Introduction to Smooth Submanifolds

Nikolai Nowaczyk

September 2008

Contents

1 Basics of Real Analysis

1.1 Theorem (Implicit Functions Theorem)**.** Let $U \times V \subset \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ be open and $F \in \mathscr{C}^r(U \times V, \mathbb{R}^m)$. Let $(a,b) \in U \times V$ such that $F(a,b) = 0$ and

$$\det(D_v F(a,b)) \neq 0$$

Then there are open neighbourhoods U' of a and V' of b and a mapping $g \in \mathscr{C}^r(U', V')$ satisfying $g(a) = b$ and

$$\forall (u,v) \in U' \times V' : F(u,v) = 0 \Leftrightarrow v = g(u)$$

1.2 Theorem (Inverse Function Theorem)**.** Let $U \subset \mathbb{R}^n$ be open and $F : U \to \mathbb{R}^n$ be of class $\mathscr{C}^r$. Let $a \in U$ satisfying

$$\det(DF(a)) \neq 0$$

Then there exist open neighbourhoods U' of a and V' of $b := f(a)$ such that $F\|_U : U' \to V'$ is a Diffeomorphism of class $\mathscr{C}^r$.

1.3 Definition (Immersion)**.** Let $W \subset \mathbb{R}^k$ be an open subset. A mapping $\psi : W \to \mathbb{R}^n$ is an *immersion of class* $\mathscr{C}^r$ if $\psi \in \mathscr{C}^r(W, \mathbb{R}^n)$ and

$$\forall w \in W : \operatorname{rg}(D\psi(w)) = k$$

2 Smooth Submanifolds of $\mathbb{R}^n$

In this section we will state four criterions a subset $M \subset \mathbb{R}^n$ can fulfill and proof their equivalence, which is the main result of this section. A *submanifold* is defined to be a subset M fulfilling at least one (and thus all) of these four conditions. We will proof their equivalence via

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)} \Rightarrow \text{(i)}$$

As a preparation it is extremely helpful to clarify some conventions of notation:

2.1 Convention.

(i) Spaces: Although $\mathbb{R}^n$ is $\mathbb{R}^n$ we are in fact dealing with two different versions of it: The submanifold M is the space itself, we are interested in. It is contained in a *surrounding space* $\mathbb{R}^n_s := \mathbb{R}^n$. We will also be dealing with a *coordinate space* $\mathbb{R}^n_c := \mathbb{R}^n$ used only to describe M. To understand the following proof it could be helpful to distinguish these spaces from one another. If it doesn't help you, then you can completely ignore it.

(ii) Integers: The dimension of the surrounding space as well as the coordinate space will be denoted by n. The dimension of the submanifold is k and it's class is r.

(iii) Canonical Isomorphism: Throughout the proof we will identify $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$.

(iv) Multidimensional Zero: We denote $0_k := (0, \dots, 0) \in \mathbb{R}^k$

(v) Neighbourhoods: We will have to deal with open neighbourhoods of points in the surrounding space as well as in the coordinate space. Our neighbourhood in $\mathbb{R}^n_s$ will be denoted with S and a corresponding neighbourhood in $\mathbb{R}^n_c$ will be denoted by C. It will also be necessary to write these neighbourhoods as cartesian products of k-dimensional and $(n-k)$-dimensional neighbourhoods. Of course this is not always possible. But since all norms are equivalent, chosing the maximum-norm and the euclidian norm we obtain open subsets $U \subset \mathbb{R}^k_s$ and $V \subset \mathbb{R}^{n-k}_s$ such that $U \times V \subset S$. Without loss of generality we can always assume that $S = U \times V$, because we will always be dealing with local questions. Similiarly we will decompose C into $C = W \times O$ where $W \subset \mathbb{R}^k_c$, $O \subset \mathbb{R}^{n-k}_c$ are both open subsets of their spaces. It will often be necessary to slightly modify a neighbourhood U for example. This will then be denoted by U'.

2.2 Theorem (Characterization of Submanifolds)**.** Let $M \subset \mathbb{R}^n_s$ be a non-empty subset and $r, k \in \mathbb{N}$, $r > 0$, $k \geq 0$. Then the following are equivalent:

(i) "M is locally immersed": For any $p \in M$ there is an open neighbourhood $S = U \times V \subset \mathbb{R}^n_s$ of p and an open subset $W \subset \mathbb{R}^k_c$ and an immersion $\psi : W \to S$ of class $\mathscr{C}^r$ such that

$$M \cap S = \psi(W)$$

and $\psi : W \to M \cap S$ is a homeomorphism. Such a ψ is called *local parametrization.*

(ii) "M is locally a graph": For any $p \in M$ there exist (after renumbering the coordinates) an open neighbourhood $U \times V \subset \mathbb{R}^k_s \times \mathbb{R}^{n-k}_s$ of p and a mapping $g : U \to V$ of class $\mathscr{C}^r$ such that: $\forall (u,v) \in U \times V : (u,v) \in M \Leftrightarrow v = g(u)$. In other words:

$$M \cap (U \times V) = \Gamma_g \cap (U \times V)$$

(iii) "M is locally a zero set": For any $p \in M$ there exists an open neighbourhood $U \times V \subset \mathbb{R}^n_s$ and a function $F : U \times V \to \mathbb{R}^{n-k}_c$ of class $\mathscr{C}^r$ such that

$$M \cap (U \times V) = \{(u,v) \in U \times V : F(u,v) = 0\} = F^{-1}(\{0\}) \cap (U \times V)$$

and for all $(u,v) \in M \cap (U \times V)$: $\operatorname{rg}(DF(u,v)) = n - k$. Such an F is called *locally defining function.*

(iv) "M is locally euclidian": M is a submanifold of dimension k and class $\mathscr{C}^r$, so: For any $p \in M$ there exists an open neighbourhood $S \subset \mathbb{R}^n_s$ of p, an open subset $C = W \times O \subset \mathbb{R}^k_c \times \mathbb{R}^{n-k}_c = \mathbb{R}^n_c$ and a $\mathscr{C}^r$-diffeomorphism $\varphi : S \to W \times O$, such that:

$$\varphi(M \cap S)) = W \times 0_{n-k}$$

Such a φ is called a *chart.*

Proof. First of all let's introduce some notation: Let $\pi_K : \mathbb{R}^n \to \mathbb{R}^k$

$$(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) \mapsto (x_1, \ldots, x_k)$$

be the projection of $\mathbb{R}^n$ onto the first k coordinates and $\pi^K : \mathbb{R}^n \to \mathbb{R}^{n-k}$

$$(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) \mapsto (x_{k+1}, \ldots, x_n)$$

be the projection of $\mathbb{R}^n$ onto the last $n-k$ coordinates.

"(i)⇒(ii)": Let $p \in M$ be arbitrary, $\psi : W \subset \mathbb{R}^k_c \to (U \times V) \subset \mathbb{R}^n_s$ be the immersion defined in the hypothesis and $x := \psi^{-1}(p)$. Since ψ is a local homeomorphism x is well defined. By definition (c.f. 1.3) the immersion ψ satisfies $\operatorname{rg}(D\psi(x)) = k$. So after renumbering the coordinates if if necessary we may assume that

$$\det\left(\left(\frac{\partial \psi_i(x)}{\partial x_j}\right)_{1 \le i,j \le k}\right) = \det(D(\pi_K \circ \psi)(x)) \neq 0$$

Define $\tilde{\psi} : W \subset \mathbb{R}^k_c \to U \subset \mathbb{R}^k_s$, $\tilde{\psi} := \pi_k \circ \psi = (\psi_1, \ldots, \psi_k)$. Then the inverse functions theorem (1.2) is applicable to $\tilde{\psi}$. So there exist open neighbourhoods $W' \subset W$ of x and $U' \subset U$ of $\pi_K(p)$ such that the restriction $\tilde{\psi} : W' \to U'$ is a class $\mathscr{C}^r$ - Diffeomorphism. Let's denote it's inverse by $\varphi : U' \to W'$. Now define $G := (G_1, \ldots, G_n) := \psi \circ \varphi : U' \subset \mathbb{R}^k_s \to \psi(W') \subset \mathbb{R}^n_s$. Then for any $u = (u_1, \ldots, u_k) \in U'$:

$$\begin{aligned}
G(u) =& (\psi \circ \varphi)(u) \\
=& (\psi_1(\varphi(u)), \ldots, \psi_k(\varphi(u)), \psi_{k+1}(\varphi(u)), \ldots \psi_n(\varphi(u))) \\
=& (\tilde{\psi}_1(\varphi(u)), \ldots, \tilde{\psi}_k(\varphi(u)), G_{k+1}(u), \ldots G_n(u) \\
=& (u_1, \ldots, u_k, G_{k+1}(u), \ldots G_n(u) \\
=& (u, (\pi^K \circ G)(u))
\end{aligned}$$

The mapping $g := \pi^K \circ G : U' \subset \mathbb{R}^k_s \to \pi^K(\psi(W')) =: V' \subset \mathbb{R}^k_s$ we just defined has all desired propertys: Let $(u, v) \in U' \times V'$ be arbitrary. Suppose $v = g(u)$. Then it follows

$$(u, v) = (u, g(u)) = G(u) = \psi(\varphi(u)) \in \psi(W') \subset \psi(W) = M \cap S \subset M$$

since $\psi(W) = M \cap S$ by hypothesis, $W' \subset W$ and $S' \subset S$. Conversely suppose that $(u, v) \in M$. Then

$$\exists! w \in W' : (u, v) = \psi(w) = (\tilde{\psi}_1(w), \ldots, \tilde{\psi}_k(w), \psi_{k+1}(w), \ldots, \psi_n(w))$$

So $u = \tilde{\psi}(w)$ and since $\tilde{\psi}$ is a diffeomorpism $\varphi(u) = \varphi(\tilde{\psi}(w)) = w$. This implies

$$(u, v) = \psi(w) = \psi(\varphi(u)) = G(u) = (u, g(u))$$

and thus $v = g(u)$.
"(ii)⇒(iii)": Let $g : U \subset \mathbb{R}^k_s \to V \subset \mathbb{R}^{n-k}_s$ as in the hypothesis and $p \in U \times V$. Define a mapping $F = (F_1, \ldots, F_{n-k}) : U \times V \subset \mathbb{R}^n_s \to \mathbb{R}^{n-k}_c$ by

$$F(u, v) := v - g(u)$$

Then F is of class $\mathscr{C}^r$ and for any $(u, v) \in U \times V$:

$$F(u, v) = 0 \Leftrightarrow v = g(u)(u, v) \Leftrightarrow \in M$$

So $M \cap (U \times V)$ is the zero set of F. Finally $D_v F(u,v) = I_{n-k} \in \mathbb{R}^{(n-k)\times(n-k)}$. Since the unit matrix I_{n-k} has rank $n-k$ and $DF(u,v) \in \mathbb{R}^{(n-k)\times n}$ it follows that for any $(u,v) \in U \times V : DF(u,v) = n-k$.

"(iii)⇒(iv)": Let $p \in M$ and $F : U \times V \to \mathbb{R}_c^{n-k}$ be as in the hypothesis. After renumbering the coordinates if necessary we may assume, that

$$\det(D_v F(p)) = \det\left(\frac{\partial(F_1,\ldots,F_{n-k})}{\partial(v_1,\ldots,v_{n-k})}(p)\right) \neq 0$$

Define $\mathrm{id}_K : \mathbb{R}_s^k \to \mathbb{R}_c^k$, $x \mapsto x$ and $\mathrm{id}_K \times F : U \times V \to \mathbb{R}_c^n$, $(u,v) \mapsto (u, F(u,v))$. Then the inequality above implies:

$$\det(D(\mathrm{id}_K \times F)(p)) = \det\begin{pmatrix} I_k & 0 \\ D_u F(p) & D_v F(p)\end{pmatrix} \neq 0$$

as well. By the inverse functions theorem (1.2) there is an open neighbourhood $S' \subset U \times V$ of p and an open neighbourhood $C' := W' \times O'$ of $q := (\mathrm{id}_K \times F)(p)$ such that the restriction $\varphi := \mathrm{id}_K \times F : S' \to W' \times O'$ is a diffeomorphism of class $\mathscr{C}^r$. It also fulfills the required propertys because for any $(u,v) \in S'$:

$$(u,v) \in M \Leftrightarrow F(u,v) = 0 \Leftrightarrow \varphi(u,v) = (\mathrm{id}_K \times F)(u,v) = (u, F(u,v)) = (u,0) \in W' \times 0_{n-k}$$

"(iv)⇒(i)": Let $p \in M$ be arbitrary and let $\varphi : S \to W \times O$ be a chart, i.e. a diffeomorphism of class $\mathscr{C}^r$ as in the hypothesis. Define $\psi : W \to S$ by

$$w \mapsto \varphi^{-1}(w, 0_{n-k})$$

Since φ^{-1} is diffeomorphism of class $\mathscr{C}^r$ as well, ψ is an immersion of class $\mathscr{C}^r$ and we have by hypothesis:

$$\psi(W) = \varphi^{-1}(W, 0_{n-k}) = M \cap S$$

□

2.3 Definition (Submanifold)**.** A subset $\varnothing \neq M \subset \mathbb{R}^n$ is a *differentiable submanifold*, if it satisfies one of the conditions listed in theorem 2.2 above. The integer r is the *class* of M. M is *smooth* if it is of class r for any $r \in \mathbb{N}$, $r > 0$. We call

$$\dim M := k \qquad\qquad \operatorname{codim} M := n-k$$

the *dimension* and *codimension* of M.

3 Tangential and Normal Spaces

3.1 Definition (Tangential Bundle)**.** Let $M \subset \mathbb{R}^n$ be a non-emtpy subset and $p \in M$ be arbitrary. A vector $v \in \mathbb{R}^n$ is a *tangent to M at p* if there exist $\varepsilon > 0$, $\gamma \in \mathscr{C}^1(]-\varepsilon,\varepsilon[, M)$, such that $\gamma(0) = p$ and $\dot\gamma(0) = v$. The set

$$T_pM := \{v \in \mathbb{R}^n | v \text{ is tangent to M at p }\}$$

is the *tangential space* of M at p. Their disjoint union

$$TM := \coprod_{p \in M} T_pM := \{(v,p) | v \in T_pM, p \in M\}$$

is the *tangential bundle* of M.

3.2 Definition (Normal Bundle)**.** Let $M \subset \mathbb{R}^n$ be a non-emtpy subset, $p \in M$ be arbitrary and $\langle \_, \_ \rangle$ be the euclidian standard scalar product on $\mathbb{R}^n$. Then

$$N_pM := \{w \in \mathbb{R}^n | \forall v \in T_pM : \langle w, v \rangle = 0\}$$

is the *normal space* of M at p. A vector $v \in N_pM$ is *normal to M at p*. Analogously

$$NM := \coprod_{p \in M} N_pM$$

is the *normal bundle*.

3.3 Theorem (Propertys of Tangential and Normal Spaces)**.** Let M be a k-dimensional submanifold of class $\mathscr{C}^r$. Then for any $p \in M$

(i) The sets T_pM and N_pM are vector spaces and there is a direct and orthogonal decomposition

$$\mathbb{R}^n = T_pM \oplus N_pM$$

(ii) We have the dimension formulae

$$\dim T_pM = \dim M = k \qquad \dim N_pM = \operatorname{codim} M = n - k$$

(iii) If $F : U \times V \subset \mathbb{R}^n_s \to \mathbb{R}^{n-k}_c$, $p \in U \times V$, is a locally defining function for M at p, then

$$T_pM = \ker dF(p) \qquad N_pM = \operatorname{im} dF(p)$$

The $(\operatorname{grad} F_1(p), \ldots, \operatorname{grad} F_{n-k}(p))$ are a basis of N_pM.

(iv) If $\psi : W \subset \mathbb{R}^k_c \to S \subset \mathbb{R}^n_s$, is a local parametrization at p such that $p \in S$, $q := \psi^{-1}(p) \in W$, T_pM is also given by:

$$T_pM = \langle \partial_1\psi(q), \ldots, \partial_k\psi(q) \rangle = \operatorname{im} d\psi(q)$$

The $(\partial_1\psi(q), \ldots, \partial_k\psi(q))$ are a basis of T_pM.

Proof.
"$\operatorname{im} d\psi(q) \subset T_pM$": Let $v \in \operatorname{im} d\psi(q)$ be arbitrary, thus

$$v = \sum_{i=1}^{k} c_i \partial_i \psi(q)$$

where $c = (c_1, \ldots, c_k) \in \mathbb{R}^k$ is a coordinate vector of v. Define a curve $\gamma :]-\varepsilon, \varepsilon[\to M$, $t \mapsto \psi(q + ct)$, where $\varepsilon > 0$ is sufficiently small, such that $\forall t \in]-\varepsilon, \varepsilon[$: $q + ct \in W$. Then $\gamma(0) = \psi(q) = p$, $\gamma \in \mathscr{C}^1(W, S)$ and by chain rule

$$\dot\gamma(0) = \psi'(\gamma(0))c = \sum_{i=1}^{k} c_i \partial_i \psi(q) = v$$

Thus $v \in T_pM$. By hypothesis $\dim(\operatorname{im} d\psi(q)) = \operatorname{rg}(d\psi(q)) = k$.
"$T_pM \subset \ker dF(p)$": Let $v \in T_pM$ and $\gamma :]-\varepsilon, \varepsilon[\to M$ be a curve satisfying $\gamma(0) = p$, $\dot\gamma(0) = v$. Then by the chain rule

$$0 = F \circ \gamma \Rightarrow 0 = dF(\gamma(0))(\dot\gamma(0)) = dF(p)(v) \Rightarrow v \in \ker dF(p)$$

So $T_pM \subset \ker dF(p)$. By hypothesis $\operatorname{rg}(dF(p)) = n - k$ and thus $\dim(\ker dF(p)) = k$.
These two parts proof, that T_pM is indeed a vector space of dimension k. It immediately follows that

$N_pM = (T_pM)^\perp$ and thus N_pM is a vector space of dimension $n-k$. So part (i) and (ii) is shown and bringing all this together it follows $\operatorname{im} d\psi(q) = T_pM$ - part (iv) - and $T_pM = \ker dF(p)$. The last part of (iii) follows very simillar:
"$\operatorname{im} dF(p) \subset N_pM$": Let $v \in T_pM \subset \ker dF(p)$. So

$$\forall 1 \leq j \leq n-k : 0 = \langle \operatorname{grad} F_j(p), v\rangle$$

wich implies

$$\operatorname{im} dF(p) = \langle \operatorname{grad} F_1(p), \dots, \operatorname{grad} F_{n-k}(p)\rangle \perp v$$

thus - since $v \in T_pM$ was arbitrary - $\operatorname{im} dF(p) \subset N_pM$.
This again implies $\operatorname{im} dF(p) = N_pM$, because by hypothesis and (ii) $\dim(\operatorname{im} dF(p)) = \dim N_pM = n-k$. □

4 Extremal Problems under Restraints

4.1 Definition (Local Extrema)**.** Let $X \subset \mathbb{R}^n$ a an arbitrary subset, $f : X \to \mathbb{R}$ be a function and $M \subset X$. A point $a \in M$ is a *local minimum of f in M* if there is a neighbourhood $U \subset M$ (with regard to the subspace topology of M seen as a metric subspace of $\mathbb{R}^n$) such that

$$\forall x \in U : f(x) \geq f(a)$$

Analogously if

$$\forall x \in U : f(x) \leq f(a)$$

we say a is a *local maximum of f in M*. We say a is a *local extremum of f in a* if it is either a local minimum or maximum of f in M.

4.2 Theorem (Extrema under restraints)**.** Let $U \subset \mathbb{R}^n$ be open, $M \subset U$ be a k-dimensional manifold of class $\mathscr{C}^1$ given by a globally defining function $g \in \mathscr{C}^1(U, \mathbb{R}^{n-k})$, i.e.

$$M = \{x \in U : g(x) = 0\} \qquad \forall x \in U : \operatorname{rg}(dg(x)) = n-k$$

Let $f \in \mathscr{C}^1(U, \mathbb{R})$ be a function having a local extremum at $p \in M$. Then

$$\operatorname{grad} f(p) \in N_pM$$

which is equivalent to:

$$\exists \lambda = (\lambda_1, \dots, \lambda_{n-k}) \in \mathbb{R}^{n-k} : \operatorname{grad} f(p) = \sum_{j=1}^{n-k} \lambda_j \operatorname{grad} g_j(p)$$

Proof. Let $v \in T_pM$ be arbitrary. Thus there is a curve $\gamma :]-\varepsilon, \varepsilon[\to M$, such that $\gamma(0) = p$, $\dot\gamma(0) = v$. Define $h :]-\varepsilon, \varepsilon[\to \mathbb{R}$ by $h := f \circ \gamma$. Since f has a local extremum at p in M, and $\operatorname{im}\gamma \subset M$, h has a local extremum at 0. So by elementary calculus and the chain rule we obtain

$$0 = h'(0) = \nabla f(\gamma(0))(\dot\gamma(0)) = \nabla f(p)v = \langle \operatorname{grad} f(p), v\rangle$$

Thus $\operatorname{grad} f(p) \perp v$. Since v was arbitrary $\operatorname{grad} f(p) \in N_pM$. Finally theorem 3.3,(iii) states that

$$N_pM = \operatorname{im} dg(p) = \langle \operatorname{grad} g_1(p), \dots, \operatorname{grad} g_{n-k}(p)\rangle$$

□

4.3 Remark.

(i) The scalars $\lambda_1, \ldots, \lambda_{n-k}$ are called *Lagrangian Multiplicators.*

(ii) One says that the condition $\operatorname{grad} f(p) \in N_p M$ is a *necassary condition* for f to have a local extremum at p under the *restraint conditions* $g_1(x) = \ldots = g_{n-k}(x) = 0$. This condition is not sufficient in general.

(iii) This can also be expressed by

$$\operatorname{grad} f(p) + \sum_{j=1}^{n-k} \lambda_j \operatorname{grad} g_j(p) = 0$$

by simply changing the signs of λ_j.

(iv) Alltogether we obtain the following non-linear system of equations as a necessary condition for p to be a local extrema of f in M:

$$\begin{aligned}
&(1): \ \partial_1 f(p) = \left(\sum_{j=1}^{n-k} \lambda_j \operatorname{grad} g_j(p) \right)_1 \\
&\ldots \\
&(n): \ \partial_n f(p) = \left(\sum_{j=1}^{n-k} \lambda_j \operatorname{grad} g_j(p) \right)_n \\
&(n+1): \ g_1(p) = 0 \\
&\ldots \\
&(n+k): \ g_{n-k}(p) = 0
\end{aligned}$$

It consists of the $n+k$ unknowns $p_1, \ldots, p_n, \lambda_1, \ldots, \lambda_{n-k}$ and the $n+k$ equations listed above.

4.4 Theorem. Every symmetric matrix $A \in \mathbb{R}^{n \times n}$ has an eigenvalue.

Proof. Define $F : \mathbb{R}^n \to \mathbb{R}$ by $x \mapsto x^T A x$. Then F is differentiable and $\nabla F(x) = 2x^T A$ since

$$\begin{aligned}
&\lim_{h \to 0} \frac{|F(x+h) - F(x) - 2x^T Ah|}{\|h\|} = \lim_{h \to 0} \frac{|(x+h)^T A(x+h) - x^T Ax - 2x^T Ah|}{\|h\|} \\
&= \lim_{h \to 0} \frac{|x^T Ax + x^T Ah + h^T Ax + h^T Ah - x^T Ax - 2x^T Ah|}{\|h\|} \\
&= \lim_{h \to 0} \frac{|h^T Ah\|}{\|h\|} = \lim_{h \to 0} \frac{|\langle Ah, h\rangle|}{\|h\|} \le \lim_{h \to 0} \frac{\|Ah\|\|h\|}{\|h\|} = \lim_{h \to 0} \|Ah\| = 0
\end{aligned}$$

where we have used the Cauchy-Schwarz-Inequality and the symmetry of A. Define $g : \mathbb{R}^n \to \mathbb{R}$ by, $x \mapsto \|x\|^2 - 1$. Then

$$M := \mathbb{S}^{n-1} = \{x \in \mathbb{R}^n | g(x) = 0\}$$

Then g is differentiable as well and

$$\nabla g(x) = 2\|x\| \frac{x^T}{\|x\|} = 2x^T$$

So g is especially continuous and thus M is the reversed image of the closed subset $\{0\} \subset \mathbb{R}$. Since it's obviously bounded, M is compact. The mapping F is continuous as well, so $F|_M : M \to \mathbb{R}$ attains it's extremal values. Let $a \in M$ be such an extremal value. By the theorem 4.2 above there is a $\lambda \in \mathbb{R}$ such that

$$\operatorname{grad} F(a) = \lambda \operatorname{grad} g \Rightarrow 2Aa = \lambda \operatorname{grad} g(a) = 2\lambda a \Rightarrow Aa = \lambda a$$

Since $\|a\| = 1$ it follows that $a \neq 0$ and thus λ is an eigenvalue of A.

$\square$

References

[1] Lee, Introduction to Smooth Manifolds

[2] Forster, Analysis II

[3] Königsberger, Analysis II