**Exercise 16.7.** Using the alternative formulation for bimodular tensor fields $\mathscr{T}^2(M)$, define what is meant by a **two tensor at a point** determined by $B \in \mathscr{T}^2(M)$.

## 16.4 Calculus for manifolds

The starting point for differentiation is with a real valued function with domain an open subset of Euclidean space $\mathbb{R}^n$. Let me further suggest starting with the case $n = 1$, so that the derivative of of a function $f : I \to \mathbb{R}$ where $I = (a, b)$ is an open interval in $\mathbb{R}$ and $a < x < b$ is given by

$$f'(x) = \frac{df}{dx}(x) = \lim_{v \to 0} \frac{f(x+v) - f(x)}{v}.$$

This, it may be recalled, may be interpreted as the limit of the average rate of change of the value of $f$ over the interval from $(x, x+v)$, that is to say the instantaneous rate of change of the value of $f$ with respect to displacement in the domain. It may be pointed out also that the value of the derivative is associated with the affine approximation of the value of $f$ and a certain linear function $df_x : \mathbb{R} \to \mathbb{R}$ through the approximation formula

$$f(x+v) \approx f(x) + f'(x)v = f(x) + df_x(v). \tag{16.6}$$

If $f$ is differentiable at each point $x \in (a, b)$ and the derivative is continuous, that is $f' \in C^0(a, b)$, then we say $f$ is **continuously differentiable** and write $f \in C^1(a, b)$.

A variety of different phenomena come into play (and come to light) when the dimension $n$ satisfies $n \geq 2$. The starting point in this context is usually with **partial derivatives**

$$D_j f(\mathbf{x}) = \frac{\partial f}{\partial x_j}(\mathbf{x}) = \lim_{v \to 0} \frac{f(\mathbf{x} + v\mathbf{e}_j) - f(\mathbf{x})}{v} \tag{16.7}$$

for $j = 1, 2, \dots, n$ which are based on, and available because of, the preferred basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ for both the vector space $\mathbb{R}^n$ and the vectors space $T_{\mathbf{x}}\mathbb{R}^n$. Here $\mathbf{x}$ is a point in some open set $U \subset \mathbb{R}^n$ providing the domain for $f : U \to \mathbb{R}$.

If all the first partial derivatives exist at a point $\mathbf{x} \in U$, then the vector of first partial derivatives is denoted by $Df(\mathbf{x})$ and is called the **gradient** of

$f$ at $\mathbf{x}$. That is, the gradient is given by

$$Df(\mathbf{x}) = \left( \frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right) \in T_{\mathbf{x}}\mathbb{R}^n.$$

Again, the value of (16.7) represents the instantaneous rate of change at $\mathbf{x}$ of the value of $f$ with respect to spatial displacement in the direction of $\mathbf{e}_j \in T_{\mathbf{x}}U = T_{\mathbf{x}}\mathbb{R}^n$. A function $f : U \to \mathbb{R}$ is said to be $C^1(U)$ if each of the partial derivatives exists at each point $\mathbf{x} \in U$, and the resulting functions

$$D_j f : U \to \mathbb{R}$$

are continuous.

This construction may be generalized to unit vectors $\mathbf{u} \in T_{\mathbf{x}}U$ by

$$D_{\mathbf{u}} f(\mathbf{x}) = \lim_{v \to 0} \frac{f(\mathbf{x} + v\mathbf{u}) - f(\mathbf{x})}{v}. \tag{16.8}$$

Again, this limit (when it exists) gives the rate of change of the value of $f$ with respect to spatial displacement in the direction $\mathbf{u}$. These derivatives are called **directional derivatives** and the partial derivatives constitute a special case and/or give examples. Existence of the partial derivatives (16.7) does not imply the existence of other directional derivatives in general.

The existence (at all points) and continuity however, that is $f \in C^1(U)$, does imply the existence of all directional derivatives at all points. The condition $f \in C^1(U)$ also implies a local approximation formula

$$f(\mathbf{x} + \mathbf{v}) \approx f(\mathbf{x}) + df_{\mathbf{x}}(\mathbf{v}) \tag{16.9}$$

where $df_{\mathbf{x}} : T_{\mathbf{x}}U \to \mathbb{R}$ is a linear function called the **differential** at $\mathbf{x}$. The approximation holds in the sense that

$$\lim_{\mathbf{v} \to \mathbf{0}} \frac{f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) - df_{\mathbf{x}}(\mathbf{v})}{|\mathbf{v}|} = 0. \tag{16.10}$$

Notice the scalar $v$ appearing in the limit (16.8) does not have anything necessarily to do with the vector $\mathbf{v}$ appearing in the limit (16.10).

**Exercise 16.8.** By setting $v = |\mathbf{v}|$ where $\mathbf{v}$ is a nonzero vector in the limit (16.10) reexpress the approximation formula (16.10) in terms of the directional derivative of $f$ according to (16.8) for an appropriate unit vector $\mathbf{u}$.

More generally, a function $f : U \to \mathbb{R}$ is said to be **differentiable** at $\mathbf{x} \in U$ if (16.10) holds for some linear function $df_{\mathbf{x}} : T_{\mathbf{x}}U \to \mathbb{R}$. In this case, the linear function $df_{\mathbf{x}}$ is determined uniquely with

$$df_{\mathbf{x}}(\mathbf{v}) = Du(\mathbf{x}) \cdot \mathbf{v} = \langle Du(\mathbf{x}), \mathbf{v} \rangle_{T_{\mathbf{x}}U}. \tag{16.11}$$

Let us pause to consider carefully the value of the differential. This value is sometimes referred to as the directional derivative in the direction $\mathbf{v}$ even when $\mathbf{v}$ is not a unit vector, and the notation

$$D_{\mathbf{v}}f(\mathbf{x}) = df_{\mathbf{x}}(\mathbf{v}) = Du(\mathbf{x}) \cdot \mathbf{v} \tag{16.12}$$

is introduced as a generalization of (16.8) even when $\mathbf{v}$ is not a unit vector. The meaning in this case however is not that of a spatial directional derivative, but rather something somewhat more complicated. The idea is that motion along the direction determined by the vector $\mathbf{v}$ is involved with some (typically) non-unit velocity $|\mathbf{v}|$. This motion is fundamentally externally driven; in particular such a motion introduces an element fundamentally independent of the geometry of the ambient space $\mathbb{R}^n$. Perhaps stated more properly, the motion envisioned is only partially constrained by the geometry of the ambient $\mathbb{R}^n$. This kind of motion can be modeled by some path $\alpha : I \to U$ with $\alpha'(t_0) = \mathbf{v}$ for some $t_0$ in the open interval $I$ with $\alpha(t_0) = \mathbf{x}$. For example,

$$\alpha(t) = \mathbf{x} + t\mathbf{v}$$

satisfies these conditions with $I = (-\epsilon, \epsilon)$ for some $\epsilon > 0$ and $t_0 = 0$. Then $df_{\mathbf{x}}(\mathbf{v})$ is not the rate of change of $f$ with respect to spatial displacement, but relative to both the direction and the velocity of motion determined by $\mathbf{v}$. In particular,

$$df_{\mathbf{x}}(\mathbf{v}) = \frac{d}{dt} f \circ \alpha(t)\Big|_{t=t_0}. \tag{16.13}$$

This is the rate of change of the value of $f$ with respect to displacement in the parameter $t$. This interpretation relies on being able to make sense of the velocity determined by $\mathbf{v}$.

In particular, the formula (16.13) may be adapted to a manifold relatively easily. The vector $\mathbf{v}$ is replaced by a filament (equivalence class) $[\alpha]$ where the path $\alpha$ is taken to have values in the manifold $M$ rather than in some set $U \subset \mathbb{R}^n$. Then for $f : M \to \mathbb{R}$, one can consider $df_P : \mathcal{L}_P M \to \mathbb{R}$ by

$$df_P([\alpha]) = \frac{d}{dt} f \circ \alpha_0(t)\Big|_{t=t_0}. \tag{16.14}$$

This will work fine and define a linear function as long as $f \in cC^1(M)$ and $\alpha_0 \in c\mathfrak{P}^1(M) \cap [\alpha]$ (at $P$) with $\alpha_0(t_0) = P$. Without a Riemannian metric tensor, however, the interpretation of $\alpha$ representing geometric motion along a path in $M$ at a certain velocity $\alpha'(t_0)$ is not available. In particular, the differential function as defined in (16.14) does not have an interpretation as a derivative of the function $f$ without the ability to measure the distance of travel in $M$ determined along the path $\alpha$ as $t$ changes.

Starting back, however, with the notion of a spatial (intrinsic) directional derivative as in (16.7) or (16.8) we can consider the notion of a derivative of a function $f \in cC^1(M)$ as long as we have a Riemannian metric (tensor). There are (at least) two ways to do this.

We can not start directly with (16.7) because there are no preferred directions and/or no preferred basis in $\mathcal{L}_P M$, which we can also denote by $T_P M$ if we make $\mathcal{L}_P M$ into a vector space by the introduction of the inner product $\mu_P$. There is also nominally a fundamental problem with (16.8) as the vector space structure of $\mathbb{R}^n$ is used in the expression $\mathbf{x} + v\mathbf{u}$ in which a vector displacement $v\mathbf{u} \in T_{\mathbf{x}}U$ is added to a point $\mathbf{x}$ in $U$. This kind of structure is not available for $M$ and we cannot add any displacement directly to $P \in M$. On the other hand, it is natural to reexpress $\mathbf{x} + v\mathbf{u}$ or $\mathbf{x} + \mathbf{v}$ as in the approximation formula (16.10) in terms of a path like the path $\alpha \in \mathfrak{P}^1(U)$ considered above with $\alpha(t) = \mathbf{x} + t\mathbf{v}$. In this way, (16.8) can be written as

$$D_{\mathbf{u}} f(\mathbf{x}) = \lim_{v \to 0} \frac{f(\mathbf{x} + v\mathbf{u}) - f(\mathbf{x})}{v} = \lim_{t \to 0} \frac{f \circ \alpha(t) - f(\mathbf{x})}{t}$$

where

$$t = \int_0^t |\alpha'(\tau)| \, d\tau$$

is a generalization of

$$\text{length} \left[ \alpha \big|_{(0,t)} \right]$$

to values of $t$ with $t < 0$. From this point of view, if $f : M \to \mathbb{R}$ satisfies $f \in cC^1(M)$ and $M$ is a Riemannian manifold, then one natural definition of a **directional derivative** of $f$ determined by the filament/vector $[\alpha] \in T_P M$ is

$$\lim_{t \to t_0} \frac{f \circ \alpha(t) - f(P)}{\int_{t_0}^t \left( \mu_{\alpha(\tau)}([\alpha]_{\alpha(\tau)}, [\alpha]_{\alpha(\tau)}) \right)^{1/2} d\tau}. \tag{16.15}$$

A more suggestive, though less general, way to express the intrinsic (spatial) derivative of the function $f : M \to \mathbb{R}$ in the direction of the nonzero vector $[\alpha]$ is

$$\lim_{t \searrow t_0} \frac{f \circ \alpha(t) - f(P)}{\text{length}\left[\alpha\big|_{(t_0,t)}\right]}.$$

We can then essentially reverse what was suggested in Exercise 16.8 by starting with this intrinsic derivative and then expressing it in terms of the differential map defined by (16.14). For this, we write (16.15) as

$$\frac{f \circ \alpha(t) - f(P)}{t - t_0} \; \frac{1}{\frac{1}{t-t_0}\int_{t_0}^{t}\big(\mu_{\alpha(\tau)}([\alpha]_{\alpha(\tau)}, [\alpha]_{\alpha(\tau)})\big)^{1/2}\, d\tau}.$$

We then observe that

$$\lim_{t \to t_0} \frac{f \circ \alpha(t) - f(P)}{t - t_0} = \frac{d}{dt} f \circ \alpha(t)\Big|_{t=t_0} = df_P([\alpha]), \tag{16.16}$$

and

$$\lim_{t \to t_0} \frac{1}{t - t_0} \int_{t_0}^{t} \big(\mu_{\alpha(\tau)}([\alpha]_{\alpha(\tau)}, [\alpha]_{\alpha(\tau)})\big)^{1/2}\, d\tau = \mu_P([\alpha], [\alpha])^{1/2} = \|\, [\alpha]\, \|_{T_P M}.$$

Thus, we can express the intrinsic derivative in terms of the differential map as

$$\frac{1}{\mu_P([\alpha], [\alpha])^{1/2}}\, df_P([\alpha]). \tag{16.17}$$

Notice the expression

$$df_P([\alpha]) = \frac{d}{dt} f \circ \alpha(t)\Big|_{t=t_0}$$

does not rely on the metric tensor and is fundamentally independent of distance measurement on $M$. The natural expression (16.17) for an intrinsic (spatial) derivative on the other hand, has a clear dependence on the value of $\mu$ at $P$ in the "direction" determined by the filament vector $[\alpha]$.

**Exercise 16.9.** Justify carefully the first equality in (16.16).

A second approach to the value of the intrinsic/spatial derivative of a function $f \in cC^1(M)$ as a directional derivative is by initially taking a normalized filament vector $[\alpha] \in T_P M$ satisfying $\mu_P([\alpha], [\alpha]) = 1$. Naturally since $T_P M = \mathcal{L}_P M$ is an inner product space with the inner product $\mu_P : \mathcal{L}_P M \times \mathcal{L}_P M \to \mathbb{R}$, we can also start with any $[\alpha] \in T_P M \backslash \{\mathbf{0}\}$ and consider the scaled filament vector

$$u = \frac{[\alpha]}{\mu_P([\alpha], [\alpha])^{1/2}} = \frac{[\alpha]}{\parallel [\alpha] \parallel}.$$

Notice we have used the Riemannian metric tensor $\mu$ to define $u$. Then we can call

$$D_u f(P) = df_P(u) = \lim_{t \searrow t_0} \frac{f \circ \alpha(t) - f(P)}{\text{length}\left[\alpha\Big|_{(t_0,t)}\right]}$$

the **directional derivative** of $f$ in the filament direction $u$ where $\alpha \in u$. In this way, we see the situation parallels to a large extent our discussion of directional derivatives and the differential in calculus. Specifically, the value of the differential $df_P([\alpha])$ gives an unambiguous intrinsic (spatial directional) derivative when $[\alpha]$ satisfies $\parallel [\alpha] \parallel_{T_P M} = 1$ with respect to the Riemannian norm (induced by the Riemannian inner product $\mu_P$). For other filament vectors we can adopt a version of (16.12) by setting/defining

$$D_{[\alpha]} f(P) = df_P([\alpha]). \tag{16.18}$$

This is first and foremost defining a convenient notation for the value of the differential. One should be careful however about the interpretation: As a derivative the value $df_P([\alpha])$ takes into account a secondary velocity or rate of change of distance on $M$ given by

$$\frac{ds}{dt} = \parallel [\alpha] \parallel$$

relative to some essentially external "driving" parameter $t$ related to $\alpha$. This value is not properly an intrinsic derivative unless $\parallel [\alpha] \parallel = 1$. Again, this is the same as in calculus. We can adopt the notation of (16.18) but we should be especially careful to keep in mind the meaning in the context of a manifold where the actual calculus related to a function $f : M \to \mathbb{R}$ is somewhat more subtle than what one encounters for $f : U \to \mathbb{R}$ for $U \subset \mathbb{R}^n$ because the simple flat geometry of $U$ can obscure the fundamental role played by length in the (underlying) manifold.

What is notably lacking in our discussion so far is the appearance of an expression parallel to the Euclidean inner product(s) appearing in (16.11) and (16.12) derived from the chain rule and using partial derivatives which in turn depend on the preferred basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and the corresponding coordinate functions in $\mathbb{R}^n$. These latter are not available on a manifold or in $T_P M$, and in short we have as yet no notion of the gradient of a function $f \in cC^1(M)$. We can take this as our next major question:

**Exercise 16.10.** We have discussed an intrinsic **directional derivative** $D_u f(P)$ of a function $f \in cC^1(M)$ in a direction $[\alpha] \in T_P M \backslash \{\mathbf{0}\}$ on a manifold $M$. We have also discussed a **differetial map** $df_P : T_P M \to \mathbb{R}$ associated with the same function. In the context of a function $f \in cC^1(M)$ where $M$ is a manifold, identify the **gradient** of $f$. (You should first contemplate what "kind" of object the gradient might be and notice that it cannot be a vector of partial derivatives. "Preferred" partial derivatives are not available, though as noted above we do have a notion of directional derivative. For example, the directional derivative of $f$ in a direction $u \in T_P M$ with $\mu_P(u, u) = 1$ is a "real number," and the differential of $f$ at $P$ is a "linear map." Once you know what kind of object the gradient should be, then explain how this object is actually determined using $f$ and explain the relations between and among the gradient, the directional derivative, and the differential.)

**Exercise 16.11.** Note that one difference between the discussion of derivatives for functions $f : I \to \mathbb{R}$ where $I = (a, b)$ is an open interval of $\mathbb{R}$ and the discussion of derivatives for functions $f : U \to \mathbb{R}$ where $U$ is an open subset of $\mathbb{R}^n$ is that in one space dimension existence of a derivative at a point implies local approximation while the existence of partial derivatives (or even all directional derivatives) at a point in dimension $n \geq 2$ does not necessarily imply local approximation. On the other hand, there are a number of constructions we discussed in higher dimensions that do have natural and meaningful "specializations" to the case $n = 1$, but these are usually not discussed and we did not discuss them.

**(a)** Give an example in which all the directional derivatives of a function $f : \mathbb{R}^2 \to \mathbb{R}$ exist at a point but no local approximation formula holds at that point.

**(b)** Go back and think about how the "differentiability" constructions associated with $f : U \to \mathbb{R}$ where $U$ is an open subset of $\mathbb{R}^n$ apply in the case $n = 1$. In particular, what form do the directional derivatives of (16.8) and (16.12) take in the case $n = 1$? (Do you find anything new?)

**(c)** Apply the discussion of derivatives/directional derivatives, differentials, and the gradient to a function $f : M \to \mathbb{R}$ where $M$ is a (Riemannian) manifold of dimension $n = 1$.

## 16.5 Follow-up questions

One perfectly good answer to the question

"What is the (intrinsic) gradient of $f \in cC^\infty(M)$?"

is that the (intrinsic) gradient $Df(P)$ is the unique vector

$$w \in \mathcal{L}_P M = T_P M$$

for which

$$\mu_P(w, u) = D_u f(P)$$

for every

$$u \in \mathbb{S}_P^{n-1} = \{z \in T_P M : \mu_P(z, z) = 1\} \subset \mathcal{L}_P M = T_P M.$$

**Exercise 16.12.** Show the following:

**(a)** Given a filament $w \in T_P M$ for which $\mu_P(w, u) = D_u f(P)$ for every $u \in \mathbb{S}_P^{n-1}$, there holds

$$\mu_P(w, v) = d_P f(v) \qquad \text{for every } v \in T_P M. \tag{16.19}$$

**(b)** If (16.19) holds, then the vector $w$ is unique.

**Exercise 16.13.** (existence of the gradient) Show that given $f \in cC^\infty(M)$, there exists a vector $w \in T_P M$ for which (16.19) holds. Hints: Take a chart $(U, \mathbf{p}) \in \mathcal{A}_*^\infty$ with $\xi(P) \in U$ and...

**(a)** Recall/show $\{v_1, v_2, \ldots, v_n\}$ is a basis for $T_P M$ where

**(i)** $v_j = [\mathbf{p} \circ \gamma_j] \in T_P M$ for $j = 1, 2, \ldots, n$ with

**(ii)** $\gamma_j(t) = \xi(P) + t\mathbf{e}_j$ for $j = 1, 2, \ldots, n$.

Note carefully: The filaments $v_1, v_2, \ldots, v_n$ have nothing to do with the filament $v$ appearing in (16.19) though they share a symbol "$v$."

**(b)** Make a formal calculation based on (16.19) under the assumption

$$w = \sum_{j=1}^{n} w_j v_j \tag{16.20}$$

with $w_1, w_2, \ldots, w_n \in \mathbb{R}$. Obtain from your calculation expressions for the coefficients $w_1, w_2, \ldots, w_n$ which are independent of $w$ though dependent on the chart. (This tells you what $w$ must be.)

**(c)** Define $w$ by (16.20) using the values for the coefficients you obtained from the formal calculation of part **(b)**, and show the resulting filament satisfies (16.19).

Having made sense of the (directional and total) derivative of a function $f : M \to \mathbb{R}$, here is an interesting question to ask:

> What are some interesting real valued functions $f : M \to \mathbb{R}$ to differentiate (and find the gradient of)?

**Exercise 16.14.** Calculate the gradient(s) of the coordinate functions $x_1$, $x_2$, and $x_3$ on

$$M = \mathbb{S}^2 = \{\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}.$$

**Exercise 16.15.** Let $\mathbf{u} \in T_P\mathbb{S}^{n-1}$ be a classical vector in $\mathbb{R}^n$ at some point $P \in \mathbb{S}^{n-1}$. Calculate the intrinsic directional derivatives $D_{\mathbf{u}} f_j(P) = \nabla_{\mathbf{u}} f_j(P)$ of the coordinate functions $f_j : \mathbb{S}^{n-1} \to \mathbb{R}$ given by

$$f_j(Q) = q_j \qquad \text{for} \qquad j = 1, 2, \ldots, n$$

where $Q = (q_1, q_2, \ldots, q_n)$. Here

$$M = \mathbb{S}^{n-1} = \{\mathbf{x} = (x_1, x_2, \ldots, x_{n-1}) \in \mathbb{R}^n : x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}.$$

Exercises 16.14 and 16.15 involve one of the most interesting examples of Riemannian manifolds, namely hypersurfaces in Euclidean space. These are also examples of **submanifolds** which share many of the same properties of hypersurfaces and from which much of the intuition for Riemannian geometry is/was derived. In fact, all of the main motivational problems of Riemannian geometry presented in these notes, namely Willmore's conjecture, Lawson's conjecture, and Wente's conjecture are conjectures about submanifolds. The exception is the Poincaré conjecture, which is not properly a problem from Riemannian geometry but just happens to have been resolved using Riemannian geometry.