

Purkiss Sym Prince

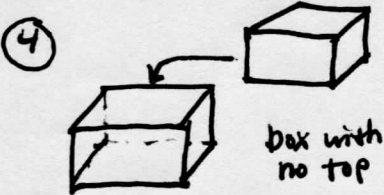
Do Sym Prnc have Sym solns?
 waterhouse

Examples

(1) rectangle perim vs area $2(x+y) = P_0$ fixed $A = xy$ is max when $x=y$

(2) Triangle inscribed in circle  min area when Δ equilateral

(3) $x, y, z, w > 0$ $xy.z.w = 16 \Rightarrow x+y+z+w \geq 8$ extrema when $x=y=z=w=2$



(5) Find min $f = (r-1)^2 + (\frac{s}{r}-1)^2 + (\frac{t}{s}-1)^2 + (\frac{4}{t}-1)^2$ [1981 Putnam]

Subj to $1 \leq r \leq s \leq t \leq 4$ ans. $\langle r, s, t \rangle = \langle \sqrt{2}, 2, 2\sqrt{2} \rangle$

Problem becomes sym with new vars $R=r, S=\frac{s}{r}, T=\frac{t}{s}, W=\frac{4}{t}$

(6) Fixed $\bar{x} = \frac{1}{n} \sum x_i$ $\rho = \sum x_i^2$ is min when all x_i are equal

(7) Arithmetic vs Geo Means: $x_i \geq 0 \forall i=1 \dots n$

$$\sum \frac{x_i}{n} \geq (\prod x_i)^{\frac{1}{n}} \quad (\text{generalized } \textcircled{3} \text{ above})$$

Bunyakowski counter example:

Bunyakowski proved the principle holds for all polys of degree ≤ 3 . So we need degree 4

$$f(x,y) = [x^2 + (y-1)^2] [(x-1)^2 + y^2] \geq 0$$

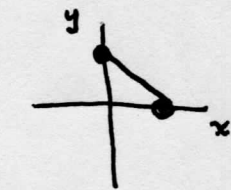
This obviously takes its ^{absolute (global)} min value at $\langle 0, 1 \rangle$ or $\langle 1, 0 \rangle$

or for a problem with a constraint:

$$f(x,y) = (x^2 + y^2 - \frac{5}{8})^2$$

consider $x, y > 0$ $x+y=1$

at $x = \frac{1}{2} = y$ $f(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{8})^2$



but there are larger values $f(\frac{1}{8}, \frac{7}{8}) = (\frac{5}{32})^2$
 and smaller $f(\frac{1}{4}, \frac{3}{4}) = 0$

rather than $x=y$
 See Hazewinkel 'Sym Problemen' dutch

So Purkiss principle must only mean a local extrema.

we want Lag Mult cond $\nabla f = \sum \lambda_i Dg_i$

or for 1 constant $g: \mathbb{R}^n \rightarrow \mathbb{R}$
 $\nabla f = \lambda \nabla g$

This is what he deals with until the end

example

$$f(x,y,z) = xy + yz + zx$$

$$g(x,y,z) = (xyz)^2$$

elementary sym polys

$$\nabla f = \begin{bmatrix} y+z \\ x+z \\ x+y \end{bmatrix}$$

$$\nabla g = 2xyz \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix}$$

plug in $\langle r, r, r \rangle = \vec{r}$

$$\nabla f_r = \begin{bmatrix} 2r \\ 2r \\ 2r \end{bmatrix}$$

$$\nabla g_r = 2r^3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

All entries are equal in ∇ so we always will have $\nabla f_r = \lambda \nabla g_r$

Lemma 1 $f(Sx) = f(x) \forall S \in \mathcal{S}_n \Rightarrow$ When all x_i are equal, all $D_i f$ are equal:
 $x_1=r, x_2=r, \dots, x_n=r$
 so $x = \langle r, r, \dots, r \rangle = \vec{r}$
 Then $D_i f_r = \alpha \forall i$

Maybe should name the fun "h" since we want to apply these lemmas to $h=f$ and $h=g$

$$\left[\begin{array}{l} \text{Thus if } D_i f_r = \alpha \forall i \\ D_i g_r = \beta \forall i \end{array} \Rightarrow \nabla f_r = \left(\frac{\alpha}{\beta}\right) \nabla g_r \right]$$

Pf. $f(Sx) = f(x)$

$$\Rightarrow Df_{Sx} DS_x = Df_x \quad \text{but } S \text{ is a perm, so linear } DS_x = S \quad \text{say } S = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

$$Df_{Sx} S = Df_x$$

let $g = \langle y, x, z \rangle = Sx$

$$\begin{bmatrix} D_1 f_g & D_2 f_g & D_3 f_g \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} D_1 f_x & D_2 f_x & D_3 f_x \end{bmatrix}$$

$$= \begin{bmatrix} 0 & D_1 & 0 \\ 0 & 0 & D_2 \\ D_3 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} D_3 f_g & D_1 f_g & D_2 f_g \end{bmatrix} = \begin{bmatrix} D_1 f_x & D_2 f_x & D_3 f_x \end{bmatrix}$$

Now if $x = r = \langle r, r, r \rangle$ then $\underbrace{Sx}_g = x$ that is to say

Fixed pt. $Sr = r$

$$\begin{bmatrix} D_3 f_r & D_1 f_r & D_2 f_r \end{bmatrix} = \begin{bmatrix} D_1 f_r & D_2 f_r & D_3 f_r \end{bmatrix}$$

$$\begin{aligned} D_3 f_r &= D_1 f_r \\ D_1 f_r &= D_2 f_r \\ D_2 f_r &= D_3 f_r \end{aligned} \quad \text{so all must be equal}$$

any perm that is a big cycle of all the elts.

so for any But in general, I think we'd need all perms $S \in \mathcal{S}_n$ no cycles, no fixed pts \square

For the 2nd deriv, we need Robinson

Robinson Lemma 1 $\nabla f_{x^*} = \lambda \nabla g_{x^*}$ on $S = g^{-1}(c)$ $\Rightarrow \left. \frac{d^2}{dt^2} f(\sigma(t)) \right|_{t=0} = v^T [D^2 f_{x^*}] v$
 σ is a curve $\sigma(0) = x^*$
 $\sigma'(0) = v$
 \uparrow $T_{x^*} M$
 so $v \in T_{x^*} M$

$\mathcal{L} = f - \lambda g$
 $v^T [D^2 f_{x^*} - \lambda D^2 g_{x^*}] v$

This gives what Waterhouse has

quadratic $Q(v) := v^T [D^2 f_p - \lambda D^2 g_p] v$ (this should be pos or neg def)

(ex) Let's see it for our example:

$f = xy + yz + zx$
 $Df_x = [(y+z) \ (x+z) \ (x+y)]$
 $g = (xyz)^2$
 $Dg_x = 2xyz [yz \ xz \ xy]$
 $= [2xy^2z^2 \ 2x^2yz^2 \ 2x^2y^2z]$

then $D^2 f_x = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

$D^2 g_x = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$

but $D^2 g_r = \begin{bmatrix} 2 & 4 & 4 \\ 4 & 2 & 4 \\ 4 & 4 & 2 \end{bmatrix}$

Lemma 2 $f(Sx) = f(x) \ \forall S \in \mathcal{S}_n$
 (or at least S is one big cycle)
 $f \in C^2$ smooth

$\Rightarrow D^2 f_r = \begin{bmatrix} \alpha & \beta & \beta \\ \beta & & \beta \\ \beta & \beta & \alpha \end{bmatrix}$

- main diag elts equal
- off diag elts equal

pf. From Lemma 1, $(f \circ S)(x) = f(x)$

$Df_{Sx} S = Df_x \quad Sr = r \quad [D_1 f_r \ D_2 f_r \ D_3 f_r] [S] = [D_1 f_r \ D_2 f_r \ D_3 f_r]$

operating from the back, S perm

Now do 2nd deriv: $D^2(f \circ S)_x(h, k) = D(Df_{Sx}(Sh))_{(k)}$ (the cols of Df)
 $= D^2 f_{Sx}(Sh, Sk) = [Sk]^T [D^2 f_{Sx}] [Sh]$

or $D^2(f \circ S)_x = S^T [D^2 f_{Sx}] S$

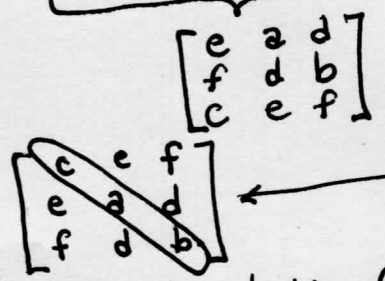
Thus $S^T [D^2 f_{Sx}] S = D^2 f_x$ plug in $x = r = \langle r_1, \dots, r_n \rangle \quad S^T D^2 f_r S = D^2 f_r$ cont 2 \rightarrow

Lemma 2 pt cont'd

Lets look at a 3x3 example we know $D^2 f_x$ is symmetric matrix

$$S^T = S^{-1} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$$

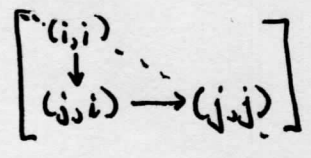
$$D^2 f_x = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$$



main diag is sent to itself and permuted. Same for non-diag elts.

How can we prove this holds for any perm and a nxn matrix?

To preserve the main diag consider elt (i,i) if row $i \rightarrow$ row j
Then we must also map col $i \rightarrow$ col j



When S acts from the back, it permutes the rows
When S^T acts from the front, it has exactly the same perm of the cols.

Consider this index vector $y^T = [1 \ 2 \ 3]$ then $[1 \ 2 \ 3] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = [3 \ 1 \ 2]$

But $S^T y = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$

And, in general, we see $(S^T y)^T = y^T S$ So what S in the back does to the cols, S^T in front does to rows

Thus main diag is preserved as a set, but permuted.

Thus $D_{ii} f_r = D_{jj} f_r$ for any i, j if we consider all perms $S \in S_n$
By a similar arg, if $(i+p, i)$ off diag

This is going to yield $D^2 f_r = \begin{bmatrix} b & c & c \\ c & b & c \\ c & c & b \end{bmatrix}$ (no relation to previous use of these letters)

□

Lemma 3 Consider quadratic form

$$Q(v) = v^T \begin{bmatrix} b & c & c \\ c & b & c \\ c & c & b \end{bmatrix} v = v^T \begin{bmatrix} b & & \\ & b & \\ & & b \end{bmatrix} v + v^T \begin{bmatrix} 0 & c & c \\ c & 0 & c \\ c & c & 0 \end{bmatrix} v$$

If $\sum v_i = 0 \Rightarrow Q(v) = (b-c) \sum v_i^2 = (b-c) \|v\|_2^2$
 pos def if $b > c$
 neg def if $b < c$

Pf. Observe $\sum v_i = [1 \ 1 \ 1] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$

Then $\nabla g_r \cdot v = 0$ is $g_r \cdot [1 \ 1 \ 1] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$ so we recover this cond. $v \perp \nabla g_r$ i.e. $v \in T_x S$

↑
should call it M

$v^T \begin{bmatrix} b & & \\ & b & \\ & & b \end{bmatrix} v = b [v_1 \ v_2 \ v_3] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = b \|v\|_2^2$ so that one is easy

$v^T \begin{bmatrix} 0 & c & c \\ c & 0 & c \\ c & c & 0 \end{bmatrix} v = [v_1 \ v_2 \ v_3] \begin{bmatrix} 0 & c & c \\ c & 0 & c \\ c & c & 0 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} = [N_1 \ N_2 \ N_3] \begin{bmatrix} c(N_2 + N_3) \\ c(N_1 + N_3) \\ c(N_1 + N_2) \end{bmatrix}$

$\sum N_i = 0 \quad N_1 + N_2 + N_3 = 0$
 $N_2 + N_3 = -N_1$

$= -c N_1^2 - c N_2^2 - c N_3^2 = -c \|N\|_2^2$

$\Rightarrow Q(v) = (b-c) \|v\|_2^2$

□

we only need subgroup of full size cycle $(1 \ 2 \dots n) \in S_n$
see example

Purkiss Prince

$S: \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear perm matrix O.N.
 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and $f(Sx) = f(x) \quad \forall S \in S_n$
 $g: \mathbb{R}^n \rightarrow \mathbb{R} \quad g(Sx) = g(x) \quad \text{" " " "}$

Let $M := g^{-1}(c_0)$

$S(M) = M$ invariant.
distances between pts unchanged

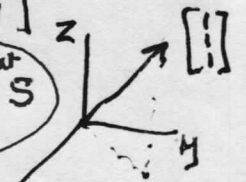
$f|_M$ has a local extrema where $M \cap E$

any pt m where $S^k m = m \quad \forall S \in S_n$
Fixed pt

E is the line where all components are equal

$E = \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \mid z \in \mathbb{R} \right\}$

E is pts that are fixed by $S \quad \forall S \in S_n$



example $f(x,y,z) = \frac{x}{y^2+1} + \frac{y}{z^2+1} + \frac{z}{x^2+1}$

Michael Penn youtube

$g(x,y,z) = x+y+z-3$ this is a plane in \mathbb{R}^3 invariant under S

The inequality is $\frac{x}{y^2+1} + \frac{y}{z^2+1} + \frac{z}{x^2+1} \leq \frac{3}{2}$

Pf of Purkiss

$P = \langle r, r, \dots, r \rangle$

assume $\nabla g(p) \neq 0$

Lemma 1 $\Rightarrow \nabla f_p = \lambda \nabla g_p$

$D^2 f_p - \lambda D^2 g_p = D^2 \mathcal{L}_p$ where $\mathcal{L} := f - \lambda g$ as in Robinson

$D^2 \mathcal{L}_p = \begin{bmatrix} b & c & c \\ c & b & c \\ c & c & b \end{bmatrix}$

$\nabla g_p = \varphi(p) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

Thus $v \perp \nabla g_p \Rightarrow \nabla g_p \cdot v = 0 \Rightarrow \varphi(p) (\sum N_i) = 0 \Rightarrow \sum N_i = 0$

Then by Lemma 3 $v^T [D^2 \mathcal{L}_p] v = (b-c) \|v\|_2^2$ pos or neg def

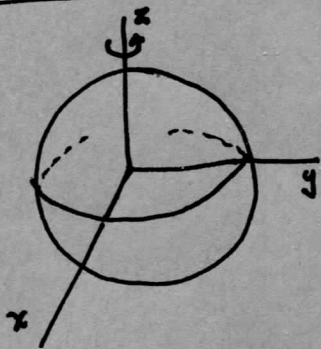
$b = D_{i,i} \mathcal{L}_p$ $c = D_{i,j} \mathcal{L}_p$ for any $i \neq j$

take $i=1$
since always same value.

$\Rightarrow P$ is an extrema of $f|_M$

□

examples of degeneracy



Rotation $Q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ take $\theta = \pi/2$

$g(x,y,z) = x^2 + y^2 + z^2 = \rho^2$ is invariant

$g(Qx) = g(x)$ $M = S^2 = g^{-1}(\rho^2)$ is invariant

$f = z^2$ is invariant (could be a better example? \hookrightarrow more non-trivial)

North pole N, S would be extrema of f , not saddle pts.

Instead of Sym wrt perms of the variables, consider more general linear

maps $T \ni f(Tx) = f(x)$

We don't need to consider continuous Sym groups (Lie Groups)

$T \in G$ some finite group

$T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ (he says $T: V \rightarrow V$)

we will call f an invariant fcn

$Tp = p$ means p fixed pt.

$Sr = r$
was FP before

Lemma 4 cl don't know what he is doing

But we want $\langle Tx, Ty \rangle = \langle x, y \rangle \forall T \in G$. Say $T = Q$ O.N. matrix.

$$Tp = p$$

$$v^T Av$$

Consider Taylor expansion of f about p

$$f(p+v) = f(p) + \underbrace{Df_p(v)}_{\nabla f_p \cdot v} + \underbrace{\frac{1}{2!} D^2 f_p(v, v)}_{q_p^f(v) \text{ quadratic form}} + o(v^3)$$

Now since $f(Tx) = f(x)$
 let $x = p+v$

$$f(p+v) = f(T(p+v)) = f(p+Tv) \stackrel{!}{=} f(p) + \langle \nabla f_p, Tv \rangle + q_p^f(Tv, Tv) + o$$

By uniqueness (of terms in Taylor expansion) we see $q_p^f(Tv) \stackrel{!}{=} q_p^f(v)$
 and $\langle \nabla f_p, v \rangle \stackrel{!}{=} \langle \nabla f_p, Tv \rangle$

In variance of IP: $\langle \nabla f_p, Tv \rangle = \langle T^{-1} \nabla f_p, v \rangle$ (This follows easily if $T=Q$ o.n. matrix)
 (or just def of adjoint)

$$\text{Thus } \langle \nabla f_p, v \rangle = \langle T^{-1} \nabla f_p, v \rangle \quad \forall v$$

$$\Rightarrow \nabla f_p \stackrel{!}{=} T^{-1} \nabla f_p$$

$$\text{or } T(\nabla f_p) = \nabla f_p \quad \text{and } \nabla f_p \text{ is fixed by } T.$$

This proves:

Lemma 5 : f is invariant under T ($f(Tx) = f(x)$) $\forall x$
 $Tp = p \quad \forall T \in \mathcal{G} \Rightarrow \left. \begin{aligned} T(\nabla f_p) &= \nabla f_p \\ q_p^f(Tv) &= q_p^f(v) \quad \forall T \in \mathcal{G} \end{aligned} \right\}$

To get the analogue of Purkiss, we need more hypotheses:

Lemma 1 had $\nabla f_p = \alpha(r) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\nabla g_p = \beta(r) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ Thus $\nabla f_p \stackrel{!}{=} \lambda \nabla g_p$
 Since both in $\text{span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

Here we require $Tx = x$ only if $x \in U_0$, a 1-dim subsp.

We need an analogue of $\sum n^i = 0$ i.e. $\nabla g_p \cdot v = \beta(r) [1 \dots 1] \begin{bmatrix} n_1 \\ \vdots \\ n_n \end{bmatrix} = 0$
 It will be an invariant subsp

Lemma 6

U invariant under T
 $T|_U : U \rightarrow U$
 Let $W := U^\perp$ wrt $\langle \cdot, \cdot \rangle$

W is also invariant under T (8)
 $T|_W : W \rightarrow W$

Pf see Schuams LA in particular, this holds for Q O.N. matrix

In particular $U_0^\perp =: W_0$

In Lemma 2 we needed the perms to move all the ^{matrix} elts, no elt mapped to itself.
 Here, the analogue is we require W_0 to have no invariant subspaces (only $\{0\}$)
 They call this irreducibility. Lemma 3 generalizes as:

Lemma 7: \mathcal{L} fin group linear maps $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$
 Each T acts irreducibly on W } \Rightarrow quadratic form
 $q(w) = \alpha \langle w, w \rangle$ for some $\alpha \in \mathbb{R}$
 $= \alpha \|w\|_2^2$
 Pos or neg definite

Pf. $q(w) = \langle Bw, w \rangle = w^T Bw$ for uniq self-adj map $B: W \rightarrow W$

We know $q_p^f(Tv) = q(v)$

in particular $q(w) = q(Tw)$

$$\langle Bw, w \rangle = \langle BTw, Tw \rangle \stackrel{\text{adjoint}}{=} \langle T^T B T w, w \rangle \quad \forall w$$

$T^T = T^{-1}$ for T ON and other cases

$$\Rightarrow B = T^{-1} B T \text{ or } T B = B T$$

The \mathbb{R} self-adj map B has a $\mathbb{R} \in W$ λ_0

~~Let X be the eigenspace~~ Let X be the eigenspace $\{ w \in W \mid Bw = \lambda_0 w \}$

~~For $w \in X$, $T \in \mathcal{L}$ we have~~

For $w \in X$, $T \in \mathcal{L}$ we have $B(Tw) = T B w = T(\lambda_0 w) = \lambda_0 (Tw)$
 Thus $Tw \in X$

Thus X is an invariant subsp of W , so by hypoth $X = W$

Thus $B = \lambda_0 (Id)$ and so $q(w) = \lambda_1 \langle w, w \rangle$ □

extended to linear maps that are not perms.

Thm Extended Purkiss Prime

\mathcal{G} finite group of linear maps $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$
 \exists 1-dim subsp U_0 where $Tu = u \ \forall T \in \mathcal{G}, u \in U_0$

There is no other non-trivial invariant subsp in U_0^\perp

$$f, g: \mathbb{R}^n \rightarrow \mathbb{R} \quad \begin{aligned} f(Tx) &= f(x) \quad \forall T \in \mathcal{G} \\ g(Tx) &= g(x) \quad \forall x \in \text{nbhd of } p \\ &\quad \text{where } T_p = p \end{aligned}$$

$$\begin{aligned} \text{let } c &:= g(p) \\ M &:= g^{-1}(c) \end{aligned}$$

$f|_M$ has local extrema at p (except in degenerate cases)

Pf. is just like previous Purkiss.

Modern Purkiss

M smooth mfd

$$\mathcal{Z}: M \rightarrow M$$

$$\mathcal{Z}: M \rightarrow M$$

\mathcal{G} finite group of smooth maps

$$\exists p \in M \quad \mathcal{Z}(p) = p \quad \forall \mathcal{Z} \in \mathcal{G}$$

$f: M \rightarrow \mathbb{R}$ smooth

$$f(\mathcal{Z}x) = f(x) \quad \forall x \in M, \text{ all } \mathcal{Z} \in \mathcal{G}$$

$d\mathcal{Z}_p: T_p M \rightarrow T_p M$ has no nontrivial invariant subspaces [irreducibility]

$$df_p = 0 \quad (\text{critical pt})$$

If non-degen, p is a local extrema of f on M

pf

f gives us $df_p: T_p M \rightarrow \mathbb{R}$

$\ker(df_p)$ is invariant under all $\mathcal{Z} \in \mathcal{G}$

Irreducibility $\Rightarrow df_p = 0$

Hence p is critical pt of f

2nd order term for " $d^2 f$ " induce quadratic form \mathcal{Q} on tangent sp $T_p M$

"clearly" $\mathcal{Q}(\mathcal{Z}v) = \mathcal{Q}(v) \quad \forall v \in T_p M, \mathcal{Z} \in \mathcal{G}$