2 The gradient of a scalar field

2.1 Differential operators and $\underline{\nabla}$

In the previous section, we saw that when we can compute the rate of change of a scalar field $f(\underline{x})$ along a curve C given by $\underline{x}(t) = x_i(t)\underline{e}_i$ (ESC), using the chain rule as

$$\frac{dF(t)}{dt} = \frac{d}{dt} f(\underline{x}(t)) \;=\; \frac{dx_1}{dt}\frac{\partial f}{\partial x_1} + \cdots + \frac{dx_n}{dt}\frac{\partial f}{\partial x_n} = \frac{dx_i}{dt}\frac{\partial f}{\partial x_i},$$

where $F(t)$ is the restriction of $f(\underline{x})$ to the curve $\underline{x}(t)$, and where again we've used ESC in the final equality.

Since this is true for all (differentiable) scalar fields, it's often useful to write this rule in terms of the derivative operators themselves, separate from the field $F(t)$. In this form, the chain rule can be given as

$$\frac{d}{dt} \;=\; \frac{dx_1}{dt}\frac{\partial}{\partial x_1} + \cdots + \frac{dx_n}{dt}\frac{\partial}{\partial x_n}$$

This is known as a *differential operator*, which can be thought of as a map which takes functions to functions using derivatives.

When using operator notation, we need to be careful about exactly what the differential operator is acting on.

Example 5. Given two real functions $f(x), g(x) : \mathbb{R} \to \mathbb{R}$, then:

- $f(x)\frac{d}{dx}$ is a differential operator which can act on $g(x)$ to give $f(x)\frac{dg(x)}{dx}$.
- $\frac{d}{dx}f(x)$ is a differential operator which can on $g(x)$ to give $\frac{d}{dx}(f(x)g(x)) = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}$ by the product rule.
- If I want an operator which multiples $g(x)$ by $\frac{df(x)}{dx}$, then I should write the operator as $\left(\frac{d}{dx}f(x)\right)$, where the brackets make it clear the the derivative is "used up", only acting on $f(x)$.

The derivative with respect to t along the curve $C : \underline{x}(t) = x_i(t)\underline{e}_i$ as above, in operator $\frac{d}{dt}$ form, can then be rewritten using the scalar product as

$$\begin{aligned}
\frac{d}{dt} &= \frac{dx_1}{dt}\frac{\partial}{\partial x_1} + \cdots + \frac{dx_n}{dt}\frac{\partial}{\partial x_n} \\
&= \left(\underline{e}_1\frac{dx_1}{dt} + \cdots + \underline{e}_n\frac{dx_n}{dt}\right) \cdot \left(\underline{e}_1\frac{\partial}{\partial x_1} + \cdots + \underline{e}_n\frac{\partial}{\partial x_n}\right) \\
&= \frac{d\underline{x}}{dt} \cdot \underline{\nabla},
\end{aligned}$$

where

$$\underline{\nabla} = \underline{e}_1\frac{\partial}{\partial x_1} + \cdots + \underline{e}_n\frac{\partial}{\partial x_n} \;=\; \underline{e}_i\frac{\partial}{\partial x_i}. \quad \text{(ESC)}$$

This differential operator $\underline{\nabla}$ is called 'del', or 'nabla', and is one of the most important objects in this course. Note that since $\underline{\nabla}$ is a vector quantity, we always write it with an underline.

If $f(\underline{x}) : \mathbb{R}^n \to \mathbb{R}$ is a scalar field, then we define its gradient ("grad f") to be given by the action of $\underline{\nabla}$ on f:

$$\underline{\nabla} f \;\equiv\; \text{grad } f = \underline{e}_1\frac{\partial f}{\partial x_1} + \underline{e}_2\frac{\partial f}{\partial x_2} + \cdots + \underline{e}_n\frac{\partial f}{\partial x_n}.$$

The gradient of a scalar field is therefore a <u>vector field</u>, with components $\frac{\partial f}{\partial x_a}$.

Example 6. In two dimensions, with $\underline{x} = x\underline{e}_1 + y\underline{e}_2$, let $f(\underline{x}) = (x^2 + y^2)/4$. Then we have

$$\begin{aligned}
\Rightarrow \quad \frac{\partial f}{\partial x} &= \frac{x}{2} \quad \frac{\partial f}{\partial y} = \frac{y}{2} \\
\Rightarrow \quad \underline{\nabla} f &= \frac{1}{2}x\underline{e}_1 + \frac{1}{2}y\underline{e}_2 \;=\; \frac{1}{2}\underline{x}.
\end{aligned}$$

The vector field can be drawn by arrows of length $|\underline{\nabla} f|$ and direction parallel to $\underline{\nabla} f$ starting at a variety of sample points:

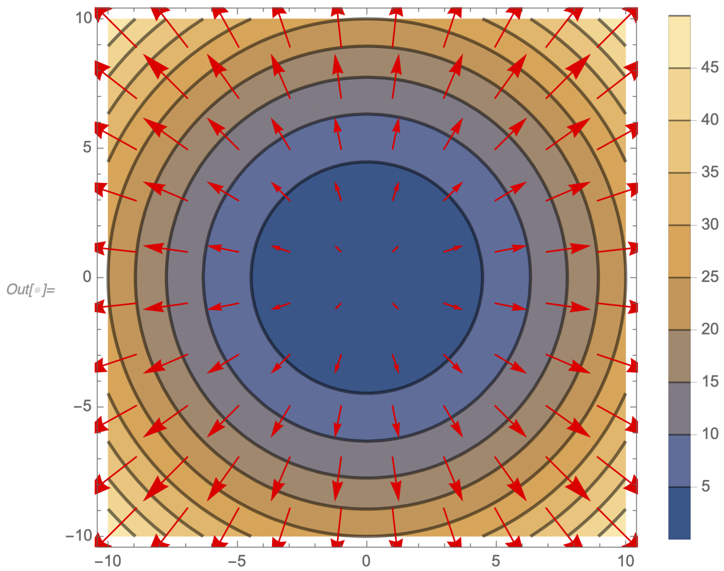

Figure 7: Level sets of the function $f(\underline{x}) = (x^2 + y^2)/4$ are shown in a contour plot. A plot of the vector field $\underline{\nabla} f(\underline{x})$ is overlaid on top of this, showing the vectors pointing away from the origin, parallel to the position vectors $\underline{x}$. The vectors in the vector field $\underline{\nabla} f(\underline{x})$ are perpendicular to the level sets of $f(\underline{x})$.

Example 7. In three dimensions, with $\underline{x} = x\underline{e}_1 + y\underline{e}_2 + z\underline{e}3$, let $\underline{a} = a_1\underline{e}_1 + a_2\underline{e}_2 + a_3\underline{e}_3$ be a constant vector, and let $f(\underline{x}) = \underline{a}.\underline{x} - \underline{x}.\underline{x}$. Then we have

$$
\begin{aligned}
& f = a_1 x + a_2 y + a_3 z - (x^2 + y^2 + z^2) \\
\Rightarrow\quad & \frac{\partial f}{\partial x} = a_1 - 2x, \qquad \frac{\partial f}{\partial y} = a_2 - 2y, \qquad \frac{\partial f}{\partial z} = a_3 - 2z \\
\Rightarrow\quad & \underline{\nabla} f = \underline{e}_1(a_1 - 2x) + \underline{e}_2(a_2 - 2y) + \underline{e}_3(a_3 - 2z) \\
& \quad = \underline{a} - 2\underline{x}.
\end{aligned}
$$

Although the picture is less easy to interpret than the 2-dimensional example, for completeness this is included as Figure 8.

2.2 Directional derivatives

Let $C : \underline{x} = \underline{x}(t)$ be a curve in $\mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ a scalar field. Then $f(\underline{x}(t)) : \mathbb{R} \to \mathbb{R}$ is f restricted to C and

$$\frac{d}{dt} f(\underline{x}(t)) \;=\; \frac{d\underline{x}}{dt}.\underline{\nabla} f \qquad \text{by chain rule.}$$

As we saw in subsection 1.2, $\frac{d\underline{x}}{dt}$ is tangent to C at $\underline{x}(t)$. If we change to a parameterisation in terms of the arc-length s, such that the tangent $\frac{d\underline{x}}{ds} = \hat{\underline{n}}$ is a unit tangent (see example 4 for a justification as to why this is possible), then we have

$$\frac{df(\underline{x}(s))}{ds} \;=\; \hat{\underline{n}}\,.\,\underline{\nabla} f. \tag{2.1}$$

Now $\frac{df(\underline{x}(s))}{ds}$ is the rate of change of f with respect to distance (arc length) in the direction $\hat{\underline{n}}$. This is called the <u>directional derivative</u> of f in the direction $\hat{\underline{n}}$ (and is sometimes written $\frac{df}{d\hat{\underline{n}}}$).

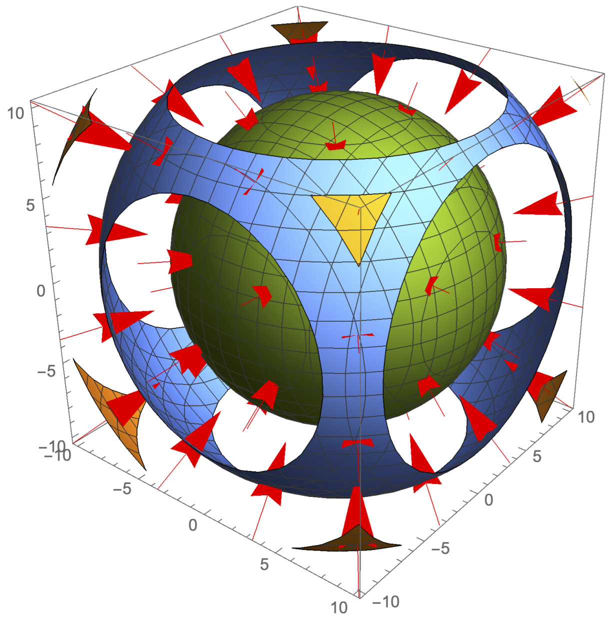

Figure 8: The contour plot of a scalar field overlaid with the gradient vector field in 3d. Parts of three level sets can be seen as spherical shells (with what center?) and representative vectors of $\underline{\nabla} f$ can be seen as red arrows. As in example 6, the vectors of $\underline{\nabla} f$ can be seen to be normal to the level sets of f. We return to this idea in the next section.

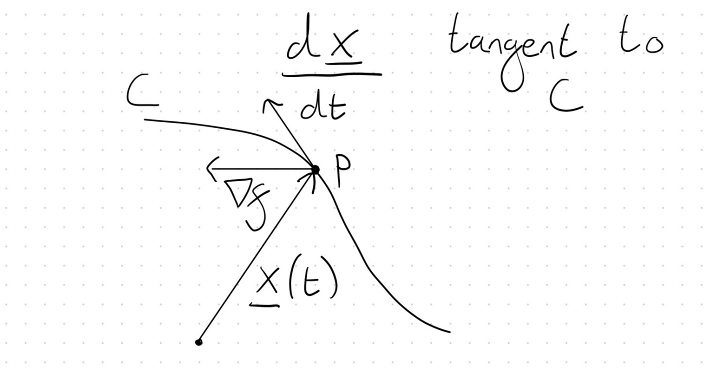

Figure 9: The tangent to a curve C at a point p is shown alongside an example of the gradient of a scalar field at the same point p.

Notice that:

$$\begin{aligned}\frac{df}{ds} &= \hat{\underline{n}} \, . \, \underline{\nabla} f \;=\; |\hat{\underline{n}}| \; |\underline{\nabla} f| \cos\theta \\ &= |\underline{\nabla} f| \cos\theta \;\leq\; |\underline{\nabla} f|.\end{aligned}$$

Therefore $|\underline{\nabla} f|$ is the greatest value of the directional derivative over all possible directions $\hat{\underline{n}}$. This value is achieved when $\theta = 0$, i.e. when $\hat{\underline{n}} \parallel \underline{\nabla} f$. Therefore $\underline{\nabla} f$ points in the direction where f increases fastest.

Note that in example 6, the vectors in the vector field $\underline{\nabla} f$ are normal to the curves of constant $|\underline{x}|$, which were the curves of constant f. This also holds more generally. In $\mathbb{R}^n$, suppose C lies entirely in the level set $f(\underline{x}) = k$ for k some constant. Call this whole level set S, so $C \subset S$.

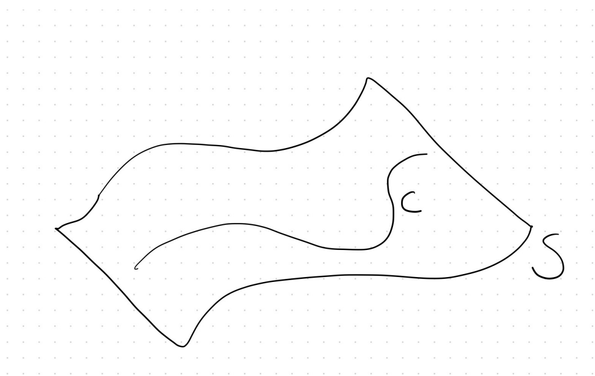

Figure 10: The curve C contained entirely within a level set of the function f. In two dimension the level sets are themselves curves, but in higher dimension spaces these level sets become surfaces, and hypersurfaces. You should think that the condition $f(\underline{x}) = k$ is a single constraint on a n-dimensional space, so the points which satisfy this constraint for an $(n-1)$-dimensional hypersurface.

So on this C, $f(\underline{x}(t)) = k$ and

$$\begin{aligned} 0 = \frac{df}{dt} &= \frac{d\underline{x}}{dt} \, . \, \underline{\nabla} f \\ \Rightarrow \quad \frac{d\underline{x}}{dt} &\perp \underline{\nabla} f, \end{aligned}$$

i.e. At all points p in a level set of f, $\underline{\nabla} f$ is orthogonal to any curve through p contained in the level set.

In $\mathbb{R}^3$, the plane through P orthogonal to $\underline{\nabla} f$ is called the tangent plane to S at P. The tangent to any curve in S through P lies in this plane. In Calculus I you already used this to find the equation of the tangent plane to a surface at a point.

2.3 Some properties of the gradient

If $f, g : \mathbb{R}^n \to \mathbb{R}$ are scalar fields, ϕ is a function $\phi : \mathbb{R} \to \mathbb{R}$, and a & b are constants then,

$$\begin{aligned} (i) \quad \underline{\nabla}(af + bg) &= a\underline{\nabla} f + b\underline{\nabla} g \\ (ii) \quad \underline{\nabla}(fg) &= (\underline{\nabla} f)g + f(\underline{\nabla} g) \\ (iii) \quad \underline{\nabla}\phi(f) &= (\underline{\nabla} f)\frac{d\phi}{df}. \end{aligned}$$

Note that as always with differential operators, the brackets are very important here. This is because $\underline{\nabla} fg$ means $\underline{\nabla}(fg)$, since derivatives ($\frac{\partial}{\partial x}$ etc.) act on everything to the right.

Example 8. Let $n = 2$ and $\underline{x} = x\underline{e}_1 + y\underline{e}_2$. If we now take $\phi(f) = f^2$ and $f(x, y) = x \sin y$,

If we now take

$$\begin{aligned} \phi(f(x, y)) &= x^2 \sin^2 y \\ \Rightarrow \quad \underline{\nabla}\phi(f) &= \underline{\nabla}(x^2 \sin^2 y) \\ &= 2x \sin^2 y \; \underline{e}_1 + 2x^2 \cos y \sin y \; \underline{e}_2, \end{aligned}$$

by direct calculation. Or:

$$\begin{aligned} \underline{\nabla} f &= \sin y \ \underline{e}_1 + x \cos y \ \underline{e}_2 \\ \Rightarrow \quad \frac{d\phi}{df} &= 2f \\ \Rightarrow \quad (\underline{\nabla} f)\frac{d\phi}{df} &= (\sin y \ \underline{e}_1 + x\cos y \ \underline{e}_2)2x\sin y \\ \Rightarrow \quad \underline{\nabla}\phi(f) &= 2x\sin^2 y \ \underline{e}_1 + 2x^2\cos y \sin y \ \underline{e}_2, \end{aligned}$$

which shows that property iii holds in this case.

If we now let $\underline{x} = x_1\underline{e}_1 + \ldots + x_n\underline{e}_n$, then these properties can be shown to hold in the general case as follows:

$$\begin{aligned} (i) \quad \underline{\nabla}(af+bg) &= \underline{e}_1\frac{\partial}{\partial x_1}(af+bg) + \ldots + \underline{e}_n\frac{\partial}{\partial x_n}(af+bg) \\ &= \underline{e}_1\left(a\frac{\partial f}{\partial x_1} + b\frac{\partial g}{\partial x_1}\right) + \ldots + \underline{e}_n\left(a\frac{\partial f}{\partial x_n} + b\frac{\partial g}{\partial x_n}\right) \\ &= a\left(\underline{e}_1\frac{\partial f}{\partial x_1} + \ldots + \underline{e}_n\frac{\partial f}{\partial y}\right) + b\left(\underline{e}_1\frac{\partial g}{\partial x_1} + \underline{e}_2\frac{\partial g}{\partial x_n}\right) \\ &= a\underline{\nabla} f + b\underline{\nabla} g. \end{aligned}$$

$$\begin{aligned} (ii) \quad \underline{\nabla}(fg) &= \underline{e}_1\frac{\partial}{\partial x_1}(fg) + \ldots + \underline{e}_n\frac{\partial}{\partial x_n}(fg) \\ &= \underline{e}_1\left(\left(\frac{\partial f}{\partial x_1}\right)g + f\frac{\partial g}{\partial x_1}\right) + \ldots + \underline{e}_n\left(\left(\frac{\partial f}{\partial x_n}\right)g + f\frac{\partial g}{\partial x_n}\right) \\ &= \left(\underline{e}_1\frac{\partial f}{\partial x_1} + \ldots + \underline{e}_n\frac{\partial f}{\partial x_n}\right)g + f\left(\underline{e}_1\frac{\partial g}{\partial x_1} + \ldots + \underline{e}_n\frac{\partial g}{\partial x_n}\right) \\ &= (\underline{\nabla} f)g + f\underline{\nabla} g. \end{aligned}$$

$$\begin{aligned} (iii) \quad \underline{\nabla}\phi &= \underline{e}_1\frac{\partial\phi}{\partial x_1} + \cdots + \underline{e}_n\frac{\partial\phi}{\partial x_n} \\ &= \underline{e}_1\frac{\partial f}{\partial x_1}\frac{d\phi}{df} + \cdots + \underline{e}_n\frac{\partial f}{\partial x_n}\frac{d\phi}{df} \\ &= (\underline{\nabla} f)\frac{d\phi}{df}. \end{aligned}$$