THE MORSE LEMMA FOR BANACH SPACES

BY RICHARD S. PALAIS[1]

Communicated April 28, 1969

THEOREM. *Let V be a Banach space, $\mathcal{O}$ a convex neighborhood of the origin in V and $f: \mathcal{O} \to \mathbf{R}$ a C^{k+2} function ($k \geqq 1$) having the origin as a nondegenerate critical point, with $f(0)=0$. Then there is a neighborhood U of the origin and a C^k diffeomorphism $\phi: U \to \mathcal{O}$ with $\phi(0)=0$ and $(D\phi)_0 =$ the identity map of V such that for $x \in U$ $f(\phi(x)) = \frac{1}{2}(D^2f)_0(x, x)$.*

REMARK. The above theorem is a classical result of Marston Morse in the case that V is finite dimensional and was generalized by the author to the case that V is a Hilbert space [**1**], [**3**]. The latter proof makes use of operator theory in Hilbert space and does not extend in any obvious way to more general Banach spaces. The proof we give below is completely elementary and works for arbitrary V of course. Recent developments in the calculus of variations from a Banach manifold point of view (see for example [**4**]) make it desirable to have the theorem in this degree of generality.

The technique behind our proof was pioneered by J. Moser in a somewhat different and finite dimensional setting [**2**]. The present paper was inspired by a recent result of A. Weinstein [**5**] where Moser's method is adapted to the Banach manifold setting.

Put $f=f^1$ and define $f^0: \mathcal{O} \to \mathbf{R}$ by $f^0(x) = \frac{1}{2}(D^2f)_0(x, x)$. Define $f^t: \mathcal{O} \to \mathbf{R}$ $0 \leqq t \leqq 1$ by $f^t = f^0 + t(f^1 - f^0)$ and note that $(t, x) \mapsto f^t(x)$ is a C^{k+2} map of $I \times \mathcal{O} \to \mathbf{R}$. Note also that if we define $\dot{f}^t: \mathcal{O} \to \mathbf{R}$ by $\dot{f}^{t_0}(x) = d/dt|_{t=t_0} f^t(x)$ then clearly $\dot{f}^t = f^1 - f^0$.

Moser's trick, suitably adapted, is to look for a smooth one-parameter family ϕ_t of C^k local diffeomorphisms, defined in a neighborhood U of 0, with ϕ_0 the identity, such that $f^t \circ \phi_t = f^0$ in U. This, of course, gives the theorem by taking $\phi = \phi_1$, since $f = f^1$. On its face we seem to have replaced our problem by a harder one; however, the nonlinear equation $f^t \circ \phi_t = f^0$ for ϕ_t is equivalent to $d/dt(f^t \circ \phi_t) = 0$ which, as we shall see, is in turn equivalent to $Df^t(X^t) = -\dot{f}^t$, where X^t is the time dependent vector field generating ϕ_t. The latter is a *linear* equation for X^t which can be solved explicitly.

Suppose then that X^t, $0 \leqq t \leqq 1$ is a C^k time dependent vector field on $\mathcal{O}$ (i.e. $(t, x) \mapsto X^t_x$ is a C^k map of $I \times \mathcal{O} \to V$) with $X^t_0 = 0$ and $(DX^t)_0 = 0$. Let $t \mapsto \phi_t(x)$ be the maximum integral curve of X^t starting at

[1] Research supported in part by Air Force grant AFOSR 68-1403.