_{\mathbf{u}} \mathrm{D}_{\mathbf{v}} \mathbf{w} := \mathrm{D}_{\mathbf{u}} \big(\mathrm{D}_{\mathbf{v}} w_x, \mathrm{D}_{\mathbf{v}} w_y \big) = \mathrm{J}_{\mathbf{w}} \mathrm{J}_{\mathbf{v}} \mathbf{u} + \mathbf{u}^T \Big(\mathrm{H}(w_x) + \mathrm{H}(w_y) \Big) \mathbf{v}$$

$$\mathrm{D}_{\mathbf{u}} \mathrm{D}_{\mathbf{v}} \mathbf{w} := \mathrm{D}_{(\mathrm{D}_{\mathbf{u}} v_x, \mathrm{D}_{\mathbf{u}} v_y)} \mathbf{w} = \mathrm{J}_{\mathbf{w}} \mathrm{J}_{\mathbf{v}} \mathbf{u}$$

where H is the Hessian matrix (please verify the computations). In the first case an expression as $\mathrm{D}^n_{\mathbf{v}}\mathbf{v}$ will soon be a nightmare, while in the second case, it holds by induction that $\mathrm{D}^n_{\mathbf{v}}\mathbf{v} = \mathrm{J}^{n-1}_{\mathbf{v}}\mathbf{v}$.

For general dimension: Generalizing the definition to dimension d, for $\Omega \subseteq \mathbf{R}^d$ it follows that the directional derivative of the function f defined over Ω in the direction of the d dimensional SVF $\mathbf{u}$ is given, from 13, by:

$$\mathrm{D}_{\mathbf{u}} f := \sum_{i=1}^{d} u_i \frac{\partial f}{\partial x_i} \ , \tag{15}$$

where u_i are the components of the SVF, i.e. continuous real valued functions defined over Ω, and x_i are the coordinates of the space. When the directional derivative is computed at the point $\mathbf{x}$, previous equation can be written as

$$\mathrm{D}_{\mathbf{u}} f(\mathbf{x}) := \sum_{i=1}^{d} u_i(\mathbf{x}) \frac{\partial f(\mathbf{x})}{\partial x_i} \ ,$$

Given two SVF, $\mathbf{u} = (u_1, u_2, ..., u_d)$ and $\mathbf{v} = (v_1, v_2, ..., v_d)$ their product results to be

$$\mathrm{D}_{\mathbf{u}}\mathbf{v} := \big(\mathrm{D}_{\mathbf{u}} v_1, \mathrm{D}_{\mathbf{u}} v_2, ..., \mathrm{D}_{\mathbf{u}} v_d\big) \ . \tag{16}$$

It results

$$\mathrm{D}_{\mathbf{u}}\mathbf{v} := \Big(\sum_{i=1}^{d} u_i \frac{\partial v_1}{\partial x_i}, \sum_{i=1}^{d} u_i \frac{\partial v_2}{\partial x_i}, \dots, \sum_{i=1}^{d} u_i \frac{\partial v_d}{\partial x_i} \Big) = \mathrm{J}_{\mathbf{v}}\mathbf{u} \ ,$$

where the Jacobian came out from the same computations previously done in dimension 2.

Note about the use of a mathematical notation: From an algebraic point of view, the directional derivative can be considered as the action (in an algebraic sense) of the algebra of vector fields defined over Ω on the ring of the infinite differentiable functions $\mathcal{C}^{\infty}(\Omega)$. To give emphasis on this operation in the notation, some authors used $\frac{\partial}{\partial x_i}$ to indicate the elements of the base. Therefore an SVF, according to this notation, can be indicated with

$$\mathbf{u} = (u_1, u_2, \dots, u_d) = \sum_{i=1}^{d} u_i \mathbf{e}_i = \sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i}$$

Where the elements of the base of the real space $\mathbf{e}_i$ have been indicated with $\frac{\partial}{\partial x_i}$. The catalogue of the things can be done with an SVF with the notation $\mathbf{u} = \sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i}$ becomes simpler:

– evaluation at the point $\mathbf{x}$ of the domain Ω:

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{d} u_i(\mathbf{x}) \frac{\partial}{\partial x_i} \Big|_{\mathbf{x}}$$

that is the vector $\mathbf{u}(\mathbf{x})$ at the point $\mathbf{x}$.

– action on the function f as partial derivative:

$$\mathrm{D}_{\mathbf{u}} f = \sum_{i=1}^{d} u_i(\mathbf{x}) \frac{\partial f}{\partial x_i}$$

– action on the function f as partial derivative at the point $\mathbf{x}$:

$$\mathrm{D}_{\mathbf{u}} f(\mathbf{x}) = \sum_{i=1}^{d} u_i(\mathbf{x}) \frac{\partial f}{\partial x_i} \Big|_{\mathbf{x}} = \sum_{i=1}^{d} u_i(\mathbf{x}) \frac{\partial f(\mathbf{x})}{\partial x_i}$$

In the SVF product, this elegant notation can create problems if we forget that $\frac{\partial}{\partial x_i}$ has the dual value of element of base and directional derivative. Using this notation, the product is computed as

$$\mathrm{D}_{\mathbf{u}} \mathbf{v} = \sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i} \Big(\sum_{j=1}^{d} v_j \frac{\partial}{\partial x_i} \Big) = \sum_{i,j=1}^{d} u_i \frac{\partial v_j}{\partial x_i} \frac{\partial}{\partial x_j}$$

where $\frac{\partial}{\partial x_i}$ acts linearly on the components of $\mathbf{v}$ belonging to $\mathcal{C}^{\infty}(\Omega)$ as partial derivative. The last result is still an SVF, and therefore can act on f, giving as result:

$$(\mathrm{D}_{\mathbf{u}} \mathbf{v}) f = \sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i} \Big(\sum_{j=1}^{d} v_j \frac{\partial}{\partial x_i} \Big) f = \sum_{i,j=1}^{d} u_i \frac{\partial v_j}{\partial x_i} \frac{\partial f}{\partial x_j}$$

But, if $\mathbf{v}$ is acting on a function f before the application of the directional derivative, we have a different results where the intermediate step is not any more an SVF:

$$\mathrm{D}_{\mathbf{u}} (\mathbf{v} f) = \sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i} \Big(\sum_{j=1}^{d} v_j \frac{\partial f}{\partial x_i} \Big) = \sum_{i,j=1}^{d} u_i \frac{\partial v_j}{\partial x_i} \frac{\partial f}{\partial x_j} + \sum_{i,j=1}^{d} u_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j} \ .$$

It follows that $(\mathrm{D}_{\mathbf{u}} \mathbf{v}) f \neq \mathrm{D}_{\mathbf{u}} (\mathbf{v} f)$.

To see this it in detail we consider the action of a single element of the base $\frac{\partial}{\partial x_i}$ over the vector field $\mathbf{u}$:

$$\begin{aligned}
\frac{\partial}{\partial x_i}(\mathbf{u}) f &= \frac{\partial}{\partial x_i}(u_1, u_2, \dots, u_d) f \\
&= (\frac{\partial u_1}{\partial x_i}, \frac{\partial u_2}{\partial x_i}, \dots, \frac{\partial u_d}{\partial x_i}) f \\
&= (\frac{\partial u_1}{\partial x_i} \frac{\partial}{\partial x_1} + \frac{\partial u_2}{\partial x_i} \frac{\partial}{\partial x_2} + \dots + \frac{\partial u_d}{\partial x_i} \frac{\partial}{\partial x_d}) f \\
&= \frac{\partial u_1}{\partial x_i} \frac{\partial f}{\partial x_1} + \frac{\partial u_2}{\partial x_i} \frac{\partial f}{\partial x_2} + \dots + \frac{\partial u_d}{\partial x_i} \frac{\partial f}{\partial x_d} \ .
\end{aligned}$$

While if $\mathbf{u}$ act as directional derivative over a function f before the derivation:

$$\begin{aligned}\frac{\partial}{\partial x_i}(\mathbf{u}f) &= \frac{\partial}{\partial x_i}\Big(u_1\frac{\partial f}{\partial x_1} + u_2\frac{\partial f}{\partial x_2} + \cdots + u_d\frac{\partial f}{\partial x_d}\Big) \\ &= \frac{\partial u_1}{\partial x_i}\frac{\partial f}{\partial x_1} + u_1\frac{\partial^2 f}{\partial x_1 x_i} + \frac{\partial u_2}{\partial x_i}\frac{\partial f}{\partial x_2} + u_2\frac{\partial^2 f}{\partial x_2 x_i} + \cdots + \frac{\partial u_d}{\partial x_i}\frac{\partial f}{\partial x_d} + u_d\frac{\partial^2 f}{\partial x_d x_i}\end{aligned}$$

So the notation $\frac{\partial}{\partial x_i}$ changes his meaning, according to the order of operation.

Going back to the formula of the chain SVF product, in consequence of what just seen, the two different definitions $\mathrm{D}_{\mathbf{u}}\mathrm{D}_{\mathbf{v}}\mathbf{w} = \mathrm{D}_{\mathbf{u}}\big(\mathrm{D}_{\mathbf{v}}w_x, \mathrm{D}_{\mathbf{v}}w_y\big)$ and $\mathrm{D}_{\mathbf{u}}\mathrm{D}_{\mathbf{v}}\mathbf{w} = \mathrm{D}_{(\mathrm{D}_{\mathbf{u}}v_x, \mathrm{D}_{\mathbf{u}}v_y)}\mathbf{w}$ differs for the order in the application of the operations.

Taylor series of SVF: Using the Occam razor, where the simpler options is the right one (but with no mathematical justification for the moment!), we define the product of two SVF as $\mathrm{D}_{\mathbf{u}}\mathrm{D}_{\mathbf{v}}\mathbf{w} := \mathrm{D}_{(\mathrm{D}_{\mathbf{u}}v_x, \mathrm{D}_{\mathbf{u}}v_y)}\mathbf{w}$. In consequence of this, as noticed before, it follows that $\mathrm{D}^n_{\mathbf{u}}\mathbf{u} = \mathrm{J}^{n-1}_{\mathbf{u}}\mathbf{u}$ and therefore the Taylor series of is expressed by:

$$\exp(t\mathbf{u}) = \sum_{j=0}^{N}\frac{t^j\mathbf{u}^j}{j!} = \mathrm{Id} + \sum_{j=1}^{N}\frac{t^j\,\mathrm{J}^{j-1}_{\mathbf{u}}\mathbf{u}}{j!} \ .$$

Numerical examples showed convergence behaviour in the linear case.

3 Homographies-based SVF

To gain more information on the validity of each method we need to have more examples where the ground truth is given, other than the linear case.

4 Memorandum

1. Computational complexity analysis.
2. Error analysis.
3. Extension to TVVF approximating the time variation with polynomials. Working on where the following equation originates [3]:

$$\frac{\mathrm{d}\phi_t}{\mathrm{d}t} = A\phi_t + ct, \quad \phi_0 = \mathbf{0} \quad \Longrightarrow \quad \phi_t = t^2\varphi_2(tA)c \quad \forall t \in \mathbb{R}\ . \tag{17}$$

   where the separation between the temporal and the spatial transformation in $A\phi_t + ct$, may be given by the fact that the dimension is stored in the fourth component of the SVF. This requires further investigations...!
4. Extension to TVVF through autonomisation of variables (see slides notes).
5. Autonomisation for scaling and squaring. See if there something is making sense.

References

1. Morita, Shigeyuki. Geometry of differential forms. Vol. 201. American Mathematical Soc., 2001.
2. Hochbruck, Marlis, and Alexander Ostermann. "Exponential integrators." Acta Numerica 19 (2010): 209-286.
3. Higham, Nicholas J., and Lin Lijing. "Matrix Functions: A Short Course." (2013).