LEMMA 2.2 (Lemma of Morse). Let p be a non-degenerate critical point for f. Then there is a local coordinate system $(y^1,\dots,y^n)$ in a neighborhood U of p with $y^i(p) = 0$ for all i and such that the identity

$$f = f(p) - (y^1)^2 - \dots - (y^\lambda)^2 + (y^{\lambda+1})^2 + \dots + (y^n)^2$$

holds throughout U, where λ is the index of f at p.

PROOF: We first show that if there is any such expression for f, then λ must be the index of f at p. For any coordinate system $(z^1,\dots,z^n)$, if

$$f(q) = f(p) - (z^1(q))^2 - \dots - (z^\lambda(q))^2 + (z^{\lambda+1}(q))^2 + \dots + (z^n(q))^2$$

then we have

$$\frac{\partial^2 f}{\partial z^i\,\partial z^j}(p) = \begin{cases} -2 & \text{if } i = j \le \lambda\,, \\ 2 & \text{if } i = j > h\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

which shows that the matrix representing f_* with respect to the basis $\frac{\partial}{\partial z^1}\Big|_p,\dots,\frac{\partial}{\partial z^n}\Big|_p$ is

$$\begin{pmatrix} -2 & & & & & \\ & \ddots & & & & \\ & & -2 & & & \\ & & & 2 & & \\ & & & & \ddots & \\ & & & & & 2 \end{pmatrix}.$$

Therefore there is a subspace of TM_p of dimension λ where f_{**} is negative definite, and a subspace V of dimension $n-\lambda$ where f_{**} is positive definite. If there were a subspace of TM_p of dimension greater than λ on which f_{**} were negative definite then this subspace would intersect V, which is clearly impossible. Therefore λ is the index of f_{**}.

We now show that a suitable coordinate system $(y^1,\dots,y^n)$ exists. Obviously we can assume that p is the origin of $\mathbf{R}^n$ and that $f(p) = f(0) =$ By 2.1 we can write

$$f(x_1,\dots,x_n) = \sum^{n} x_i g_i(x_1,\dots,x_n)$$

for $(x_1,\dots,x_n)$ in some neighborhood of 0. Since 0 is assumed to be a critical point:

$$g_j(0) = \frac{\partial f}{\partial x^j}(0) = 0\,.$$

Therefore, applying 2.1 to the g_j we have

$$g_j(x_1,\dots,x_n) = \sum_{i=1}^{n} x_i h_{ij}(x_1,\dots,x_n)$$

for certain smooth functions h_{ij}. It follows that

$$f(x_1,\dots,x_n) = \sum_{i,j=1}^{n} x_i x_j h_{ij}(x_1,\dots,x_n)\,.$$

We can assume that $h_{ij} = h_{ji}$, since we can write $\bar h_{ij} = \frac{1}{2}(h_{ij} + h_{ji})$, and then have $\bar h_{ij} = \bar h_{ji}$ and $f = \sum x_i x_j \bar h_{ij}$. Moreover the matrix $(\bar h_{ij}(0))$ is equal to $\left(\frac{1}{2}\frac{\partial^2 f}{\partial x^i\,\partial x^j}(0)\right)$, and hence is non-singular.

There is a non-singular transformation of the coordinate functions which gives us the desired expression for f, in a perhaps smaller neighborhood of 0. To see this we just imitate the usual diagonalization proof for quadratic forms. (See for example, Birkhoff and MacLane, "A survey of modern algebra," p. 271.) The key step can be described as follows.

Suppose by induction that there exist coordinates $u_1,\dots,u_n$ in a neighborhood U_1 of 0 so that

$$f = \pm (u_1)^2 \pm \dots \pm (u_{r-1})^2 + \sum_{i,j\ge r} u_i u_j H_{ij}(u_1,\dots,u_n)$$

throughout U_1; where the matrices $(H_{ij}(u_1,\dots,u_n))$ are symmetric. After a linear change in the last $n-r+1$ coordinates we may assume that $H_{rr}(0) \ne 0$. Let $g(u_1,\dots,u_n)$ denote the square root of $|H_{rr}(u_1,\dots,u_n)|$. This will be a smooth, non-zero function of $u_1,\dots,u_n$ throughout some smaller neighborhood $U_2 \subset U_1$ of 0. Now introduce new coordinates $v_1,\dots,v_n$ by

$$v_i = u_i \qquad \text{for } i \ne r$$

$$v_r(u_1,\dots,u_n) = g(u_1,\dots,u_n)\Big[u_r + \sum_{i>r} u_i H_{ir}(u_1,\dots,u_n)/H_{rr}(u_1,\dots,u_n)\Big].$$

It follows from the inverse function theorem that $v_1,\dots,v_n$ will serve as coordinate functions within some sufficiently small neighborhood U_3 of 0. It is easily verified that f can be expressed as

$$f = \sum_{i\le r} \pm (v_i)^2 + \sum_{i,j>r} v_i v_j H'_{ij}(v_1,\dots,v_n)$$