Calculus III Part 2

Name: Solutions

1. (cf Cal3 exam part 1 problem 6) Normal $\langle 1,1,1\rangle$, so plane $x+y+z=1$

2. (cf Cal3 exam part 1 problem 7) We solve

$$\begin{cases} y^2-x^2=0 \\ y(x-y^2)=0 \end{cases}$$

   Critical points: $(0,0)$, $(1,\pm 1)$. Test

$$f_{xx}f_{yy}-f_{xy}^2=144(3xy^2-x^2-y^2)$$

$$f_{xx}=-12x$$

   So $(1,\pm 1)$ are local maxima, and $(0,0)$ is inconclusive.

3. (i) The integral is evaluated in polar

$$\int_0^{2\pi} d\theta \int_0^{\infty} e^{-r^2/2} r dr = 2\pi \int_0^{\infty} e^{-r^2/2} d(r^2/2) = 2\pi$$

   (ii) So

$$\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{\left(\int_{-\infty}^{\infty} e^{-x^2/2}dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2/2}dy\right)} = \sqrt{2\pi}$$

4. Fundamental theorem for line integrals says the path from point A to point B

$$\int_A^B \nabla f \cdot d\vec{r} = f(B)-f(A)$$

   Let's see why this makes sense.

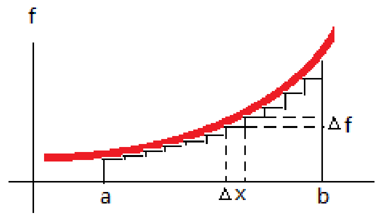
   Recall fundamental theorem of calculus of one variable is that

$$\int_a^b \left(\frac{df}{dx}\right) dx = f(b)-f(a)$$

   We can approximate the integral on the left by

$$\sum \left(\frac{df}{dx}\right) \Delta x$$

Since df/dx is the slope of the tangent, $\left(\frac{df}{dx}\right)\Delta x = \Delta f$, and adding all the Δf gives $f(b)-f(a)$. Graphically



For line integral, let's approximate the integral as

$$\int_A^B \nabla f \cdot d\vec{r} \approx \sum (\nabla f) \cdot \Delta\vec{r}$$

Recall in Cal3 exam part 1 problem 6, we said that

$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z = (\nabla f)\cdot\Delta\vec{r}$$

so adding all the Δf (they are all scalar now) gives

$$\int_A^B \nabla f \cdot d\vec{r} \approx \sum \Delta f = f(B) - f(A)$$

For our problem

$$\vec{F} = \nabla f, \quad f = \frac{x}{y^2+1} + e^{yz} + z^2$$

f is unique up to an additive constant, independent of x,y and z. So

$$\int \vec{F}\cdot d\vec{r} = f(1,0,2\pi) - f(1,0,0) = 4\pi^2$$

5. Stokes says

$$\oint_C \vec{F}\cdot d\vec{r} = \iint_A curl\,\vec{F}\cdot d\vec{S}$$

where A is any area whose boundary is C.

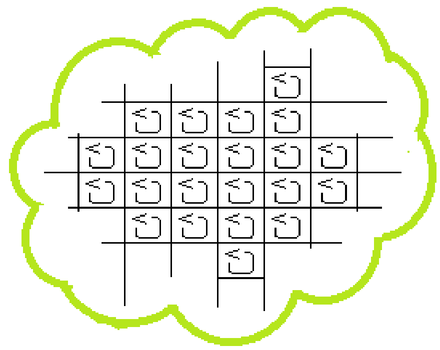
If C lies entirely on the xy plane, then A can be chosen to be on the xy plane, so $d\vec{S}$ is along $\hat{z}$ direction only, so only the $\hat{z}$ component of $curl\,\vec{F}$ matters. Therefore

$$\oint_C \vec{F}\cdot d\vec{r} = \iint_A \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS$$

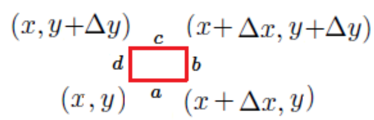
which is the Green's theorem. We don't have to memorize Green, for it is a special case of Stokes.

However Stokes is derived from Green, because assuming Green is correct, then for any given C and A, we just project them onto the xy plane, yz plane, and zx plane, we get 3 Green equations. Adding the 3 equations, the area integrand becomes curl of $\vec{F}$, and the 3 equations turn into dot product, which in turn gives Stokes.

Let's see why Green makes sense. Suppose C is on the xy plane, and let A be the area enclosed by C on xy plane and assume that we can chop S into many and many tiny rectangles with sides Δx, Δy.



Let us look at one of the tiny rectangles whose four corners are (x, y), $(x+\Delta x, y)$, $(x, y+\Delta y)$, and $(x + \Delta x, y + \Delta y)$.



Let us compute $\oint_{tiny\,\square} \vec{F} \cdot d\vec{r}$ for this tiny rectangle counterclockwise. Denote $\vec{F}_a$, $\vec{F}_b$, $\vec{F}_c$, $\vec{F}_d$ to be the vector $\vec{F}$ on the four sides (and F_{ax} means the x component of $\vec{F}_a$, etc.), then

$$\begin{aligned}\oint_{tiny\,\square} \vec{F} \cdot d\vec{r} &= F_{ax}\Delta x + F_{by}\Delta y - F_{cx}\Delta x - F_{dy}\Delta y \\ &= (F_{ax} - F_{cx})\Delta x + (F_{by} - F_{dy})\Delta y \\ &= -\frac{\partial F_x}{\partial y}\Delta y \Delta x + \frac{\partial F_y}{\partial x}\Delta x \Delta y \\ &= \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\Delta x \Delta y\end{aligned}$$

Now summing all tiny rectangles on S,

$$\sum \oint_{tiny\,\square} \vec{F} \cdot d\vec{r} = \sum \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\Delta x \Delta y$$

The left hand side becomes

$$\oint_C \vec{F} \cdot d\vec{r}$$

because all internal line integrals cancel, for internal edges of the rectangles are common edges of the two adjacent rectangles, and the line integrals are running in opposite directions. Only the boundary C survives; while the right hand side gives

$$\iint_A \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dS$$

showing Green is right.

For our problem, the integral is

$$\iint 2ydA = 2\int_{-1}^{1} dx \int_{1}^{3} ydy = 16$$

6. Question: Shall we use Stokes or Gauss's?

   - Let us see the Stokes' way.

Consider

$$\text{surface area of } A = \iint_A \hat{n} \cdot d\vec{S}$$

where A is the surface $x^2 + y^2 + z = 2$ above the xy plane, and $\hat{n}(x, y, z)$ is the unity vector whose direction is equal to $d\vec{S}$, i.e.

$$\hat{n} = \frac{\langle 2x, 2y, 1\rangle}{\sqrt{4x^2 + 4y^2 + 1}}$$

Now we put $curl\, \vec{F} = \hat{n}$ for some $\vec{F}$. We have

$$\text{surface area of } A = \iint_A curl\, \vec{F} \cdot d\vec{S} = \oint_C \vec{F} \cdot d\vec{r}$$

where $C = \{(x, y, z)|x^2 + y^2 = 2, z = 0\}$.

We know from Stokes

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_{A'} curl\, \vec{F} \cdot d\vec{S}$$

where A' is any area whose boundary is C. For convenience we choose $A' = \{(x, y, z)|x^2 + y^2 \leq 2, z = 0\}$. We find

$$\text{surface area of } A = \iint_{A'} \hat{n} \cdot d\vec{S} \quad = \quad \iint_{A'} \frac{1}{\sqrt{4x^2 + 4y^2 + 1}} dS = 2\pi \int_0^{\sqrt{2}} \frac{r}{\sqrt{4r^2 + 1}} dr = \pi$$

Everything looks good except this is *not* correct. It is not correct because there is no such $\vec{F}$ so that

$$curl\, \vec{F} = \hat{n} = \frac{\langle 2x, 2y, 1\rangle}{\sqrt{4x^2 + 4y^2 + 1}}$$

To see why, let's take divergence to both sides

$$\nabla \cdot (\nabla \times \vec{F}) = \nabla \cdot \hat{n}$$

Since divergence of curl is identically equal to zero, but

$$\nabla \cdot \hat{n} = 4\frac{2x^2+2y^2+1}{(4x^2+4y^2+1)^{3/2}} \neq 0$$

What can we do? Stokes is certainly not applicable. We should have checked that if

$$\nabla \cdot \hat{n} = 0$$

before we went on to do all the calculations. Fortunately as we'll see, most of the hard calculations above are recyclable in Gauss's way. We now switch to Gauss's way.

- Gauss's way

We take a closed surface

$$A \cup A'$$

where A is the surface $x^2+y^2+z=2$ above the xy plane, and $A' = \{(x,y,z)|x^2+y^2 \leq 2, z=0\}$, as defined earlier.

Applying Gauss to $A \cup A'$, we get

$$\oiint_{A\cup A'} \hat{n} \cdot d\vec{S} = \iiint_V (\nabla \cdot \hat{n}) dV$$

where $\hat{n}$ is also defined earlier.

Left hand side is

$$\text{surface area of } A + \iint_{A'} \hat{n} \cdot d\vec{S} = \text{surface area of } A - \pi$$

The minus sign in the front of π is due to the fact that we define outward normal flux to be positive as in Gauss'. (We will show in the next problem why this is a natural way to define the direction of a closed surface.)

Right hand side is computed in cylindrical coordinate

$$8\pi \int_0^{\sqrt{2}} \rho d\rho \int_0^{2-\rho^2} dz \frac{2\rho^2+1}{(4\rho^2+1)^{3/2}} = \frac{10\pi}{3}$$

So

$$\text{surface area of } A = \frac{13\pi}{3}$$

In summary

(a) We have seen that $\exists \vec{A}$ such that $\vec{B} = curl\,\vec{A}$ iff $div\,\vec{B} = 0$, then we can apply Stokes to the vector field $\vec{B}$. (Those who have taken general physics II may have recognized that the physical prototype here is that $\vec{B}$ is the magnetic field, and $\vec{A}$ is the vector potential. That $div\, \vec{B} = 0$ is equivalently to say there is no magnetic monopole (i.e. no free stand positive or negative magnetic charges) so magnetic field lines are always closed, i.e. curly looking.) Important to notice, closed vector lines (such as magnetic fields) have 0 divergence, however 0 divergent fields may not consist of closed lines, nor it contains any curly lines. cf. problem 9(viii)

(b) Here is a parallel statement to above, that $\exists$ a scalar function ϕ such that $\vec{E} = \nabla\phi$ iff $curl\, \vec{E} = 0$. (One may have recognized that the physical prototype for $\vec{E}$ is the electrostatic field, and ϕ is the electrostatic potential. So $\vec{E}$ is conservative, and electrostatic field lines are always coming out of positive charges and ending at negative charges, so if we put the negative charges at ∞, the field produced by a single positive charge looks likes diverging to ∞.)

(c) We now show the connection between the fundamental theorem for line integrals, Stokes' and Gauss'. Say starting from Gauss'

$$\oiint_A \vec{F} \cdot d\vec{S} = \iiint_V (\nabla \cdot \vec{F}) dV,$$

we change $\vec{F}$ to $\vec{F} + \vec{F'}$ for any arbitrary vector field $\vec{F'}$ such that $\nabla \cdot \vec{F'} = 0$. The right hand side of Gauss is clearly unchanged, but what about the left hand side? The LHS is now

$$\oiint_A \vec{F} \cdot d\vec{S} + \oiint_A \vec{F'} \cdot d\vec{S}$$

Because $\vec{F'} = curl\, \vec{A}$ for some vector $\vec{A}$, (by (a) above) Applying Stokes to the second term shows it is zero.

Physically this means that in computing the flux, adding some other flux $\vec{F'}$, which has $\nabla \cdot \vec{F'} = 0$ inside of V, makes no difference. Because that $\nabla \cdot \vec{F'} = 0$ inside of V means no sources or sinks that produce or terminate $\vec{F'}$ are inside of V, all fluxes that go into V must come out of V. (divergence of magnetic field is always zero everywhere means there is no magnetic charges.)

Recall in the solution to problem 6, we first calculated

$$\iint_{A'} \hat{n} \cdot d\vec{S} = \pi$$

in the Stokes way, but then we did the same calculation in Gauss way, we got

$$\iint_{A'} \hat{n} \cdot d\vec{S} = -\pi$$

This flip of sign is what makes Stokes' and Gauss' consistent to each other.

Now we write down Stokes

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_A curl\, \vec{F} \cdot d\vec{S}$$

and add $\vec{F'}$ to $\vec{F}$ for any arbitrary $\vec{F'}$ such that $curl\,\vec{F'} = 0$, then the RHS is clearly unchanged, while the LHS is also unchanged due to the fundamental theorem for line integrals, because $\vec{F'} = \nabla f$ for some scalar function f. (see (b) above) Physically this means that the additional work done by $\vec{F'}$

$$\oint_C \vec{F'} \cdot d\vec{r}$$

is zero, because $\vec{F'}$ is conservative.

[Extra Credits for those who know surface parametrization: calculate the surface area, using the following parametrization

$$\vec{\gamma}(u,v) = \left\langle u\cos v, u\sin v, 2-u^2 \right\rangle \quad u \in [0,\sqrt{2}],\; v \in [0,2\pi]$$

and compute

$$\vec{\gamma}_u,\; \vec{\gamma}_v,\; |\vec{\gamma}_u \times \vec{\gamma}_v|$$

Then compute

$$\text{surface area of } A = \int dv \int du |\vec{\gamma}_u \times \vec{\gamma}_v|$$

Should give the same answer, and much faster.]

7. Let's see why Gauss' makes sense. The discussion is almost parallel to Green's. Suppose V is the volume in Gauss' and assume that we can chop V into many and many tiny cubes with sides Δx, Δy, Δz.

   Let us look at one of the tiny cubes whose 8 vertices are (x,y,z), $(x+\Delta x, y, z)$, $(x, y+\Delta y, z)$, ... , and $(x+\Delta x, y+\Delta y, z+\Delta z)$.

   Let us compute $\oiint_{one\,tiny\,cube} \vec{F} \cdot d\vec{S}$ for this tiny cube. Denote $\vec{F_l}$, $\vec{F_r}$, $\vec{F_u}$, $\vec{F_d}$, $\vec{F_f}$, and $\vec{F_b}$ to be the vector $\vec{F}$ on the left, right, up, down, front, and back sides (and F_{lx} means the x component of $\vec{F_l}$), and our coordinate system is chosen such that $x+$ to the front, $y+$ to the right, and $z+$ up, and choose outward normal to be the positive flux direction, then

$$\begin{aligned}\oiint_{one\,tiny\,cube} \vec{F} \cdot d\vec{S} &= (F_{ry} - F_{ly})\Delta x \Delta z + (F_{fx} - F_{bx})\Delta y \Delta z + (F_{uz} - F_{dz})\Delta x \Delta y \\ &= (\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z})\Delta x \Delta y \Delta z\end{aligned}$$

   Notes:

   (a) The end result is celebrated divergence. This shows why we choose outward normal to be positive flux. If we chose mix direction on each side of cube, we would mess up. The same principle applies to Stokes, where right hand rule is the right choice.

   (b) In the calculation above we used one value (e.g. F_{lx}, etc) for each face. You may wonder what if we use better approximation such as $F_{lx}(x^*) + F'_{lx}(x^*)\Delta x + ...$ for each face, and do integration on each face. But the higher order corrections eventually give 4th order correction at the end. They are not compatible to our 3rd order correction, so we don't have to keep them. (cf Cal3 exam part 1 solution problem 7 b). The same logic applies to our "proof" of Stokes.

Now summing all tiny cubes in V,

$$\sum \oiint_{one\,tiny\,cube} \vec{F} \cdot d\vec{S} = \sum (\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z})\Delta x \Delta y \Delta z$$

The left hand side becomes

$$\oiint_A \vec{F} \cdot d\vec{S}$$

because all internal flux integrals cancel, for internal faces of the rectangles are common faces of the two adjacent cubes, and the flux integrals are running in opposite directions. Only the boundary A survives; while the right hand side gives

$$\iiint_V \left(\nabla \cdot \vec{F}\right) dV$$

showing Gauss is right.

Notes:

(a) In the "proof" above, to greatly simplify the left hand side, we use a quite peculiar argument that the internal flux integrals cancel, only boundary survives. Because of this argument, we can easily extend Gauss to region V not a solid, but has holes in it, and we just have to do additional flux integral evaluation on the boundary of the internal holes. The same principle applies to Stokes and Greens, where regions A, not necessary simple connected, with holes, are too allowed.

(b) However there are examples like Möbius strip and Klein bottle. We can chop the region into many and many tiny loops/cubes (the main idea of Calculus) and we are able to define the direction of positive line/flux integral for each little element, but what happens at the local level may not descent the global structure of the shape. Möbius strip does not work in Stokes, because it has no well-defined orientation. We know the functioning of right hand rule is crucial for the internal line integrals to cancel (see our "proof" above). For Klein bottle, we cannot distinguish the internal and external, i.e. the direction of outward normal positive flux is not well-defined, but we need this for the internal Gauss flux integrals to cancel.

(c) On the right hand side, the step where we turn $\sum(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z})\Delta x \Delta y \Delta z$ into is very common, but if we pause for a minute to think about it, we see that the step going from

$$\Delta x \Delta y \Delta z \to dV$$

is too very interesting. Because $\Delta x \Delta y \Delta z$, (as well as our "proof") refers to Cartesian coordinates, but dV has no explicit reference to any coordinates, namely we can evaluate

$$\iiint_V \left(\nabla \cdot \vec{F}\right) dV$$

in any coordinates we like. The same principle applies to Stokes.

For our problem, the integral is

$$4\pi \int_0^1 r^4 dr = \frac{4\pi}{5}$$

8. (i) vector field on x axis is $\langle 0, -x\rangle$, so it is (h);

   (ii) vector field on y axis is $\langle 0, 1\rangle$, so it is (b);

   (iii) vector field on y axis is $\langle 0, y\rangle$ and on x axis is $\langle x, 0\rangle$, so it must be one of (e), (d), or (g), but only (e) is symmetric with respect to x and y, so it is (e);

   (iv) Similar to above, for large $x > 0$ value the x component is much bigger than y component. So it is (d);

   (v) This is to change the y component of (iii) from positive to negative, so it is (f).

9. (i) No.

   (ii) Yes.

   (iii) Yes.

   (iv) Should read C is a simple closed "counterclockwise" curve... ANS: Yes, by Green.

   (v) Yes.

   (vi) Yes.

   (vii) Yes.

   Let us do (ix) before (viii).

   (ix) Yes, we can just compute

$$curl\, \vec{F} = curl\ \langle x, -y\rangle = 0$$

   Therefore by Green's, taking a closed loop of any size at anywhere, the line integral is zero, although it is no so clear just by looking.

   (viii) Yes. Can we compute $div\, \vec{F}$ now?

   We don't know the exact form of $\vec{F}$, but we know it has no y component, and the x component depends only on y, namely

$$\vec{F} = \langle P(y), 0\rangle$$

   so $div\, \vec{F} = 0$.

   There is an alternative way to solve this problem. In (ix), it is hard to tell line integral is zero by looking. However it is clear that here the flux integral is zero of closed surfaces of any sizes at anywhere. Let me be more clear what I mean by "closed surfaces" in this 2 dimensional context.

We know Green is the 2-dimensinal version of Stokes, analogously there is a 2-dimensional version of Gauss. Let's see what 2-dimensional version of Gauss is.

Consider a generic 2 dimensional vector, and we elevate it to 3 dimensional

$$\vec{F} = \langle P(x,y), Q(x,y), 0 \rangle$$

Now apply usual Gauss' to it, and choose V the volume in the integral to be a pillbox, whose top and bottom surfaces are parallel to the xy plane. So no flux on these two surfaces, and the flux on the side of the pillbox are constant in z.

For our problem, if we take such pillbox of any size at anywhere, then the flux integral is clearly zero. Hence divergence is 0.

(x) No, can be saddle points.

Counterexample: find max or min of

$$V = y - x^3$$

subject to

$$y = x^5$$

Lagrange gives critical points $x = \pm\sqrt{3/5}$, and $x = 0$. The first two are local max and min, and the last one is a saddle point.

Let's see why Lagrange makes sense.

Warning: you are headed six pages of long proof!! You are advised not to read it if you are pressed for time, since it is not required for Cal 3. However this proof is the most elementary version I can show you and use only the material up to Cal 3. (That is why it is so long.) If you wish, skip straight to the conclusion, last paragraph on the last page.

Let us consider one constraint problem

$$V = V(x,y)$$

subject to

$$f(x,y) = 0$$

Proof. We can parametrize the curve given by $f(x,y) = 0$ as

$$\vec{\gamma}(t) = \langle x(t), y(t) \rangle$$

then our task becomes to find t so that $V(t) = V(x(t), y(t))$ is optimal, hence

$$\frac{dV}{dt} = 0 \implies \frac{\partial V}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial t} = 0$$

Moreover taking t derivative of the equation $f(x, y) = 0$, we get

$$\frac{df}{dt} = 0 \implies \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} = 0$$

The two equations above look very much like a system of two linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0 \quad (1a) \\ a_{21}x_1 + a_{22}x_2 = 0 \quad (1b) \end{cases} \tag{1}$$

where $a_{11} = \frac{\partial V}{\partial x}$, $a_{12} = \frac{\partial V}{\partial y}$, $a_{21} = \frac{\partial f}{\partial x}$, $a_{22} = \frac{\partial f}{\partial y}$, $x_1 = \frac{\partial x}{\partial t}$, and $x_2 = \frac{\partial y}{\partial t}$. We can easily guess that one solution to the linear system is

$$x_1 = x_2 = 0$$

Unfortunately this is not what we want, because we have a lot of freedom in choosing parametrization. More specifically the optimal point is independent of what parametrization is chosen. Most of time we choose regular parametrization, namely

$$\vec{\gamma}'(t) = \left\langle \frac{\partial x}{\partial t}, \frac{\partial y}{\partial t} \right\rangle$$

is never $\vec{0}$. So we insist that there must be other solutions to the linear system (1a) & (1b).

We claim that there exists other solutions to the linear system (1a) & (1b), iff

$$\langle a_{11}, a_{12} \rangle = \lambda \langle a_{21}, a_{22} \rangle \tag{2}$$

for some constant λ (λ could be 0).

Let's prove equation (2).

Let us assume at least one of $a_{21,}$ a_{22} is not 0. (Because if both are 0, then the constraint $f(x, y)$ is just a constant function, which means there is no constraint. So to optimize V is of course to set

$$a_{11} = a_{12} = 0$$

then equation (2) is automatically satisfied. This slows why Lagrange multiplier is written as $\nabla V = \lambda \nabla f$, not $\lambda \nabla V = \nabla f$, i.e. $\nabla f = \vec{0} \implies \nabla V = \vec{0}$, not the another way around.)

Let's assume $a_{21} \neq 0$. (If in fact $a_{21} = 0$ and $a_{22} \neq 0$, we can exchange the role of $a_{21} \leftrightarrow a_{22}$, $x_1 \leftrightarrow x_2$, and $a_{11} \leftrightarrow a_{12}$, so the following analysis will work exactly the same and equation (2) will too be the same.)

Now replace (1a) by a new equation obtained by multiplying (1b) by a_{11}/a_{21} then subtract it to (1a)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + a_{22}x_2 = 0 \end{cases} \rightarrow \begin{cases} 0 + a'_{12}x_2 = 0 & (3a) \\ a_{21}x_1 + a_{22}x_2 = 0 & (3b) \end{cases} \tag{3}$$

where $a'_{12} = a_{12} - a_{11}a_{22}/a_{21}$.

From (3a), if $a'_{12} \neq 0$, then $x_2 = 0$, then by (3b) since $a_{21} \neq 0$, $x_1 = 0$. This shows $a'_{12} = 0$, hence

$$a_{12} - a_{11}a_{22}/a_{21} = 0$$

This shows if $a_{11} = \lambda a_{21,}$ then $a_{12} = \lambda a_{22}$, proving equation (2). □

Let us consider two constraint problem of three variables

$$V = V(x, y, z)$$

subject to

$$f(x, y, z) = 0, \; g(x, y, z) = 0$$

Proof. We can parametrize the interaction of $f(x, y, z) = 0, \; g(x, y, z) = 0$ as

$$\vec{\gamma}(t) = (x(t), y(t), z(t))$$

(this requires that the interaction is a curve, not a surface. This in turn requires that ∇f is not parallel to ∇g. This includes the case that neither $\nabla f = \vec{0}$ nor $\nabla g = \vec{0}$.)

Then our task becomes to find t so that $V(t) = V(x(t), y(t), z(t))$ is optimal, hence

$$\frac{dV}{dt} = 0 \implies \frac{\partial V}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial V}{\partial z}\frac{\partial z}{\partial t} = 0$$

Moreover taking t derivative of $f(x, y, z) = 0, \; g(x, y, z) = 0$, we get two more equations

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} = 0, \quad \frac{dg}{dt} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial t} = 0$$

We can write above as system of three linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 & (4a) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 & (4b) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0 & (4c) \end{cases} \tag{4}$$

where $a_{11} = \frac{\partial V}{\partial x}$, $a_{21} = \frac{\partial f}{\partial x}$, $a_{31} = \frac{\partial g}{\partial x}$, $x_1 = \frac{\partial x}{\partial t}$, etc. Because we choose regular curves, there must be other solutions to the linear system. We will show this implies

$$\langle a_{11}, a_{12}, a_{13}\rangle = \lambda\,\langle a_{21}, a_{22}, a_{23}\rangle + \mu\,\langle a_{31}, a_{32}, a_{33}\rangle \tag{5}$$

for some constants λ, μ (either or both could be zero).

Let's see why.

Assume $a_{21} \neq 0$ (because $\nabla f \neq \vec{0}$ and likewise we can exchange the role of $a_{21} \leftrightarrow a_{22}$, $x_1 \leftrightarrow x_2$, $a_{11} \leftrightarrow a_{12}$, $a_{31} \leftrightarrow a_{32}$ if $a_{21} = 0$ and $a_{22} \neq 0$.) Then replace (4a) by a new equation obtained by multiplying (4b) by a_{11}/a_{21} then subtract it to (4a), and replace (4c) by a new equation obtained by multiplying (4b) by a_{31}/a_{21} then subtract it to (4c)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0 \end{cases} \to \begin{cases} 0 + a'_{12}x_2 + a'_{13}x_3 = 0 & (6a) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 & (6b) \\ 0 + a'_{32}x_2 + a'_{33}x_3 = 0 & (6c) \end{cases} \tag{6}$$

where $a'_{12} = a_{12} - a_{11}a_{22}/a_{21}$, $a'_{32} = a_{32} - a_{11}a_{22}/a_{21}$, etc.

Pay attention to (6a) (6c)

$$\begin{cases} a'_{12}x_2 + a'_{13}x_3 = 0 & (6a) \\ a'_{32}x_2 + a'_{33}x_3 = 0 & (6c) \end{cases}$$

Clearly at least one of a'_{32} and a'_{33} is not 0, because $\nabla f \nparallel \nabla g$. We also know $x_2 = x_3 = 0$ should not be the only solution to the linear system (6a) & (6c), because if $x_2 = x_3 = 0$ is the only solution, then by (6b) since $a_{21} \neq 0$, $x_1 = 0$, so $x_1 = x_2 = x_3 = 0$ is the only solution. That is not good. Therefore the conditions for the linear system (6a) & (6c) agree exactly to the hypotheses of the linear system (1a) & (1b), hence

$$\langle a'_{12}, a'_{13}\rangle = \nu\,\langle a'_{32}, a'_{33}\rangle \tag{7}$$

for some constant ν.

Vectorially if we denote the 1st row $\vec{R}_1 = \langle a_{11}, a_{12}, a_{13}\rangle$ and similarly denote $\vec{R}_2$, $\vec{R}_3$, equation (7) says

$$\vec{R}_1 - \frac{a_{11}}{a_{21}}\vec{R}_2 = \nu\left(\vec{R}_3 - \frac{a_{31}}{a_{21}}\vec{R}_2\right)$$

Or

$$\vec{R}_1 = \left(\frac{a_{11}}{a_{21}} - \nu\frac{a_{31}}{a_{21}}\right)\vec{R}_2 + \nu\vec{R}_3$$

showing equation (5) is true with $\lambda = \frac{a_{11}}{a_{21}} - \nu\frac{a_{31}}{a_{21}}$, and $\mu = \nu$. □

Use the same idea, one can show Lagrange multiplier for any number of constraints as long as the constraints are linearly independent and the intersection of the constraints is a curve. For example

$$V = V(x, y, z, w)$$

subject to

$$f(x, y, z, w) = 0,\ g(x, y, z, w) = 0,\ h(x, y, z, w) = 0$$

then Lagrange multiplier is

$$\nabla V = \lambda \nabla f + \mu \nabla g + \nu \nabla h$$

To show this, first write down a system of 4 linear equations with 4 unknowns. Then reduce the problem to 3 equations with 3 unknowns, and check the conditions match the hypothesis of 3 equation with 3 unknowns of the previous problem.

Lastly let us consider problem of three variables

$$V = V(x, y, z)$$

subject to one constraint

$$f(x, y, z) = 0$$

Hence the constraint is not a curve but a surface.

Proof. We parametrize the surface of $f(x, y, z) = 0$ by

$$\vec{S}(u, v) = \langle x(u, v), y(u, v), z(u, v) \rangle$$

Then the optimal V is reached when

$$\frac{\partial V}{\partial u} = 0, \quad \frac{\partial V}{\partial v} = 0$$

Take partial u and v of equation $f(x, y, z) = 0$,

$$\frac{\partial f}{\partial u} = 0, \quad \frac{\partial f}{\partial v} = 0$$

Hence we obtain 4 linear equations. We can view them as a system of 2 linear equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 & (8a) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 & (8b) \end{cases} \tag{8}$$

where $a_{11} = \frac{\partial V}{\partial x}$, $a_{12} = \frac{\partial V}{\partial y}$, $a_{13} = \frac{\partial V}{\partial z}$, $a_{21} = \frac{\partial f}{\partial x}$, $a_{22} = \frac{\partial f}{\partial y}$, $a_{23} = \frac{\partial f}{\partial z}$, $x_1 = \frac{\partial x}{\partial u}$ or $\frac{\partial x}{\partial v}$, $x_2 = \frac{\partial y}{\partial u}$ or $\frac{\partial y}{\partial v}$, $x_3 = \frac{\partial z}{\partial u}$ or $\frac{\partial z}{\partial v}$.

Recall surface parametrization is regular, i.e. at all time

$$\frac{\partial \vec{S}}{\partial u} = \left\langle \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right\rangle \neq 0, \quad \frac{\partial \vec{S}}{\partial v} = \left\langle \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right\rangle \neq 0$$

and

$$\left\langle \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right\rangle \nparallel \left\langle \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right\rangle$$

namely the Jacobian of the surface parametrization

$$\left| \frac{\partial \vec{S}}{\partial u} \times \frac{\partial \vec{S}}{\partial v} \right|$$

is never zero. (This is similar to the definition of regular curve parametrization, i.e. the Jacobian of the curve parametrization $\left|\frac{d\vec{\gamma}}{dt}\right|$ is never 0.) □

Therefore we require system of equations (8a) & (8b) to have two sets of non-$\vec{0}$ solutions, and the two sets are not parallel to each other. We will show this implies

$$\langle a_{11}, a_{12}, a_{13} \rangle = \lambda \langle a_{21}, a_{22}, a_{23} \rangle \tag{9}$$

Let us see why.

Assume $a_{21} \neq 0$. (because if $\nabla f = \vec{0}$, equation (9) is automatically true. So assume $\nabla f \neq \vec{0}$, and likewise exchange indices if necessary), then replace (8a) by a new equation obtained by multiplying (8b) by a_{11}/a_{21} then subtract it to (8a)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 \end{cases} \rightarrow \begin{cases} 0 + a'_{12}x_2 + a'_{13}x_3 = 0 & (10a) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 & (10b) \end{cases} \tag{10}$$

where $a'_{12} = a_{12} - a_{11}a_{22}/a_{21}$, $a'_{13} = a_{13} - a_{11}a_{23}/a_{21}$.

We will show both a'_{12}, $a'_{13} = 0$. Suppose one of them is not 0, say $a'_{12} \neq 0$.

Then

$$x_2 = -\frac{a'_{13}}{a'_{12}}x_3, \quad x_1 = -\frac{a_{22}x_2 + a_{23}x_3}{a_{21}} = -\frac{-\frac{a'_{13}}{a'_{12}} + a_{23}}{a_{21}}x_3$$

Hence if (x_1, x_2, x_3) and (y_1, y_2, y_3) are two sets of non-$\vec{0}$ solutions (i.e. $x_3 \neq 0$, $y_3 \neq 0$), then

$$(x_1, x_2, x_3) = \nu(y_1, y_2, y_3)$$

where $\nu = x_3/y_3$. This violates the requirement that the two solutions are not parallel to each other. Therefore

$$a'_{12} = a'_{13} = 0$$

Hence equation (9) is true.

Use the same idea, one can show Lagrange multiplier for any number of constraints whose intersection is a surface. For example

$$V = V(x, y, z, w)$$

subject to

$$f(x, y, z, w) = 0,\ g(x, y, z, w) = 0$$

then Lagrange multiplier is

$$\nabla V = \lambda \nabla f + \mu \nabla g$$

To show this, first write down a system of 3 linear equations with 4 unknowns. Then reduce the problem to 2 equations with 3 unknowns, and check the conditions match the hypothesis of 2 equation with 3 unknowns of the previews problem.

In summary we have shown the formula of Lagrange multiplier

$$\nabla V = \lambda \nabla f + \mu \nabla g + \nu \nabla h + ... \tag{11}$$

of any number of linearly independent constraints as long as the intersection of the constraints is a regular curve, regular surface, or regular whatever. As one can see from the derivation of Lagrange multiplier, only first derivatives are considered, so Lagrange will give local max/min as well as saddle points. See counterexample on page 10.