7. Vector and Covector Fields and their Flows

(7.1) Definition. *Let M be a manifold. A* **vector field** *on M is a rule which assigns to each point in M a tangent vector at this point. In other words, a vector field on M is a mapping $X : M \to TM$ such that $X(p) \in T_pM$ for all $p \in M$. We call X a C^k-vector field if X is a mapping of class C^k (which only makes sense if both M and TM are C^k-manifolds, which is the case if M is a C^{k+1}-manifold).*

(7.2) Definition. *Let M be a manifold. A* **covector field** *or a* **one-form** *on M is a rule which assigns to each point in M a cotangent vector at this point, i.e., a linear form on the tangent space at this point. In other words, a covector field on M is a mapping $\omega : M \to T^\star M$ such that $\omega(p) \in (T_pM)^\star$ for all $p \in M$. We call ω a C^k-vector field if ω is a mapping of class C^k (which, again, only makes sense if M is of class C^{k+1}).*

If we identify a vector field X with its graph $\{(p, X(p)) \mid p \in M\}$, we can symbolically view a vector field on M as a "section" of the tangent bundle. Similarly, a one-form is a "section" of the cotangent bundle $T^\star M$.

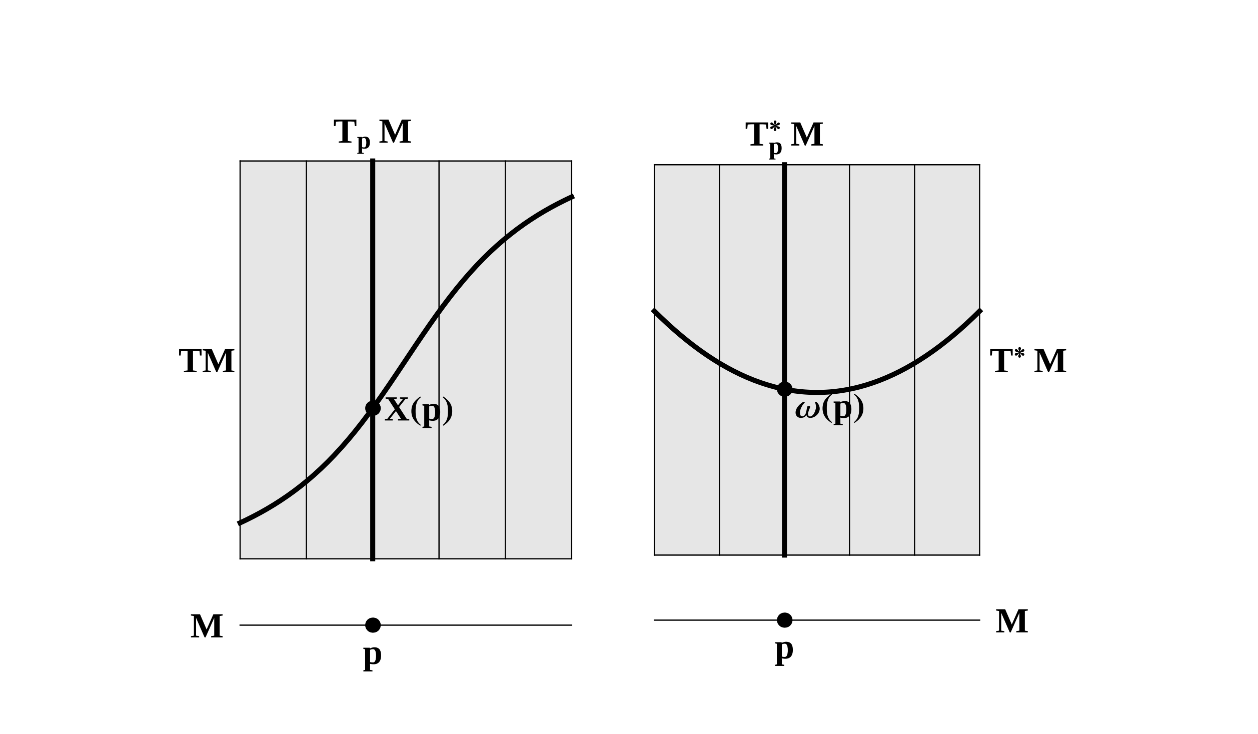

Fig. 7.1: Vector field as section of the tangent bundle; covector field as section of the cotantent bundle.

Intuitively, we visualize a vector field on M as a collection of arrows attached to the points of M, each tangent to M, which we may interpret as a force field on M. For example, we can interpret a surface M in three-dimensional space as a deformed metal sheet on which a magnetic field is generated. If this field is tangent to M (which implies that a charged particle placed on M can only move on M, never away from M, under the influence of the magnetic field) then this determines a vector field on M.

The C^k-vector fields on a C^{k+1}-manifold M form a real vector space; i.e., given two such vector fields X and Y and two real numbers a and b, we can form a new vector field $aX + bY$ via

$$(aX + bY)(p) := a\,X(p) + b\,Y(p) \qquad (p \in M).$$

It is a trivial exercise to check that the vector space axioms hold. Moreover, the C^k-vector fields on M also form a module over the ring $C^k(M)$ of all C^k-functions $f : M \to \mathbb{R}$ if we define

$$(fX)(p) := f(p)X(p) \qquad (p \in M)\,.$$

It is trivial to verify the module axioms $f_2(f_1X) = (f_2f_1)X$, $(f_1 + f_2)X = f_1X + f_2X$ and $f(X_1 + X_2) = fX_1 + fX_2$. To move on to less trivial properties of vector fields, we note that, given a C^k-vector field X on M and a point $p \in M$, the initial value problem

$$\dot{\alpha}(t) = X\big(\alpha(t)\big), \quad \alpha(0) = p$$

possesses a unique maximal solution $t \mapsto \alpha(t;p)$ (integral curve) defined on some open interval $I = I_p \subseteq \mathbb{R}$ containing 0.

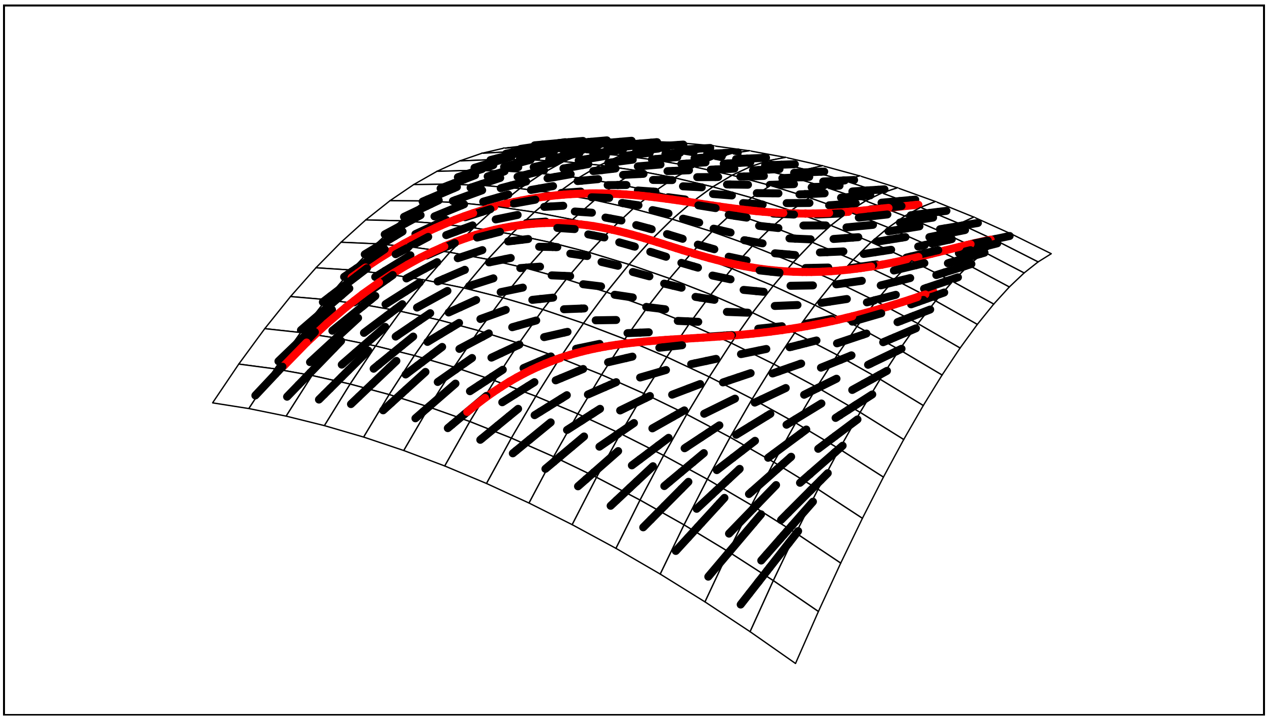

Fig. 7.2: Vector field on a manifold with associated integral curves.

Emphasizing the dependence on p rather than t, we write $\alpha(t;p) = \varphi_t(p)$. If we interpret X as the velocity field of a flowing liquid, we can interpret $\varphi_t(p)$ as the position of the particle in this flow at time t which was at point p at initial time zero. This interpretation immediately implies the semigroup properties

$$\varphi_0 = \mathrm{id}_M, \qquad \varphi_t \circ \varphi_s = \varphi_{t+s}, \qquad \varphi_{-t} = \varphi_t^{-1}.$$

Uniqueness of solutions of initial value problems implies that each φ_t is a bijection, and smooth dependence of solutions of differential equations on initial conditions implies that each φ_t is even a C^k-diffeomorphism if X is of class C^k. We call the family (φ_t) of diffeomorphisms of M the **local flow** associated with the vector field X. We note that, in general, $\varphi_t(p)$ is only defined on some open interval $a_p < t < b_p$ containing 0, not necessarily on all of $\mathbb{R}$. However, each point p possesses a neighborhood U such that $\varphi_t(q)$ is defined on some fixed interval $(-\varepsilon, \varepsilon)$ uniformly for all $q \in U$, and since we will only be interested in local information in the vicinity of a given point we can always replace M by such a neighborhood U (without mentioning this fact each time). We call the vector field X **complete** if all trajectories $t \mapsto \varphi_t(p)$ are defined for all $t \in \mathbb{R}$. To see an example, consider a fixed matrix $A \in \mathbb{R}^{n\times n}$ and the vector field on $\mathbb{R}^n$ given by $x \mapsto Ax$; the associated flow is $\varphi_t = \exp(tA)$, because $x(t) := \exp(tA)p$ is the unique solution of the initial value problem $\dot{x}(t) = Ax(t)$, $x(0) = p$. Next, we show that a vector field can be transported from one manifold to another by a diffeomorphism.

(7.3) Definition. *Let M and N be C^{k+1}-manifolds and let $\Phi : M \to N$ be a C^{k+1}-diffeomorphism. Given a C^k-vector field X on M with local flow (φ_t), the* **push-forward** *of X via Φ is the C^k-vector field $\Phi_\star X$ on N defined by*

$$(\Phi_\star X)(q) \;:=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi\big(\varphi_t(p)\big) \;=\; \Phi'(p)X(p) \quad \textit{where } p := \Phi^{-1}(q)\,.$$

Clearly, if $p = \Phi^{-1}(p)$ then $(\Phi_\star X)(q) = (\mathrm{d}/\mathrm{d}t)|_{t=0}\,(\Phi \circ \varphi_t \circ \Phi^{-1})(q)$ which implies that $\Phi \circ \varphi_t \circ \Phi^{-1}$ is the local flow of $\Phi_\star X$ if φ_t is the local flow of X.

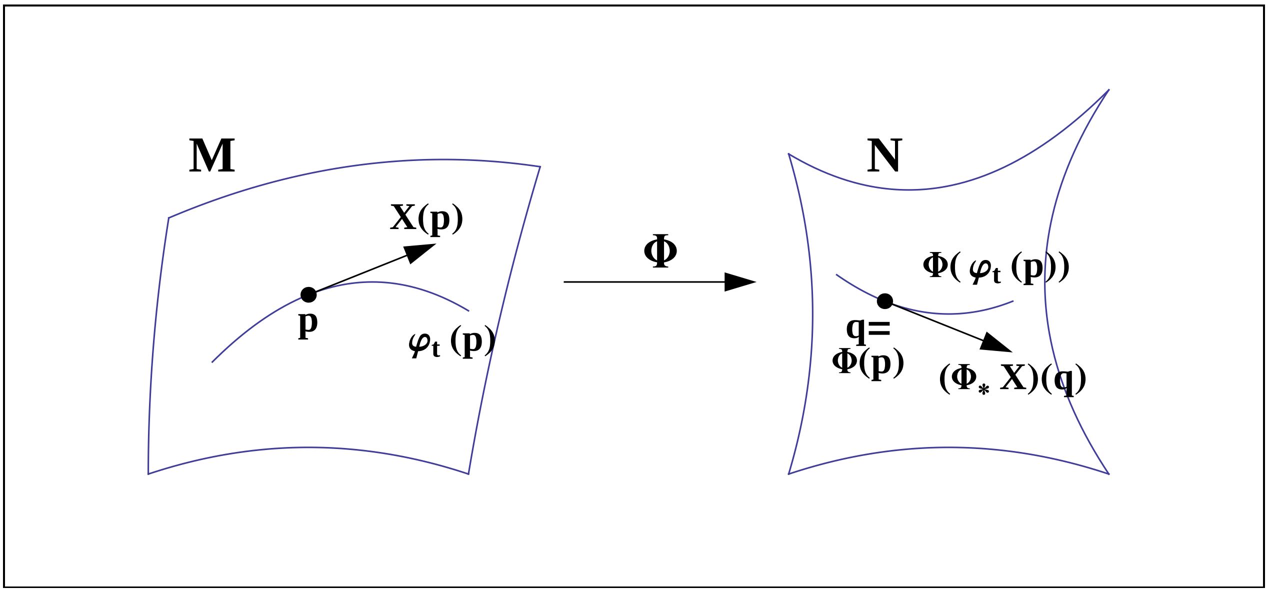

Fig. 7.3: Push-forward of a vector field.

It is now an important observation that a C^k-vector field X can be interpreted as an operator acting on functions on M, more precisely, as a mapping $X : C^k(M) \to C^{k-1}(M)$, by defining

$$(Xf)(p) \;:=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f\big(\varphi_t(p)\big) \;=\; f'(p)X(p)$$

where (φ_t) is the local flow of X. In other words, we identify X with the operator which assigns to a function $f : M \to \mathbb{R}$ its directional derivative in the direction of X (so that $(Xf)(p)$ is the directional derivative of f in the direction $X(p)$). In local coordinates we have

$$(Xf)(p) \;=\; \sum_{i=1}^{d} X_i\big(x(p)\big)\frac{\partial f}{\partial x_i}(p)$$

which we simply write as $X = \sum_{i=1}^d X_i(x)(\partial/\partial x_i)$ and which is fully consistent with the notation

$$X(p) \;=\; \sum_{i=1}^{d} X_i\big(x(p)\big)\;\frac{\partial}{\partial x_i}\Big|_p$$

which we introduced before. This shows that a vector field, if interpreted as an operator acting on functions, is a linear differential operator. Applying this operator to the coordinate functions $f = x_i$ we see that $Xx_i = X_i \circ x$, which shows that a vector field X is uniquely determined by the way it acts on functions. Thus the interpretation of vector fields as differential operators is unambiguous.

(7.4) Proposition. *If X is a vector field on a manifold M and if $\Phi : M \to N$ is a diffeomorphism, then the push-forward $\Phi_\star X$ operates on functions $g : N \to \mathbb{R}$ via*

$$(\Phi_\star X)(g) \;=\; X(g\circ\Phi)\circ\Phi^{-1}$$

.

Proof. For all $q \in N$ (with $p := \Phi^{-1}(q)$) we have

$$\begin{aligned}\big((\Phi_\star X)(g)\big)(q) \;&=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g\big((\Phi\circ\varphi_t\circ\Phi^{-1})(q)\big) \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (g\circ\Phi)\big(\varphi_t(p)\big)\\ &=\; X(g\circ\Phi)(p) \;=\; \big(X(g\circ\Phi)\circ\Phi^{-1}\big)\,(q),\end{aligned}$$

and since q was arbitrary the claim follows. ■

(7.5) Proposition. *Let X be a vector field on M with local flow φ_t. For a given function $f : M \to \mathbb{R}$ and a fixed value of t we define $f_t : M \to \mathbb{R}$ by $f_t(p) := f\big(\varphi_t(p)\big)$.*

(a) *We have $(Xf_t)(p) = (Xf)_t(p) = (\mathrm{d}/\mathrm{d}t)f_t(p)$.*

(b) *If f is of class C^k then*

$$f\circ\varphi_t \;=\; f + t\,Xf + \frac{t^2}{2!}X^2 f + \cdots + \frac{t^k}{k!}\,X^k f \,+\, \mathrm{o}(t^k)\,.$$

Proof. (a) On the one hand we have

$$\begin{aligned}(Xf_t)(p) &= \frac{\partial}{\partial s}\Big|_{s=0} f_t\big(\varphi_s(p)\big) = \frac{\partial}{\partial s}\Big|_{s=0} f\big(\varphi_t\big(\varphi_s(p)\big)\big) \\ &= \frac{\partial}{\partial s}\Big|_{s=0} f\big(\varphi_{t+s}(p)\big) = \frac{\mathrm{d}}{\mathrm{d}t} f\big(\varphi_t(p)\big) = \frac{\mathrm{d}}{\mathrm{d}t} f_t(p).\end{aligned}$$

On the other hand we have

$$\begin{aligned}(Xf)_t(p) &= (Xf)\big(\varphi_t(p)\big) = \frac{\partial}{\partial s}\Big|_{s=0} f\big(\varphi_s\big(\varphi_t(p)\big)\big) \\ &= \frac{\partial}{\partial s}\Big|_{s=0} f\big(\varphi_{s+t}(p)\big) = \frac{\mathrm{d}}{\mathrm{d}t} f\big(\varphi_t(p)\big) = \frac{\mathrm{d}}{\mathrm{d}t} f_t(p).\end{aligned}$$

(b) Write $\alpha(t) := f\big(\varphi_t(p)\big)$. Then part (a) implies first that $\alpha'(t) = (Xf)_t(p)$, then second that $\alpha''(t) = (X^2 f)_t(p)$, and so on; in general, we have $\alpha^{(m)}(t) = (X^m f)_t(p)$ and in particular $\alpha^{(m)}(0) = (X^m f)(p)$. Hence Taylor's theorem (applied to the real-valued function $t \mapsto \alpha(t)$) implies

$$\alpha(t) = \sum_{m=0}^{k} \frac{\alpha^{(m)}(0)}{m!} t^m + \mathrm{o}(t^k) = \sum_{m=0}^{k} (X^m f)(p) \frac{t^m}{m!} + \mathrm{o}(t^k).$$

■

We are now ready to introduce the most important operation on vector fields, namely, the formation of Lie brackets.

(7.6) Definition. *Let X and Y be C^k-vector fields on a manifold M with local flows (φ_t) and (ψ_t), respectively. Then the* **Lie bracket** *of X and Y is defined by*

$$[X, Y] := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\varphi_{-t})_\star Y.$$

Hence $[X, Y]$ can be considered as the infinitesimal version of the push-forward of Y via the local flow defined by X. Explicitly, the definition reads

$$[X, Y](p) = \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} \varphi_{-t}\left(\psi_s\big(\varphi_t(p)\big)\right).$$

The fact that φ_{-t} rather than φ_t is used in the definition of the Lie bracket is merely a convention (to be explained soon). To interpret $[X, Y]$ as an operator on functions, we note that

$$\begin{aligned}
[X,Y]f &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \big((\varphi_{-t})_\star Y\big)f \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} Y(f\circ\varphi_{-t})\circ\varphi_t \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} Y\big(f - tXf + \mathrm{o}(t)\big)\circ\varphi_t \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \underbrace{\big(Yf - tYXf + \mathrm{o}(t)\big)}_{=:\,g}\circ\varphi_t \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g(\varphi_t) \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \big(g + tXg + \mathrm{o}(t)\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \big(Yf - tYXf + tXYf + \mathrm{o}(t)\big) \;=\; (XY-YX)f
\end{aligned}$$

so that $[X,Y] = XY - YX$. (This explains why $-t$ instead of t was used in the definition of the Lie bracket; otherwise we would have obtained $YX - XY$ and not the commutator $XY - YX$ in the operator interpretation of vector fields.) In local coordinates (in which $X = \sum_{i=1}^d X_i(\partial/\partial x_i)$ und $Y = \sum_{j=1}^d Y_j(\partial/\partial x_j)$) we have

$$\begin{aligned}
(XY-YX)f &= \sum_{i=1}^d X_i \frac{\partial}{\partial x_i}\left[\sum_{j=1}^d Y_j \frac{\partial f}{\partial x_j}\right] - \sum_{j=1}^d Y_j \frac{\partial}{\partial x_j}\left[\sum_{i=1}^d X_i \frac{\partial f}{\partial x_i}\right] \\
&= \sum_{i,j=1}^d X_i \left[\frac{\partial Y_j}{\partial x_i}\cdot\frac{\partial f}{\partial x_j} + Y_j \frac{\partial^2 f}{\partial x_i \partial x_j}\right] - \sum_{i,j=1}^d Y_j \left[\frac{\partial X_i}{\partial x_j}\frac{\partial f}{\partial x_i} + X_i \frac{\partial^2 f}{\partial x_j \partial x_i}\right] \\
&= \sum_{i,j=1}^d X_i \frac{\partial Y_j}{\partial x_i}\cdot\frac{\partial f}{\partial x_j} - \sum_{i,j=1}^d Y_j \frac{\partial X_i}{\partial x_j}\frac{\partial f}{\partial x_i} \quad \text{(by Schwarz' Theorem)} \\
&= \sum_{i,j=1}^d X_j \frac{\partial Y_i}{\partial x_j}\cdot\frac{\partial f}{\partial x_i} - \sum_{i,j=1}^d Y_j \frac{\partial X_i}{\partial x_j}\frac{\partial f}{\partial x_i} \\
&= \sum_{i=1}^d \left(\sum_{j=1}^d X_j \frac{\partial Y_i}{\partial x_j} - Y_j \frac{\partial X_i}{\partial x_j}\right)\frac{\partial f}{\partial x_i}
\end{aligned}$$

which implies that $[X,Y]$ is again a vector field on M (the components of which are given by $[X,Y]_i = \sum_{j=1}^d \big(X_j(\partial Y_i/\partial x_j) - Y_j(\partial X_i/\partial x_j)\big)$.

(7.7) Example. If we identify vector fields on an open subset $\Omega \subseteq \mathbb{R}^n$ with mappings $\Omega \to \mathbb{R}^n$, the coordinate expression for the Lie bracket just derived takes the form $[f,g]_i = \sum_{j=1}^n \big(f_j(\partial_j g_i) - g_j(\partial_j f_i)\big)$, which may be written in the form

$$[f,g] = \begin{bmatrix} \partial_1 g_1 & \cdots & \partial_n g_1 \\ \vdots & & \vdots \\ \partial_1 g_n & \cdots & \partial_n g_n \end{bmatrix}\begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} - \begin{bmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_n & \cdots & \partial_n f_n \end{bmatrix}\begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}$$

which means that we have $[f,g](x) = g'(x)f(x) - f'(x)g(x)$ for all $x \in \Omega$ if $[f,g]$ is again to be interpreted as a function $\Omega \to \mathbb{R}^n$.

We next show that Lie brackets are respected by push-forwards.

(7.8) Proposition. *Let X, Y be vector fields on M and let $\Phi : M \to N$ be a diffeomorphism. Then*

$$\Phi_\star[X,Y] \;=\; [\Phi_\star X, \Phi_\star Y]\,.$$

Proof. Given a function $g : N \to \mathbb{R}$, we have $(\Phi_\star Y)(g) = Y(g \circ \Phi) \circ \Phi^{-1} =: \gamma$ and then $(\Phi_\star X)(\gamma) = X(\gamma \circ \Phi) \circ \Phi^{-1} = X\big(Y(g\circ\Phi)\big) \circ \Phi^{-1}$ so that $(\Phi_\star X)(\Phi_\star Y)g = XY(g\circ\Phi)\circ\Phi^{-1}$. Analogously, $(\Phi_\star Y)(\Phi_\star X)g = YX(g\circ\Phi)\circ\Phi^{-1}$. Consequently,

$$\begin{aligned} [\Phi_\star X, \Phi_\star Y]g &= \big((\Phi_\star X)(\Phi_\star Y) - (\Phi_\star Y)(\Phi_\star X)\big)(g) \\ &= (XY - YX)(g\circ\Phi)\circ\Phi^{-1} \\ &= [X,Y](g\circ\Phi)\circ\Phi^{-1} \;=\; \big(\Phi_\star[X,Y]\big)g. \end{aligned}$$

■

(7.9) Proposition. *Let X and Y be vector fields with local flows φ_t and ψ_t. Then the following conditions are equivalent:*
(1) $[X,Y] = 0$;
(2) $t \mapsto (\varphi_{-t})_\star Y$ is constant;
(3) $\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$ for all s, t.

Proof. If (1) holds then

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}(\varphi_{-t})_\star Y \;=\; \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} (\varphi_{-(t+\varepsilon)})_\star Y \;=\; \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} (\varphi_{-t}\circ\varphi_{-\varepsilon})_\star Y \\ &= \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} (\varphi_{-t})_\star(\varphi_{-\varepsilon})_\star Y \;=\; (\varphi_{-t})_\star \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} (\varphi_{-\varepsilon})_\star Y \;=\; (\varphi_{-t})_\star[X,Y] \;=\; 0; \end{aligned}$$

hence (1) implies (2). Assume that (2) holds. Fix t and let $\Phi_s := \varphi_{-t}\circ\psi_s\circ\varphi_t$; then (Φ_s) is a local flow with

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\Phi_s &= \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \varphi_{-t}\circ\psi_{s+\varepsilon}\circ\varphi_t \;=\; \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \varphi_{-t}\circ\psi_\varepsilon\circ\psi_s\circ\varphi_t \\ &= \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \varphi_{-t}\circ\psi_\varepsilon\circ\varphi_t\circ\Phi_s \;=\; \big((\varphi_{-t})_\star Y\big)\circ\Phi_s \;=\; Y(\Phi_s). \end{aligned}$$

This shows that Φ_s satisfies the defining condition of the flow ψ_s; hence $\Phi_s = \psi_s$, which is (3). Hence (2) implies (3). Finally, if (3) holds then we have

$$XYf \;=\; \frac{\partial^2}{\partial s\partial t}\Big|_{s=t=0} f\circ\varphi_t\circ\psi_s \;=\; \frac{\partial^2}{\partial s\partial t}\Big|_{s=t=0} f\circ\psi_s\circ\varphi_t \;=\; YXf$$

for all f and consequently $0 = XY - YX = [X,Y]$. Hence (3) implies (1). ■

(7.10) Proposition. *Let X and Y be vector fields with local flows φ_t and ψ_t and let $p \in M$. Then*

$$[X,Y](p) \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\psi_{-\sqrt{t}} \circ \varphi_{-\sqrt{t}} \circ \psi_{\sqrt{t}} \circ \varphi_{\sqrt{t}})(p).$$

Proof. Given a function $f : M \to \mathbb{R}$ we see, using (7.5)(b), that $f \circ \psi_{-t} \circ \varphi_{-t} \circ \psi_t \circ \varphi_t$ equals

$$\left(\mathbf{1} + tX + \frac{t^2}{2}X^2\right)\left(\mathbf{1} + tY + \frac{t^2}{2}Y^2\right)\left(\mathbf{1} - tX + \frac{t^2}{2}X^2\right)\left(\mathbf{1} - tY + \frac{t^2}{2}Y^2\right) f \;+\; \mathrm{o}(t^3).$$

After multiplying out and sorting terms, this reduces to $\big(\mathbf{1} + t^2(XY - YX)\big)f + \mathrm{o}(t^3)$. Assuming $t > 0$, replacing t by $\sqrt{t}$ and writing $\alpha(t) := (\psi_{-\sqrt{t}} \circ \varphi_{-\sqrt{t}} \circ \psi_{\sqrt{t}} \circ \varphi_{\sqrt{t}})(p)$, this yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f\big(\alpha(t)\big) \;=\; ([X,Y]f)(p),$$

i.e., $f'(p)\alpha'(0) = f'(p)\big([X,Y](p)\big)$ if we interpret $[X,Y]$ not as an operator, but as an assignment of tangent vectors. Since f was arbitrarily chosen, this yields $\alpha'(0) = [X,Y](p)$, and we are done. ∎

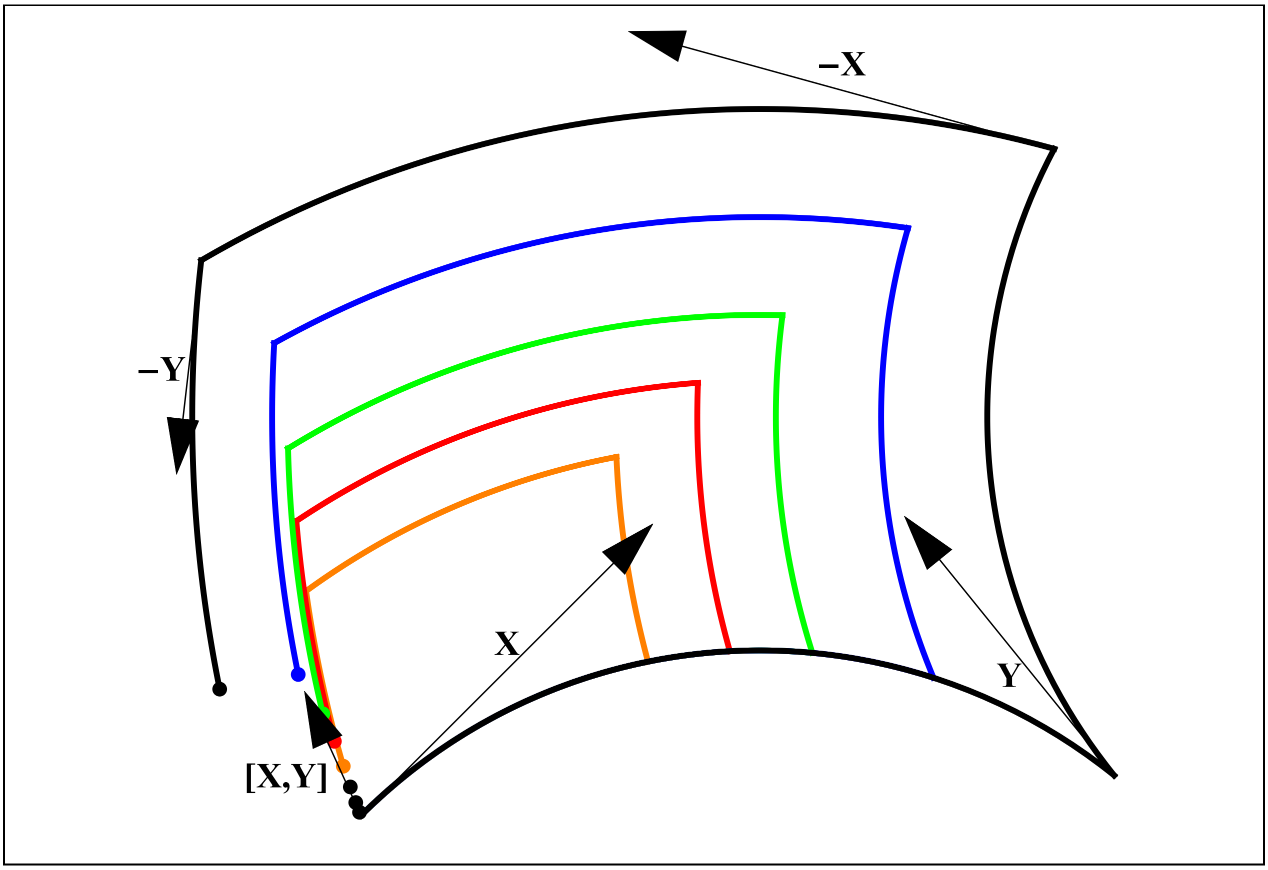

Fig. 7.4: Geometric interpretation of the Lie bracket of two vector fields.

This last result is of considerable importance in control theory. Namely, if X and Y are vector fields on the state space of a system and hence represent possible directions into which the system can be steered, this result shows that the system can also be steered in the direction $[X,Y]$ (even though this may not be intuitively clear).

(7.12) Example. Let us consider a simplified model of a car with front-wheel drive. Let A and B be the centers of the rear and the front axle, respectively, and let $\ell = \overline{AB}$ be the length of the tie rod connecting the two axles. Then the configuration of the car can be specified by four numbers: the coordinates (x,y) of the point A, the angle φ (counted from a fixed "horizontal" direction) in which the car is headed, and the steering angle θ; in other words, the configuration space M of the car is a four-dimensional manifold. (We may assume $M = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ or, more realistically, $M = \mathbb{R}^2 \times \mathbb{S}^1 \times (-\theta_{\max}, \theta_{\max})$ with a maximal steering angle $\theta_{\max}$.) Assume that the steering wheel is in a fixed position θ. During an infinitesimally small time interval $[t, t+\mathrm{d}t]$, the car moves in the direction in which the rear wheels are pointing, so that

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = s \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$$

where s is the speed of the car at time t.

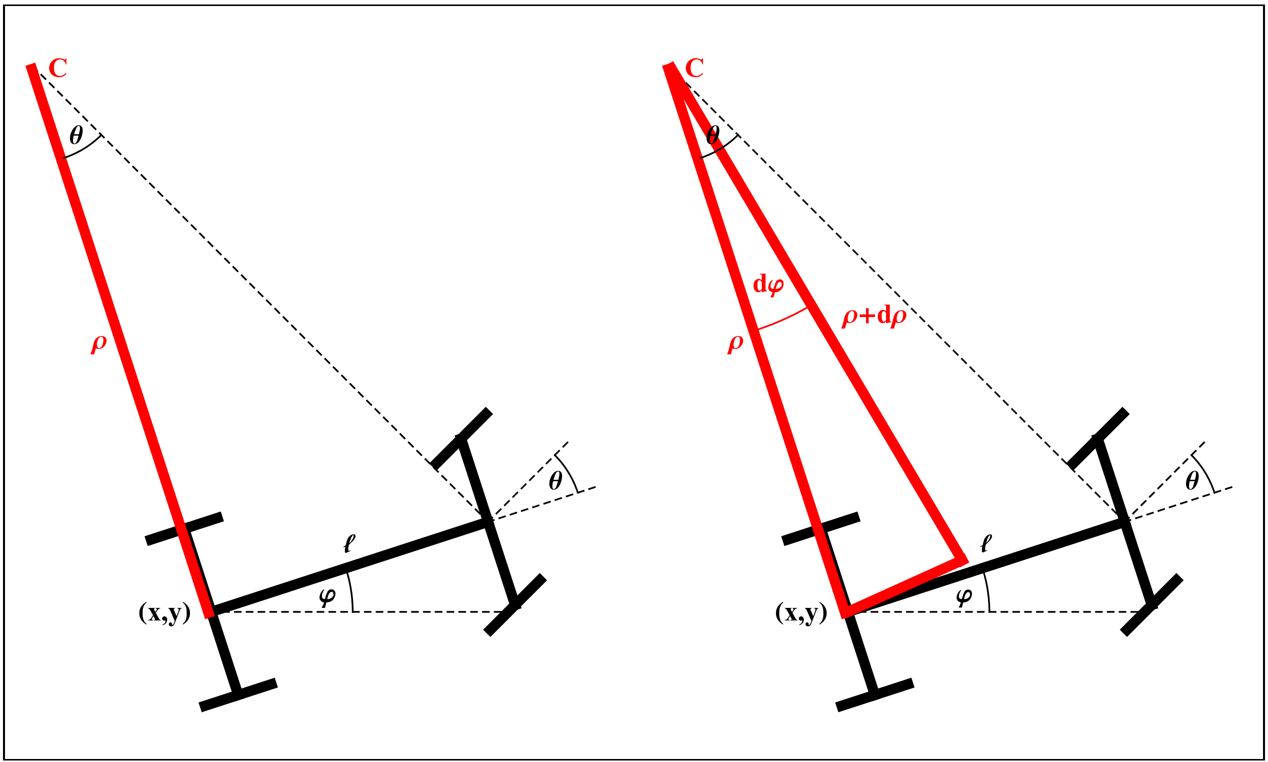

Fig. 7.5: Visualization of the instantaneous motion of a car with front-wheel drive.

The instantaneous motion of the car is a rotation about the point C shown in the diagram; the turning radius ρ satisfies $\tan(\theta) = \ell/\rho$. During the time interval $[t, t+\mathrm{d}t]$, the car moves by the amount $s\,\mathrm{d}t = \rho\,\mathrm{d}\varphi$ so that $\dot{\varphi} = s/\rho = s\tan(\theta)/\ell$. Altogether, we find that

$$\dot{x} = s\cos(\varphi), \qquad \dot{y} = s\sin(\varphi), \qquad \dot{\varphi} = \frac{s\tan(\theta)}{\ell}, \qquad \dot{\theta} = 0.$$

Let $\pm\omega_{\max} := \pm s\tan(\theta_{\max})/\ell$ be the angular velocities associated with the extremal possible steering angles. To discuss the manœuvrability of the car, it is enough to study these extremal steering angles and hence the vector fields

$$f(x,y,\varphi) = \begin{bmatrix} s\cos\varphi \\ s\sin\varphi \\ \omega_{\max} \end{bmatrix} \qquad \text{and} \qquad g(x,y,\varphi) = \begin{bmatrix} s\cos\varphi \\ s\sin\varphi \\ -\omega_{\max} \end{bmatrix}.$$

A common strategy (not very good for the tires!) to get out of a tight parking space is as follows: Move forwards with the steering wheel in its leftmost position, then turn the steering wheel to its rightmost position and move backwards. This means following first the vector field f, then the vector field $-g$, hence approximately following the vector field

$$f + (-g) \;=\; \begin{bmatrix} 0 \\ 0 \\ 2\omega_{\max} \end{bmatrix},$$

which describes a pure rotation about the midpoint of the rear axle. (See the left-hand side of Figure 7.6.)

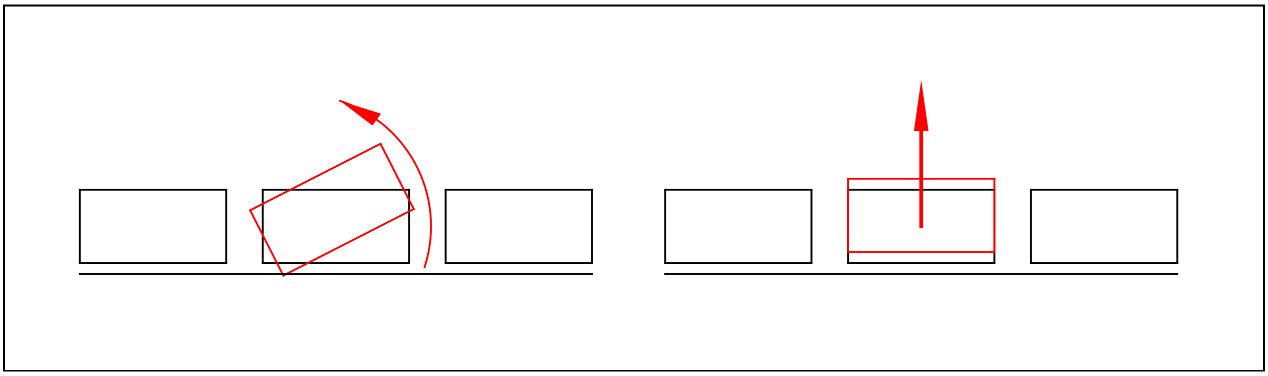

Fig. 7.6: Two different strategies to get out of a parking space.

This works as long as there is enough space to perform the rotation, i.e., as long as the parking space is not too tight. If it is too tight, then a different strategy works, namely that of following first f, then g, then $-f$, then $-g$, which according to (7.10) means to steer the system roughly into the direction of the Lie bracket $[f,g]$. This Lie bracket is given by $[f,g](\xi) = g'(\xi)f(\xi) - f'(\xi)g(\xi)$, i.e.,

$$[f,g](x,y,\varphi) \;=\; \begin{bmatrix} 0 & 0 & -s\sin\varphi \\ 0 & 0 & s\cos\varphi \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s\cos\varphi \\ s\sin\varphi \\ \omega_{\max} \end{bmatrix} - \begin{bmatrix} 0 & 0 & -s\sin\varphi \\ 0 & 0 & s\cos\varphi \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s\cos\varphi \\ s\sin\varphi \\ -\omega_{\max} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & -s\sin\varphi \\ 0 & 0 & s\cos\varphi \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2\omega_{\max} \end{bmatrix} \;=\; 2s\omega_{\max} \begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix},$$

which points in the direction perpendicular to the current orientation of the car. Hence following first f, then g, then $-f$ and then $-g$ for a short time has the net effect of moving the car in a direction close to that of $[f,g]$ and hence in a direction almost perpendicular to the car's current orientation. This shows that a car parked in a tight spot along the edge of a street can be moved perpendicularly away from the kerb, which will clearly get us out of any parking space, no matter how tight it is. (See the right-hand side of Figure 7.6.)

(7.13) Example. Similarly, let us consider a simplified model of a car with a rigid rear axle and a movable front axle. Again, let A and B be the centers of the rear and the front axle, respectively, and let $\ell = \overline{AB}$ be the length of the tie rod connecting the two axles. Let us use as configuration parameters the coordinates (x,y) of the point B (not A as before!), the angle φ (counted from a fixed "horizontal" direction) in which the car is headed, and the steering angle θ. To operate the car, one has two basic options available: one can steer the car (turn the steering wheel without using the accelerator), and one can drive (move ahead without steering). These two actions specify two vector fields on M which indicate two possible changes in the car's current configuration. We want to express these two vector fields in the local coordinates (x,y,φ,θ). For steering, we simply have the vector field

$$S \;=\; \frac{\partial}{\partial\theta} \quad (\text{"steer"})\,.$$

To obtain the analogous expression for D, we assume we drive the car by an infinitesimal amount δ (in the direction in which the car is currently headed) and ask for the infinitesimal changes $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}\varphi$, $\mathrm{d}\theta$ this has on the car's coordinates. We denote by A and B the axle midpoints before and by $A^\star$ and $B^\star$ the axle midpoints after the infinitesimal motion; hence if $B=(x,y)$ then $B^\star = (x+\mathrm{d}x, y+\mathrm{d}y)$. As is clear from the following diagram, we have $\mathrm{d}\theta = 0$, $\mathrm{d}x = \delta\cos(\varphi+\theta)$ and $\mathrm{d}y = \delta\sin(\varphi+\theta)$. To determine $\mathrm{d}\varphi$, we introduce the orthogonal projection P of the point $B^\star$ onto the line AB. In the right triangle $BPB^\star$ we see that $BP^\star = BB^\star\sin(\theta) = \delta\sin(\theta)$. Consequently, in the right triangle $A^\star PB^\star$ we have $A^\star B^\star = AB = \ell$ (since the tie rod is rigid and cannot change its length) and hence $A^\star P = \sqrt{\ell^2-\delta^2\sin^2\theta} = \ell\sqrt{1-\delta^2\sin(\theta)^2/\ell^2} \approx \ell\big(1-\delta^2\sin(\theta)^2/(2\ell^2)\big) \approx \ell$; hence in first-order approximation, the triangle $PA^\star B^\star$ is equilateral. Consequently, $\mathrm{d}\varphi = \angle(PA^\star B^\star) \approx \delta\sin(\theta)/\ell$. Thus

$$\mathrm{d}x = \delta\cos(\varphi+\theta), \quad \mathrm{d}y = \delta\sin(\varphi+\theta), \quad \mathrm{d}\varphi = \delta\frac{\sin(\theta)}{\ell}, \quad \mathrm{d}\theta = 0\,.$$

If we choose units such that $\ell = 1$ this yields

$$D \;=\; \cos(\varphi+\theta)\frac{\partial}{\partial x} + \sin(\varphi+\theta)\frac{\partial}{\partial y} + \sin(\theta)\frac{\partial}{\partial\varphi} \quad (\text{"drive"})\,.$$

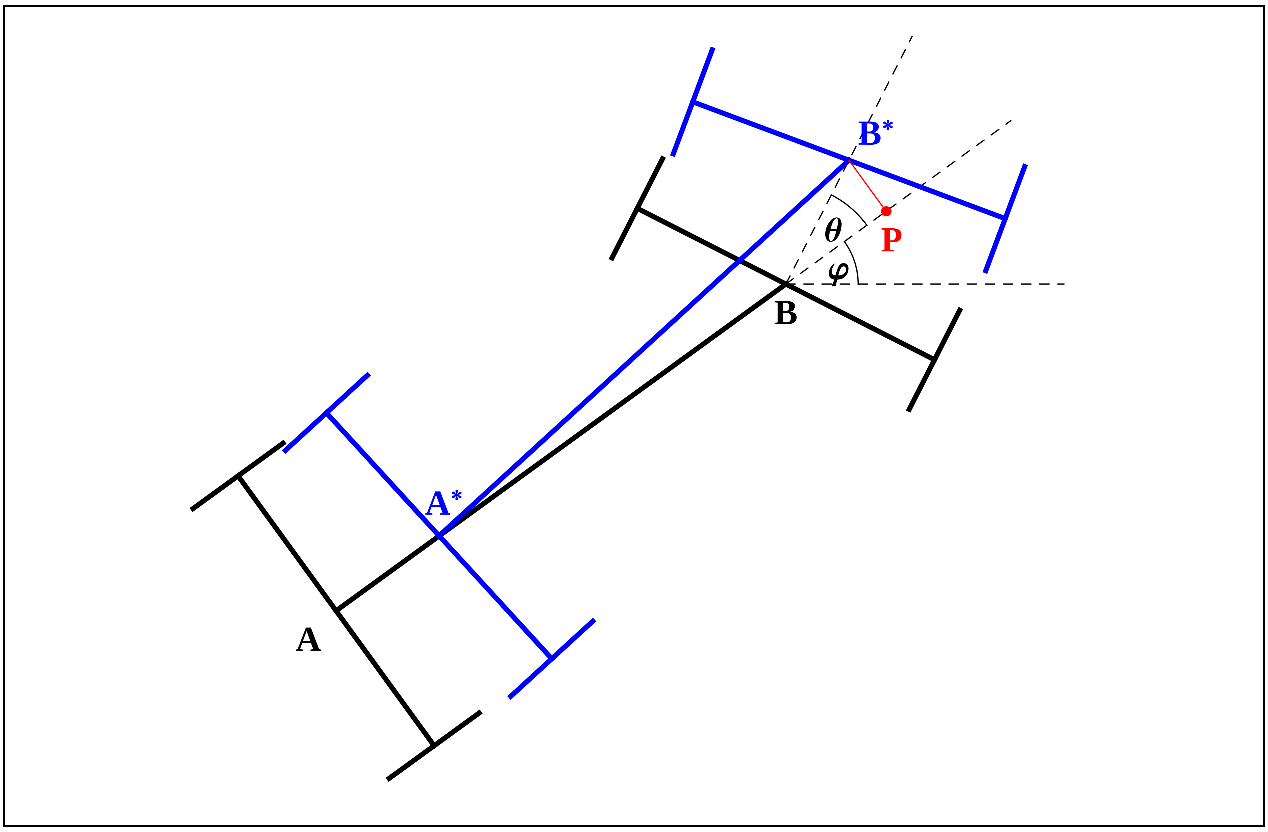

Fig. 7.7: Visualization of the vector field D ("drive").

We introduce the following vector fields:

$$
\begin{aligned}
S &= \frac{\partial}{\partial\theta} \quad (\text{"steer"}),\\
D &= \cos(\varphi+\theta)\frac{\partial}{\partial x}+\sin(\varphi+\theta)\frac{\partial}{\partial y}+\sin(\theta)\frac{\partial}{\partial\varphi} \quad (\text{"drive"}),\\
W &= -\sin(\varphi+\theta)\frac{\partial}{\partial x}+\cos(\varphi+\theta)\frac{\partial}{\partial y}+\cos(\theta)\frac{\partial}{\partial\varphi} \quad (\text{"wriggle"}),\\
G &= -\sin(\varphi)\frac{\partial}{\partial x}+\cos(\varphi)\frac{\partial}{\partial y} \quad (\text{"glide"}),\\
R &= \frac{\partial}{\partial\varphi} \quad (\text{"rotate"}).
\end{aligned}
$$

The following commutator relations are readily verified:

$$[S,D]=W, \quad [S,W]=-D, \quad [W,D]=G, \quad [G,S] = [G,D] = [G,W] = 0.$$

The fact that G is in the Lie algebra generated by S and F and that $W = G + R$ at configurations at which $\theta = 0$ implies that the car can be maneuvered between any two given configurations using the only practically available controls S and D.