

Counter examples and special cases that I sometimes need to show my maths students

Dr C. G. Lester, Peterhouse

February 28, 2025

Contents

1	Unusual saddles	1
2	Examples of initial value problems with non-unique solutions	2
3	What orthogonal matrices are not	3
4	Why you should be careful with your complex number manipulations	3
5	Why the difference between conditional- and absolute-convergence is important	4
6	Why you should be careful when changing the order of integration	4
7	Second derivatives need not commute (trivial example)	5
8	Second derivatives need not commute (second sort of example)	6
9	Second derivatives need not commute (totally different sort of example)	7
10	Why you should not extrapolate based on the first few terms of a sequence	8
11	Why you should be careful with Lagrange Multipliers	8
12	Cauchy Riemann Equations and Wirtinger Derivatives	9
	12.1 Wirtinger derivatives	9
	12.2 Extension to fields	10
	12.3 Cauchy Riemann	11
13	L'Hopital's rule	11
14	$SU(3)$ multiplet multiplicities	11
15	Playing the Cello (in equal temperament)	11
16	Another way of doing one of the sums from the Fourier part of the course	12
17	Inverses of Lorentz transformations in tensor/component form require sideways index movements.	13

1 Unusual saddles

Stationary points of functions of two variables are often called saddle points if they are neither maxima nor minima. Contours through such saddles frequently take the form of two lines crossing at some (non-zero) angle, at least locally, however this is not always the case. Four examples of unusual saddles are shown in Figure 1.

These are examples of stationary points which are neither local maxima, nor local minima, but which do not have the "classic" saddle contour shape →  

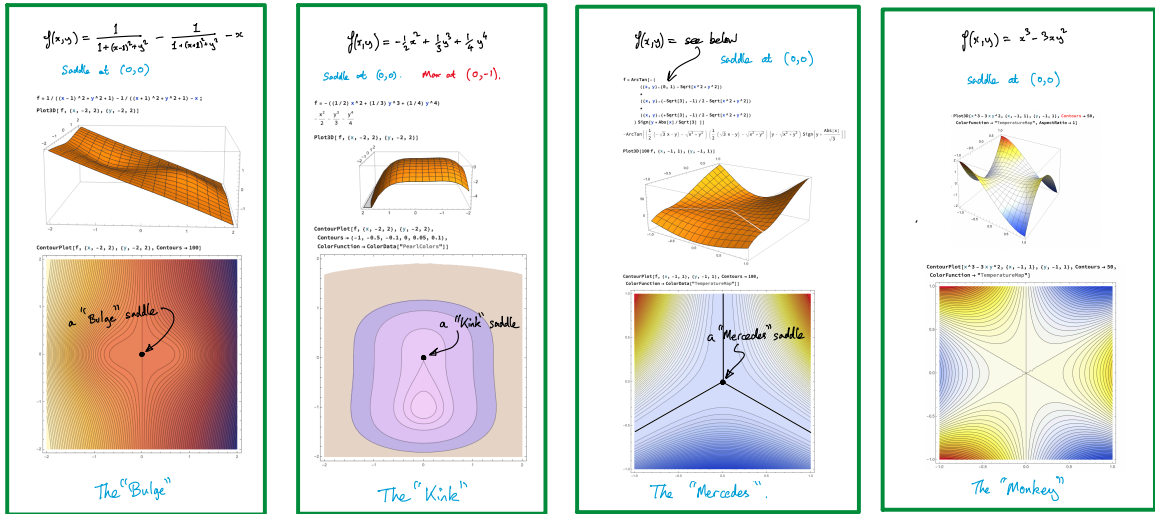


Figure 1: Unusual saddles.

2 Examples of initial value problems with non-unique solutions

Consider the initial value problem

$$\frac{dy}{dx} = 3y^{2/3} \quad \text{with} \quad y(0) = 0 \quad (1)$$

for a real function $y(x)$ of a non-negative real variable $x \geq 0$. One may show by direct substitution that this initial value problem has **infinitely many** solutions which can all be written in terms of a real parameter $a \geq 0$ as follows:

$$y(x) = \begin{cases} 0 & x \leq a, \\ (x - a)^3 & x \geq a. \end{cases} \quad (2)$$

That there can be **infinitely many** solutions to a simple initial value problem can appear surprising the first time one sees that it can happen!

One could demonstrate that (2) solves (1) by simply substituting the former into the latter. A more discursive construction of the solution is provided in section 2 for those who might be interested. Skip if not interested!

A discursive derivation of (2)

By inspection we can see that (1) has at least one solution, namely:

$$y(x) = 0 \quad \text{for all } x.$$

Are there any other solutions? Let us try solving the differential equation in a region in which $y(x) \neq 0$. In such a region we can rewrite it as

$$\frac{1}{3}y^{-2/3} \frac{dy}{dx} = 1$$

and so can say

$$\begin{aligned}
 & \frac{1}{3} \int_{x_0}^x y^{-2/3} \frac{dy}{dx} dx = \int_{x_0}^x dx \\
 \implies & \frac{1}{3} \int_{y_0}^y y^{-2/3} dy = x - x_0 \\
 \implies & \left[y^{1/3} \right]_{y_0}^y = x - x_0 \\
 \implies & y^{1/3} - y_0^{1/3} = x - x_0. \tag{3}
 \end{aligned}$$

Although we derived (3) under the assumption that $y \neq 0$, it is tempting to investigate whether (3) can be used even when $y = 0$. In that spirit, if we apply the boundary condition of (1) to (3) (i.e. if we set x_0 and y_0 to 0) we find:

$$y^{1/3} = x$$

which, since x and y are real, is the same as

$$y = x^3. \tag{4}$$

Quite clearly (4) satisfies the initial value in (1). We may also show that (4) satisfies the differential equation in (1) since

$$(4) \implies \frac{dy}{dx} = 3x^2$$

and so

$$((4) \text{ and } x \geq 0) \implies \frac{dy}{dx} = y^{3/2}$$

in which the last step has made use of the non-negative assumption about x which featured in the statement of the problem.

Thus far we have found two solutions of (1), namely:

$$y(x) = \begin{cases} y_1(x) \equiv 0 \\ y_2(x) \equiv x^3 \end{cases}$$

however we are not done. We generate even more solutions to (1) by sticking together solutions of the form $y_1(x)$ to suitably translated solutions of the form $y_2(x)$. Specifically, the following is a solution of (1) for any $a \geq 0$:

$$y(x) = \begin{cases} 0 & x \leq a, \\ (x - a)^3 & x \geq a. \end{cases}$$

3 What orthogonal matrices are not

The matrix

$$M = \begin{pmatrix} \cosh 7 & i \sinh 7 \\ -i \sinh 7 & \cosh 7 \end{pmatrix} \tag{5}$$

is not orthogonal, even though it satisfies $MM^T = M^T M = \mathbf{1}$. It is, however, a Lester Matrix, since Lester Matrices are those matrices M which satisfy $MM^T = M^T M = \mathbf{1}$. Orthogonal Matrices are required to be real, whereas Lester Matrices are not. Orthogonal Matrices are therefore the real sub-set subset of the Lester Matrices.

4 Why you should be careful with your complex number manipulations

Complex numbers are well founded, and can always be manipulated safely. However, it is easy to be sloppy when manipulating them. For example, there is an error in the following argument but it is hard

to spot:

$$\begin{array}{ll}
 1 &= \sqrt{1} && \text{hopefully this is obvious} \\
 &= \sqrt{(-1)(-1)} && \text{since } 1 = (-1)(-1) \\
 &= \sqrt{-1}\sqrt{-1} && \text{since } (ab)^c = a^c b^c \\
 &= ii && \text{by definition of } i \\
 &= i^2 && \text{since } aa = a^2 \\
 &= -1 && \text{since } i \text{ is the square root of } -1
 \end{array} \tag{6}$$

Here is a variation on the same bad argument which avoids use of any explicit “ i ”s:

$$\begin{array}{ll}
 1 &= \sqrt{1} && \text{hopefully this is obvious} \\
 &= \sqrt{(-1)(-1)} && \text{since } 1 = (-1)(-1) \\
 &= \sqrt{-1}\sqrt{-1} && \text{since } (ab)^c = a^c b^c \\
 &= (\sqrt{-1})^2 && \text{since } aa = a^2 \\
 &= ((-1)^{\frac{1}{2}})^2 && \text{since } \sqrt{a} = a^{\frac{1}{2}} \\
 &= (-1)^1 && \text{since } (a^b)^c = a^{bc} \\
 &= -1 && \text{since } a^1 = a \text{ for all } a.
 \end{array} \tag{7}$$

5 Why the difference between conditional- and absolute-convergence is important

Add up all the numbers in the table below and what do you get? It appears to be “0” or “-2” depending on whether you add up the column totals or the row totals! Why?

Row and column totals	-1	$-\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{1}{8}$	$-\frac{1}{16}$...	column totals sum to “-2”
0	-1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$...	
0	0	-1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$...	
0	0	0	-1	$\frac{1}{2}$	$\frac{1}{4}$...	
0	0	0	0	-1	$\frac{1}{2}$...	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	
row totals sum to “0”							

6 Why you should be careful when changing the order of integration

Define $f(x, y)$ (which is well behaved for all real values of x and y) by

$$f(x, y) = \frac{(x - y)}{((x - y)^2 + 1)^2}. \tag{8}$$

You should be able to satisfy yourself that

$$\int_{y=0}^{\infty} \int_{x=0}^{\infty} f(x, y) dx dy = +\frac{\pi}{4}$$

yet

$$\int_{x=0}^{\infty} \int_{y=0}^{\infty} f(x, y) dy dx = -\frac{\pi}{4}.$$

Furthermore, since the function $\exp(-x^2 - y^2)$ integrates over the same region to $\frac{\pi}{4}$ irrespective of which order of integration is used, the function $g(x, y) = f(x, y) + \exp(-x^2 - y^2)$ has therefore the property that:

$$\int_{y=0}^{\infty} \int_{x=0}^{\infty} g(x, y) dx dy = \frac{\pi}{2}$$

and

$$\int_{x=0}^{\infty} \int_{y=0}^{\infty} g(x, y) dy dx = 0.$$

For a different interesting example, try

$$h(x, y) = \begin{cases} e^{-(y-x)} & \text{if } y > x \\ -2e^{2(y-x)} & \text{otherwise,} \end{cases}$$

and compare

$$\begin{aligned} \int_{y=0}^{\infty} \int_{x=0}^{\infty} h(x, y) dx dy &= \int_{y=0}^{\infty} \left(\int_{x=0}^y h(x, y) dx + \int_{x=y}^{\infty} h(x, y) dx \right) dy \\ &= \int_{y=0}^{\infty} \left(\int_{x=0}^y e^{-(y-x)} dx + \int_{x=y}^{\infty} -2e^{2(y-x)} dx \right) dy \\ &= \int_{y=0}^{\infty} \left(\left[e^{-(y-x)} \right]_{x=0}^y + \left[e^{2(y-x)} \right]_{x=y}^{\infty} \right) dy \\ &= \int_{y=0}^{\infty} ((1 - e^{-y}) + (0 - 1)) dy \\ &= \int_{y=0}^{\infty} -e^{-y} dy \\ &= \left[e^{-y} \right]_{y=0}^{\infty} \\ &= 0 - 1 \\ &= -1 \end{aligned}$$

with

$$\begin{aligned} \int_{x=0}^{\infty} \int_{y=0}^{\infty} h(x, y) dy dx &= \int_{x=0}^{\infty} \left(\int_{y=0}^x h(x, y) dy + \int_{y=x}^{\infty} h(x, y) dy \right) dx \\ &= \int_{x=0}^{\infty} \left(\int_{y=0}^x -2e^{2(y-x)} dy + \int_{y=x}^{\infty} e^{-(y-x)} dy \right) dx \\ &= \int_{x=0}^{\infty} \left(\left[-e^{2(y-x)} \right]_{y=0}^x + \left[-e^{-(y-x)} \right]_{y=x}^{\infty} \right) dx \\ &= \int_{x=0}^{\infty} ((-1 - (-e^{-2x})) + (0 - (-1))) dx \\ &= \int_{x=0}^{\infty} e^{-2x} dx \\ &= \left[-\frac{1}{2} e^{-2x} \right]_{x=0}^{\infty} \\ &= 0 - \left(-\frac{1}{2}\right) \\ &= \frac{1}{2}. \end{aligned}$$

Determining the reason for the discrepancy is left as an exercise for the reader. [Hint: consider *convergence* versus *absolute convergence* for series.]

7 Second derivatives need not commute (trivial example)

If x and y are expressed in terms of r and θ as follows

$$x = r \cos \theta \tag{9}$$

$$y = r \sin \theta \tag{10}$$

then note that

$$\frac{\partial^2 y}{\partial x|_y \partial \theta|_r} \neq \frac{\partial^2 y}{\partial \theta|_r \partial x|_y} \tag{11}$$

(i.e. note that second derivatives need not commute) since

$$\frac{\partial^2 y}{\partial x|_y \partial \theta|_r} \equiv \frac{\partial}{\partial x|_y} \left(\frac{\partial}{\partial \theta|_r} (y) \right) \quad (12)$$

$$= \frac{\partial}{\partial x|_y} \left(\frac{\partial}{\partial \theta|_r} (r \sin \theta) \right) \quad (13)$$

$$= \frac{\partial}{\partial x|_y} (r \cos \theta) \quad (14)$$

$$= \frac{\partial}{\partial x|_y} (x) \quad (15)$$

$$= 1 \quad (16)$$

while

$$\frac{\partial^2 y}{\partial \theta|_r \partial x|_y} \equiv \frac{\partial}{\partial \theta|_r} \left(\frac{\partial}{\partial x|_y} (y) \right) \quad (17)$$

$$= \frac{\partial}{\partial \theta|_r} (0) \quad (18)$$

$$= 0. \quad (19)$$

8 Second derivatives need not commute (second sort of example)

In this example, we illustrate non-commuting where the the non-commutation comes only from ‘the thing being held constant’. This time we show that

$$\frac{\partial^2 x}{\partial r|_\theta \partial r|_y} \neq \frac{\partial^2 x}{\partial r|_y \partial r|_\theta}.$$

We prove the above by noting that:

$$\frac{\partial^2 x}{\partial r|_\theta \partial r|_y} = \frac{\partial}{\partial r|_\theta} \left(\frac{\partial}{\partial r|_y} (x) \right) \quad (20)$$

$$= \frac{\partial}{\partial r|_\theta} \left(\frac{\partial}{\partial r|_y} (\sqrt{r^2 - y^2}) \right) \quad (21)$$

$$= \frac{\partial}{\partial r|_\theta} \left(\frac{r}{\sqrt{r^2 - y^2}} \right) \quad (22)$$

$$= \frac{\partial}{\partial r|_\theta} \left(\frac{r}{x} \right) \quad (23)$$

$$= \frac{\partial}{\partial r|_\theta} \left(\frac{r}{r \cos \theta} \right) \quad (24)$$

$$= \frac{\partial}{\partial r|_\theta} \left(\frac{1}{\cos \theta} \right) \quad (25)$$

$$= 0 \quad (26)$$

while

$$\frac{\partial^2 x}{\partial r|_y \partial r|_\theta} = \frac{\partial}{\partial r|_y} \left(\frac{\partial}{\partial r|_\theta} (x) \right) \quad (27)$$

$$= \frac{\partial}{\partial r|_y} \left(\frac{\partial}{\partial r|_\theta} (r \cos \theta) \right) \quad (28)$$

$$= \frac{\partial}{\partial r|_y} (\cos \theta) \quad (29)$$

$$= \frac{\partial}{\partial r|_y} \left(\frac{r \cos \theta}{r} \right) \quad (30)$$

$$= \frac{\partial}{\partial r|_y} \left(\frac{x}{r} \right) \quad (31)$$

$$= \frac{\partial}{\partial r|_y} \left(\frac{\sqrt{r^2 - y^2}}{r} \right) \quad (32)$$

$$= \frac{r \left(\frac{r}{\sqrt{r^2 - y^2}} \right) - \sqrt{r^2 - y^2}}{r^2} \quad (33)$$

$$= \frac{\frac{r^2}{\sqrt{r^2 - y^2}} - \frac{r^2 - y^2}{\sqrt{r^2 - y^2}}}{r^2} \quad (34)$$

$$= \frac{y^2}{r^2 \sqrt{r^2 - y^2}} \quad (35)$$

$$= \frac{y^2}{r^2 x} \quad (36)$$

$$= \frac{\sin^2 \theta}{r \cos \theta}. \quad (37)$$

9 Second derivatives need not commute (totally different sort of example)

The everywhere-smooth and everywhere-continuous function $f(x, y)$ defined by

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{otherwise,} \end{cases}$$

has the following everywhere-continuous first derivatives:

$$\frac{\partial f}{\partial x|_y} = \begin{cases} \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial f}{\partial y|_x} = \begin{cases} \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{otherwise,} \end{cases}$$

and has the following second derivatives valid away from the origin:

$$\frac{\partial^2 f}{\partial x|_y \partial y|_x} = \frac{\partial^2 f}{\partial y|_x \partial x|_y} = A(x, y) \equiv \frac{(x^2 - y^2)(x^4 + 10x^2y^2 + y^4)}{(x^2 + y^2)^3}.$$

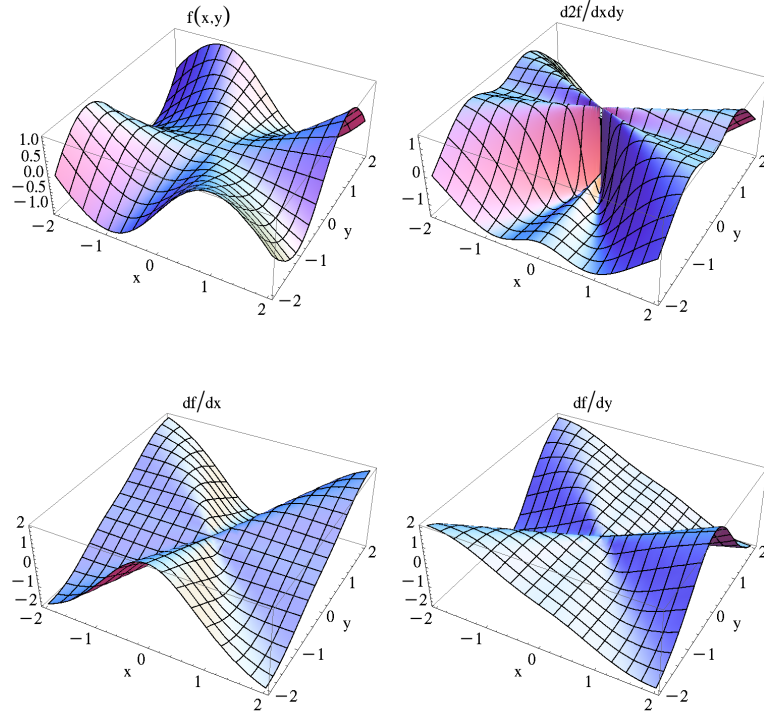
But $A(x, y)$ is itself not continuous at the origin because $A(x \neq 0, y = 0) = +1$ and $A(x = 0, y \neq 0) = -1$.

When calculating the second derivatives from first principles one finds that

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{(0,0)} = +1 \quad \text{and} \quad \left. \frac{\partial^2 f}{\partial y \partial x} \right|_{(0,0)} = -1.$$

Note that this means that the function g defined by $g = f + xy$ will satisfy:

$$\left. \frac{\partial^2 g}{\partial x \partial y} \right|_{(0,0)} = 2 \quad \text{and} \quad \left. \frac{\partial^2 g}{\partial y \partial x} \right|_{(0,0)} = 0.$$



10 Why you should not extrapolate based on the first few terms of a sequence

$$\begin{aligned}
 \int_{-\infty}^{\infty} \frac{\sin x}{x} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} \cdot \frac{\sin \frac{x}{7}}{\frac{x}{7}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} \cdot \frac{\sin \frac{x}{7}}{\frac{x}{7}} \cdot \frac{\sin \frac{x}{9}}{\frac{x}{9}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} \cdot \frac{\sin \frac{x}{7}}{\frac{x}{7}} \cdot \frac{\sin \frac{x}{9}}{\frac{x}{9}} \cdot \frac{\sin \frac{x}{11}}{\frac{x}{11}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} \cdot \frac{\sin \frac{x}{7}}{\frac{x}{7}} \cdot \frac{\sin \frac{x}{9}}{\frac{x}{9}} \cdot \frac{\sin \frac{x}{11}}{\frac{x}{11}} \cdot \frac{\sin \frac{x}{13}}{\frac{x}{13}} dx &= \pi \\
 \int_{-\infty}^{\infty} \frac{\sin x}{x} \cdot \frac{\sin \frac{x}{3}}{\frac{x}{3}} \cdot \frac{\sin \frac{x}{5}}{\frac{x}{5}} \cdot \frac{\sin \frac{x}{7}}{\frac{x}{7}} \cdot \frac{\sin \frac{x}{9}}{\frac{x}{9}} \cdot \frac{\sin \frac{x}{11}}{\frac{x}{11}} \cdot \frac{\sin \frac{x}{13}}{\frac{x}{13}} \cdot \frac{\sin \frac{x}{15}}{\frac{x}{15}} dx &= \frac{467807924713440738696537864469}{467807924720320453655260875000} \pi
 \end{aligned}$$

All equalities above are *exact*. Note that the last fraction is *approximately*

$$0.999999999852937186\pi.$$

11 Why you should be careful with Lagrange Multipliers

By inspection, we can see that for any finite value of the fixed parameter m , the extremum of the function

$$f(x, y, z) = x^2 + y^2 + z^2 + mx$$

subject to the constraint

$$x^2 + y^2 = 0$$

is just zero, since the constraint simply implies that $x = y = 0$ and so f reduces to $f_c = z^2$ which has a global minimum of zero at $x = y = z = 0$. Note that an “unthinking” application of the method of Lagrange multipliers fails: If you define

$$\mathcal{L} = x^2 + y^2 + z^2 + mx - \lambda(x^2 + y^2)$$

then the Euler-Lagrange equation corresponding to x is:

$$2x + m - 2\lambda x = 0$$

and it is clear that this equation is NOT satisfied by the desired solution $x = y = z = 0$ (unless we happen to be in the special case where $m = 0$).¹ This is due to the gradient of the constraint function being a null vector at the place where the constraint is satisfied. It is not clear to me how in general one avoids getting caught in this trap when the maths is more obscure (as it frequently is in real problems). Is it sufficient to simply take grad of any constraint and check for inequality with the null vector at all places satisfying the constraint? Is it acceptable to perturb the constraint from null - eg by replacing the constraint with $x^2 + y^2 = \epsilon^2$?

12 Cauchy Riemann Equations and Wirtinger Derivatives

12.1 Wirtinger derivatives

Physicists use $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \bar{z}}$ quite a lot, but often don't say how these seemingly innocent derivatives are defined. Often people seem to assume that they can be interpreted as ordinary partial derivatives in which “the other” variable is kept constant: viz $\frac{\partial}{\partial z}\Big|_{\bar{z}}$ and $\frac{\partial}{\partial \bar{z}}\Big|_z$. However, after some thought one notices that it is not possible to vary z while keeping \bar{z} constant, or to vary \bar{z} while keeping z constant. If one changes, then so does the other. Is this a problem? It need not be ... after all, if $f = x + y$ is considered to define f as a function of two variables x and y , and if we later discover that y itself is a function of x , such as $y = x^2$, this later knowledge (which tells us that x can't really be varied without varying y) doesn't actually prevent us still considering things like $\frac{\partial f}{\partial x}\Big|_y$. Such derivatives intentionally blind themselves to any functional dependence that y might have – asking us to look at f explicitly as a function of just x and y , regardless of the internal workings of either x or y .

Nonetheless, even though one *can* view $\frac{\partial}{\partial z}$ as being a sloppy notation for $\frac{\partial}{\partial z}\Big|_{\bar{z}}$, many people still find such an approach disturbing. Persons in that camp can sometimes find it helpful to see an alternative approach (Wirtinger) which is to **define** these derivatives as follows:

$$\frac{\partial}{\partial z} \equiv \frac{1}{2} \left(\frac{\partial}{\partial x}\Big|_y - i \frac{\partial}{\partial y}\Big|_x \right), \quad (38)$$

$$\frac{\partial}{\partial \bar{z}} \equiv \frac{1}{2} \left(\frac{\partial}{\partial x}\Big|_y + i \frac{\partial}{\partial y}\Big|_x \right) \quad (39)$$

acting on the space of functions on the complex plane $\{z \mid z = x + iy\}$. Among the useful consequences of this definition are:

- These operators really are allowed to be called “derivatives” (or “derivations” by mathematicians) as they satisfy the product rule: $\frac{\partial(fg)}{\partial z} = f \frac{\partial g}{\partial z} + \frac{\partial f}{\partial z} g$ and $\frac{\partial(fg)}{\partial \bar{z}} = f \frac{\partial g}{\partial \bar{z}} + \frac{\partial f}{\partial \bar{z}} g$,
- $\frac{\partial}{\partial \bar{z}}(z^n) = \frac{\partial}{\partial \bar{z}}((x + iy)^n) = \frac{1}{2}n(x + iy)^{n-1}(1 + i^2) = 0$ and thus $\frac{\partial f(z)}{\partial \bar{z}} = 0$ for functions that have a Taylor or Laurent Series in powers of z ,
- $\frac{\partial}{\partial z}(z^n) = \frac{\partial}{\partial z}((x + iy)^n) = \frac{1}{2}n(x + iy)^{n-1}(1 - i^2) = nz^{n-1}$

Taking the above properties together, one sees that one really can ‘treat’ the derivatives “ $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \bar{z}}$ ” almost ‘as if’ they really did mean “ $\frac{\partial}{\partial z}\Big|_{\bar{z}}$ and $\frac{\partial}{\partial \bar{z}}\Big|_z$ ”, even though strictly this notation is not self-consistent!

¹Note, however, that this argument presupposes λ is well defined. If the condition is perturbed by a small amount ϵ (see later) then λ may be seen to grow as an inverse power of ϵ – in effect saying that λ needs to go to infinity as the constraint becomes unperturbed.

12.2 Extension to fields

In text books on Quantum Field Theory one often sees similar methods being applied to allow a complex scalar field ϕ , its complex conjugate $\bar{\phi}$, its time derivative $\dot{\phi}$ and the time derivative of its complex conjugate $\dot{\bar{\phi}}$, to *appear* to be treated as independent quantities in lagrangians of the form $\mathcal{L}(\phi, \bar{\phi}, \dot{\phi}, \dot{\bar{\phi}})$. As before, that view point is not wrong, but if you wish to do so, you can alternatively take the view that the “actual” independent degrees of freedom are the fields $u(x, t)$, $\dot{u}(x, t)$, $v(x, t)$ and $\dot{v}(x, t)$ where $\phi = u + iv$, and where the Wirtinger derivatives are being defined as:

$$\frac{\partial}{\partial \phi} \equiv \frac{1}{2} \left(\frac{\partial}{\partial u} \Big|_{v, \dot{u}, \dot{v}} - i \frac{\partial}{\partial v} \Big|_{u, \dot{u}, \dot{v}} \right), \quad (40)$$

$$\frac{\partial}{\partial \bar{\phi}} \equiv \frac{1}{2} \left(\frac{\partial}{\partial u} \Big|_{v, \dot{u}, \dot{v}} + i \frac{\partial}{\partial v} \Big|_{u, \dot{u}, \dot{v}} \right), \quad (41)$$

$$\frac{\partial}{\partial \dot{\phi}} \equiv \frac{1}{2} \left(\frac{\partial}{\partial \dot{u}} \Big|_{\dot{v}, u, v} - i \frac{\partial}{\partial \dot{v}} \Big|_{\dot{u}, u, v} \right), \quad (42)$$

$$\frac{\partial}{\partial \dot{\bar{\phi}}} \equiv \frac{1}{2} \left(\frac{\partial}{\partial \dot{u}} \Big|_{\dot{v}, u, v} + i \frac{\partial}{\partial \dot{v}} \Big|_{\dot{u}, u, v} \right). \quad (43)$$

As before, these definitions may safely be used “as if” they were equivalent to derivatives at “constant” values of the other “independent” field components, vis $\frac{\partial}{\partial \phi} \equiv \frac{\partial}{\partial \phi} \Big|_{\bar{\phi}, \dot{\phi}, \dot{\bar{\phi}}}$, etc. Note that none of these definitions would be much use if things like the E-L equations did not retain a simple form in the new derivatives. Fortunately they do. Inverting the relations above, we see that

$$\frac{\partial \mathcal{L}}{\partial u} \Big|_{v, \dot{u}, \dot{v}} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}} \Big|_{\dot{v}, u, v} \right)$$

becomes

$$\frac{\partial \mathcal{L}}{\partial \phi} + \frac{\partial \mathcal{L}}{\partial \bar{\phi}} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} + \frac{\partial \mathcal{L}}{\partial \dot{\bar{\phi}}} \right) \quad (44)$$

and

$$\frac{\partial \mathcal{L}}{\partial v} \Big|_{u, \dot{u}, \dot{v}} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{v}} \Big|_{\dot{u}, u, v} \right)$$

becomes

$$i \left(\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial \mathcal{L}}{\partial \bar{\phi}} \right) = i \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\partial \mathcal{L}}{\partial \dot{\bar{\phi}}} \right) \quad (45)$$

and so taking (44) and (45) together we recover:

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) \quad \text{and} \quad (46)$$

$$\frac{\partial \mathcal{L}}{\partial \bar{\phi}} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\bar{\phi}}} \right). \quad (47)$$

Note that if \mathcal{L} is real, then these two equations are not independent (one being the complex conjugate of the other) and so in this case it is sufficient to use just one of them.

12.3 Cauchy Riemann

We are now able to derive the Cauchy Riemann equations for functions meeting the conditions above:

$$0 = \frac{\partial f(z)}{\partial \bar{z}} \quad (\text{provided } f(z) \text{ is of the appropriate type!}) \quad (48)$$

$$= \frac{\partial(u(x, y) + iv(x, y))}{\partial \bar{z}} \quad (49)$$

$$= \frac{1}{2} \left(\frac{\partial}{\partial x} \Big|_y + i \frac{\partial}{\partial y} \Big|_x \right) (u(x, y) + iv(x, y)) \quad (50)$$

$$= \frac{1}{2} \left(\frac{\partial u}{\partial x} \Big|_y + i \frac{\partial u}{\partial y} \Big|_x + i \frac{\partial v}{\partial x} \Big|_y + i^2 \frac{\partial v}{\partial y} \Big|_x \right) \quad (51)$$

$$= \frac{1}{2} \left(\frac{\partial u}{\partial x} \Big|_y - \frac{\partial v}{\partial y} \Big|_x \right) + \frac{i}{2} \left(\frac{\partial u}{\partial y} \Big|_x + \frac{\partial v}{\partial x} \Big|_y \right) \quad (52)$$

and therefore, since u , v , x and y are all real, we must have

$$\frac{\partial u}{\partial x} \Big|_y = + \frac{\partial v}{\partial y} \Big|_x$$

and

$$\frac{\partial u}{\partial y} \Big|_x = - \frac{\partial v}{\partial x} \Big|_y.$$

13 L'Hopital's rule

If $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = a$ and $[a = 0 \text{ or } a = \pm\infty]$ and $\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}$.

Your attention is drawn to the last precondition in the statement above. It is important, as the example $f(x) = x + \sin x$ and $g(x) = x$ with $x \rightarrow \infty$ illustrates.

14 $SU(3)$ multiplet multiplicities

If p and q are the number of gaps across two adjacent edges of the multiplet weight diagram, then it is called the (p, q) -multiplet and contains $\frac{1}{2}(p+1)(q+1)(p+q+2)$ states within it. N.B., this number of states can also be written as $\frac{1}{2}n_a n_b (n_a + n_b)$ where n_a and n_b are the number of states on each of two adjacent edges.

Multiplicity is not a good way of labelling multiplets, as different multiplets can have the same multiplicity. For example.

There are 15 states in both the $(2, 1)$ and the $(4, 0)$ multiplets.

There are 105 states in both the $(13, 0)$ and the $(6, 2)$ multiplets.

There are 120 states in each of the $(9, 1)$, $(14, 0)$ and $(5, 3)$ multiplets.

840 and 960 can also be realised in three different ways (find them!) and there are many other ways of realising multiplicities in two different ways. I know of no multiplicity that can be realised in four or more ways.

15 Playing the Cello (in equal temperament)

$$N = \left\lfloor \frac{8x}{a} + \frac{1}{2} \right\rfloor - 3 \quad (3 \text{ for the key of C major}) \quad (53)$$

$$n = N \bmod 7 \quad (0=\text{tonic}, 6=\text{sub-tonic}) \quad (54)$$

$$\sigma = \lfloor N/7 \rfloor \quad (\text{which octave we are in}) \quad (55)$$

$$f = (130.8 \text{ Hz} \times 2^\sigma) \{1, 2^{\frac{2}{12}}, 2^{\frac{4}{12}}, 2^{\frac{5}{12}}, 2^{\frac{7}{12}}, 2^{\frac{9}{12}}, 2^{\frac{11}{12}}\} [n] \quad (\text{note frequency}) \quad (56)$$

$$L = \frac{1}{2f} \sqrt{\frac{T}{\rho}} \quad (T=\text{tension}, \rho=\text{mass per unit length}) \quad (57)$$

16 Another way of doing one of the sums from the Fourier part of the course

To perform the sum

$$S = \sum_{n=1}^{\infty} \frac{1}{1+n^2}$$

one can try to create a function of a complex number with poles at the integers whose residues are equal to the terms being summed, so that the sum can be converted to a contour integral. [Thanks to Herschel Chawdhry for suggesting this approach.] We could start by noting that the function

$$f(z) = \frac{\pi \cot \pi z}{1+z^2}$$

has simple poles at $z = \pm i$ and $z = n$ for integer n . We will need to know the residues r_z of $f(z)$ at each of those poles. From the periodicity of \cot , the residues of $\cot \pi z$ at each of the positions $z = n$ will be identical. From the Taylor series:

$$\pi \cot \pi z = \pi \frac{1 + O((\pi z)^2)}{\pi z + O((\pi z)^3)} = \frac{1}{z} + O(1)$$

we see that the residue of $\pi \cot \pi z$ at the origin, and thus also at each of the positions $z = n$, is 1. The residues of $f(z)$ at these integer positions are therefore:

$$r_n = 1 \cdot \frac{1}{1+n^2} = \frac{1}{1+n^2}$$

which sum over all n to $2S + 1$. The residues at $z = \pm i$ are best seen by factorising $f(z)$

$$f(z) = (\pi \cot \pi z) \frac{1}{z-i} \frac{1}{z+i}$$

and thereby noting that

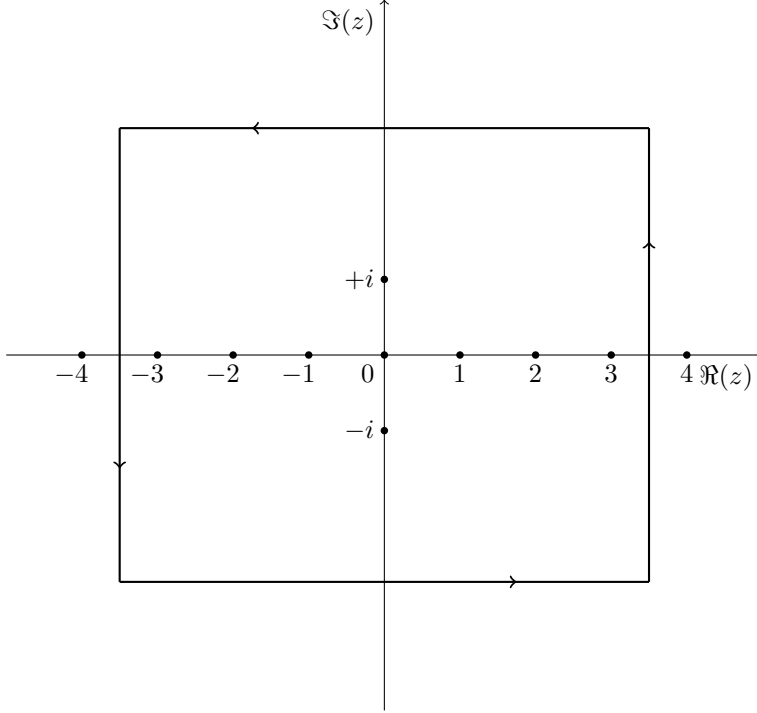
$$r_{\pm i} = \frac{\pi \cot \pi i}{2i} = -\frac{\pi \coth \pi}{2}.$$

We are now in a position to state by Cauchy's Theorem that:

$$\int_{\Gamma} f(z) dz = 2\pi i \{ (2S + 1) - (\pi \coth \pi) \}$$

where first quantity in round brackets is the sum of the poles on the real axis, the second is the sum of the residues of the other two poles, and Γ is an infinitely big version of the curve shown in the plot below.²

²Technically, we ought to use a sequence of ever larger but nonetheless finite curves, and therefore should work with finite sums associated with finite subsets of the poles on the real line. We will leave those details as an exercise for the reader.



The curve Γ has two vertical sections and two horizontal sections. The integrals on the two horizontal parts of Γ go to zero as H goes to infinity since

$$\cot(x + iy) = \frac{\sin 2x - i \sinh 2y}{\cosh 2y - \cos 2x}$$

and so

$$\lim_{y \rightarrow \pm\infty} \cot(x + iy) = \mp i.$$

As this is constant and bounded, the $\frac{1}{1+z^2}$ part of $f(x)$ will ensure that the integral along the top and bottom of Γ will go to zero as the box size increases if the aspect ratio of the box remains approximately unity. On the vertical parts of Γ the complex number z takes values where its real part is half-integer. Taking $z = n + \frac{1}{2} + iy$ yields $\cot \pi z = -i \tanh \pi y$ which is bounded between -1 and $+1$. As this also is bounded, the remaining $\frac{1}{1+z^2}$ part of $f(x)$ will ensure that the integral along the vertical part of Γ will go to zero as the box size increases if the aspect ratio of the box remains approximately one. Accordingly, we can see that as the curve Γ grows in size, the integral around it will tend to zero. In other words, we may state:

$$0 = 2\pi i \{(2S + 1) - (\pi \coth \pi)\}$$

which if solved for S gives

$$S = \sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{1}{2} (\pi \coth \pi - 1).$$

17 Inverses of Lorentz transformations in tensor/component form require sideways index movements.

If primed coordinates are related to unprimed coordinates via a Lorentz Transformation in the following way:

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \quad (58)$$

then

$$x'^{\mu} y'_{\nu} = x^{\mu} y_{\nu} \quad \forall x, y \quad (59)$$

and the true-by-definition identity:

$$\delta^{\mu}_{\nu} = (\Lambda^{-1})^{\mu}_{\sigma} \Lambda^{\sigma}_{\nu} \quad (60)$$

are together sufficient to gurantee that:

$$(\Lambda^{-1})^\mu{}_\nu = (\Lambda)_\nu{}^\mu \quad (61)$$

or equivalently that

$$(\Lambda^{-1})_\nu{}^\mu = (\Lambda)^\mu{}_\nu. \quad (62)$$

Another way of describing the last two results is that: **a Lorentz transformation written with indices can be inverted by sliding all its indices sideways** \leftrightarrow . A trivial consequence of the above result is that all/any of the following are also identities:

$$\delta_\nu^\mu = \Lambda_\sigma^\mu \Lambda_\nu^\sigma \quad (\text{note summation in **first** slot (row)}), \quad (63)$$

$$\delta_\nu^\mu = \Lambda^\mu{}_\sigma \Lambda_\nu^\sigma \quad (\text{note summation in **second** slot (column)}). \quad (64)$$

Students looking for matrix views of the Lorentz transformations are reminded always to take the left-most index as a row index and the right-most index as a column index, regardless of raised/lowered status and regardless of the index label name.

It is tempting to think of the ‘sideways index sliding’ as being a transposing operation (and in some sense it sort of is!). However, that idea is also dangerous/problematic/confusing since it is well kown that the inverse of this 1D Lorentz boost:

$$\Lambda = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \quad (65)$$

is this:

$$\Lambda^{-1} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \quad (66)$$

which is clearly not the same thing as Λ^T .

The reason for the conflict is that ‘sideways index sliding’ is not **just** transposing (i.e. interchanging rows and columns), it is transposing coupled to a pair of index raising and lowering operations. The latter are necessary because the tensor which is the end result doesn’t ‘fit into the old tensor’s body’.

More concretely: one could legitimately call the following operation ‘transposing’:

$$T_{\mu\nu} \rightarrow T_{\nu\mu} \quad (67)$$

since both start and end take the form $T_{\blacksquare\blacksquare}$. One could also legitimately call the following operation ‘transposing’

$$T^{\mu\nu} \rightarrow T^{\nu\mu} \quad (68)$$

since start and endpoints are both of the form $T^{\blacksquare\blacksquare}$. However: when we perform the operation

$$T_\nu^\mu \rightarrow T_\nu{}^\mu \quad (69)$$

we are moving from $T^{\blacksquare\blacksquare}$ to $T_{\blacksquare\blacksquare}$, in addition to changing labels μ and ν , and it is this extra raising and lowering that changes change the sign of β in the 1D example given in (66). This is why inverting a Lorentz transformation is about more than ‘just’ transposing.