Constrained critical points and Lagrange multipliers.

Let $M \subset \mathbb{R}^n$ be a manifold.

Recall that a C^k function $f : M \to \mathbb{R}$ is a function s.t. $\forall\, b \in M$, $\exists$ an open $U \subset \mathbb{R}^n$ s.t. $b \in U$ and $\exists\ \tilde{f} : U \to \mathbb{R}$ s.t. $\tilde{f}\big|_{U \cap M} = f\big|_U$ and $\tilde{f}$ is C^k.

$\tilde{f}$ is called a local extension of f.

Also recall that, when $k \geq 1$, the derivative of f at b is the restriction of $D\tilde{f}(b)$ to $T_b M$.

$$Df(b) := D\tilde{f}(b)\Big|_{T_b M}$$

$$\begin{tikzcd}
D\tilde{f}(b) : \mathbb{R}^n \arrow[r] & \mathbb{R} \\
T_b M \arrow[u, hook] \arrow[ur, "Df(b)"'] &
\end{tikzcd}$$

or $\quad Df(b) : T_b M \hookrightarrow \mathbb{R}^n \xrightarrow[D\tilde{f}(b)]{} \mathbb{R}$

<u>Definition 3.7.1</u>: If $M \subset \mathbb{R}^n$ is a manifold and $f : M \to \mathbb{R}$ is C^1. Then $b \in M$ is a critical point of f if $Df(b) = 0$

<u>Note</u>: This means that, for a local extension $\tilde{f}$ of f

$$D\tilde{f}(b)\Big|_{T_b M} = 0 \quad \text{or} \quad T_b M \subset \ker D\tilde{f}(b).$$

Like before, we have

<u>Theorem 3.7.2:</u> $M \subset \mathbb{R}^n$ a manifold, $f: M \to \mathbb{R}$ C^1.
If b is a local extremum of f, then b is a critical point of f.

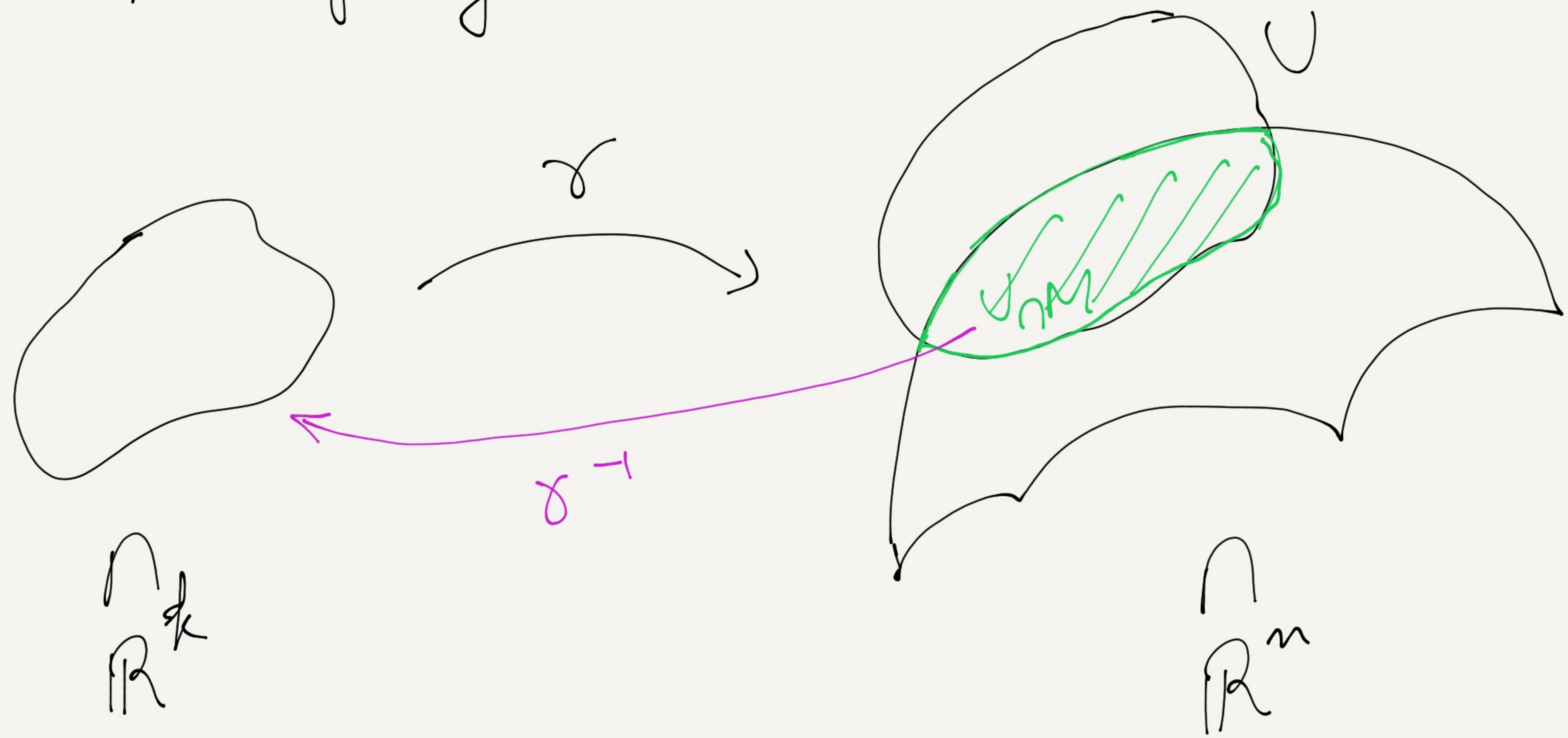
<u>Note:</u> if we know how to locally parameterize M at every point, $\forall\, b$: $\exists$ $V \underset{\text{open}}{\subset} \mathbb{R}^k$ and

$$\gamma : V \xrightarrow[C^1]{} M \quad \text{s.th} \quad \gamma(V) \ni b \quad (\dim M = k)$$

$$\gamma(V) = M \cap U \text{ for some open } U \subset \mathbb{R}^n$$

γ injective and $d\gamma(t)$ injective $\forall\, t \in V$.

We can find the critical points of f by finding those of $g = f \circ \gamma : V \longrightarrow \mathbb{R}$, we can also classify them using the quadratic form at each critical point for g.



In theory, we can always find local parametrizations (basically, by the definition of a manifold).

However, in practice, it can be very difficult to write down a local parametrization and use it to compute.

This is why we use Lagrange multipliers.

<u>Theorem and definition 3.7.5</u>:

$U \subset \mathbb{R}^n$ open $\quad F : U \to \mathbb{R}^m$ C^1 s.t.

if we define $\quad M = Z(F) = \text{zeros of } F = F^{-1}(0)$,

then $\forall\, b \in M \quad DF(b) : \mathbb{R}^n \to \mathbb{R}^m$ is onto.

(By the implicit function theorem, M is a manifold)

Let $f: U \longrightarrow \mathbb{R}$ be C^1. Then, $b \in M$ is a critical point of f iff $\exists\ \lambda_1, \dots, \lambda_m \in \mathbb{R}$ s.t.

$$Df(b) = \lambda_1 DF_1(b) + \lambda_2 DF_2(b) + \dots + \lambda_m DF_m(b).$$

The numbers $\lambda_1, \dots, \lambda_m$ are called Lagrange multipliers.

<u>Duality</u>: $\mathbb{R}^n$ is a vector space

The coordinates $x_1, \dots, x_n$ are linear maps on $\mathbb{R}^n$:

$$b = (b_1, \dots, b_n) \qquad x_i(b) = b_i, \ x_i : \mathbb{R}^n \to \mathbb{R}$$

$x_1, \dots, x_n$ span the set of all linear maps $\mathbb{R}^n \to \mathbb{R}$.

$$(\mathbb{R}^n)^\vee := \left\{ \begin{array}{l} l: \mathbb{R}^n \longrightarrow \mathbb{R} \text{ is linear if } l(\lambda b) = \lambda l(b) \\ \quad \forall \lambda \in \mathbb{R},\ b \in \mathbb{R}^n, \text{ and } l(b+c) = l(b) + l(c) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall\, b, c \in \mathbb{R}^n \end{array} \right\}$$

$\forall \ \ell : \mathbb{R}^n \longrightarrow \mathbb{R}$ linear, $\exists \ \lambda_1, \dots, \lambda_n \in \mathbb{R}$

s.t. $\ell = \lambda_1 x_1 + \dots + \lambda_n x_n$

So $(\mathbb{R}^n)^\vee$ is the set of all linear polynomials.

Any linear subspace of $\mathbb{R}^n$ is the set of zeros of some linear polynomials or linear forms.

e.g.: in $\mathbb{R}^3$: $\{ax + by + cz = 0\} \subset \mathbb{R}^3$ is a plane through the origin, a vector subspace.

or you can look at $Z\left\{\begin{array}{l} a_1 x + b_1 y + c_1 z = 0 \\ a_2 x + b_2 y + c_2 z = 0 \end{array}\right\} \subset \mathbb{R}^3$

can be a line or a plane (or everything if the coefficients are 0).

This can also be thought of as the kernel of a matrix: $\{ax+by+cz=0\}$ is the kernel of $\begin{pmatrix} a & b & c \end{pmatrix}$

$\left\{\begin{array}{l} a_1x+b_1y+c_1z=0 \\ a_2x+b_2y+c_2z=0 \end{array}\right\}$ is the kernel of $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$

We have a duality:

$$\mathbb{R}^n \times (\mathbb{R}^n)^\vee \longrightarrow \mathbb{R}$$
$$(v, l) \longmapsto l(v)$$

$\ker l = \{v \mid l(v)=0\} \subset \mathbb{R}^n$, we call this a hyperplane

(it is a manifold of dim. $n-1$)

Note $(\mathbb{R}^n)^\vee$ is a "copy" of $\mathbb{R}^n$:

$$\left\{ l = \lambda_1 x_1 + \cdots + \lambda_n x_n \right\} \longleftrightarrow \left\{ (\lambda_1, \ldots, \lambda_n) \right\}.$$

$$\ker l = \left\{ (b_1, \ldots, b_n) \;\middle|\; \sum_{i=1}^{n} \lambda_i b_i = 0 \right\}$$

Back to Lagrange multipliers :

$$M = Z(F) \subset U \qquad F : U \longrightarrow \mathbb{R}^m$$

$b \in M$. $\qquad T_b M = \ker DF(b)$

$$DF(b) = \begin{pmatrix} DF_1(b) \\ \vdots \\ DF_m(b) \end{pmatrix} \qquad DF_i(b) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

is a linear form.

$$T_b M = \ker DF(b) = \ker DF_1(b) \cap \cdots \cap \ker DF_m(b).$$

$Df(b) : \mathbb{R}^n \to \mathbb{R}$ is also a linear form.

Each time we have a linear subspace $V \subset \mathbb{R}^n$, we have $V^\perp \subset (\mathbb{R}^n)^\vee$

$$V^\perp := \{ \ell \in (\mathbb{R}^n)^\vee \mid \ell(v) = 0 \quad \forall\, v \in V \}.$$

$W \subset (\mathbb{R}^n)^\vee$: $W^\perp \subset \mathbb{R}^n$

$$W^\perp := \{ v \in \mathbb{R}^n \mid \ell(v) = 0 \quad \forall\, \ell \in W \}$$

e.g.: If $V \subset \mathbb{R}^3$ is a line,

$V^\perp$ is the set of equations of all the planes containing that line.

If V and W are both linear, then

$\dim V + \dim V^\perp = n$ and $\dim W + \dim W^\perp = n$.