# Notes of “Funktionen mehrerer Variablen”

*Naoki Pross* – `naoki.pross@ost.ch`

Spring Semseter 2021

## 1 Preface

These are just my personal notes of the FuVar course, and definitively not a rigorously constructed mathematical text. The good looking LaTeX typesetting may trick you into thinking it is rigorous, but really, it is not.

## 2 Derivatives of vector valued scalar functions

**Definition 1** (Partial derivative)**.** A vector valued function $f : \mathbb{R}^m \to \mathbb{R}$, with $\mathbf{v} \in \mathbb{R}^m$, has a partial derivative with respect to $v_i$ defined as

$$\partial_{v_i} f(\mathbf{v}) = \frac{\partial f}{\partial v_i} = \lim_{h\to 0} \frac{f(\mathbf{v} + h\mathbf{e}_i) - f(\mathbf{v})}{h}$$

**Theorem 1.** (Schwarz's theorem, symmetry of partial derivatives) Under some generally satisfied conditions (continuity of $n$-th order partial derivatives) Schwarz's theorem states that it is possible to swap the order of differentiation.

$$\partial_x \partial_y f(x, y) = \partial_y \partial_x f(x, y)$$

**Application 1** (Find the slope of an implicit curve)**.** Let $f(x, y) = 0$ be an implicit curve. Its slope at any point where $\partial_y f \neq 0$ is $m = -\partial_x f / \partial_y f$

**Definition 2** (Total differential)**.** The total differential $df$ of $f : \mathbb{R}^m \to \mathbb{R}$ is

$$df = \sum_{i=1}^{m} \partial_{x_i} f \cdot dx.$$

That reads, the *total* change is the sum of the change in each direction. This implies

$$\frac{df}{dx_k} = \frac{\partial f}{\partial x_k} + \sum_{i\in\{1\leq i\leq m : i\neq k\}} \frac{\partial f}{\partial x_i} \cdot \frac{dx_i}{dx_k},$$

i.e. the change in direction $x_k$ is how $f$ changes in $x_k$ (ignoring other directions) plus, how $f$ changes with respect to each other variable $x_i$ times how they ($x_i$) change with respect to $x_k$.

**Application 2** (Linearization)**.** A function $f : \mathbb{R}^m \to \mathbb{R}$ has a linearization $g$ at $\mathbf{x}_0$ given by

$$g(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{m} \partial_{x_i} f(\mathbf{x}_0)(x_i - x_{i,0}),$$

if all partial derivatives are defined at $\mathbf{x}_0$. With the gradient (defined below) $g(\mathbf{x}) = f(\mathbf{x}_0) + \boldsymbol{\nabla} f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$.

**Application 3** (Propagation of uncertanty)**.** Given a measurement of $m$ values in a vector $\mathbf{x} \in \mathbb{R}^m$ with values given in the form $x_i = \bar{x}_i \pm \sigma_{x_i}$, a linear approximation of the error of a dependent variable $y = f(\mathbf{x})$ is computed with

$$y = \bar{y} \pm \sigma_y \approx f(\bar{\mathbf{x}}) \pm \sqrt{\sum_{i=1}^{m} \left(\partial_{x_i} f(\bar{\mathbf{x}})\sigma_{x_i}\right)^2}$$

**Definition 3** (Gradient vector)**.** The *gradient* of a function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^m$ is a column vector[1] containing the partial derivatives in each direction.

$$\boldsymbol{\nabla} f(\mathbf{x}) = \sum_{i=1}^{m} \partial_{x_i} f(\mathbf{x})\mathbf{e}_i = \begin{pmatrix} \partial_{x_1} f(\mathbf{x}) \\ \vdots \\ \partial_{x_m} f(\mathbf{x}) \end{pmatrix}$$

**Theorem 2.** The gradient vector always points towards *the direction of steepest ascent*, and thus is always perpendicular to contour lines.

**Definition 4** (Directional derivative)**.** A function $f(\mathbf{x})$ has a directional derivative in direction $\mathbf{r}$ (with $|\mathbf{r}| = 1$) of

$$\frac{\partial f}{\partial \mathbf{r}} = \nabla_{\mathbf{r}} f = \mathbf{r} \cdot \boldsymbol{\nabla} f = \sum_{i=1}^{m} r_i \partial_{x_i} f$$

**Definition 5** (Jacobian Matrix)**.** The *Jacobian* $\mathbf{J}_f$ (sometimes written as $\frac{\partial(f_1,\dots f_m)}{\partial(x_1,\dots,x_n)}$) of a function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ is a matrix $\in \mathbb{R}^{m\times n}$ whose entry at the $i$-th row and $j$-th column is given by $(\mathbf{J}_f)_{i,j} = \partial_{x_j} f_i$, so

$$\mathbf{J}_f = \begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_m} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_n & \cdots & \partial_{x_m} f_n \end{pmatrix} = \begin{pmatrix} (\boldsymbol{\nabla} f_1)^t \\ \vdots \\ (\boldsymbol{\nabla} f_m)^t \end{pmatrix}$$

**Remark 1.** In the scalar case ($n = 1$) the Jacobian matrix is the transpose of the gradient vector.

**Definition 6** (Hessian matrix)**.** Given a function $f : \mathbb{R}^m \to \mathbb{R}$, the square matrix whose entry at the $i$-th row and $j$-th column is the second derivative of $f$ first with respect to $x_j$ and then to $x_i$ is known as the *Hessian* matrix. $(\mathbf{H}_f)_{i,j} = \partial_{x_i}\partial_{x_j} f$ or

$$\mathbf{H}_f = \begin{pmatrix} \partial_{x_1}\partial_{x_1} f & \cdots & \partial_{x_1}\partial_{x_m} f \\ \vdots & \ddots & \vdots \\ \partial_{x_m}\partial_{x_1} f & \cdots & \partial_{x_m}\partial_{x_m} f \end{pmatrix}$$

Because (almost always) the order of differentiation does not matter, it is a symmetric matrix.

[1] In matrix notation it is also often defined as row vector to avoid having to do some transpositions in the Jacobian matrix and dot products in directional derivatives

# 3 Methods for maximization and minimization problems

## 3.1 Analytical methods

**Method 1** (Find stationary points)**.** Given a function $f : D \subseteq \mathbb{R}^m \to \mathbb{R}$, to find its maxima and minima we shall consider the points

- that are on the boundary[2] of the domain $\partial D$,
- where the gradient $\boldsymbol{\nabla} f$ is not defined,
- that are stationary, i.e. where $\boldsymbol{\nabla} f = \mathbf{0}$.

**Method 2** (Determine the type of stationary point for 2 dimensions)**.** Given a scalar function of two variables $f(x, y)$ and a stationary point $\mathbf{x}_s$ (where $\boldsymbol{\nabla} f(\mathbf{x}_s) = \mathbf{0}$), we define the *discriminant*

$$\Delta = \partial_x^2 f \partial_y^2 f - \partial_y \partial_x f$$

- if $\Delta > 0$ then $\mathbf{x}_s$ is an extrema, if $\partial_x^2 f(\mathbf{x}_s) < 0$ it is a maximum, whereas if $\partial_x^2 f(\mathbf{x}_s) > 0$ it is a minimum;
- if $\Delta < 0$ then $\mathbf{x}_s$ is a saddle point;
- if $\Delta = 0$ we need to analyze further.

**Remark 2.** The previous method is obtained by studying the second directional derivative $\nabla_{\mathbf{r}} \nabla_{\mathbf{r}} f$ at the stationary point in direction of a vector $\mathbf{r} = \mathbf{e}_1 \cos(\alpha) + \mathbf{e}_2 \sin(\alpha)$.

**Method 3** (Determine the type of stationary point in higher dimensions)**.** Given a scalar function of multiple variables $f(\mathbf{x})$ and a stationary point $\mathbf{x}_s$ ($\boldsymbol{\nabla} f(\mathbf{x}_s) = \mathbf{0}$), we compute the Hessian matrix $\mathbf{H}_f(\mathbf{x}_s)$ and its eigenvalues $\lambda_1, \ldots, \lambda_m$, then

- if all $\lambda_i > 0$, the point is a minimum;
- if all $\lambda_i < 0$, the point is a maximum;
- if there are both positive and negative eigenvalues, it is a saddle point.

In the other cases, when there are $\lambda_i \leq 0$ and/or $\lambda_i \geq 0$ further analysis is required.

**Remark 3.** Recall that to compute the eigenvalues of a matrix, one must solve the equation $(\mathbf{H} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$. Which can be done by solving the characteristic polynomial $\det(\mathbf{H} - \lambda \mathbf{I}) = 0$ to obtain $\dim(\mathbf{H})$ $\lambda_i$, which when plugged back in result in a overdetermined system of equations.

**Method 4** (Quickly find the eigenvalues of a $2 \times 2$ matrix)**.** This is a nice trick. For a square matrix $\mathbf{H}$, let

$$m = \frac{1}{2} \operatorname{tr} \mathbf{H} = \frac{a + d}{2}, \quad p = \det \mathbf{H} = ad - bc,$$

then $\lambda_{1,2} = m \pm \sqrt{m^2 - p}$.

[2]If it belongs to $f$.

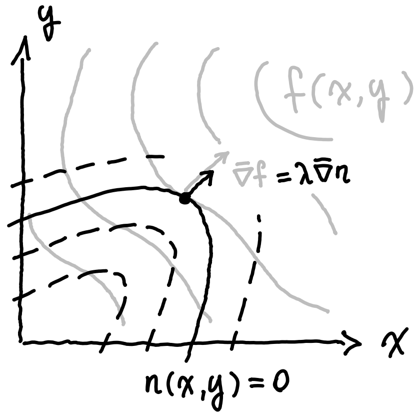


Figure 1: Intuition for the method of Lagrange multipliers. Extrema of a constrained function are where $\boldsymbol{\nabla} f$ is proportional to $\boldsymbol{\nabla} n$.

**Method 5** (Search for a constrained extremum in 2 dimensions)**.** Let $n(x, y) = 0$ be a constraint in the search of the extrema of a function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$. To find the extrema we look for points

- on the boundary[2] $\mathbf{u} \in \partial D$ where $n(\mathbf{u}) = 0$;
- $\mathbf{u}$ where the gradient either does not exist or is $\mathbf{0}$, and satisfy $n(\mathbf{u}) = 0$;
- that solve the system of equations

$$\begin{cases} \partial_x f(\mathbf{u}) \cdot \partial_y n(\mathbf{u}) = \partial_y f(\mathbf{u}) \cdot \partial_x n(\mathbf{u}) \\ n(\mathbf{u}) = 0 \end{cases}$$

**Method 6** (Search for a constrained extremum in higher dimensions, method of Lagrange multipliers)**.** We wish to find the extrema of $f : D \subseteq \mathbb{R}^m \to \mathbb{R}$ under $k < m$ constraints $n_1 = 0, \cdots, n_k = 0$. To find the extrema we consider the following points:

- Points on the boundary[2] $\mathbf{u} \in \partial D$ that satisfy $n_i(\mathbf{u}) = 0$ for all $1 \leq i \leq k$,
- Points $\mathbf{u} \in D$ where either
  - any of $\boldsymbol{\nabla} f, \boldsymbol{\nabla} n_1, \ldots, \boldsymbol{\nabla} n_k$ do not exist, or
  - $\boldsymbol{\nabla} n_1, \ldots, \boldsymbol{\nabla} n_k$ are linearly *dependent*,

  and that satisfy $0 = n_1(\mathbf{u}) = \ldots = n_k(\mathbf{u})$.
- Points that solve the system of $m + k$ equations

$$\begin{cases} \boldsymbol{\nabla} f(\mathbf{u}) = \sum_{i=1}^{k} \lambda_i \boldsymbol{\nabla} n_i(\mathbf{u}) & (m\text{-dimensional}) \\ n_i(\mathbf{u}) = 0 & \text{for } 1 \leq i \leq k \end{cases}$$

  The $\lambda$ values are known as *Lagrange multipliers*.

The calculation of the last point can be written more compactly by defining the *Lagrangian*

$$\mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = f(\mathbf{u}) - \sum_{i=0}^{k} \lambda_i n_i(\mathbf{u}),$$

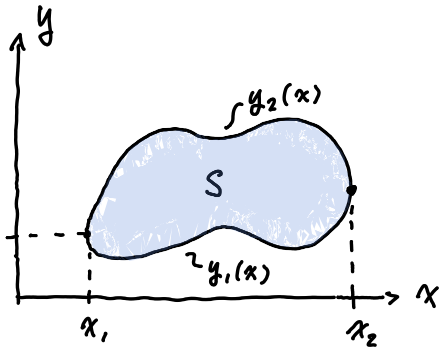


Figure 2: Double integral.

| | Volume $dv$ | Surface $d\mathbf{s}$ |
|---|---|---|
| Cartesian | $-$ | $dx\,dy$ |
| Polar | $-$ | $r d\,r d\phi$ |
| Curvilinear | $-$ | $\lvert\mathbf{J}_f\rvert\,du\,dv$ |
| Cartesian | $dx\,dy\,dz$ | $\hat{\mathbf{z}}\,dx\,dy$ |
| Cylindrical | $r\,dr\,d\phi\,dz$ | $\hat{\mathbf{z}}r\,dr\,d\phi$ |
| | | $\hat{\boldsymbol{\phi}}\,dr\,dz$ |
| | | $\hat{\mathbf{r}}r\,d\phi\,dz$ |
| Spherical | $r^2\sin\theta\,dr\,d\theta\,d\phi$ | $\hat{\mathbf{r}}r^2\sin\theta\,d\theta\,d\phi$ |
| Curvilinear | $\lvert\mathbf{J}_f\rvert\,du\,dv\,dw$ | $-$ |

Table 1: Differential elements for integration.

where $\boldsymbol{\lambda} = \lambda_1, \dots, \lambda_k$ and then solving the $m+k$ dimensional equation $\nabla\mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{0}$ (this is generally used in numerical computations and not very useful by hand).

## 3.2 Numerical methods

**Method 7** (Newton's method)**.** For a function $f : \mathbb{R}^m \to \mathbb{R}$ we wish to numerically find its stationary points (where $\nabla f = \mathbf{0}$).

1. Pick a starting point $\mathbf{x}_0$.
2. Set the linearisation[3] of $\nabla f$ at $\mathbf{x}_k$ to zero and solve for $\mathbf{x}_{k+1}$.

$$\begin{aligned}\nabla f(\mathbf{x}_k) + \mathbf{H}_f(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) &= \mathbf{0}\\ \mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_f^{-1}(\mathbf{x}_k)\nabla f(\mathbf{x}_k)\end{aligned}$$

3. Repeat the last step until the magnitude of the error $|\boldsymbol{\epsilon}| = |\mathbf{H}_f^{-1}(\mathbf{x}_k)\nabla f(\mathbf{x}_k)|$ is sufficiently small.

**Method 8** (Gradient ascent / descent)**.** Given $f : \mathbb{R}^m \to \mathbb{R}$ we wish to numerically find the stationary points (where $\nabla f = \mathbf{0}$).

1. Define an arbitrarily small length $\eta$ and a starting point $\mathbf{x}_0$
2. Compute $\mathbf{v} = \pm\nabla f(\mathbf{x}_k)$ (positive for ascent, negative for descent), then $\mathbf{x}_{k+1} = \mathbf{x}_k + \eta\mathbf{v}$ if the rate of change $\epsilon$ is acceptable ($\epsilon = |\nabla f(\mathbf{x}_{k+1})| > 0$) else recompute $\mathbf{v} := \pm\nabla f(\mathbf{x}_{k+1})$.
3. Stop when the rate of change $\epsilon$ stays small enough for many iterations.

# 4 Integration of vector valued scalar functions

**Theorem 3** (Change the order of integration for double integrals)**.** For a double integral over a region $S$ (see Fig. 2) we need to compute

$$\iint_S f(x,y)\,ds = \int_{x_1}^{x_2}\int_{y_1(x)}^{y_2(x)} f(x,y)\,dydx.$$

[3]The gradient becomes a hessian matrix.

If $y_1(x)$ and $y_2(x)$ are bijective we can swap the order of integration by finding the inverse functions $x_1(y)$ and $x_2(y)$. If they are not bijective (like in Fig. 2), the region must be split into smaller parts. If the region is a rectangle it is always possible to change the order of integration.

**Theorem 4** (Transformation of coordinates in 2 dimensions)**.** Given two "nice" functions $x(u,v)$ and $y(u,v)$, that means are a bijection from $S$ to $S'$ with continuous partial derivatives and nonzero Jacobian determinant $|\mathbf{J}_f| = \partial_u x\partial_v y - \partial_v x\partial_u y$, which transform the coordinate system. Then

$$\iint_S f(x,y)\,ds = \iint_{S'} f(x(u,v), y(u,v))|\mathbf{J}_f|\,ds.$$

**Theorem 5** (Transformation of coordinates)**.** The generalization of theorem 4 is quite simple. For an $m$-integral of a function $f : \mathbb{R}^m \to \mathbb{R}$ over a region $B$, we let $\mathbf{x}(\mathbf{u})$ be "nice" functions that transform the coordinate system. Then as before

$$\int_B f(\mathbf{x})\,ds = \int_{B'} f(\mathbf{x}(\mathbf{u}))|\mathbf{J}_f|\,ds.$$

**Application 4** (Physics)**.** Given the mass $m$ and density function $\rho$ of an object, its *center of mass* is calculated with

$$\mathbf{x}_c = \frac{1}{m}\int_V \mathbf{x}\rho(\mathbf{x})\,dv \overset{\rho\text{ const.}}{=} \frac{1}{V}\int_V \mathbf{x}\,dv.$$

The (scalar) *moment of inertia* $J$ of an object is given by

$$J = \int_V \rho(\mathbf{r})r^2\,dv.$$

# 5 Parametric curves, line and surface integrals

**Definition 7** (Parametric curve)**.** A parametric curve is a vector function $\mathcal{C} : \mathbb{R} \to W \subseteq \mathbb{R}^n, t \mapsto \mathbf{f}(t)$, that takes a parameter $t$.

**Theorem 6** (Derivative of a curve)**.** The derivative of a curve is

$$\begin{aligned}\mathbf{f}'(t) &= \lim_{h\to 0}\frac{\mathbf{f}(t+h)-\mathbf{f}(t)}{h}\\ &= \sum_{i=0}^{n}\left(\lim_{h\to 0}\frac{f_i(t+h)-f_i(t)}{h}\right)\mathbf{e}_i\\ &= \sum_{i=0}^{n}\frac{df_i}{dt}\mathbf{e}_i = \left(\frac{df_1}{dt},\dots,\frac{df_m}{dt}\right)^t.\end{aligned}$$

**Theorem 7** (Multivariable chain rule)**.** Let $\mathbf{x}:\mathbb{R}\to\mathbb{R}^m$ and $f:\mathbb{R}^m\to\mathbb{R}$, so that $f\circ\mathbf{x}:\mathbb{R}\to\mathbb{R}$, then the multivariable chain rule states:

$$\frac{d}{dt}f(\mathbf{x}(t)) = \boldsymbol{\nabla} f(\mathbf{x}(t))\cdot\mathbf{x}'(t) = \nabla_{\mathbf{x}'(t)}f(\mathbf{x}(t)).$$

**Theorem 8** (Signed area enclosed by a planar parametric curve)**.** A planar (2D) parametric curve $(x(t),y(t))^t$ with $t\in[r,s]$ that does not intersect itself encloses a surface with area

$$A = \int_r^s x'(t)y(t)\,dt = \int_r^s x(t)y'(t)\,dt.$$

**Definition 8** (Line integral in a scalar field)**.** Let $\mathcal{C}:[a,b]\to\mathbb{R}^n, t\mapsto\mathbf{x}(t)$ be a parametric curve. The *line integral* in a field $f(\mathbf{x})$ is the integral of the signed area under the curve traced in $\mathbb{R}^n$, and is computed with

$$\int_{\mathcal{C}} f(\mathbf{x})\,d\ell = \int_{\mathcal{C}} f(\mathbf{x})\,|d\mathbf{x}| = \int_a^b f(\mathbf{x}(t))|\mathbf{x}'(t)|\,dt.$$

**Application 5** (Length of a parametric curve)**.** By computing the line integral of the function $1(\mathbf{x})$ we get the length of the parametric curve $\mathcal{C}:[a,b]\to\mathbb{R}^n$.

$$\int_{\mathcal{C}} d\ell = \int_{\mathcal{C}}|d\mathbf{x}| = \int_a^b\sqrt{\sum_{i=1}^n x_i'(t)^2}\,dt$$

The special case with the scalar function $f(x)$ results in $\int_a^b\sqrt{1+f'(x)^2}\,dx$.

**Definition 9** (Line integral in a vector field)**.** The line integral in a vector field $\mathbf{F}(\mathbf{x})$ is the "sum" of the projections of the field's vectors on the tangent of the parametric curve $\mathcal{C}$.

$$\int_{\mathcal{C}}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = \int_a^b\mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt$$

**Theorem 9** (Line integral in the opposite direction)**.** By integrating while moving backwards $(-t)$ on the parametric curve gives

$$\int_{-\mathcal{C}}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = -\int_{\mathcal{C}}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}.$$

**Definition 10** (Conservative field)**.** A vector field is said to be *conservative* the line integral over a closed path is zero.

$$\oint_{\mathcal{C}}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = 0$$

**Theorem 10.** For a twice partially differentiable vector field $\mathbf{F}(\mathbf{x})$ in $n$ dimensions without "holes", i.e. in which each closed curve can be contracted to a point (simply connected open set), the following statements are equivalent:

- $\mathbf{F}$ is conservative,
- $\mathbf{F}$ is path-independent,
- $\mathbf{F}$ is a *gradient field*, i.e. there is a function $\phi$ called *potential* such that $\mathbf{F}=\boldsymbol{\nabla}\phi$,
- $\mathbf{F}$ satisfies the condition $\partial_{x_j}F_i=\partial_{x_i}F_j$ for all $i,j\in\{1,2,\dots,n\}$. In the 2D case $\partial_xF_y=\partial_yF_x$, and in 3D
$$\begin{cases}\partial_yF_x=\partial_xF_y\\ \partial_zF_y=\partial_yF_z\\ \partial_xF_z=\partial_zF_x\end{cases}$$

**Theorem 11.** In a conservative field $\mathbf{F}$ with gradient $\phi$, using the multivariable the chain rule:

$$\begin{aligned}\int_{\mathcal{C}}\mathbf{F}\cdot d\mathbf{r} &= \int_{\mathcal{C}}\mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt\\ &= \int_{\mathcal{C}}\boldsymbol{\nabla}\phi(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt\\ &= \int_{\mathcal{C}}\frac{d\phi(\mathbf{r}(t))}{dt}\,dt = \phi(\mathbf{r}(b))-\phi(\mathbf{r}(a)).\end{aligned}$$

**Definition 11** (Parametric surface)**.** A parametric surface is a vector function $\mathcal{S}:W\subseteq\mathbb{R}^2\to\mathbb{R}^3$.

**Theorem 12** (Area of a parametric surface)**.** The area spanned by a parametric surface $\mathbf{s}(u,v)$, with continuous partial derivatives and that satisfy $\partial_u\mathbf{s}\times\partial_v\mathbf{s}\neq\mathbf{0}$, is given by

$$A = \int_{\mathcal{S}} ds = \iint|\partial_u\mathbf{s}\times\partial_v\mathbf{s}|\,dudv.$$

**Definition 12** (Scalar surface integral)**.** Let $f:\mathbb{R}^3\to\mathbb{R}$ be a function on a parametric surface $\mathbf{s}:W\subseteq\mathbb{R}^2\to\mathbb{R}^3$. The surface integral of $f$ over $\mathcal{S}$ is

$$\int_{\mathcal{S}} f\,ds = \iint_W f(\mathbf{s}(u,v))\cdot|\partial_u\mathbf{s}\times\partial_v\mathbf{s}|\,dudv.$$

# 6 Vector analysis

**Definition 13** (Flux)**.** In a vector field $\mathbf{F}:\mathbb{R}^m\to\mathbb{R}^n$ we define the *flux* through a parametric surface $\mathcal{S}$ as

$$\Phi = \int_{\mathcal{S}}\mathbf{F}\cdot d\mathbf{s} = \int_{\mathcal{S}}\mathbf{F}\cdot\hat{\mathbf{n}}\,ds.$$

If $\mathcal{S}$ is a closed surface we write $\mathring{\Phi}=\oint_{\mathcal{S}}\mathbf{F}\cdot d\mathbf{s}$.

If we now take the normalized flux on the surface of an arbitrarily small volume $V$ (limit as $V\to 0$) we get the *divergence*

$$\boldsymbol{\nabla}\cdot\mathbf{F} = \lim_{V\to 0}\frac{1}{V}\oint_{\partial V}\mathbf{F}\cdot d\mathbf{s}.$$

**Theorem 13** (Formula for divergence)**.** Let $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$ be a vector field. The divergence of $\mathbf{F} = (F_{x_1}, \ldots, F_{x_m})^t$ is

$$\boldsymbol{\nabla}\cdot \mathbf{F} = \sum_{i=1}^{m} \partial_{x_i} F_{x_i},$$

as suggested by the (ab)use of the dot product notation.

**Theorem 14** (Divergence theorem, Gauss's theorem)**.** Because the flux on the boundary $\partial V$ of a volume $V$ contains information of the field inside of $V$, it is possible relate the two with

$$\int_V \boldsymbol{\nabla}\cdot \mathbf{F}\, dv = \oint_{\partial V} \mathbf{F} \cdot d\mathbf{s}.$$

**Definition 14** (Circulation, Vorticity)**.** The result of a closed line integral can be interpreted as a macroscopic measure how much the field rotates around a given point, and is thus sometimes called *circulation* or *vorticity*.

As before, if we now make the area $A$ enclosed by the parametric curve for the circulation arbitrarily small, normalize it, and use Gauss's theorem we get a local measure called *curl*.

$$\boldsymbol{\nabla}\times \mathbf{F} = \lim_{A\to 0} \frac{\hat{\mathbf{n}}}{A} \oint_{\partial A} \mathbf{F} \cdot d\mathbf{s}$$

Notice that the curl is a vector, normal to the enclosed surface $A$.

**Theorem 15** (Formula for curl)**.** Let $\mathbf{F}$ be a vector field. In 2 dimensions

$$\boldsymbol{\nabla}\times \mathbf{F} = (\partial_x F_y - \partial_y F_x)\, \hat{\mathbf{z}}.$$

And in 3D

$$\boldsymbol{\nabla}\times \mathbf{F} = \begin{pmatrix} \partial_y F_z - \partial_z F_y \\ \partial_z F_x - \partial_x F_z \\ \partial_x F_y - \partial_y F_x \end{pmatrix} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix}.$$

**Theorem 16** (Stokes' theorem)**.**

$$\int_{\mathcal{S}} \boldsymbol{\nabla}\times \mathbf{F} \cdot d\mathbf{s} = \oint_{\partial\mathcal{S}} \mathbf{F} \cdot d\mathbf{r}$$

**Theorem 17** (Green's theorem)**.** The special case of Stokes' theorem in 2D is knowns as Green's theorem.

$$\int_{\mathcal{S}} \partial_x F_y - \partial_y F_x \, ds = \oint_{\partial\mathcal{S}} \mathbf{F} \cdot d\mathbf{r}$$

**Definition 15** (Laplacian operator)**.** A second vector derivative is so important that it has a special name. For a scalar function $f : \mathbb{R}^m \to \mathbb{R}$ the divergence of the gradient

$$\nabla^2 = \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} f) = \sum_{i=1}^{m} \partial_{x_i}^2 f_{x_i}$$

is called the *Laplacian operator*.

**Definition 16** (Vector Laplacian)**.** The Laplacian operator can be extended on a vector field $\mathbf{F}$ to the *Laplacian vector* by applying the Laplacian to each component:

$$\boldsymbol{\nabla}^2\, \mathbf{F} = (\nabla^2\, F_x)\hat{\mathbf{x}} + (\nabla^2\, F_y)\hat{\mathbf{y}} + (\nabla^2\, F_z)\hat{\mathbf{z}}.$$

The vector Laplacian can also be defined as

$$\boldsymbol{\nabla}^2\, \mathbf{F} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot \mathbf{F}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times \mathbf{F}).$$

**Theorem 18** (Product rules and second derivatives)**.** Let $f, g$ be sufficiently differentiable scalar functions $D \subseteq \mathbb{R}^m \to \mathbb{R}$ and $\mathbf{A}, \mathbf{B}$ be sufficiently differentiable vector fields in $\mathbb{R}^m$ (with $m = 2$ or $3$ for equations with the curl).

- Rules with the gradient

$$\begin{aligned} \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot \mathbf{A}) &= \boldsymbol{\nabla}\times \boldsymbol{\nabla}\times \mathbf{A} + \boldsymbol{\nabla}^2\, \mathbf{A} \\ \boldsymbol{\nabla}(f \cdot g) &= (\boldsymbol{\nabla} f) \cdot g + f \cdot \boldsymbol{\nabla} g \\ \boldsymbol{\nabla}(\mathbf{A} \cdot \mathbf{B}) &= (\mathbf{A} \cdot \boldsymbol{\nabla})\mathbf{B} + (\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{A} \\ &\quad + \mathbf{A} \times (\boldsymbol{\nabla}\times \mathbf{B}) + \mathbf{B} \times (\boldsymbol{\nabla}\times \mathbf{A}) \end{aligned}$$

- Rules with the divergence

$$\begin{aligned} \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} f) &= \nabla^2 f \\ \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times \mathbf{A}) &= 0 \\ \boldsymbol{\nabla}\cdot(f \cdot \mathbf{A}) &= (\boldsymbol{\nabla} f) \cdot \mathbf{A} + f \cdot (\boldsymbol{\nabla}\cdot \mathbf{A}) \\ \boldsymbol{\nabla}\cdot(\mathbf{A} \times \mathbf{B}) &= (\boldsymbol{\nabla}\times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\boldsymbol{\nabla}\times \mathbf{B}) \end{aligned}$$

- Rules with the curl

$$\begin{aligned} \boldsymbol{\nabla}\times(\boldsymbol{\nabla} f) &= \mathbf{0} \\ \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times \mathbf{A}) &= \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot \mathbf{A}) - \boldsymbol{\nabla}^2\, \mathbf{A} \\ \boldsymbol{\nabla}\times(\boldsymbol{\nabla}^2\, \mathbf{A}) &= \boldsymbol{\nabla}^2(\boldsymbol{\nabla}\times \mathbf{A}) \\ \boldsymbol{\nabla}\times(f \cdot \mathbf{A}) &= (\boldsymbol{\nabla} f) \times \mathbf{A} + f \cdot \boldsymbol{\nabla}\times \mathbf{A} \\ \boldsymbol{\nabla}\times(\mathbf{A} \times \mathbf{B}) &= (\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{A} - (\mathbf{A} \cdot \boldsymbol{\nabla})\mathbf{B} \\ &\quad + \mathbf{A} \cdot (\boldsymbol{\nabla}\cdot \mathbf{B}) - \mathbf{B} \cdot (\boldsymbol{\nabla}\cdot \mathbf{A}) \end{aligned}$$

# License