Chapter 8

on manifolds and multipliers

In this chapter we show the application of the most difficult results in this course, namely the implicit and inverse mapping theorems. Our first application is in the construction of manifolds as graphs or level sets. Then once we have a convenient concept of a manifold we discuss the idea of Lagrange multipliers. The heart of the method combines orthogonal complements from linear algebra along side the construction of tangent spaces in this course. Hopefully this chapter will help you understand why the implicit and inverse mapping theorems are so useful and also why we need manifolds to make sense of our problems. The patching definition for a manifold is not of much use in this chapter although we will mention how it connects to the other two formulations of a manifold in $\mathbb{R}^m$ in the context of a special case.

8.1 surfaces in $\mathbb{R}^3$

Manifolds or surfaces play a role similar to functions in this course. Our goal is not the study of manifolds alone but it's hard to give a complete account of differentiation unless we have some idea of what is a tangent plane. This subsection does break from the larger pattern of thought in this chapter. I include it here to try to remind how surfaces and tangent planes are described in $\mathbb{R}^3$. We need some amount of generalization beyond this section because the solution of max/min problems with constraints will take us into higher dimensional surfaces even for problems that only involve two or three spatial dimensions. We treat those questions in the next chapter.

There are three main methods to describe surfaces:

1. As a **graph**: $S = \{(x, y, z) \mid z = f(x, y) \text{ where } (x, y) \in dom(f)\}$.
2. As a **level surface**: $S = \{(x, y, z) \mid F(x, y, z) = 0\}$
3. As a **parametrized surface**: $S = \{X(u, v) \mid (u, v) \in dom(X)\}$

Let me remind you we found the tangent plane at $(x_o, y_o, z_o) \in S$ for each of these formalisms as follows (continuing to use the same notation as above):

1. For the **graph**: $z = z_o + f(x_o, y_o) + f_x(x_o, y_o)(x - x_o) + f_y(x_o, y_o)(y - y_o)$.
2. For the **level surface**: plane through (x_o, y_o, z_o) with normal $(\nabla F)(x_o, y_o, z_o)$
3. For the **parametrized surface**: find (u_o, v_o) with $X(u_o, v_o) = (x_o, y_o, z_o)$, the tangent plane goes through $X(u_o, v_o)$ and has normal $N(u_o, v_o) = X_u(u_o, v_o) \times X_v(u_o, v_o)$.

Perhaps you recall that the normal vector field to the surface S was important in the formulation of surface integrals to calculate the flux of vector fields.

Example 8.1.1. *The plane through the point $\vec{r}_o$ with normal $\vec{n} =< a, b, c >$ can be described as:*

1. *all $\vec{r} \in \mathbb{R}^3$ such that $(\vec{r} - \vec{r}_o) \cdot \vec{n} = 0$.*
2. *all $(x, y, z) \in \mathbb{R}^3$ such that $a(x - x_o) + b(y - y_o) + c(z - z_o) = 0$*
3. *if $c \neq 0$, the graph $z = f_3(x, y)$ where $f_3(x, y) = z_o + \frac{a}{c}(x - x_o) + \frac{b}{c}(y - y_o)$*
4. *if $b \neq 0$, the graph $y = f_3(x, z)$ where $f_2(x, z) = y_o + \frac{a}{b}(x - x_o) + \frac{c}{b}(z - z_o)$*
5. *if $a \neq 0$, the graph $x = f_1(y, z)$ where $f_1(y, z) = x_o + \frac{b}{a}(y - y_o) + \frac{c}{a}(z - z_o)$*
6. *given any two linearly independent vectors $\vec{a}, \vec{b}$ in the plane, the plane is the image of the mapping $X : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $X(u, v) = \vec{r}_o + u\vec{a} + v\vec{b}$*

Example 8.1.2. *The sphere of radius R centered about the origin can be described as:*

1. *all $(x, y, z) \in \mathbb{R}^3$ such that $F(x, y, z) = x^2 + y^2 + z^2 = R^2$*
2. *the graphs of $z = f_\pm(x, y)$ where $f_\pm(x, y) = \pm\sqrt{R^2 - x^2 - y^2}$*
3. *for $(u, v) \in [0, 2\pi] \times [0, \pi]$, $X(u, v) = (R \cos u \sin v, R \sin u \sin v, R \cos v)$*

You may recall that the level surface concept allowed by far the easiest computation of the normal of the tangent plane for a particular point. For example, $\nabla F =< 2x, 2y, 2z >$ in the preceding example. Contrast that to calculation of $X_u \times X_v$ where the $\times$ denotes the dreaded cross-product. Of course each formalism has its place in calculus III.

Remark 8.1.3.

In this warm-up section we have hopefully observed this much about surfaces in $\mathbb{R}^3$:

1. the tangent plane is always 2-dimensional, it is really a plane in the traditional sense of the term.
2. the normal to the tangent plane is always 1-dimensional, the normal through a particular point on the surface is just a line which is orthogonal to all possible tangents through the point.
3. the dimension of the tangent plane and normal give the total dimension of the ambient space; $2 + 1 = 3$.

8.2 manifolds as level sets

We will focus almost exclusively on the level surface formulation of a manifold in the remainder of this chapter. We say $M \subseteq \mathbb{R}^n$ is a **manifold** of dimension $p \leq n$ if M has a p-dimensional tangent plane for each point on M. In other words, M is a p-dimensional manifold if it can be locally approximated by $\mathbb{R}^p$ at each point on M. Moreover, the set of all vectors normal to the tangent space will be $n - p$ dimensional.

These are general concepts which encompasses lines, planes volumes and much much more. Let me illustrate by example:

Example 8.2.1. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be defined by $g(x, y) = y - x - 1$ note that $g(x, y) = 0$ gives the line $y - x - 1 = 0$ commonly written as $y = x + 1$; note that the line has direction vector $< -1, 1 >$. Furthermore, $\nabla g =< 1, -1 >$ which is orthogonal to $< -1, 1 >$.*

Example 8.2.2. *Let $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $g(x, y, z) = y - x - 1$ note that $g(x, y, z) = 0$ gives the plane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0 >$ which gives the normal to the plane $g = 0$.*

Example 8.2.3. *Let $g : \mathbb{R}^4 \to \mathbb{R}$ be defined by $g(x, y, z, t) = y - x - 1$ note that $g(x, y, z, t) = 0$ gives the hyperplane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0, 0 >$ which gives the normal to the hyperplane $g = 0$. What does that mean? It means that if I take any vector in the hyperplane it is orthogonal to $< 1, -1, 0, 0 >$. Let $\vec{r}_1, \vec{r}_2$ be points in the solution set of $g(x, y, z, t) = 0$. Denote $\vec{r}_1 = (x_1, y_1, z_1, t_1)$ and $\vec{r}_1 = (x_2, y_2, z_2, t_2)$, we have $y_1 = x_1 + 1$ and $y_2 = x_2 + 1$. The vector in the hyperplane is found from the difference of these points:*

$$\vec{v} = \vec{r}_2 - \vec{r}_1 = (x_2, x_2 + 1, z_2, t_2) - (x_1, x_1 + 1, z_1, t_1) = (x_2 - x_1, x_2 - x_1, z_2 - z_1, t_2 - t_1).$$

It's easy to see that $\vec{v} \cdot \nabla g = 0$ hence ∇g is perpendicular to an arbitrary vector in the hyperplane

If you've begun to develop an intuition for the story we're telling this last example ought to bug you a bit. Why is the difference of points a tangent vector? What happened to the set of all tangent vectors pasted together or the differential or the column space of the derivative? All those concepts still apply but since we were looking at a linear space the space itself matched the tangent hyperplane. The point of the triple of examples above is just to constrast the nature of the equation $g = 0$ in various contexts. We find the dimension of the ambient space changes the dimension of the level set. Basically, we have one equation $g = 0$ and n-unknowns then the inverse image of zero gives us a $(n-1)$-dimensional manifold. If we wanted to obtain a $n-2$ dimensional manifold then we would need two equations which were independent. Before we get to that perhaps I should give a curvy example.

Example 8.2.4. *Let $g : \mathbb{R}^4 \to \mathbb{R}$ be defined by $g(x,y,z,t) = t+x^2+y^2-2z^2$ note that $g(x,y,z,t) = 0$ gives a three dimensional subset of $\mathbb{R}^4$, let's call it M. Notice $\nabla g = < 2x, 2y, -4z, 1 >$ is nonzero everywhere. Let's focus on the point $(2,2,1,0)$ note that $g(2,2,1,0) = 0$ thus the point is on M. The tangent plane at $(2,2,1,0)$ is formed from the union of all tangent vectors to $g = 0$ at the point $(2,2,1,0)$. To find the equation of the tangent plane we suppose $\gamma : \mathbb{R} \to M$ is a curve with $\gamma' \neq 0$ and $\gamma(0) = (2,2,1,0)$. By assumption $g(\gamma(s)) = 0$ since $\gamma(s) \in M$ for all $s \in \mathbb{R}$. Define $\gamma'(0) = < a, b, c, d >$, we find a condition from the chain-rule applied to $g \circ \gamma = 0$ at $s = 0$,*

$$\begin{aligned} \frac{d}{ds}\big(\ g\circ\gamma(s)\ \big) = \big(\nabla g\big)(\gamma(s))\cdot\gamma'(s) = 0 \quad &\Rightarrow \quad \nabla g(2,2,1,0)\cdot < a,b,c,d >= 0 \\ &\Rightarrow \quad < 4,4,-4,1 > \cdot < a,b,c,d >= 0 \\ &\Rightarrow \quad 4a+4b-4c+d = 0 \end{aligned}$$

Thus the equation of the tangent plane is $4(x-2)+4(y-2)-4(z-1)+t = 0$. In invite the reader to find a vector in the tangent plane and check it is orthogonal to $\nabla g(2,2,1,0)$. However, this should not be surprising, the condition the chain rule just gave us is just the statement that $< a,b,c,d > \in Null(\nabla g(2,2,1,0)^T)$ and that is precisely the set of vector orthogonal to $\nabla g(2,2,1,0)$.

One more example before we dive into the theory of Lagrange multipliers. (which is little more than this section applied to word problems plus the powerful orthogonal complement theorem from linear algebra)

Example 8.2.5. *Let $G : \mathbb{R}^4 \to \mathbb{R}^2$ be defined by $G(x,y,z,t) = (z+x^2+y^2-2, z+y^2+t^2-2)$. In this case $G(x,y,z,t) = (0,0)$ gives a two-dimensional manifold in $\mathbb{R}^4$ let's call it M. Notice that $G_1 = 0$ gives $z+x^2+y^2 = 2$ and $G_2 = 0$ gives $z+y^2+t^2 = 2$ thus $G = 0$ gives the intersection of both of these three dimensional manifolds in $\mathbb{R}^4$ (no I can't "see" it either). Note,*

$$\nabla G_1 = < 2x, 2y, 1, 0 > \qquad \nabla G_2 = < 0, 2y, 1, 2t >$$

It turns out that the inverse mapping theorem says $G = 0$ describes a manifold of dimension 2 if the gradient vectors above form a linearly independent set of vectors. For the example considered here the gradient vectors are linearly dependent at the origin since $\nabla G_1(0) = \nabla G_2(0) = (0,0,1,0)$.

In fact, these gradient vectors are colinear along along the plane $x = t = 0$ *since* $\nabla G_1(0, y, z, 0) = \nabla G_2(0, y, z, 0) =< 0, 2y, 1, 0 >$. *We again seek to contrast the tangent plane and its normal at some particular point. Choose* $(1, 1, 0, 1)$ *which is in* M *since* $G(1, 1, 0, 1) = (0 + 1 + 1 - 2, 0 + 1 + 1 - 2) = (0, 0)$. *Suppose that* $\gamma : \mathbb{R} \to M$ *is a path in* M *which has* $\gamma(0) = (1, 1, 0, 1)$ *whereas* $\gamma'(0) =< a, b, c, d >$. *Note that* $\nabla G_1(1, 1, 0, 1) =< 2, 2, 1, 0 >$ *and* $\nabla G_2(1, 1, 0, 1) =< 0, 2, 1, 1 >$. *Applying the chain rule to both* G_1 *and* G_2 *yields:*

$$\begin{array}{lcl} (G_1 \circ \gamma)'(0) = \nabla G_1(\gamma(0)) \cdot < a, b, c, d >= 0 & \Rightarrow & < 2, 2, 1, 0 > \cdot < a, b, c, d >= 0 \\ (G_2 \circ \gamma)'(0) = \nabla G_2(\gamma(0)) \cdot < a, b, c, d >= 0 & \Rightarrow & < 0, 2, 1, 1 > \cdot < a, b, c, d >= 0 \end{array}$$

This is two equations and four unknowns, we can solve it and write the vector in terms of two free variables correspondant to the fact the tangent space is two-dimensional. Perhaps it's easier to use matrix techiques to organize the calculation:

$$\begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We calculate, $rref \begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 1 & 1/2 & 1/2 \end{bmatrix}$. *It's natural to chose* c, d *as free variables then we can read that* $a = d/2$ *and* $b = -c/2 - d/2$ *hence*

$$< a, b, c, d >=< d/2, -c/2 - d/2, c, d >= \tfrac{c}{2} < 0, -1, 2, 0 > + \tfrac{d}{2} < 1, -1, 0, 2 >$$

We can see a basis for the tangent space. In fact, I can give parametric equations for the tangent space as follows:

$$X(u, v) = (1, 1, 0, 1) + u < 0, -1, 2, 0 > + v < 1, -1, 0, 2 >$$

Not surprisingly the basis vectors of the tangent space are perpendicular to the gradient vectors $\nabla G_1(1, 1, 0, 1) =< 2, 2, 1, 0 >$ *and* $\nabla G_2(1, 1, 0, 1) =< 0, 2, 1, 1 >$ *which span the* **normal plane** N_p *to the tangent plane* T_p *at* $p = (1, 1, 0, 1)$. *We find that* T_p *is orthogonal to* N_p. *In summary* $T_p^{\perp} = N_p$ *and* $T_p \oplus N_p = \mathbb{R}^4$. *This is just a fancy way of saying that the normal and the tangent plane only intersect at zero and they together span the entire ambient space.*

Remark 8.2.6.

The reason I am bothering with these seemingly bizarre examples is that the method of Lagrange multipliers comes down to the observation that both the constraint and objective function's gradient vectors should be normal to the tangent plane of the constraint surface. This means they must both reside in the normal to the tangent plane and hence they will either be colinear or for several constraints they will be linearly dependent. The geometry we consider here justifies the method. Linear algebra supplies the harder part which is that if two vectors are both orthogonal to the tangent plane then they must both be in the orthogonal complement to the tangent plane. The heart of the method of Lagrange multipliers is the orthogonal complement theory from linear algebra. Of course, you can be heartless and still sucessfully apply the method of Lagrange.

8.3 Lagrange multiplier method for one constraint

PROBLEM: FIND EXTREME VALUES OF $f(x_1, x_2, \dots, x_n)$ on the level set $g(x_1, x_2, \dots, x_n) = 0$.

We suppose the objective function f and the constraint g share some common domain $U \subseteq \mathbb{R}^n$ and $f, g : U \to \mathbb{R}$.

1.) If $p \in U$ gives max/min $f(p)$ then if we take any smooth path $\gamma : \mathbb{R} \to M$ (where $M = g^{-1}(\{0\})$) then $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ and we see $f \circ \gamma$ should have maximum when $\gamma(t) = p$. Usually we set up $\gamma : \mathbb{R} \to M$ with $\gamma(0) = p$, and smooth $\Rightarrow \gamma'(0) \neq 0$. Consider then

$$\frac{d}{dt}\Big[(f \circ \gamma)(t)\Big] = \big[\nabla f(\gamma(t))\big] \cdot \gamma'(t)$$

But when $t = 0$ we have a critical pt. for $f \circ \gamma$

Hence $\underbrace{\boxed{(\nabla f)(p) \cdot \gamma'(0) = 0}}$.

this eq$^{\underline{n}}$ says $(\nabla f)(p)$ is orthogonal to an arbitrary tangent vector to constraint surface at p.

2.) Likewise, consider the path $\gamma : \mathbb{R} \to M$ composed with g; $g \circ \gamma : \mathbb{R} \to \mathbb{R}$. By def$^{\underline{n}}$ of M we have $(g \circ \gamma)(t) = 0 \quad \forall t$.

$$0 = \frac{d}{dt}\Big((g \circ \gamma)(t)\Big) = \big[(\nabla g)(\gamma(t))\big] \cdot \gamma'(t)$$

$$\therefore \quad \boxed{(\nabla g)(p) \cdot \gamma'(0) = 0}$$

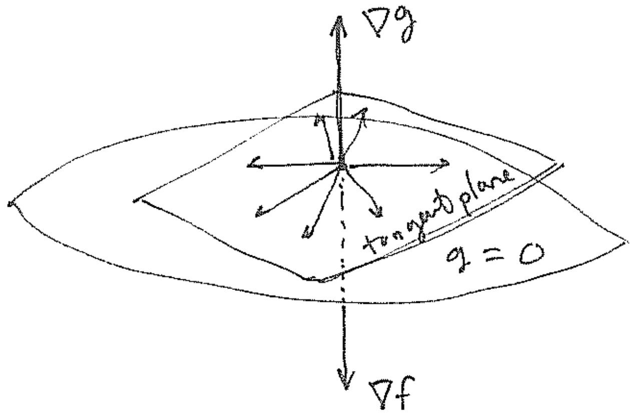
Again this suggests $(\nabla g)(p)$ is likewise orthogonal to all tangent vectors in the tangent plane to $g = 0$.

Continuing:

Think about this, $g(x_1, x_2, \dots, x_n) = 0$ gives 1-eq$^{\underline{n}}$ in n-unknowns. It turns out $(\nabla g)(P) \neq 0$ makes $g(x_1, x_2, \dots, x_n) = 0$ an $(n-1)$-dimensional surface. This means there is only 1-dimension left. Thus

$$(\nabla f)(P) \cdot \gamma'(0) = 0 \quad \text{AND} \quad (\nabla g)(P) \cdot \gamma'(0) = 0$$

$$\Rightarrow \boxed{\nabla f(P) = \lambda \nabla g(P)}$$ for some constant "Lagrange Multiplier" λ.



8.4 Lagrange multiplier method for several constraints

PROBLEM: FIND EXTREMA FOR $f : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ subject to constraints $G_1 = 0$, $G_2 = 0$, .., $G_m = 0$ where $G_j : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ for $j = 1, 2, \ldots, m$.

Observation: $G = (G_1, G_2, .., G_m) = (0, 0, .., 0)$ imposes m- constraints at once. A convenient def$^{\underline{n}}$ for the constraint surface is $G^{-1}\{(0,0,..,0)\} = M$. This will give an $(n-m)$-dimensional surface or "manifold" in U provided a condition on $\nabla G_1, \nabla G_2, .., \nabla G_m$ is met. (need rank $(G'(P)) = m$)

1.) Let $\gamma : I \subseteq \mathbb{R} \longrightarrow M$ with $\gamma(0) = P$ and $\gamma'(0) \neq 0$. Suppose f has extrema at P; $f \circ \gamma : \mathbb{R} \longrightarrow \mathbb{R}$ has critical pt (max/min maybe) at $t = 0$.

$$0 = \frac{d}{dt}\Big[(f \circ \gamma)(t)\Big] = \big[\nabla f(\gamma(t))\big] \cdot \gamma'(t)$$

Again $\boxed{\nabla f(P) \cdot \gamma'(0) = 0}$

2.) Note $G_j \circ \gamma : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is identically zero since γ is a curve whose image $\gamma(I) \subset M$. Thus

$$\frac{d}{dt}(0) = \frac{d}{dt}\Big[(G_j \circ \gamma)(t)\Big] = \big[\nabla G_j(\gamma(t))\big] \cdot \gamma'(t) = 0$$

This holds for $j = 1, 2, .., m$ hence

$$\boxed{(\nabla G_j(P)) \cdot \gamma'(0) = 0 \quad \text{for } j = 1, 2, .., m}$$

Continuing:

Provided $\text{rank}(G'(p)) = m$ it follows the set of tangent vectors at p span a $(n-m)$-dimensional subspace translated to the point p. Moreover, we note $(\nabla f)(p) \cdot \gamma'(0) = 0$ says $(\nabla f)(p)$ is in the orthogonal complement to the tangent space. On the other hand $\text{rank}(G'(p)) = m$ implies $\nabla G_1(p), \nabla G_2(p), \dots, \nabla G_m(p)$ gives a basis for the complement as each $\nabla G_j(p)$ is in the orthogonal complement of the tangent space by $\underline{\nabla G_j(p) \cdot \gamma'(0) = 0}$. It follows that ∇f must be a linear combination of the basis for the complement;

$$\boxed{\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2 + \cdots + \lambda_m \nabla G_m}$$

Indeed, if p gives a critical point for $f \circ \gamma$ it must be subject to the boxed condition above for reasonably posed constraint $G = 0$.

Example: find points on circle $x^2+y^2=1$ and parabola $y^2 = 2(4-x)$ which are closest.

Let $f(x,y,u,v) = (x-u)^2 + (y-v)^2$ and construct $G(x,y,u,v) = (x^2+y^2-1,\ v^2+2u-8)$. Max/Min f subject to constraints $G=0$.

$$\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$$

$$\langle 2(x-u),\ 2(y-v),\ -2(x-u),\ -2(x-v)\rangle =$$
$$= \lambda_1 \langle 2x, 2y, 0, 0\rangle + \lambda_2 \langle 0, 0, 2, 2v\rangle$$

Yields,

$$\left.\begin{aligned} 2(x-u) &= 2\lambda_1 x \\ 2(y-v) &= 2\lambda_1 y \end{aligned}\right\} \longrightarrow \frac{y}{x} = \frac{y-v}{x-u} \quad \text{I}$$

$$-2(x-u) = 2\lambda_2 \longrightarrow \underbrace{-\lambda_2 = x-u}$$

$$-2(y-v) = 2\lambda_2 v \longrightarrow y - v = (x-u)v \quad \text{II}$$

Combining I & II

$$\frac{y}{x} = \frac{y-v}{x-u} = \frac{\cancel{(x-u)}v}{\cancel{(x-u)}} = v \quad \therefore \quad y = xv \quad \text{III}$$

Given together with II $\longrightarrow$ $xv - v = (x-u)v$

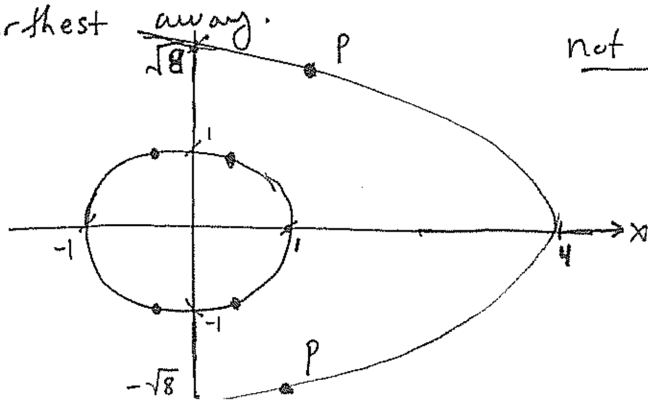
$$\Rightarrow xv - v = xv - uv$$
$$\Rightarrow v = uv$$
$$\Rightarrow u = 1.$$

But, $v^2 = 2(4-u) = 2(3) = 6 \quad \therefore \quad v = \pm\sqrt{6}$.

Hence, $y = (\pm\sqrt{6})x \Rightarrow x^2+y^2 = \underbrace{x^2+6x^2=1}_{7x^2=1} \quad \therefore \quad x = \pm 1/\sqrt{7}$

We find points $(1, \pm\sqrt{6})$ on the parabola and $(\pm 1/\sqrt{7}, \pm\sqrt{6}/\sqrt{7})$ on the circle are closest or furthest away.

not to scale!



Example: find points on $x^2+y^2+z^2=1$ and plane $x+y+z=3$ which are closest.

$g_1(x,y,z) = (1-x^2-y^2-z^2) = 0$ gives sphere as $g_1^{-1}\{0\}$.

$g_2(u,v,w) = 3-u-v-w = 0$ gives plane as $g_2^{-1}\{0\}$.

Consider $f(\vec{x},\vec{u}) = \|\vec{x}-\vec{u}\|^2$. We'd like to find min/max for f subject the constraints

$$G(\vec{x},\vec{u}) = \langle g_1(\vec{x}), g_2(\vec{u})\rangle = \langle 0,0\rangle.$$

To say $G=0$ is to place $\vec{x}$ on the sphere and $\vec{u}$ on the plane. Let $G_1(\vec{x},\vec{u}) = g_1(\vec{x})$ and $G_2(\vec{x},\vec{u}) = g_2(\vec{u})$ then

$$\nabla G_1 = \langle \nabla g_1, 0\rangle = \langle g_{1x}, g_{1y}, g_{1z}, 0, 0, 0\rangle$$
$$\nabla G_2 = \langle 0, \nabla g_2\rangle = \langle 0,0,0, g_{2u}, g_{2v}, g_{2w}\rangle$$

Likewise, $f(\vec{x},\vec{u}) = \sum_{j=1}^{n}(x_j-u_j)^2 = (x-u)^2+(y-v)^2+(z-w)^2$

$$\nabla f(x,y,z,u,v,w) = \langle 2(x-u), 2(y-v), 2(z-w), -2(x-u), -2(y-v), -2(z-w)\rangle$$

Then $\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$ yields,

$$\left.\begin{aligned} 2(x-u) &= \lambda_1 g_{1x} = -2\lambda_1 x \\ 2(y-v) &= \lambda_1 g_{1y} = -2\lambda_1 y \\ 2(z-w) &= \lambda_1 g_{1z} = -2\lambda_1 z \end{aligned}\right\} \quad \frac{x-u}{x} = \frac{y-v}{y} = \frac{z-w}{z}$$

$$\left.\begin{aligned} -2(x-u) &= \lambda_2 g_{2u} = -\lambda_2 \\ -2(y-v) &= \lambda_2 g_{2v} = -\lambda_2 \\ -2(z-w) &= \lambda_2 g_{2w} = -\lambda_2 \end{aligned}\right\} \quad x-u = y-v = z-w$$

Thus, $\dfrac{x-u}{x} = \dfrac{x-u}{y} = \dfrac{x-u}{z} \Rightarrow \underline{x=y=z} \Rightarrow \underline{u=v=w}$

$$\underbrace{3x^2=1}_{x=\pm 1/\sqrt{3}} \qquad \underbrace{3u=3}_{u=1}$$

We obtain the point $(1,1,1)$ on the plane is closest to $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ and furthest from $(-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3})$ on the sphere.

<u>Remark</u>: p. 116 of Edwards derives this on basis of the <u>amazing</u> Example 10 of Edwards.

$\boxed{\text{Example: Find max/min of } w = x+z \text{ where } x^2+y^2+z^2 \leq 1.}$

Notice $f(x,y,z) = x+z$ has $\nabla f = \langle 1, 0, 1 \rangle$ thus f has no critical pts. It follows f must attain max/min on boundary $g(x,y,z) = x^2+y^2+z^2-1 = 0$. Use method of Lagrange,

$$\nabla f = \lambda \nabla g \quad \text{where } g = 0$$

$$\langle 1, 0, 1 \rangle = \lambda \langle 2x, 2y, 2z \rangle$$

$$\left.\begin{array}{l} 1 = 2\lambda x \\ 0 = 2\lambda y \\ 1 = 2\lambda z \end{array}\right\} \longrightarrow y = 0 \;\; \& \;\; 1 = \frac{x}{z} \Rightarrow z = x$$

$$\Rightarrow x^2 + 0^2 + x^2 = 1$$

$$\Rightarrow x^2 = \frac{1}{2}$$

$$\Rightarrow \underline{x = \pm 1/\sqrt{2}}.$$

Thus $(1/\sqrt{2}, 0, 1/\sqrt{2})$ or $(-1/\sqrt{2}, 0, -1/\sqrt{2})$ yield extremal values of f on $g = 0$.

$$f(1/\sqrt{2}, 0, 1/\sqrt{2}) = 1/\sqrt{2} + 1/\sqrt{2} = 2/\sqrt{2} = \sqrt{2}.$$

$$f(-1/\sqrt{2}, 0, -1/\sqrt{2}) = -1/\sqrt{2} - 1/\sqrt{2} = -2/\sqrt{2} = -\sqrt{2}.$$

The max. value is $\sqrt{2}$ reached at $(1/\sqrt{2}, 0, 1/\sqrt{2})$ and the min. value is $-\sqrt{2}$ reached at $(-1/\sqrt{2}, 0, -1/\sqrt{2})$.

Example: A plane wave in the z-direction has the form

$$\vec{E} = E_o \cos(kz - \omega t)\,\hat{z}$$

where E_o, k, ω are constants and $\hat{z} = \frac{\nabla z}{|\nabla z|} = \langle 0, 0, 1\rangle$.
Here $\vec{E} : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ is a time-dependent vector field.
One interesting function built from $\vec{E}$ is the square of its length;
define $f(t, x, y, z) = \vec{E} \cdot \vec{E} = E_o^2 \cos^2(kz - \omega t)$.

PROBLEM: find critical points for f and find its max/min relative to $t = 0$ or $x^2 + y^2 + z^2 = R^2$ or $x^2+y^2+z^2 = (Rt)^2$

I choose the ordering $x_0 = t$, $x_1 = x$, $x_2 = y$, $x_3 = z$ for this problem. With respect to this ordering,

$$\nabla f = \left\langle \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$$

$$= \left\langle -2E_o^2 \cos(kz - \omega t)\sin(kz - \omega t)(-\omega),\ 0,\ 0,\ -2E_o^2 \cos(kz-\omega t)\sin(kz-\omega t)\,k \right\rangle$$

$$= E_o^2 \left\langle \omega \sin(2[kz - \omega t]),\ 0,\ 0,\ -k\sin(2[kz - \omega t]) \right\rangle$$

Critical points must have $\nabla f = 0$ since f is everywhere differentiable. We see that both of the sine terms should vanish if we are to obtain $\nabla f = 0$. Hence,

$$\nabla f = 0 \iff n\pi = 2(kz - \omega t), \quad n \in \mathbb{Z}.$$

$$\iff kz - \omega t = \frac{n\pi}{2}, \quad n \in \mathbb{Z}.$$

$$\iff z = \frac{\omega t + n\pi/2}{k}, \quad n \in \mathbb{Z}.$$

Think about it, for a fixed time there are planes of critical points every $\frac{\pi}{2k}$ units in the z-direction. On the other hand, if we imagine time flowing then you can envision these planes of critical points flowing upward with a speed of $\frac{dz}{dt} = \frac{\omega}{k}$.

Now let's think about the constraints

(i.) $t = 0$, (ii) $x^2+y^2+z^2 = R^2$, (iii) $x^2+y^2+z^2 = (Rt)^2$

see over

(i.) $t = 0$ gives $g_1(t,x,y,z) = t = 0$ study $\nabla f = \lambda \nabla g$,

$$\nabla f = \lambda \nabla g \longrightarrow E_o^2 \langle w\sin(2kz), 0, 0, -k\sin(2kz)\rangle = \lambda\langle 1,0,0,0\rangle$$

$$\hookrightarrow \quad E_o^2 w \sin(2kz) = \lambda$$
$$-E_o^2 w \sin(2kz) = 0$$

These equations are inconsistent unless $2kz = n\pi$ for some $n \in \mathbb{Z}$. However, in that case $\nabla f = 0$ thus Lagrange multipliers don't apply. We can still analyze the function despite the failure of the method, for $t = 0$ and $2kz = n\pi$ we have

$$f\left(0, x, y, \frac{n\pi}{2k}\right) = E_o^2 \cos^2\left[k\left(\frac{n\pi}{2k} - w(0)\right)\right] = E_o^2\cos^2\left(\frac{n\pi}{2}\right)$$

Note for $n \in 2\mathbb{Z} + 1$ (odd n) we have $f = 0$ whereas for $n \in 2\mathbb{Z}$ (even n) we have $f = E_o^2$.

Clearly these give max/min for f relative to the constraint $t = 0$.

$$f\left(0, x, y, \frac{n\pi}{2k}\right) = \begin{cases} 0 & \text{if } n \text{ odd (minimum)} \\ E_o^2 & \text{if } n \text{ even (maximum)} \end{cases}$$

The planes of critical pts. are now separated into max/min

(ii.) $x^2+y^2+z^2 = R^2$ encoded by $g(x,y,z) = x^2+y^2+z^2-R^2 = 0$

study $\nabla f = \lambda \nabla g$, where $g = 0$

$$E_o^2 \langle w\sin[2(kz-wt)], 0, 0, -k\sin 2(kz-wt)\rangle = \lambda\langle 0, 2x, 2y, 2z\rangle$$

$$E_o^2 w \sin(2(kz-wt)) = 0 \quad \Rightarrow \quad kz - wt = \frac{n\pi}{2}, \; n \in \mathbb{Z}$$

$$\left.\begin{array}{l} 0 = 2\lambda x \\ 0 = 2\lambda y \end{array}\right\} \Longrightarrow \underline{x = y = 0} \Rightarrow \underline{z^2 = R^2}.$$

$$-E_o^2 k \sin(2(kz-wt)) = 2\lambda z \quad \Rightarrow \quad 0 = 2\lambda z$$

$$\Rightarrow \underbrace{\lambda = 0} \text{ or } \underbrace{z = 0}$$

Either $\lambda = 0$, $kz - wt = \frac{n\pi}{2}$.

or $x = y = z = 0$ and $-2wt = \frac{n\pi}{2}$. ← not possible.

$$\boxed{z = \frac{wt + n\pi/2}{k} \quad \text{and} \quad x^2+y^2 = R^2 - z^2}$$

(iii) $x^2+y^2+z^2 = (Rt)^2$ ← spherical wave front

$g(t,x,y,z) = R^2t^2 - x^2 - y^2 - z^2$

$\nabla f = \lambda \nabla g \quad , \quad g = 0$

$$E_0^2 \langle \omega \sin 2(kz-\omega t), 0, 0, -k \sin 2(kz-\omega t) \rangle = \lambda \langle 2R^2 t, -2x, -2y, -2z \rangle$$

$$\begin{cases} E_0^2 \omega \sin 2(kz-\omega t) = 2\lambda R^2 t \\ 0 = -2x\lambda \\ 0 = -2y\lambda \\ -E_0^2 k \sin 2(kz - \omega t) = -2\lambda z \end{cases}$$

$$\underbrace{\frac{R^2 t}{\omega} = \frac{z}{k}} = \frac{E_0^2 \sin 2(kz-\omega t)}{2\lambda} \quad \& \quad x = y = 0.$$

$$\underline{z = \left(\frac{kR^2}{\omega}\right) t} \quad \text{or} \quad \underline{t = \left(\frac{\omega}{kR^2}\right) z}.$$

where $x = y = 0$. Hence, $z^2 = R^2t^2$

$$\Rightarrow \left(\frac{kR^2}{\omega} t\right)^2 = R^2t^2$$

$$\Rightarrow \left(\frac{k^2R^4}{\omega^2} - R^2\right) t^2 = 0$$

$$\Rightarrow \left(\frac{k^2R^2}{\omega^2} - 1\right) R^2t^2 = 0$$

$$\therefore \ R^2 = \frac{\omega^2}{k^2} \quad \text{or} \quad t = 0.$$

The spherical wave-front $x^2+y^2+z^2 = (Rt)^2$ collides with a max/min of the plane-wave $\vec{E}$ provided that $R = \frac{\omega}{k}$ (we assume $\omega, k > 0$ for physical reasons)

<u>Remark</u>: the two waves match at time zero then flow upward with same speed of $\frac{\omega}{k}$ along z-axis.