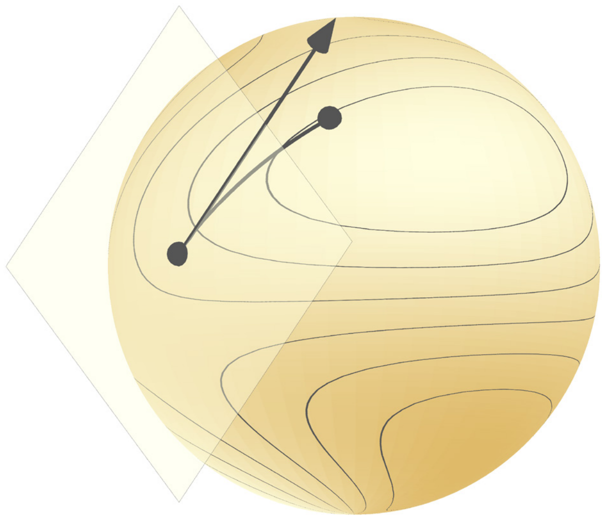
112

Comparing tangent vectors: Parallel transport



Spring 2023

Optimization on manifolds, MATH 512 @ EPFL

Instructor: Nicolas Boumal

Three reasons why it's useful to compare tangent vectors at different points

① Algorithms: RCG: $p_{k+1} = {}$ " $\underbrace{-\operatorname{grad} f(x_{k+1})}_{T_{x_{k+1}}\mathcal{M}} + \beta_k \underbrace{p_k}_{T_{x_k}\mathcal{M}}$ "

(BFGS)

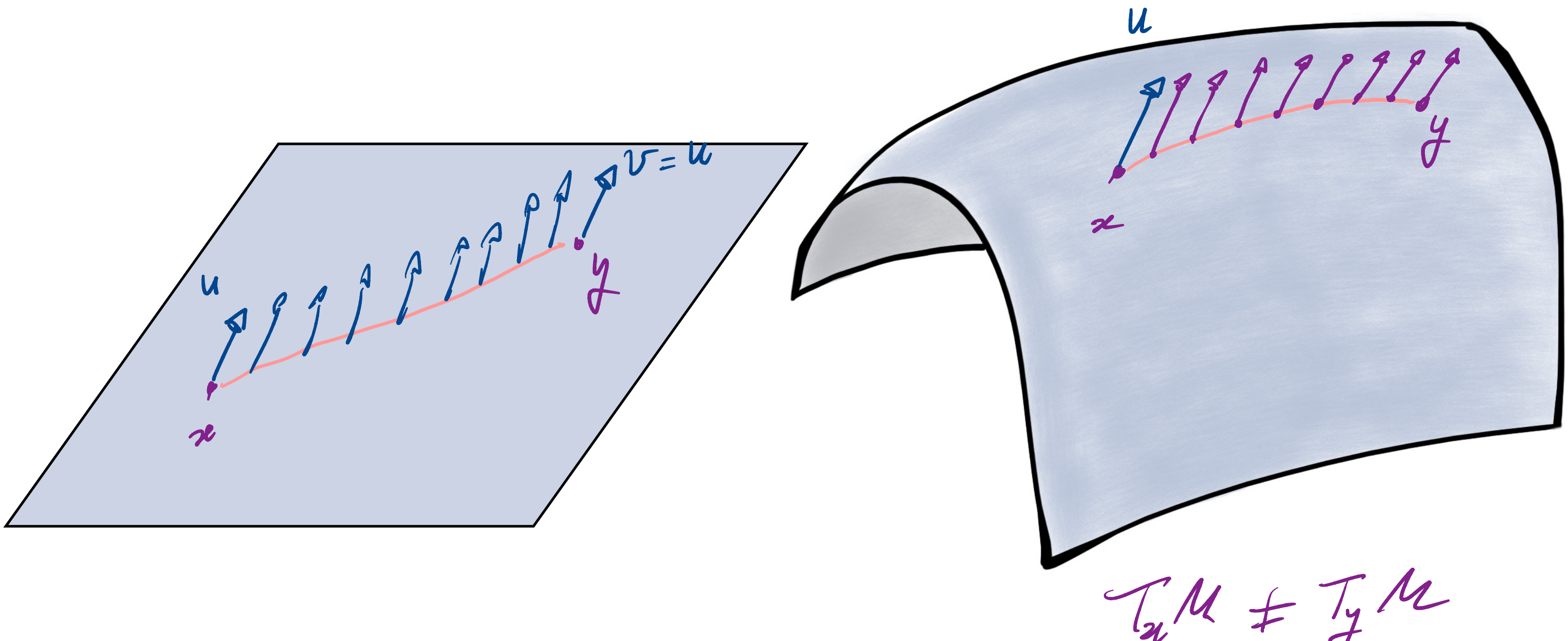
② Finite differences: in $\mathbb{R}^n$, $\operatorname{Hess} f(x)[u] = \lim_{t \to 0} \dfrac{\operatorname{grad} f(x+tu) - \operatorname{grad} f(x)}{t}$

$$= \frac{d}{dt} \operatorname{grad} f(x+tu)\Big|_{t=0}$$

③ Lipschitz continuous gradients:

In $\mathbb{R}^n$: $\left[\|\operatorname{grad} f(x) - \operatorname{grad} f(y)\| \leq L\|x-y\|\right] \Longrightarrow \left|f(x+u) - [f(x) + \langle \operatorname{grad} f(x), u\rangle]\right| \leq \frac{L}{2}\|u\|^2$

Given $u \in \mathrm{T}_x\mathcal{M}$ and $v \in \mathrm{T}_y\mathcal{M}$, what should it mean if we say they are “kind of the same”?



Let c be a smooth curve on $\mathcal{M}$ with covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$.

Def.: A smooth vector field Z along c is parallel if $\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = 0$ for all t.

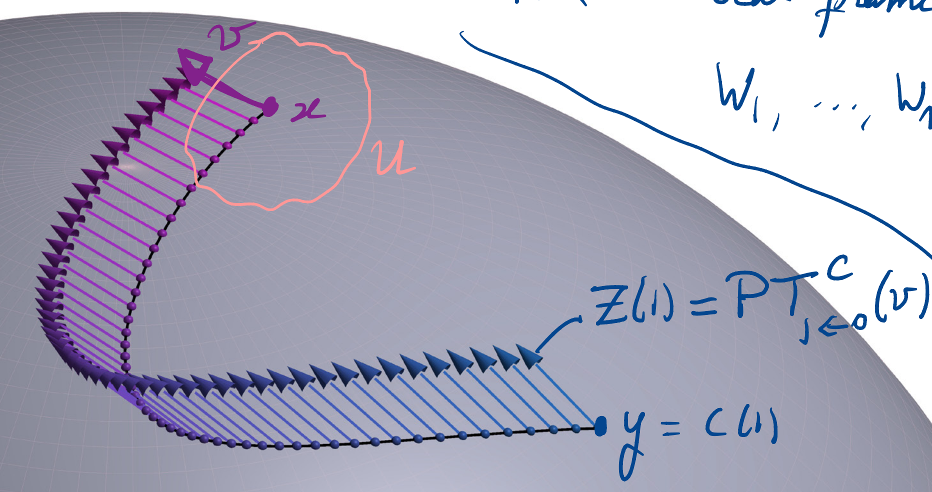
Let $x = c(0)$ and pick any $v \in \mathrm{T}_x\mathcal{M}$.

Fact: There exists a unique $Z \in \mathfrak{X}(c)$ parallel s.t. $Z(0) = v$.

Pick a local frame $W_1, \dots, W_n$ around x:

$W_1, \dots, W_n \in \mathfrak{X}(\mathcal{U})$, $W_1(y) \dots W_n(y)$ form a basis for $T_y\mathcal{M}$.

$$Z(t) = \sum_{i=1}^{n} \underbrace{a_i(t)}_{\text{real, smooth}} W_i(c(t)).$$



① $v = Z(0) = \sum_{i=1}^{n} a_i(0) W_i(x) \Rightarrow$ $a(0) \in \mathbb{R}^n$ is uniquely defined by v.

② $0 = \frac{D}{dt} Z(t) = \frac{D}{dt}\left[\sum_{i=1}^{n} a_i(t) W_i(c(t))\right]$

$$= \sum_{i=1}^{n}\left[a_i'(t) W_i(c(t)) + a_i(t) \underbrace{\frac{D}{dt}(W_i \circ c)(t)}_{= \nabla_{c'(t)} W_i}\right]$$

$$\forall j \quad 0 = \sum_{i=1}^{n}\left[a_i'(t) \langle W_i(c(t)), W_j(c(t))\rangle_{c(t)} + a_i(t) \langle \nabla_{c'(t)} W_i, W_j(c(t))\rangle_{c(t)}\right]$$

Def.: Parallel transport along c from time t_1 to t_2 is the linear map

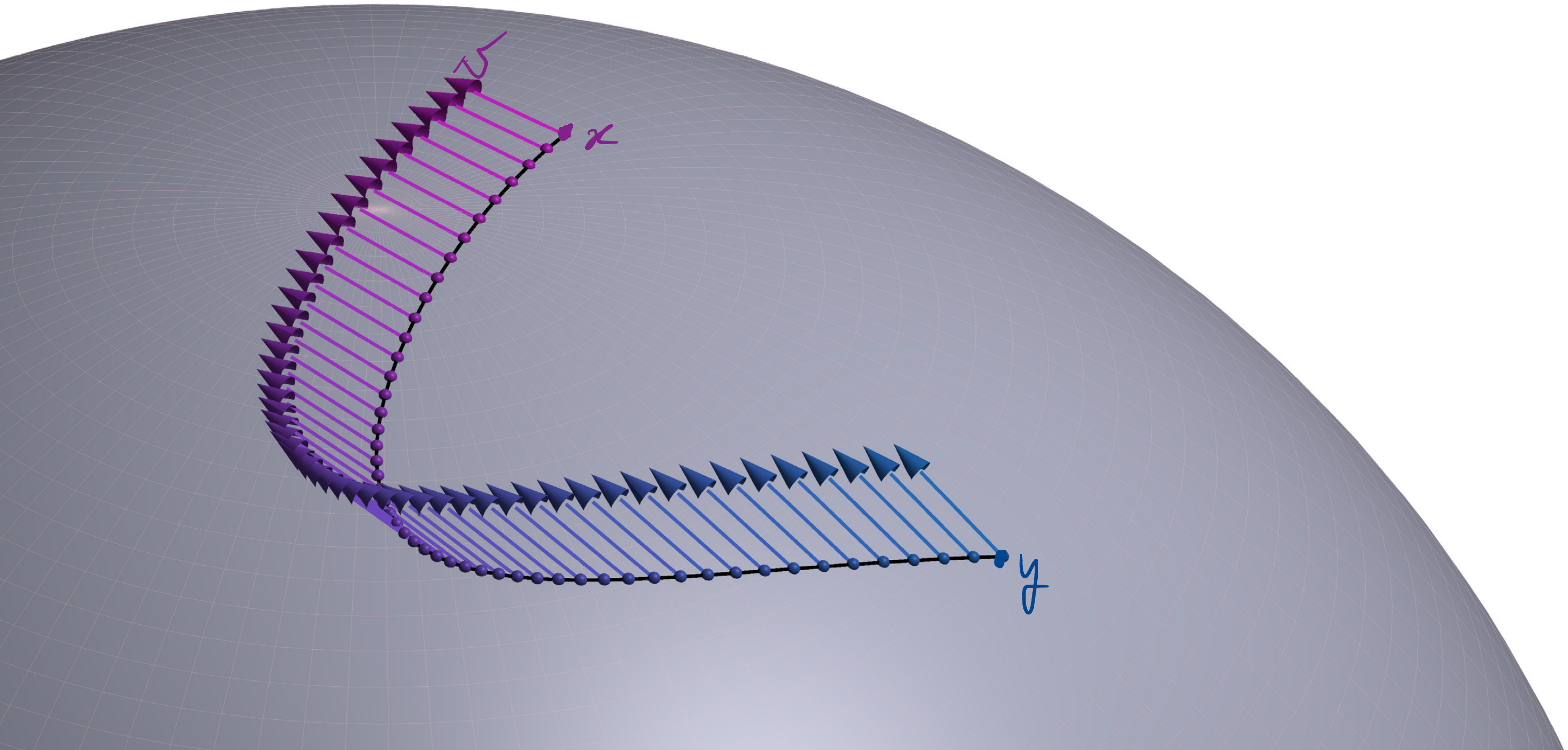
$$\mathrm{PT}^c_{t_2 \leftarrow t_1} : \mathrm{T}_{c(t_1)}\mathcal{M} \to \mathrm{T}_{c(t_2)}\mathcal{M}$$

defined by $\mathrm{PT}^c_{t_2 \leftarrow t_1}(v) = Z(t_2)$,

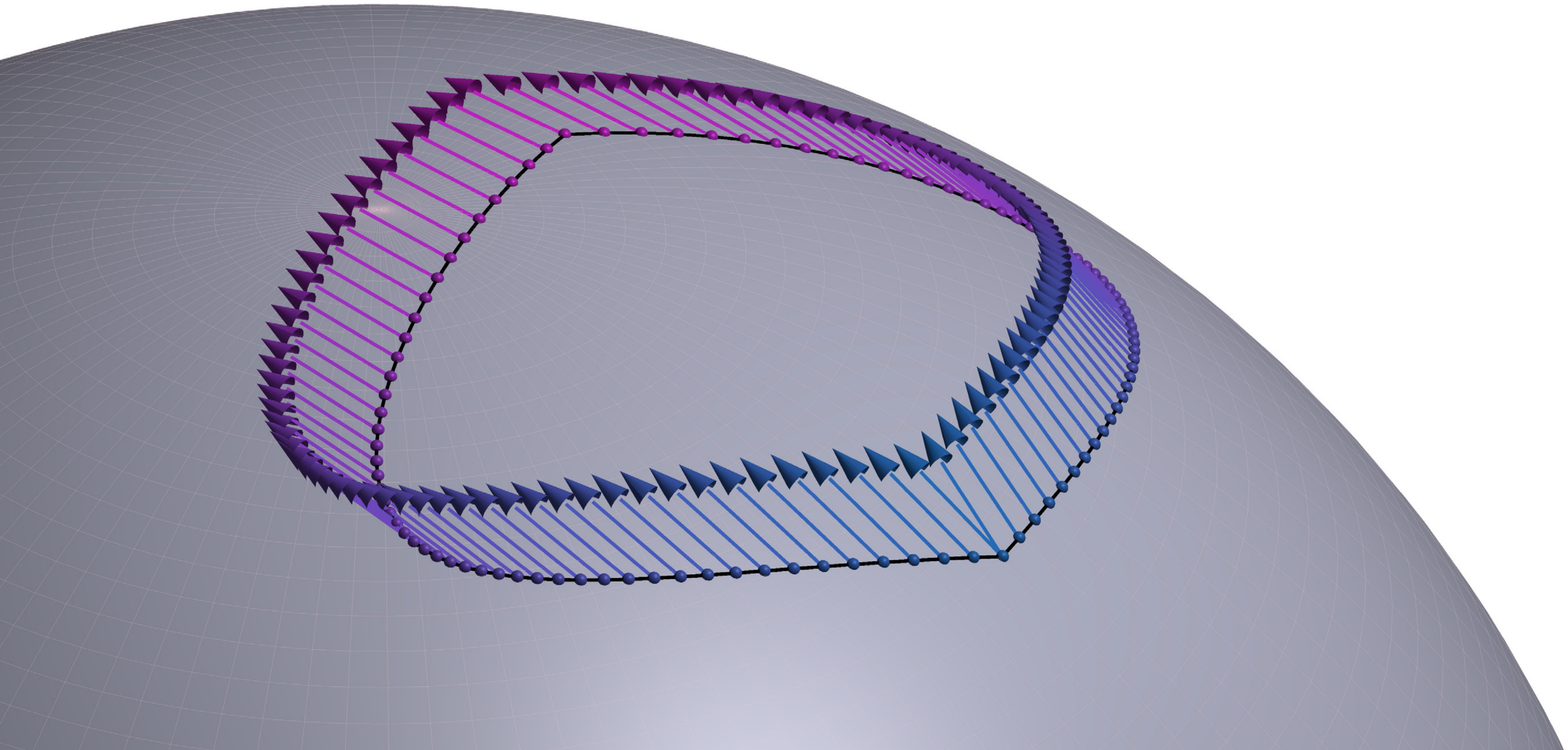
where Z is the parallel field along c such that $Z(t_1) = v$.

$$\left(\mathrm{PT}^c_{t_2 \leftarrow t_1}\right)^{-1} = \mathrm{PT}^c_{t_1 \leftarrow t_2}.$$

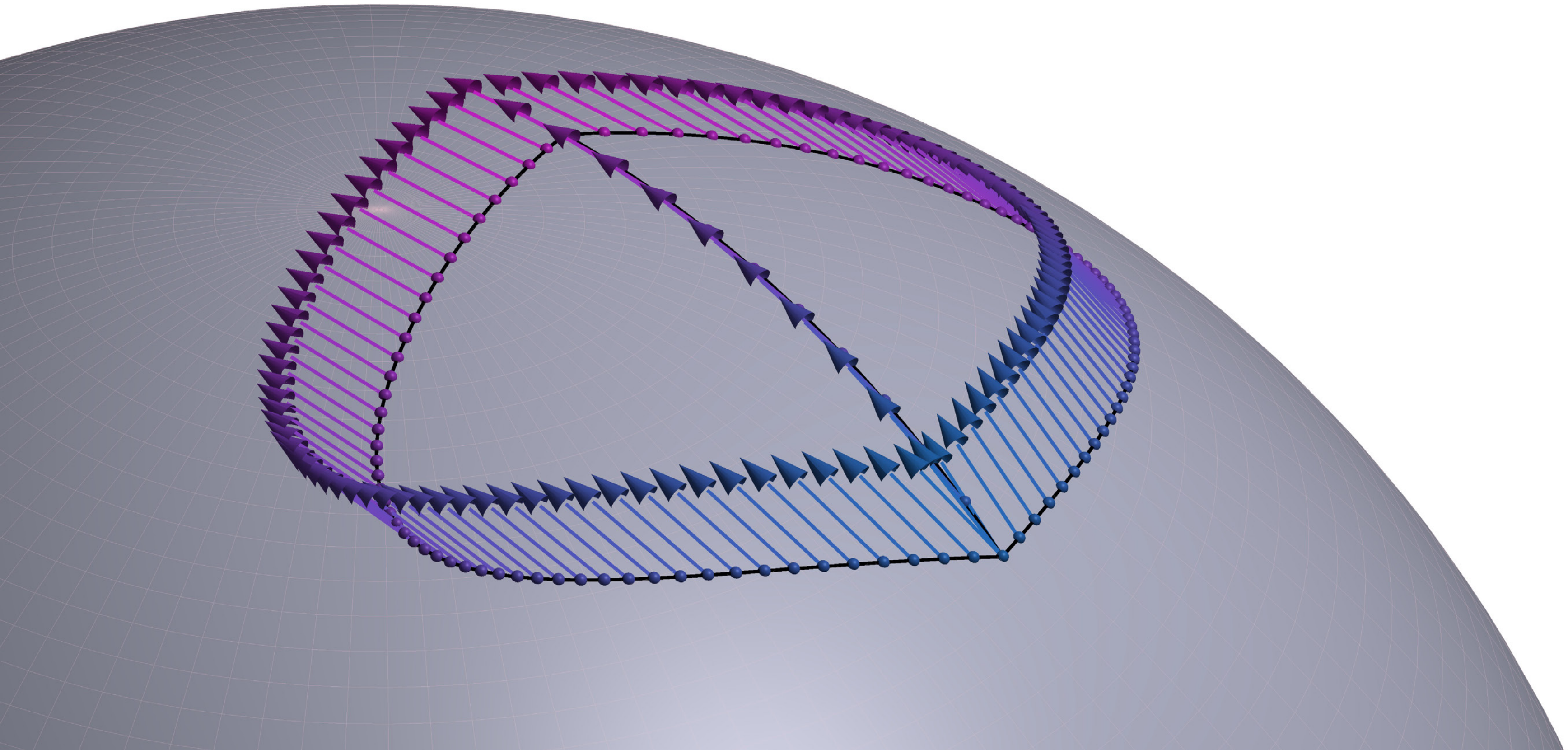
Parallel transport depends on the curve c



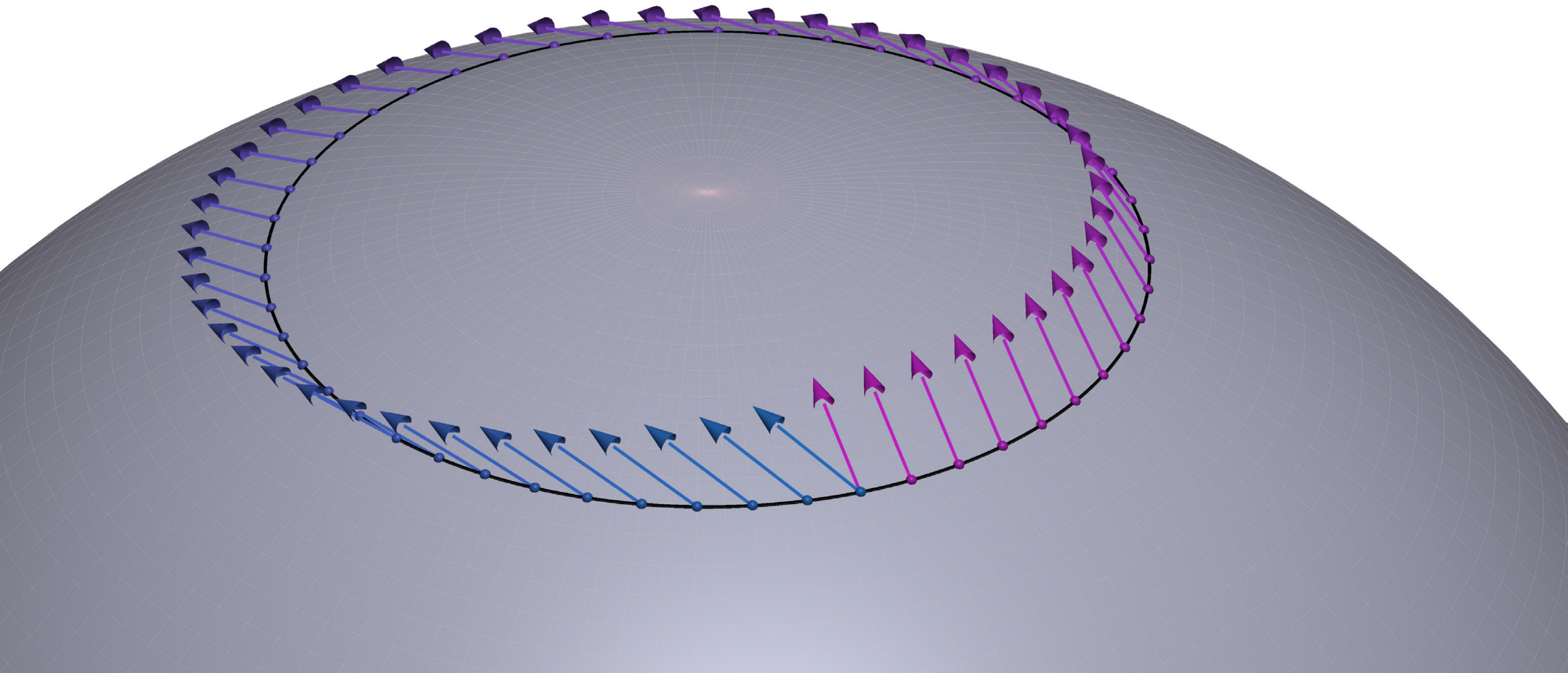
Parallel transport depends on the curve c



Parallel transport depends on the curve c



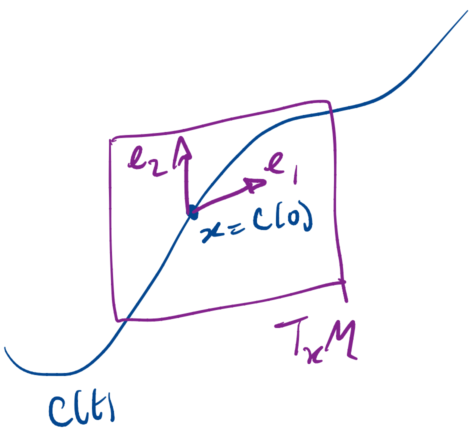
Parallel transport along a loop $\neq$ identity



Use case: re-interpreting $\frac{\mathrm{D}}{\mathrm{d}t}$ as a limit

Pick a curve $c : \mathbb{R} \to M$, let $x = c(0)$.

Pick a smooth vector field $Z \in \mathfrak{X}(c)$. $\frac{D}{dt} Z$?



Choose $e_1, \ldots, e_n$: basis for $T_x M$.

Let $E_i(t) = PT^{c}_{t \leftarrow 0}(e_i)$.

b/c invertible

Notice : $E_1(t), \ldots, E_n(t)$ form a basis for $T_{c(t)} M$.

So : $Z(t) = \sum_{i=1}^{n} a_i(t) E_i(t)$.

① $\frac{D}{dt} Z(t) = \sum_{i=1}^{n} \left[a_i'(t) E_i(t) + a_i(t) \underbrace{\cancel{\frac{D}{dt} E_i(t)}}_{0} \right]$

② $PT_{0 \leftarrow t}^{c} \left(Z(t) \right) = \sum_{i=1}^{n} a_i(t) \, PT_{0 \leftarrow t}^{c} \left(PT_{t \leftarrow 0}^{c} (e_i) \right)$

$= \sum_{i=1}^{n} a_i(t) \, e_i$

① $\frac{D}{dt} Z(0) = \sum_{i=1}^{n} a_i'(0) \, e_i = \frac{d}{dt} \left[PT_{0 \leftarrow t}^{c} \left(Z(t) \right) \right] \Big|_{t=0}$

Fact: For $Z \in \mathcal{X}(c)$:

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(0) = \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{PT}_{0 \leftarrow t}^{c}\big(Z(t)\big)\Big|_{t=0} = \lim_{t \to 0} \frac{\mathrm{PT}_{0 \leftarrow t}^{c}\big(Z(t)\big) - Z(0)}{t}$$

Corollary: $\mathrm{Hess} f(x)[u] = \nabla_u \, \mathrm{grad} f$

$c(0) = x$

$c'(0) = u$

$$= \frac{D}{dt} \left(\mathrm{grad} f \circ c \right)(0)$$

$$= \lim_{t \to 0} \frac{PT^{c}_{0 \leftarrow t}\left(\mathrm{grad} f(c(t)) \right) - \mathrm{grad} f(x)}{t}$$