Classification of Surfaces via Morse Theory

Abhijit Champanerkar Ajit Kumar S. Kumaresan

Introduction

In this article we classify all compact surfaces up to homeomorphism using Morse theory. The single most important tool is the gradient like flow associated with a Morse function. While it is well-known (and worked out in detail in [4] and [7]) how the homotopy type changes as when one passes a critical point, the change in homeomorphism type is perhaps not so well-known. Even in the books where it is done, technical details are inadequate and a beginning graduate student may have difficulty in filling them. The highlights of this exposition are the two theorems (Theorem 13 and theorem 16) which tell us how the homeomorphism type changes as we pass a critical point of index 0, 1 and 2. Once we prove these theorems the classification follows easily.

We have taken most of the ideas from [2]. We have modified and supplied details for most of the proofs and changed the style of presentation to make the exposition rigorous and lucid.

Section 1 deals with basic definitions and Morse Theorem. In Section 2 we introduce the gradient flow and use it to prove that the diffeomorphism type of a manifold does not change between the two levels which does not contain a critical point. This is the first application of the gradient flow and all the details have been painstakingly worked out. Section 3 includes modification of Morse functions and the proofs are an excellent illustration of the use of bump functions. Section 4 deals with one of the two important theorems regarding crossing of critical levels of index 0 and 2. In Section 5 we prove theorems regarding passing of critical level of index 1. For better understanding of the results proved, we have included classification of closed 1-manifolds and concrete examples such as sphere, torus and projective plane in Sections 4 and 5. Section 6 includes connected sums. We have not dealt with the technicalities of unambiguity and associativity of connected sums. In the last section we finish the classification using the results from earlier sections. The noteworthy point of the classification is that the orientable as well as non orientable cases are treated simultaneously.

Since we have aimed the article at fresh graduate students, we have supplied all details for most of the proofs. We also illustrate the theorems using examples and give simple applications of the theorem proved. We have included pictures wherever necessary to assist the geometric understanding of ideas and the results. We hope that this article will introduce the readers to some of the basic techniques and ideas of differential topology.

1 Critical Points and Morse Lemma

We assume all surfaces to be compact, connected, without boundary (closed) unless stated otherwise.

Definition 1. Let M be a smooth manifold and $f: M \to \mathbb{R}$, a smooth function on M. A point $p \in M$ is said to be a critical point of f if Df(p) is singular on $T_p(M)$. The image of p under f, that is, f(p) is called a critical value of f. Any real number which is not a critical value is called a regular value of f.

Example 1. If f is a constant function on M then all points of M are critical points.

Example 2. Let $M = S^2 \subseteq \mathbb{R}^3$ and f(x, y, z) = z. Then N = (0, 0, 1) and S = (0, 0, -1) are the two critical points of f.

Example 3. If M is a compact manifold then there exist at least two critical points for any nonconstant function f on M, namely the maximum and minimum of f.

Example 4. Let $M = T_1$ - the 2-dimensional torus and f be the height function on M. Then there are four critical points X_1, X_2, X_3, X_4 . See Fig. 1.

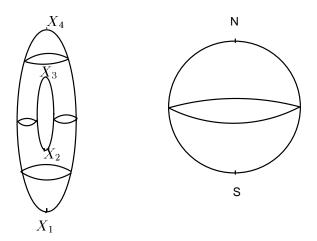


Figure 1: Critical Points on the torus and the sphere

Let M be a smooth manifold of dimension n. Let $\phi: U \to M$ be a parameterizations, where $U \subseteq \mathbb{R}^n$ is an open set containing 0. If $\phi(0) = p$, then we say that ϕ is centered at p.

Definition 2. Let M be a smooth manifold of dimension n and $f: M \to \mathbb{R}$, a smooth function on M. Let (ϕ, U) be a parameterization centered at p. Define $g := f \circ \phi$. The Hessian of f with respect to ϕ is a matrix defined as

$$H_{\phi}(f) = H(f \circ \phi) := (\frac{\partial^2 g}{\partial x_i \partial x_j}), 1 \le i, j \le n.$$

Let M be a surface. Let (ϕ, U) and (ψ, V) be two parameterization centered at p. Let (x, y) and (u, v) be coordinates w.r.t. ϕ and ψ respectively. Let $\theta(u, v) = (x(u, v), y(u, v))$ be the corresponding change of coordinates. Define $h(u, v) := g \circ \theta(u, v)$. Then $h = f \circ \psi$ and the partial derivatives of h are as follows:

$$\frac{\partial h}{\partial u} = g_x x_u + g_y y_u,
\frac{\partial h}{\partial v} = g_x x_v + g_y y_v,
\frac{\partial^2 h}{\partial u^2} = g_{xx} x_u^2 + 2g_{xy} x_u y_u + g_{yy} y_u^2 + g_x x_{uu} + g_y y_{uu},
\frac{\partial^2 h}{\partial v^2} = g_{xx} x_v^2 + 2g_{xy} x_v y_v + g_{yy} y_v^2 + g_x x_{vv} + g_y y_{vv},
\frac{\partial^2 h}{\partial u \partial v} = g_{xx} x_u x_v + g_{xy} x_u y_v + g_{xx} x_{uv} + g_{yx} x_v y_u + g_{yy} y_u y_v + g_y y_{uv}.$$

If p is a critical point of f then $g_x = 0 = g_y$ and hence the Hessians of f with respect to ϕ and ψ are related as follows:

$$H_{\psi}(f) = J^{t}(\theta) \circ H_{\phi}(f) \circ J(\theta) \tag{1}$$

where $J(\theta)$ denotes the Jacobian of θ .

Definition 3. Let f be a smooth function on a surface M. A critical point p of f is said to be non degenerate if $H_{\phi}(f)(p)$ is non singular for any parameterization ϕ centered at p.

It follows from Eq. 1 that this definition is independent of the parameterization.

Example 5. All critical points in Example 1 are degenerate. All critical points in Example 2 and 4 are non degenerate.

Definition 4. The index of a non degenerate critical point p of smooth function f on a surface M is the dimension of the maximal subspace of T_pM on which H(f) is negative definite.

The index of a critical point is independent of the parameterization follows from Sylvester's Law.

Remark 1. Concepts of non degeneracy and index also hold for any smooth manifold.

Example 6. In Example 4 X_1 is a critical point of index 0, X_2 and X_3 are critical points of index 1 and X_4 is of index 2.

Theorem 1 (Morse, 1932). Let M be a surface and $f: M \to \mathbb{R}$, a smooth function on M. Let $p \in M$ be a non degenerate critical point of f. Then there exists a parameterization (ϕ, U) centered at p and coordinates (X, Y) such that

$$f \circ \phi(X, Y) = f(p) + g_i(X, Y), \qquad 0 \le i \le 2,$$

where i is the index of p and g_i 's are defined as follows:

$$g_0(X,Y) = X^2 + Y^2$$
, $g_1(X,Y) = X^2 - Y^2$ and $g_2(X,Y) = -X^2 - Y^2$.

Define a map $g := f \circ \phi \colon U \to \mathbb{R}$ where ϕ is a parameterization centered at p. Then g is smooth and (0,0) is a non degenerate critical point of g. It is enough to prove the following form of the above theorem.

Theorem 2. Let $U \subseteq \mathbb{R}^2$ be a neighborhood of (0,0) and $f: U \to \mathbb{R}$ be a smooth function. Assume that (0,0) is a non degenerate critical point of f of index i for $0 \le i \le 2$. Then there exists a diffeomorphism $\phi: V \to \phi(V) \subseteq U$, where V is an open set containing (0,0) in \mathbb{R}^2 , $\phi(0,0) = (0,0)$ and a system of coordinates (X,Y) on U such that

$$f \circ \phi(X, Y) = f(0, 0) + g_i(X, Y).$$

Proof. By Taylor series expansion f near origin is of the form:

$$f(x,y) - f(0,0) = R(x,y)x^{2} + 2S(x,y)xy + T(x,y)y^{2}$$
(2)

where R, S and T are smooth functions defined as:

$$R(x,y) = \int_0^1 (1-t) f_{xx}(tx,ty) dt, \qquad R(0,0) = r,$$

$$S(x,y) = \int_0^1 (1-t) f_{xy}(tx,ty) dt, \qquad S(0,0) = s,$$

$$T(x,y) = \int_0^1 (1-t) f_{yy}(tx,ty) dt, \qquad T(0,0) = t.$$

Note that $R(0,0) = \frac{1}{2} f_{xx}(0,0)$, $S(0,0) = \frac{1}{2} f_{xy}(0,0)$ and $T(0,0) = \frac{1}{2} f_{yy}(0,0)$. In classical notation we say that (0,0) is a non degenerate critical point of f if $rt - s^2 \neq 0$.

Case 1. Let $rt - s^2 > 0$ and r > 0. Since R is continuous, there exists a neighbourhood $U_1 \subseteq U$ of (0,0) in which R and $RT - S^2$ remain positive. So we can write Eq. 2 as:

$$f(x,y) - f(0,0) = R(x,y)\left[x + y\frac{S(x,y)}{R(x,y)}\right]^2 + y^2\frac{R(x,y)T(x,y) - S^2(x,y)}{R(x,y)}.$$
 (3)

Define

$$\begin{array}{rcl} X & = & X(x,y) := & \sqrt{R(x,y)}[x+y\frac{S(x,y)}{R(x,y)}], \\ Y & = & Y(x,y) := & y\sqrt{\frac{R(x,y)T(x,y)-S^2(x,y)}{R(x,y)}}. \end{array}$$

Define $\theta \colon U_1 \to \mathbb{R}^2$ as $\theta(x,y) = (X(x,y),Y(x,y))$. Then $\theta(0,0) = (0,0)$ and Jacobian $J(\theta)(0,0) = \sqrt{rt-s^2} \neq 0$. Hence by inverse mapping theorem θ is invertible in some neighbourhood V of (0,0). Define $\phi(X,Y) = \theta^{-1}(X,Y) = (x(X,Y),y(X,Y))$. Then Eq. 3 becomes:

$$f\circ\phi(X,Y)=f(0,0)+X^2+Y^2=f(0,0)+g_0(X,Y).$$

Since r is positive, t is also positive and hence the index of (0,0) is 0.

Case 2. Let $rt - s^2 > 0$ and r < 0. Then we write Eq. 3 as:

$$f(x,y) - f(0,0) = -(-R(x,y))[x + y\frac{S(x,y)}{R(x,y)}]^2 - y^2 \frac{R(x,y)T(x,y) - S^2(x,y)}{-R(x,y)}.$$

Then define

$$X = X(x,y) := \sqrt{-R(x,y)} [x + y \frac{S(x,y)}{R(x,y)}],$$

$$Y = Y(x,y) := y \sqrt{\frac{R(x,y)T(x,y) - S^2(x,y)}{-R(x,y)}}.$$

Again proceeding as before we get

$$f \circ \phi(X,Y) = f(0,0) - X^2 - Y^2 = f(0,0) + g_2(X,Y).$$

Since r is negative, t is negative and hence index of (0,0) is 2.

Case 3. Let $rt - s^2 < 0$ and r > 0. Using similar arguments as before we get:

$$f \circ \phi(X,Y) = f(0,0) + X^2 - Y^2 = f(0,0) + g_1(X,Y).$$

If r is negative, then we can write Eq. 3 as:

$$f(x,y) - f(0,0) = -(-R(x,y))[x + y\frac{S(x,y)}{R(x,y)}]^2 + y^2 \frac{S^2(x,y) - R(x,y)T(x,y)}{-R(x,y)}.$$

Then define

$$\begin{array}{rcl} X & = & X(x,y) := & y\sqrt{\frac{S^2(x,y) - R(x,y)T(x,y)}{-R(x,y)}}, \\ Y & = & Y(x,y) := & \sqrt{-R(x,y)}[x + y\frac{S(x,y)}{R(x,y)}] \end{array}$$

Proceeding as above we get

$$f \circ \phi(X,Y) = f(0,0) + X^2 - Y^2 = f(0,0) + g_1(X,Y).$$

If t is non zero then the same arguments go through.

Lastly suppose both of them are zero (for example the case of hyperbola f(x,y)=xy). Consider a map $\phi(x,y)=(u(x,y),v(x,y))$, where u=x+y and v=x-y. Now define a map $g(u,v)=f(\frac{u+v}{2},\frac{u-v}{2})$. Then

$$g_u = \frac{1}{2}f_x + \frac{1}{2}f_y$$
 and $g_{uu} = \frac{1}{4}f_{xx} + \frac{1}{2}f_{xy} + \frac{1}{4}f_{yy}$.

Hence $g_{uu}(0,0) = \frac{1}{2}f_{xy}(0,0) = s \neq 0$. Otherwise the $rt - s^2 = 0$, a contradiction. Again proceeding as above for the function g we get the required result. Check that in any case the index of (0,0) is 1.

Remark 2. The converse of the above theorem is also true.

Example 7. Let $M = \mathbb{P}^2$ be the projective plane obtained by identifying antipodal points of S^2 . Consider the map $f: M \to \mathbb{R}$ defined by $f[(x, y, z)] = x^2 + 2y^2 + 3z^2$. Check that this is a smooth function on \mathbb{P}^2 . [(1,0,0)],[(0,1,0)] and [(0,0,1)] are the only critical points with index 0, 1 and 2 respectively.

Let us illustrate this example. Let $U_1 = \{[(x, y, z)] : x \neq 0\}, U_2 = \{[(x, y, z)] : y \neq 0\}$ and $U_3 = \{[(x, y, z)] : z \neq 0\}$. Define maps

$$\phi_1 : \mathbb{R}^2 \to U_1 \text{ as } \phi_1(u, v) = \frac{[(1, u, v)]}{\sqrt{1 + u^2 + v^2}},$$

$$\phi_2 : \mathbb{R}^2 \to U_2 \text{ as } \phi_2(u, v) = \frac{[(u, 1, v)]}{\sqrt{1 + u^2 + v^2}} \text{ and }$$

$$\phi_3 : \mathbb{R}^2 \to U_3 \text{ as } \phi_3(u, v) = \frac{[(u, v, 1)]}{\sqrt{1 + u^2 + v^2}}.$$

It is easy to check that (ϕ_i, \mathbb{R}^2) for i = 1, 2, 3 are parameterizations and that $(U_i, {\phi_i}^{-1})$ for i = 1, 2, 3 form a chart on \mathbb{P}^2 . We will find critical points and their indices using the remark 2. Define a map $g = f \circ \phi_1 : \mathbb{R}^2 \to \mathbb{R}$ as

$$g(u,v) = f \circ \phi_1(u,v) = \frac{1}{1+u^2+v^2} + \frac{2u^2}{1+u^2+v^2} + \frac{3v^2}{1+u^2+v^2}$$

Its is easy to check that g_u, g_v vanishes only at (u, v) = (0, 0). Hence [(1, 0, 0)] is the only critical point of f in U_1 . Define

$$U(u,v) := \frac{u}{\sqrt{1+u^2+v^2}}$$
 and $V(u,v) := \frac{v\sqrt{2}}{\sqrt{1+u^2+v^2}}$

It is easy to check that these are coordinates in some neighbourhood of (0,0). Let θ be the inverse of (U,V). Then $g \circ \theta(U,V) = 1 + U^2 + V^2$. Hence by remark 2 [(1,0,0)] is a non degenerate critical point of index 0. Similarly one can check for the other critical points.

Remark 3. Morse Theorem is true for any smooth manifold.

Corollary 3. Non degenerate critical points of a smooth function are isolated.

The function g_i , i = 0, 1, 2 defined in the theorem 1 are called the model functions.

Model neighbourhood of g_i 's are neighbourhoods U(s), s > 0 of (0,0) in \mathbb{R}^2 defined as follows:

For i = 0 and i = 2, U(s) is a disc of radius \sqrt{s} .

$$U(s) = \{(X, Y) \in \mathbb{R}^2 : X^2 + Y^2 \le s\}$$

For i = 1,

$$U(s) = \{(X, Y) \in \mathbb{R}^2 : |X^2 - Y^2| < s, |XY| < s\}.$$

This is an octagon which is homeomorphic to a rectangle. See Fig. 2.

Definition 5. Let $f: M \to \mathbb{R}$ be a smooth function on a closed surface M. Let p be a non degenerate critical point of f. Let (ϕ, U) be a parameterization centered at p such that f in U is of the form

$$f(\phi(X,Y)) = f(p) + q_i(X,Y)$$

where i is the index of p. Let $V \subseteq U$ be a model neighbourhood for g_i . Then the neighbourhood $\phi(V)$ of p is called a canonical neighbourhood of p and (ϕ, V) is called a canonical parameterization. See Fig. 3.

We will assume that boundary is included in canonical neighbourhoods.

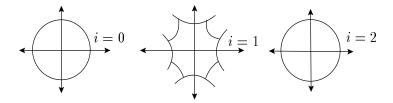


Figure 2: Model Neighbourhoods

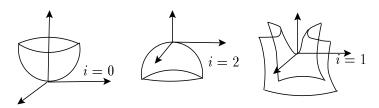


Figure 3: Canonical Neighbourhoods

2 Morse Functions and the Gradient Flow

Definition 6. Let M be a smooth manifold. A smooth one-parameter group of diffeomorphisms on M is a smooth map $\phi \colon \mathbb{R} \times M \to M$ defined as $(t, x) \mapsto \phi_t(x)$ satisfying the following properties:

- \bullet $\phi_0 = id_M$
- For each $t \in \mathbb{R}$ the map $\phi_t \colon M \to M$ as $t \mapsto \phi_t(x)$ is a diffeomorphism.
- For $r, s \in \mathbb{R}$, $\phi_{r+s}(x) = \phi_r \circ \phi_s(x)$ for all $x \in M$.

Since the map ϕ is smooth, for each $x \in \mathbb{R}$ the map $t \mapsto \phi_t(x)$ is a smooth curve in M. Hence

$$\frac{d}{dt}(\phi_t(x))|_{t=0} = X(x) \in T_x M.$$

That is, the map $X: M \to TM$ defined as $x \mapsto \frac{d}{dt}(\phi_t(x))|_{t=0}$ is a smooth vector field on M. This vector field is said to be generated by the smooth one-parameter group of diffeomorphism ϕ on M. Conversely, any smooth one-parameter group of diffeomorphism arises this way on any compact manifold. More precisely:

Theorem 4. Let M be a smooth compact manifold. Let $X \in \chi(M)$ be a smooth vector field on M. Then there exists a unique smooth one-parameter group of diffeomorphism $\phi \colon \mathbb{R} \times M \to M$ which generates X. That is,

$$\frac{d}{dt}(\phi_t(x))|_{t=0} = X(x) \text{ and } \phi_0(x) = x.$$

This theorem follows from a well known basic theorem in ODE.

Theorem 5. Let $U \subset \mathbb{R}^n$ be open. Let $X: U \to \mathbb{R}^n$ be a smooth map. Given $x_0 \in U$ there exists an open neighbourhood Ω of x_0 in U, an $\epsilon > 0$ and a smooth map $F: (\epsilon, \epsilon) \times \Omega \to U$ such that F(y, 0) = y for all $y \in \Omega$ and $\frac{d}{dt}(F(t, x)|_{t=0} = X(x)$.

If (δ, V, G) is another solution satisfying conditions similar to above then F = G on $(\eta, \eta) \times V$ where $\eta := \min\{\delta, \epsilon\}$.

Proof. Refer to [1] or [5].
$$\Box$$

We will use one-parameter group of diffeomorphisms to prove many important results in this article.

Definition 7. Let f be a smooth function on a surface M. Then f is said to be a Morse function if all its critical points are non degenerate.

Example 8. 1. All model functions are Morse functions.

- 2. Height functions on the sphere S^2 and the torus are Morse functions.
- 3. The function defined in Example 7 is a Morse function.

There always exist Morse functions on any closed manifold. (We will not get into the technicalities of this result. Interested readers can refer [7].)

Let f be a smooth function on any surface M. Let a and b be real numbers such that a < b. We will use the following notations:

$$\begin{array}{rclcrcl} M(a) & = & \{x \in M : f(x) \leq a\} & = & f^{-1}(-\infty,a], \\ M'(a) & = & \{x \in M : f(x) \geq a\} & = & f^{-1}[a,\infty), \\ V(a) & = & \{x \in M : f(x) = a\} & = & f^{-1}(a), \\ W(a,b) & = & \{x \in M : a < f(x) < b\} & = & f^{-1}[a,b]. \end{array}$$

These sets are illustrated in Fig. 4.

Theorem 6. Let M and N be a dimensional smooth manifolds with boundary. Let $f: M \to N$ be a diffeomorphism. Them f maps interior of M onto interior of N and the boundary of M onto the boundary of N.

Theorem 7. Let M be a compact surface and f be a smooth function on M. Let a and b be regular values of f. Then M(a) and W(a,b) are compact surfaces with V(a) as the boundary of M(a) and disjoint union of V(a) and (b) as the boundary of W(a,b). Also V(a) is a closed 1-manifold.

The proofs of theorem 6 and theorem 7 are simple applications of Implicit Function Theorem and Inverse Mapping Theorem.

Any closed manifold M can be embedded in \mathbb{R}^N for some large N. Hence for each $x \in M$, the tangent space T_xM inherits an inner product from \mathbb{R}^N .

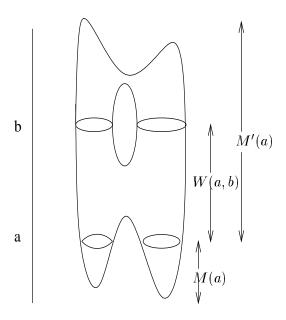


Figure 4: The sets M(a), M'(a), W(a,b) and V(a)

Theorem 8. Let M be a surface and $f: M \to \mathbb{R}$, a smooth function on M. Assume that a and b are regular values of f with a < b such that W(a,b) does not contain any critical point of f. Then M(b) is diffeomorphic to M(a), V(b) is diffeomorphic to V(a) and W(a,b) is diffeomorphic to $V(a) \times [a,b]$.

Proof. The idea of the proof is to push M(a) to M(b) using the one-parameter group of diffeomorphisms. Since we want V(a) to be mapped diffeomorphically to V(b), the integral curves of the vector field should be transversal to the level curves of f. In particular the modified gradient vector field of f may do the job. The details are given below.

Let $\epsilon > 0$ be small enough such that $f^{-1}(a - \epsilon, b + \epsilon)$ does not contain any critical points of f. Let $\alpha \colon M \to \mathbb{R}$ be a non negative smooth function such that α is 1 on W(a, b) and it is 0 on the complement of $W(a - \epsilon, b + \epsilon)$. See Fig. 5. Define a vector field

$$Y(x) = \begin{cases} \frac{\alpha(x)}{\|grad(f(x))\|^2} grad(f(x)) & x \in W(a - \epsilon, b + \epsilon), \\ 0 & \text{otherwise.} \end{cases}$$

This is a smooth vector field transversal to the level curves of f in W(a, b). Let ϕ be the one-parameter group of diffeomorphisms associated with the vector field Y. For each $x \in M$ consider the map $t \mapsto f(\phi_t(x)) = \psi(t)$. This is a smooth function from \mathbb{R} to \mathbb{R} . Its derivative

$$\frac{d}{dt}\psi(t) = Df(\phi_t(x))\frac{d}{dt}\phi_t(x) = Df(\phi_t(x))Y(\phi_t(x)) = \alpha(\phi_t(x)).$$

That is, if $\phi_t(x) \in W(a,b)$ then $\frac{d}{dt}\psi(t)=1$. Hence ψ is linear as long as $\phi_t(x)$ lies in

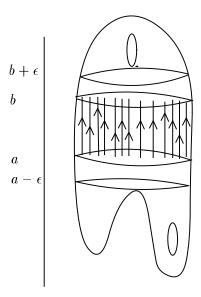


Figure 5: M(b) is diffeomorphic to M(a)

W(a,b). This means that $\psi(t)=t+A$ for some constant A. But $\psi(0)=f(\phi_0(x))=f(x)$. Hence $\psi(t)=t+f(x)$ for all x such that $\phi_t(x)\in W(a,b)$.

We claim that ϕ_{b-a} maps M(a) diffeomorphically onto M(b). Let us fix $x \in M(a)$. By Mean-Value Theorem

$$\left| \frac{f(\phi_{b-a}(x)) - f(\phi_0(x))}{b - a} \right| = \left| \frac{\psi(b-a) - \psi(0)}{b - a} \right| \le \sup_{t \in [0, b-a]} |\psi'(t)| \le 1.$$

This implies that $|f(\phi_{b-a}(x)) - f(x)| \le b-a$. That is, $f(\phi_{b-a}(x)) \le b-a+f(x) \le b-a+a = b$. Thus ϕ_{b-a} maps M(a) into M(b).

To prove that ϕ_{b-a} is onto as a map from M(a) to M(b), it is enough to prove that if f(x) > a then $f(\phi_{b-a}(x)) > b$. Let us assume this and prove that ϕ_{b-a} is onto. Let $x \in M(b)$. Since ϕ_{b-a} is a diffeomorphism of M, there exists $y \in M$ such that $\phi_{b-a}(y) = x$. If f(y) > a, then by assumption $f(\phi_{b-a}(y)) = f(x) > b$. Which is a contradiction to the fact that $x \in M(b)$. Hence $f(y) \le a$. This proves that ϕ_{b-a} is onto as a map from M(a) to M(b). Also ϕ_{b-a} is one-one and hence ϕ_{b-a} is a diffeomorphism as a map from M(a) to M(b).

Now let us prove the claim. Let $f(\phi_{b-a}(x)) \leq b$ for f(x) > a. Since $\psi'(t) \geq 0$, ψ is increasing. That is, $\psi(b-a) \geq \psi(0)$. This implies that for all $x \in M$, $f(\phi_{b-a}(x)) \geq f(\phi_0(x)) = f(x) > a$. Thus $a \leq f(\phi_t(x)) \leq b$ for all $t \in [0, b-a]$. Hence $f \circ \phi_t(x)$ is linear in [0, b-a]. Hence $f(\phi_{b-a}(x)) = b-a+f(x) > b-a+a > b$. This is a contradiction.

 ϕ_{b-a} maps V(a) diffeomorphically onto V(b) by theorem 6.

For the last part we define a map

$$\theta \colon V(a) \times [a,b] \to W(a,b)$$
 as $\theta(x,t) = \phi_{t-a}(x)$.

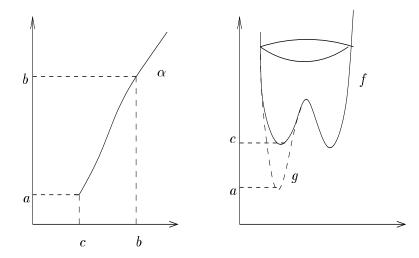


Figure 6: Modification at a local minimum

We claim that θ is a diffeomorphism. Since ψ is increasing along the integral curves,

$$a = f(\phi_0(x)) \le f(\phi_{t-a}(x)) \le f(\phi_{b-a}(x)) = b.$$

Hence $\theta(x,t) \in W(a,b)$. For $y \in W(a,b)$, define $\beta(y) = (\phi_{a-f(y)}(y), f(y))$. Check that θ and β are inverse of each other.

3 Modification of a Morse Function

Let f be a Morse function on a surface M. The idea is to improve f to a new Morse function g having the same critical points with same indices and which coincides with f outside some canonical neighbourhood of a critical point. We need following lemmas at our disposal.

Lemma 9. Let $f: M \to \mathbb{R}$ be a Morse function. Let $p \in M$ be a critical point of f of index 0. Assume that f(p) = c. Let $a \le c$ be any real number. Then there exists a Morse function g on M having the same critical points with same indices as f and which coincides with f outside some canonical neighbourhood of p and is such that g(p) = a.

Proof. If a = c, then there is nothing to prove.

Let a < c. Let (ϕ, U) be a canonical parameterization centered at p. Let $\phi(U) = V$. Let $f(\partial V) = b$. Let $\alpha \colon [c, b] \to \mathbb{R}$ be a smooth map such that $\alpha' > 0$, $\alpha(c) = a$ and $\alpha(x) = x$ for x near b. Now define $g \colon M \to \mathbb{R}$ as:

$$g(x) = \begin{cases} \alpha(f(x)), & x \in V \\ f(x), & x \notin V \end{cases}$$

See Fig. 6. It is easy to check that g is smooth. We claim that g has the required properties.

First of all $g(p) = \alpha(f(p)) = \alpha(c) = a$. For $x \in V$, $Dg(x) = \alpha'(f(x))Df(x)$. Since the derivative of α is positive, derivative of g in V is zero only when x = p. Thus p is the only critical point of g in V. Also $g \circ \phi(X,Y) = \alpha(f(p) + X^2 + Y^2)$ on U, the Hessian of g at 0 is given by

 $H(g) = \begin{pmatrix} 2\alpha'(c) & 0 \\ 0 & 2\alpha'(c) \end{pmatrix}.$

Hence the index of p is 0.

as

Lemma 10. Let $f: M \to \mathbb{R}$ be a Morse function. Let $p \in M$ be a critical point of f of index 2. Assume that f(p) = c. Let $a \ge c$ be any real number. Then there exists a Morse function g on M having the same critical points with same indices as f and which coincides with f outside some canonical neighbourhood of p and is such that g(p) = a.

Proof. Use similar arguments as above.

Lemma 11. Let $f: M \to \mathbb{R}$ be a Morse function. Let $p \in M$ be a critical point of f of index 1. Assume that f(p) = c. Let $\alpha \in \mathbb{R}$ be any real number. Then there exists a Morse function g on M having the same critical points with same indices as f and which coincides with f outside some canonical neighbourhood V of p and is such that $g(p) = c + \alpha$.

Proof. The idea is to change the model function g_1 suitably. We will show that for any real number α , there exists a bounded neighbourhood say U_1 of (0,0) in \mathbb{R}^2 and a Morse function $h: \mathbb{R}^2 \to \mathbb{R}$ such that $h(x,y) = x^2 - y^2$ for $(x,y) \notin U_1$, $h(0,0) = \alpha$ and (0,0) is the only critical point of h and index of (0,0) is 1. Let us assume this claim and complete the proof.

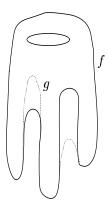


Figure 7: Modification at a saddle point

Choose a canonical parameterization (ϕ, U) such that $U_1 \subseteq U$. Now define $g: M \to \mathbb{R}$

$$g(x) = \begin{cases} f(p) + h \circ \phi^{-1}(x) & x \in \phi(U) = V, \\ f(x) & x \notin V. \end{cases}$$

Then $g(p) = c + h(0,0) = c + \alpha$. It is easy to see that g is smooth. Also critical points of g in V are critical points of $h \circ \phi^{-1}$ in U. Hence p is the only critical point of g in V and its index is 1. See Fig. 7.

Now let us prove the claim made earlier. Let $\alpha>0$ be fixed. Let $w:\mathbb{R}\to\mathbb{R}$ be a smooth non negative function such that $w(0)=1, w'(0)=0, yw'(y)\leq 0$ for all y and w(y)=0 for $y\notin (-a,a)$ for some positive number a. Also define a smooth non negative function $\lambda\colon\mathbb{R}\to\mathbb{R}$ such that $\lambda(0)=\alpha, \lambda'(0)=0, \lambda(x)=0$ for $x\notin (-a,a)$ for some positive number a and such that

$$2x + \lambda'(x) > 0 \quad \text{if } x > 0,$$

$$2x + \lambda'(x) < 0 \quad \text{if } x < 0.$$

Note that since α is fixed, this will put some bounds on a.

Construction of w. Consider $w_1 = 1 - \frac{x^2}{a^2}$. Clearly $w_1(0) = 1$. Now smoothen w_1 at its roots $\pm a$ to get w. Since $w'(x) = \frac{-2x}{a^2}$, $xw'(x) \leq 0$.

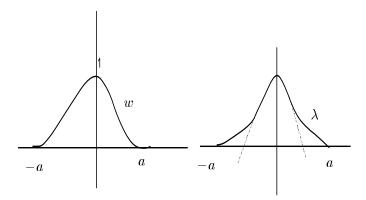


Figure 8: Graphs of w and λ

Construction of λ . Consider $\lambda_1(x) = \alpha - \frac{\alpha x^2}{a^2}$. Now smoothen λ_1 at its roots $\pm a$ to get λ . Clearly $\lambda(0) = \alpha$, and $\lambda(x) = 0$ for $x \notin (-a, a)$. We want $2x + \lambda'(x) >$ for x > 0, that is, $2x - 2x\frac{\alpha}{a^2} > 0$ for x > 0 and $2x - 2x\frac{\alpha}{a^2} < 0$ for x < 0. So, choose $a^2 > \alpha$. Then λ has the required properties. See Fig. 8

Now define

$$h(x, y) = x^2 - y^2 + \lambda(x)w(y).$$

Then $h(0,0) = \alpha$ and $h(x,y) = x^2 - y^2$ for $(x,y) \notin ([-a,a] \times [-a,a])$. Partial derivatives of h are

$$h_x = 2x + \lambda'(x)w(y) \& h_y = -2y + \lambda(x)w'(y).$$

Hence (0,0) is a critical point of h. Let $(x,y) \neq 0$. Since $0 \leq w \leq 1$,

$$h_x = 2x + \lambda'(x)w(y) > 2x + \lambda'(x) > 0$$
 if $x > 0$,
 $h_x = 2x + \lambda'(x)w(y) < 2x + \lambda'(x) < 0$ if $x < 0$.

Hence if $x \neq 0$, $h_x \neq 0$. If x = 0, then

$$h_y = -2y + \lambda(x)w'(y) < 0$$
 if $y > 0$, as $w'(y) < 0$
 $h_y = -2y + \lambda(x)w'(y) > 0$ if $y < 0$, as $w'(y) > 0$.

Hence (0,0) is the only critical point of h. Also

$$h_{xx} = 2 + \lambda''(x)w(y), h_{yy} = -2 + \lambda(x)w''(y), \text{ and } h_{xy} = \lambda'(x)w'(y).$$

The Hessian

$$H(h)(0,0) = \begin{pmatrix} 2 + \lambda''(0) & 0 \\ 0 & -2 + \alpha w''(0) \end{pmatrix}.$$

Using the construction of w and λ it is easy to show that index of (0,0) is 1.

For
$$\alpha < 0$$
, take $h(x,y) = x^2 - y^2 - w(x)\lambda(y)$ and proceed as before.

Remark 4. In Lemma 9, Lemma 10 and Lemma 11 we may pass critical levels. (See Fig. 6 and Fig. 7.)

Definition 8. Let $f: M \to \mathbb{R}$ be a Morse function. Let X_i, Y_j and Z_k be critical points of f of indices 0, 1 and 2 respectively. We say that f is an ordered Morse function if $f(X_i) < f(Y_j) < f(Z_k)$ for all i, j, k and if it separates critical points, that is, $f(x) \neq f(y)$ for any two distinct critical points x and y.

Theorem 12. On any surface M there exists an ordered Morse function.

Proof. Let f be a Morse function on M. Let $a, b \in \mathbb{R}$, a < b be real numbers such that all the critical points of index 1 of f lie in $f^{-1}(a, b)$. By Lemma 9 we can find a Morse function such that all the critical values of index 0 are less than a. Similarly by Lemma 10 we can find a Morse function such that all the critical values of index 2 are greater than b. Similarly we can separates critical points.

4 Crossing Critical Levels of Index 0 or 2

Let a and b be regular values of f such that a < b. Let W(a, b) contain a critical point say p of f. Then Theorem 8 need not be true. In this section we see what happens when we cross a critical level.

Theorem 13. Let $f: M \to \mathbb{R}$ be a Morse function on M and $p \in M$, a critical point of f. Let f(p) = c. Let a and b be regular values such that W(a, b) does not contain any critical point other than p. Then the following hold:

• If p is a critical point of index 0, then M(b) is diffeomorphic to a disjoint union of M(a) with a disc D which is a canonical neighbourhood of p. Also V(b) is diffeomorphic to a disjoint union of V(a) with a circle which is the boundary of D.

• If p is of index 2, then M(b) is diffeomorphic to M(a) with a disc D attached along one of the components of V(a) and the attaching map is injective as map from the boundary of D to V(a). Also V(a) is diffeomorphic to a disjoint union of V(b) with a circle which is the boundary of D.

Proof. It is enough to prove the lemma for regular values α, β such that $a < \alpha < c < \beta < b$. This follows from theorem 8. Choose an $\epsilon > 0$ such that $W(c - 2\epsilon, c + 2\epsilon)$ does not contain any critical point of f other than f. Hence f and f are canonical neighbourhoods of f. For convenience let us assume that f and f and f and f are canonical neighbourhoods of f and f are canonical neighbourhoods.

Let p be of index 0. In this case $M(b) \cap U(\epsilon) = U(\epsilon)$, diffeomorphic to a disc. Let us denote $U(\epsilon)$ by D. $V(b) \cap U(\epsilon)$ is the boundary of the disc D. Since p is of index 0 and $U(\epsilon)$, a canonical neighbourhood of p, $f(x) \geq f(p) = c$ for all $x \in U(\epsilon)$. Hence $M(a) \cap U(\epsilon) = \emptyset = V(a) \cap U(\epsilon)$. See Fig. 9.

Choose a smooth non negative function $\alpha \colon M \to \mathbb{R}$ such that α vanishes in $U(\epsilon)$ and is equal to 1 outside $U' = U(2\epsilon)$. Also consider a smooth function $\beta \colon M \to \mathbb{R}$ such that β vanishes outside $W(c - 2\epsilon, c + 2\epsilon)$ and is 1 in $W(c - \epsilon, c + \epsilon)$. Now define a vector field

$$Y(x) = \begin{cases} \frac{\alpha(x)\beta(x)}{\|grad(f(x))\|^2} grad(f(x)) & x \in W(c - 2\epsilon, c + 2\epsilon) \setminus U(\epsilon) \\ 0 & \text{otherwise.} \end{cases}$$

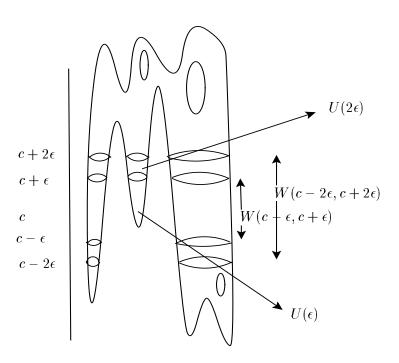


Figure 9: Crossing a local minimum

Then Y is transversal to level curves of f.

Let ϕ be the one-parameter group of diffeomorphism associated with the vector field Y. We claim that the map ϕ_{b-a} maps M(a) to $M(b) \setminus D$. For $x \in M$, define $g(t) := f \circ \phi_t(x)$. Then $g'(t) = \alpha(\phi_t(x))\beta(\phi_t(x)) \leq 1$. Using Mean Value Theorem we can show that $\phi_{b-a}(M(a)) \subseteq M(b)$. (Recall arguments in Theorem 8.) Since Y(x) = 0 for $x \in U(\epsilon)$, $\phi_t(x) = x$ for all $t \in \mathbb{R}$ and $x \in U(\epsilon)$. That is, the integral curves starting in $U(\epsilon)$ are constant. Hence, if $x \notin U(\epsilon)$, $\phi_t(x) \notin U(\epsilon)$ for all $t \in \mathbb{R}$. Thus ϕ_{b-a} maps M(a) into $M(b) \setminus D$.

We now show that ϕ_{b-a} is onto as a map from M(a) to $M(b) \setminus D$. Let $x \in M(b) \setminus D$. Since ϕ_{b-a} is a diffeomorphism on M, there exists $y \in M$ such that $\phi_{b-a}(y) = x$. We claim that $f(y) \leq a$. If f(y) > a, then we claim that $f(\phi_{b-a}(y)) > b$. Let f(y) > a and $f(\phi_{b-a}(x)) \leq b$. Since g is increasing, $a < f(\phi_{b-a}(y)) \leq b$. But then $f(\phi_{b-a}(y)) = b - a + f(y) > b$. This is a contradiction. This proves that ϕ_{b-a} is onto as a map from M(a) to $M(b) \setminus D$. Also ϕ_{b-a} is identity on D. Hence $\phi_{b-a} : M(a) \coprod D \to M(b) \setminus D = M(b)$ is a diffeomorphism.

Lastly we claim that M(b) is diffeomorphic to disjoint union of M(a) and D. Since M(a) is compact, $\phi_{b-a}(M(a)) = M(b) \setminus D$ is compact and hence closed in M(b). D is compact implies $M(b) \setminus D$ is open in M(b). Hence D is a component of M(b). This proves that M(b) is homeomorphic to disjoint union of M(a) and a disc D. The last part of the first assertion follows from Theorem 6.

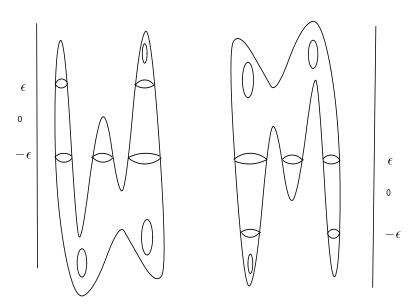


Figure 10: Crossing a local maximum

Let index of p be 2. For convenience let us assume that c = 0, $a = -\epsilon$ and $b = \epsilon$. Define g := -f. Then p is a critical point of g of index 0. Hence by the first part of the theorem $\widetilde{M}(\epsilon) \setminus D$ is diffeomorphic to $\widetilde{M}(-\epsilon)$, where

$$\begin{array}{lcl} \widetilde{M}(-\epsilon) & = & \{x \in M: -f(x) \leq -\epsilon\} & = & \{x \in M: f(x) \geq \epsilon\} & = & M'(\epsilon), \\ \widetilde{M}(\epsilon) & = & \{x \in M: -f(x) \leq \epsilon\} & = & \{x \in M: f(x) \geq -\epsilon\} & = & M'(-\epsilon), \end{array}$$

and D is a canonical neighbourhood $U(\epsilon)$ of p. See Fig. 10.

Hence int $M'(-\epsilon)$ is diffeomorphic to int $(M'(\epsilon) \coprod D)$ via the map ϕ_{b-a} (theorem 6). Since ϕ_{b-a} is a diffeomorphism on M, $(\operatorname{int} M'(-\epsilon))^c = M(-\epsilon)$ and $(\operatorname{int} (M'(\epsilon) \coprod D))^c = M(\epsilon) \setminus \operatorname{int} D$ are smooth manifolds. $M(-\epsilon)$ is diffeomorphic to $M(\epsilon) \setminus \operatorname{int} D$. Hence $M(\epsilon)$ is diffeomorphic to $M(-\epsilon) \cup D$ where D is attached to $M(-\epsilon)$ via an injection on ∂D to $V(-\epsilon)$. The last part of the second assertion follows from theorem 6.

Remark 5. Analogue of theorem 13 is true in the case M is of dimension n and p is a critical point of index 0 or n. In this case disc D is of dimension n.

Proposition 14. Let f be an ordered Morse function on a surface M. Suppose that f has no critical points of index 1. Then M is homeomorphic to a sphere.

Proof. Let X_1, X_2, \ldots, X_m be critical points of index 0 and Z_1, Z_2, \ldots, Z_n be critical points of index 2. First of all we claim that n = m.

Let a be a regular value of f such that $f(X_i) < a < f(Z_j)$ for all i and j. Then using induction on number of critical points of index 0 we can show that M(a) is a disjoint union of m discs, say, $\{D_i\}_{1 \le i \le m}$ and its boundary is a disjoint union of m circles. Similarly M'(a) is disjoint union of n discs, say, $\{D_i'\}_{1 \le i \le n}$ and its boundary is disjoint union of n circles. Since V(a) is the common boundary of M(a) and M(a), m = n.

Let n = 1. Since f has one critical point of index 0 and one critical point of index 2, M(a) is a disc say D and M'(a) is also a disc say D'. M is obtained by attaching D' to D along their boundaries. Hence M is homeomorphic to a sphere.

Lastly we claim that n=1. Let a be as above. Then M is obtained by attaching n discs $\{D_i'\}_{1\leq i\leq n}$ (homeomorphic to M'(a)) to n discs $\{D_i\}_{1\leq i\leq n}$ (homeomorphic to M(a)) along their boundaries. Since the attaching map is injective, \overline{D}_i' is attached to some $D_{j(i)}$ resulting in a sphere. Hence M is homeomorphic to n disjoint spheres. Thus if n>1, then M will be disconnected. This proves that n=1.

Theorem 15. Any closed 1-manifold is homeomorphic to a circle S^1 .

Proof. Let M be a closed 1-manifold and f be an ordered Morse function on M. Let X_1, X_2, \ldots, X_m be critical points of index 0 and Y_1, Y_2, \ldots, Y_n be critical points of index 1. Let a be a regular value of f such that $f(X_i) < a < f(Y_j)$ for all i and j. Then M(a) is a disjoint union of m arcs $\{I_j\}_{1 \le j \le m}$ and M'(a) is disjoint union of n arcs $\{J_k\}_{1 \le k \le n}$. This follows from the analogue of theorem 13 (the case when the index is zero, see the remark 5).

The boundary of M(a) consists of 2m points and the boundary of M'(a) consists on 2n points. Since V(a) is the common boundary of M(a) and M'(a), m = n.

We proceed by induction on n. If n = 1, then M is homeomorphic to a circle. This can be seen in the same way as Proposition 14. Let the result be true for any k < n. Let X_1 be such the $f(X_1) > f(X_i)$ for all $2 \le i \le n$. Let the arc I be the component of M(a) containing X_1 . Let A and B be boundary points of I. Let the arc I be the component of M'(a) containing I as one of the boundary point. Let I be a canonical neighbourhood of a critical point say

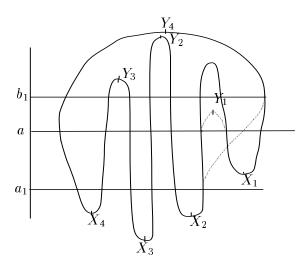


Figure 11: Classification of one manifolds

 Y_1 of index 1. We can find a Morse function g on M having the same critical points with the same indices as f and is such that $g(X_i) < a < g(Y_1) < g(Y_j)$ for all i and $j \neq 1$. Let a_1, b_1 be regular values of g such that $g(X_i) < a_1 < g(X_1) < g(Y_1) < b_1 < g(Y_j)$ for all $i \neq 1$ and $j \neq 1$. Let K be the component of $g^{-1}[a_1, b_1]$ containing X_1 and Y_1 . We now modify the function g to a Morse function g which agrees with g outside a coordinate neighbourhood of g and such that g has no critical points in that neighbourhood g. See Fig. 11. Then g has only g 1 critical points of index 0. Hence by induction g 1 is homeomorphic to a circle.

Remark 6. If f is a smooth function on a surface M and a is a regular value of f, then by Theorem 7 and Theorem 15, V(a) is a disjoint union of circles. Hence the boundary of any compact surface is homeomorphic to a disjoint union of circles.

5 Crossing a Critical Level of Index 1

Theorem 16. Let $f: M \to \mathbb{R}$ be a Morse function on a surface M. Let p be a critical point of f of index 1. Let a, b with a < b be regular values of f such that W(a, b) contains no critical points of f other that p. Then M(b) is homeomorphic to M(a) with a rectangle attached to two disjoint segments of V(a) along pair of opposite sides of the rectangle.

Proof. Choose ϵ such that $W(c-3\epsilon,c+3\epsilon)$ does not contain any critical point of f other than p. For convenience let $a=c-2\epsilon$ and $b=c+2\epsilon$. Let (ϕ,W) be a canonical parameterization centered at p. Let $U(2\epsilon)=V\subseteq W$ be a model neighbourhood of (0,0). See Fig. 12. Let $\phi(V)=U$. Let I'=B'C' and J'=F'G' be as in the Fig. 12. Let $I=BC=\phi(I')$ and $J=FG=\phi(J')$. Then $I,J\subseteq V(a)$. Let $K=\overline{V(a)\setminus (I\cup J)}$. Define $T=\overline{W(a,b)\setminus U}$. Note that T need not be connected. We claim that T is diffeomorphic to $K\times [a,b]$.

We choose a compact tubular neighbourhood T_1 of T as follows: Let $V_1 \subseteq W$ be a neighbourhood of (0,0) bounded by the curves

$$X^{2} - Y^{2} = a - c - \epsilon, X^{2} - Y^{2} = b - c + \epsilon, \text{ and } XY = b - c - \epsilon, XY = a - c + \epsilon.$$

See Fig. 12. Let $U_1 = \phi(V_1)$. Now define $T_1 := \overline{W(a - \epsilon, b + \epsilon) \setminus U_1}$.

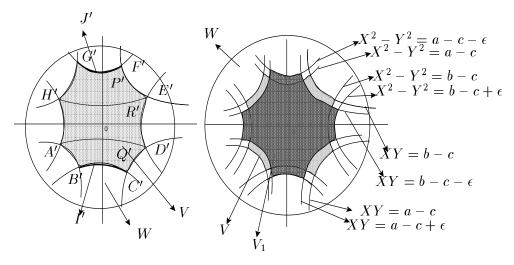


Figure 12: Illustration in \mathbb{R}^2 for the proof of theorem 16

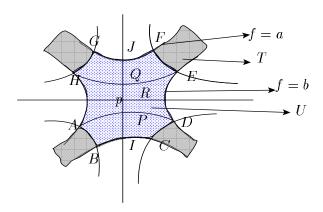


Figure 13: Illustration in the surface for the proof of theorem 16

Consider a smooth non negative map $\alpha \colon M \to \mathbb{R}$ such that α is identically 1 in T and it vanishes outside T_1 . Define a vector field Y on M as

$$Y(x) = \begin{cases} \frac{\alpha(x)}{\|grad(f(x))\|^2} grad(f(x)) & x \in T_1 \\ 0 & x \notin T_1 \end{cases}$$

Let φ be the one-parameter group of diffeomorphisms generated by the vector field Y.

Define a map $\Phi: K \times [a, b] \to T$ as $\Phi(x, t) = \varphi_{t-a}(x)$. We claim that this is the required diffeomorphism.

First of all let us show that $\Phi(x,t) \in T$ for all $x \in K$ and $t \in [a,b]$. Define a map $g \colon \mathbb{R} \to \mathbb{R}$ as $g(t) = f \circ \varphi_{t-a}(x)$. Since $g'(t) = \alpha(\varphi_{t-a}(x))$, we have $0 \le g'(t) \le 1$. Note that g(a) = a. By Mean Value Theorem $g(t) - g(a) \le t - a$. Hence $g(t) \le g(a) + t - a \le b$. Since g is increasing, $g(t) \ge g(a) = a$ for all $t \in [a,b]$. Hence $\varphi_{t-a}(x) \in W(a,b)$ for all $t \in [a,b]$ and for all $x \in K$. Next we claim that $\Phi(x,t) \notin \text{int } U$ for all $t \in [a,b]$ and $x \in K$. Let A'B', E'F', C'D', G'H' be as in the Fig. 12. Since (ϕ, W) be a canonical parameterization, their images AB, EF, CD, GH are integral curves of the vector field Y. See Fig. 13. Assume for some $x \in K$ and $t \in [a,b], \varphi_{t-a}(x) \in \text{int } U$. Then $\varphi_{t-a}(x)$ will intersect one of the above integral curves AB, EF, CD, GH as they are the common boundary of T and U. Hence it is one of these curves and cannot lie in the int U, a contradiction to our assumption. Hence $\varphi_{t-a}(x) \in T$ for all $t \in [a,b], x \in K$.

 Φ is one-one follows from the fact that g(t) is linear for $t \in [a, b]$ and φ_{t-a} is a diffeomorphism as a map from M to M.

Next we show that Φ is onto. Let $y \in T$. Then the curve $t \mapsto \varphi_t(y)$ is an integral curve starting at y. Define a map $h \colon \mathbb{R} \to \mathbb{R}$ defined as $h(t) = f \circ \varphi_t(y)$. Then h is smooth, $h'(t) \geq 0$ and h'(t) = 1, if $\varphi_t(y) \in T$. Note that $h(0) = f(y) \geq a$. We claim that there exists $t_0 \leq 0$ such that $h(t_0) = a$. If h(0) = a, then $y \in K$ and $\Phi(y, a) = y$. Let h(0) > a. Assume h(t) > a for all t < 0. Since h is increasing $h(t) \leq h(0) \leq b$ for all t < 0. Hence $\varphi_t(y) \in W(a,b)$ for all t < 0. Also for $y \in T$, $t \mapsto \varphi_t(y)$ is the integral curve starting at y in T. By previous argument $\varphi_t(y) \notin \text{int } U$ for all t < 0. Hence h is linear, which implies that h(t) = t + f(y) > a for all t < 0, a contradiction. Hence there exists $t_0 < 0$ such that $h(t_0) = a$. Since h is increasing for all $t \in [t_0, 0]$, $h(t_0) \leq h(t) \leq h(0)$. Hence h(t) is linear for $t \in [t_0, 0]$. This implies that $a = h(t_0) = t_0 + f(y) \leq t_0 + b$ and hence $a - t_0 \in [a, b]$. $h(t_0) = f \circ \varphi_{t_0}(y) = a$ implies that $\varphi_{t_0}(y) \in V(a)$. If $\varphi_{t_0}(y) \in V(a) \setminus K$, then $\varphi_t(y)$ has to intersect one of the integral curves. We reach a contradiction by above arguments. Hence $\varphi_{t_0}(y) \in K$. Also

$$\Phi(\varphi_{t_0}(y), a - t_0) = \varphi_{a - t_0 - a} \circ \varphi_{t_0}(y) = \varphi_0(y) = y.$$

This proves that Φ is a diffeomorphism.

Divide U into three parts P,Q and R (see Fig. 13). Each of them is homeomorphic to a rectangle. Let us define a homeomorphism $\Phi_1\colon I\times[a,b]\to P$ such that $\Phi_1(B\times t)=\Phi(B,t)$ and $\Phi_1(C\times t)=\Phi(C,t)$. Similarly define a homeomorphism $\Phi_2\colon J\times[a,b]\to Q$ such that $\Phi_2(G\times t)=\Phi(G,t)$ and $\Phi_2(F\times t)=\Phi(F,t)$. Observe that $T\cup P\cup Q=\overline{W(a,b)\setminus R}$. Also $K\cup I\cup J=V(a)$ and hence $M(a)\cup T\cup P\cup Q=\overline{M(b)\setminus R}$. See Fig. 14. Define a map $\Psi\colon V(a)\times[a,b]\to T\cup P\cup Q$ as:

$$\Psi(x,t) = \begin{cases} \Phi(x,t), & x \in K \\ \Phi_1(x,t), & x \in I \\ \Phi_2(x,t), & x \in J \end{cases}$$

 Ψ is well defined and is a homeomorphism. Hence Φ is homeomorphic to $M(a) \cup (V(a) \times [a,b])$ along a homeomorphism from V(a) to $V(a) \times a$. Note that by Remark 6, V(a) is

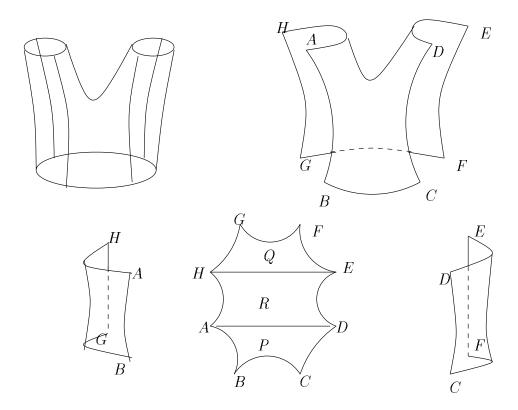


Figure 14: Attaching a rectangle

homeomorphic to disjoint union of circles and $V(a) \times [a, b]$ is homeomorphic to disjoint union of cylinders.

Now $M(a) \cup T \cup P \cup Q$ is homeomorphic to M(a) with cylinders attached in the above manner which is homeomorphic to M(a). This implies that $M(b) \setminus R$ is homeomorphic to M(a). Now M(b) is obtained by attaching the rectangle R to M(a) along opposite sides ED and AH.

Analysis of crossing critical Point of index 1

Let p be a critical point of index 1 of a Morse function f on M. Let a, b, a < b be regular values of f such that $f^{-1}[a,b]$ contains no critical points of f other than p. Then we have shown that M(b) is homeomorphic to M(a) with a rectangle R attached to two disjoint segments I and J of V(a) along opposite sides of the rectangle. Let us analyze the different ways of attaching the rectangle R to V(a) and the component of W(a,b) containing p.

Case 1. Let I and J lie in 2 different components of V(a). See Fig. 15.

In this case the number of components of V(b) – the number of components of V(a) = -1. That is, the boundary components have reduced by one after passing this level.

Case 2. Let I and J lie in the same component of V(a). We attach the rectangle R

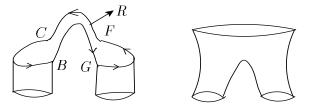


Figure 15: Case 1

straight. See the Fig. 16.

In this case the number of components of V(b) — the number of components of V(a) = 1. That is, number of components increases by one after passing this critical level.

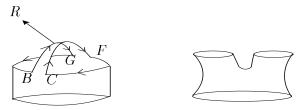


Figure 16: Case 2

Thus in the two cases the component of W(a,b) containing p is homeomorphic to a disc with two holes.

Case 3. If I and J lie in the same component of V(a) and the rectangle R is attached with a twist. See the Fig. 17.

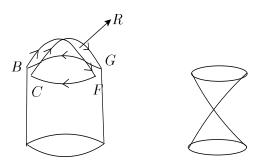


Figure 17: Case 3

In this case the number of components do not change. The component of W(a,b) containing p is homeomorphic to a Möbius band without a disc. In this case we have a homeomorphic copy of a Möbius band inside M.

Remark 7. We can attach the rectangle R with more than 2 twists but this reduces to the earlier cases. (Exercise.)

Definition 9. Let f be a Morse function on a surface M. Let a and b be regular values of f with a < b. Let W(a, b) contain only one critical point say p of index 1. Let the number of components of V(a) and V(b) be m and n respectively. Then we say that

- 1. p is of type I of the first kind if m n = 1.
- 2. p is of type I of the second kind if m n = -1.
- 3. p is of type II if m = n.

See Fig. 18

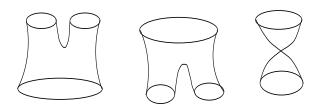


Figure 18: Critical points of index 1 and their types

Corollary 17. Let f be an ordered Morse function on M. Let a, b, a < b be regular values of f such that W(a,b) contains p critical points of index 0 and p critical points of index 1 of the type I of the first kind. Then M(a) is homeomorphic to M(b).

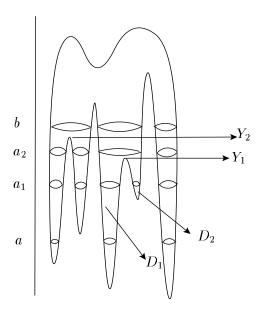


Figure 19: Illustration for Corollary 17

Proof. Since f is an ordered Morse function, M(a) contains only critical points of index 0. Let the number of critical points of index 0 in M(a) be q. Then M(a) is disjoint union of q discs. We have to show that M(b) is also homeomorphic to disjoint union of q discs. Let $\{X_i\}_{1\leq i\leq p+q}$ be the critical points of index 0 and $\{Y_j\}_{1\leq j\leq p}$, the critical points of index 1 of type I of the first kind in M(b). Let $a_1, a_2, \ldots, a_{p-1}, a_p$ be regular values of f such that

$$f(X_i) < a < f(X_j) < a_1 < f(Y_1) < a_2 < f(Y_2) < a_3 < \dots < a_p < f(Y_p) < b$$

for all 1 < i < q and q + 1 < j < p + q.

Then $M(a_1)$ is homeomorphic to the disjoint union of p+q discs, say, $\{D_i\}_{1\leq i\leq p+q}$. Since Y_1 is a critical point of type I of the first kind, there exist two discs say D_1 and D_2 in $M(a_1)$ which are attached to each other by a rectangle resulting in a disc again. See Fig. 19. Hence $M(a_2)$ is a disjoint union of p+q-1 discs. Proceeding in this fashion we reduce one disc when we pass a critical Y_j . Hence M(b) is the disjoint union of p+q-p=q discs. This proves that M(a) is homeomorphic to M(b).

Example 9. Let us illustrate the passing the critical levels in Example 7. Take the unit disc model of \mathbb{P}^2 . See Fig. 20.

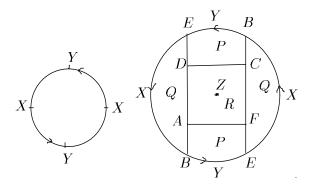


Figure 20: Disc Model for \mathbb{P}^2

Here X, Y, Z are critical points of index 0, 1, 2 respectively. (Project $S^2 \subseteq \mathbb{R}^3$ onto unit disc in \mathbb{R}^2 and we obtain Fig. 20 and Fig. 21 and the required critical points of the function.) P, Q, R are the canonical neighbourhoods of X, Y, Z respectively.

Now Q is attached to P along segments I and J in a twisted manner as in the Fig. 22. Now R is attached to $P \cup Q$ along the boundary AFCD giving \mathbb{P}^2 . Here Y is a critical point of type II. See Fig. 22.

Proposition 18. Let $f: M \to \mathbb{R}$ be an ordered Morse function. Let f have only one critical point of index 0, 1 and 2 each. Then M is homeomorphic to a projective plane.

Proof. Let X, Y and Z be critical points of index 0,1 and 2 respectively. Let a and b are regular values of f such that f(X) < a < f(Y) < b < f(Z). Then M(a) and M'(b) are

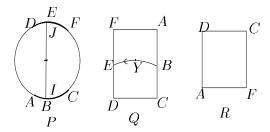


Figure 21: Canonical neighbourhoods of critical points

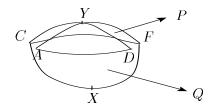


Figure 22: Projective plane minus a disc after passing the minimum and the saddle point

discs. Since V(a) and V(b) have only one component each, Y has to be of type II. Hence by case 3 of analysis, M(b) is a Möbius band. M is obtained by attaching a disc to the Möbius band M(b), hence M(a) is homeomorphic to a projective plane.

Proposition 19. Let f have only one critical point of index 0, only one critical point of index 2 and two critical points of index 1 of type I. Then M is homeomorphic to a torus.

Proof. Let X be the critical point of f of index 0, Y_1, Y_2 , critical points of index 1 and Z, the critical point of index 2. Let a and b be regular values of f which separate critical points of index 0, 1 and 2. That is, $f(X) < a < f(Y_i) < b < f(Z)$. Then M(a) and M'(b) are discs D and D' respectively. Let c be a regular value of f which separates Y_1 and Y_2 . Without loss of generality assume that $f(Y_1) < c < f(Y_2)$.

By Theorem 16, M(c) is obtained by attaching a rectangle to M(a). When we pass a critical point of index 1 of type I of the first kind we attach a rectangle to disjoint segments I and J belonging to different components of V(a). Since V(a) has only one component, Y_1 is of first kind. That is, M(c) is homeomorphic to a cylinder. Now M(b) has one boundary component and is obtained by attaching a rectangle to M(c). This forces Y_2 to be of the second kind. Thus M(b) is homeomorphic to a rectangle attached to a cylinder along two disjoint segments in different components of the boundary of the cylinder. (See Fig. 23.) Hence M(b) is homeomorphic to a torus but for a disc. M is obtained by attaching a disc to M(b), is homeomorphic to a torus.

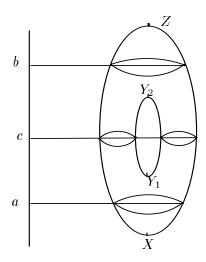


Figure 23: Critical points on a torus

6 Connected Sums and All that

Definition 10. A compact surface M (with or without boundary) is said to be non orientable if it contains a homeomorphic copy of a Möbius band.

Proposition 20. The Möbius strip and the projective plane are non orientable.

Definition 11. Let M_1, M_2 be closed surfaces and M'_i for i = 1, 2 be the space obtained by removing a disc from M_i . Then the boundary of M'_i is a circle. The connected sum of M_1 and M_2 is the space obtained attaching M'_1 to M'_2 via a homeomorphism from $\partial M'_1$ to $\partial M'_2$. We will denote this by $M_1 \# M_2$.

Remark 8. The connected sum is independent of the discs removed and the homeomorphism on the boundary.

Let T_n denote the connected sum of n tori and P_n denote the connected sum of n projective planes. Let V_n denote T_n but for a disc and U_n denote P_n but for a disc. By convention T_1 is a torus and $P_1 = \mathbb{P}^2$ is a projective plane.

Let M and N be surfaces with same number of boundary components. When we say that M is attached to N along the boundary we mean that they are attached via a homeomorphism from ∂M to ∂N .

Let us admit the following facts about the connected sum of two surfaces. For more details refer [4] or [6].

Observation 1. \mathbb{P}^2 is a disc attached to a Möbius band along the boundary.

Observation 2. Let M be a closed surface. Then $M \# \mathbb{P}^2$ is homeomorphic to the space obtained by removing a disk from M and attaching a Möbius band along the boundary.

Observation 3. P_q is homeomorphic to the space obtained by removing q disjoint discs from S^2 and attaching q Möbius bands along the boundary.

Observation 4. T_q is homeomorphic to the space obtained by removing 2q disjoint discs from S^2 and attaching q handles (cylinders) along the boundary.

Proposition 21. The connected sum of T_m and P_n is homeomorphic to P_{2m+n} .

Proof. First of all we claim that the connected sum of T_1 and \mathbb{P}^2 is homeomorphic to P_3 . We denote homeomorphism by \simeq . Let us assume this claim and prove the result. To prove the result let us first prove that $T_m \# P_1 \simeq P_{2m+1}$.

We will proceed by induction on m. The case m=1 follows from the claim above. Let us assume the result for all k < m. Then

$$T_m \# P_1 \simeq T_{m-1} \# T_1 \# P_1$$

 $\simeq T_{m-1} \# P_3$ (first claim)
 $\simeq T_{m-1} \# P_1 \# P_2$
 $\simeq P_{2m-2+1} \# P_2$ (by induction)
 $\simeq P_{2m+1}$.

Now let us prove that $T_m \# P_n \simeq P_{2m+n}$. We apply induction on n. The case n=1 follows from above. Assume the result for k < n. Then

$$T_m \# P_n \simeq T_m \# P_{n-1} \# P_1$$

 $\simeq P_{2m+n-1} \# P_1$ (by induction)
 $\simeq P_{2m+n}$.

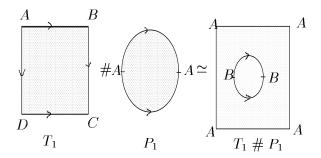


Figure 24: Connected sum of P_1 and T_1

We now prove our claim. The idea is to cut 3 Möbius bands from $T_1 \# P_1$ and then glue the remaining part to get S^2 with 3 holes. See Fig. 24, Fig. 25 and Fig. 26.

Torus is homeomorphic to a rectangle with AB identified to DC and AD identified to BC. Also P_1 is homeomorphic to D^2 with antipodal points identified. See Fig.24

In Fig. 25 M_1 , M_2 and M_3 are the three Möbius bands which are removed from $T_1 \# P_1$. Quotienting the remaining region after identifying R_i for $1 \le i \le 7$ we obtain S^2 with three holes. See Fig. 26.

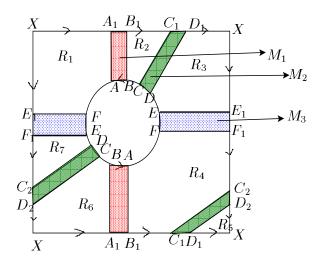


Figure 25: 3 Möbius bands on $T_1 \# P_1$

Proposition 22. T_n (respectively P_n) is not homeomorphic to $T_{n'}$ (respectively $P_{n'}$) for $n \neq n'$. Also T_n is not homeomorphic to P_m for any positive integers n and m.

Proof. Refer [4] or [6].
$$\Box$$

Ex 1. Let V'_1 denote torus without two disjoint discs. The space obtained by attaching V'_1 to V_n along the boundary of V_n is homeomorphic to V_{n+1} .

Similarly, let U'_1 denote a projective plane without two disjoint discs. The space obtained by attaching U'_1 to U_n along the boundary of U_n is homeomorphic to U_{n+1} .

7 Classification of Surfaces

Henceforth we assume that M is a closed surface and $f: M \to \mathbb{R}$ an ordered Morse function on M. Let $\{X_i\}_{1 \le i \le n(0)}$ be critical points of index 0, $\{Y_i\}_{1 \le i \le n(1)}$, critical points of index 1 and $\{Z_i\}_{1 \le i \le n(2)}$, critical points of index 2. Let a and b be regular values of f such that $f(X_i) < a < f(Y_j) < b < f(Z_k)$ for all $1 \le i \le n(0)$, $1 \le j \le n(1)$ and $1 \le k \le n(2)$.

Lemma 23. Let the notations be as above. Then there exist two regular values c, d, c < d and an ordered Morse function g having the same critical points with the same indices as f such that:

- $g^{-1}(-\infty, c]$ contains n(0) 1 critical points of index 1 and $g^{-1}(-\infty, c]$ is homeomorphic to a disc.
- $g^{-1}[d,\infty)$ contains n(2)-1 critical points of index 1 and $g^{-1}[d,\infty)$ is homeomorphic to a disc.
- $g^{-1}[c,d]$ contains n(1) n(0) n(2) + 2 critical points of index 1.

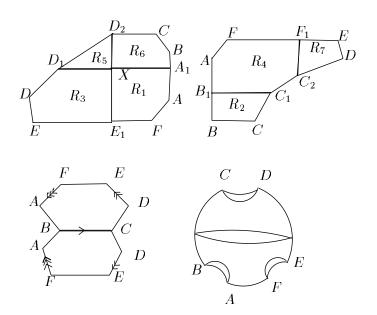


Figure 26: $T_1 \# P_1$ without 3 Möbius bands

Proof. Since all the critical points of index 1 are below the level b, it is enough to improve the function f below the level b. Also, second assertion follows from the first by taking -f in place of f.

M(a) is disjoint union of n(0) discs, say, $\{D_i\}_{1 \leq i \leq n(0)}$ and M'(b) is disjoint union n(2) discs say $\{D'_j\}_{1 \leq j \leq n(2)}$. That M(b) is connected follows from the fact that M is obtained by attaching M'(b) (disjoint union of discs) to M(b) along their boundaries.

First of all we prove the existence of the level c

If n(0) = 1, then take g = f and c = a.

If $n(0) \geq 2$, we proceed recursively. If all critical points of index 1 are of type II, then M(b) is disconnected. Hence there exists a critical point of index 1 of type I. If all critical points of index 1 of type I are of the second kind then again M(b) is disconnected. Hence there exists $j \in \{1, 2, \ldots, n(1)\}$ such that Y_j is of type I of first kind. Without loss of generality assume that j=1. Now choose an ordered Morse function g_1 having the same critical points with same indices as f and a regular value a of g_1 such that $g_1(X_i) < a < g_1(Y_1) < a_1 < g_1(Y_j)$ for all $1 \leq i \leq n(0)$ and $j \neq 1$. By definition 9, $V(a_1)$ has one less components than V(a). If n(0) - 1 = 1, then we are done. Otherwise proceeding in similar way we get another ordered Morse function g_2 , a level a_2 and a critical point Y_2 of type I of first kind such that $g_2(X_i) < a < g_2(Y_1) < a_1 < g_2(Y_2) < a_2 < g_2(Y_j)$ for all $1 \leq i \leq n(0)$ and $j \neq 1, 2$. Then the number of components of $V(a_2)$ is n(0) - 2. Recursively we get an ordered Morse function $g_{n(0)-1}$, a level $a_{n(0)-1}$ such the $g_{n(0)-1}(X_i) < a < g_{n(0)-1}(Y_j) < a_{n(0)-1}$ for all $1 \leq i \leq n(0)$ and $1 \leq j \leq n(0) - 1$. Define $c := a_{n(0)-1}$ and $g := g_{n(0)-1}$. This proves the first part of the first assertion of the lemma.

 $g^{-1}(-\infty,c]$ is homeomorphic to a disc follows from Corollary 17.

Corollary 24. Let the notations be as in the Lemma 23. Then the number of critical points of type I between the levels c and d is even, half of which are of the first kind and half are of the second kind and outside these levels all critical points are of type I.

Theorem 25 (Classification). Let $f: M \to \mathbb{R}$ be an ordered Morse function as in the claim of Lemma 23. Let f have q critical points of index 1 in W(c, d). Then M is homeomorphic to P_q or T_p , where q = 2p.

Proof. Case 1. Let us first assume that all critical points of index 1 in W(c, d) are of type II. In this case we claim that M is homeomorphic to P_q .

Let us proceed by induction on the number of critical points in W(c, a) where a > c is a regular value of f. We claim that if W(c, a) contains n critical points of index 1 of type II then M(a) is homeomorphic to U_n .

Let $a \in \mathbb{R}$ be a regular value of f such that W(c, a) contains only one critical point of index 1. Then M(a) is homeomorphic to a Möbius band, (Proposition 18), which is same as U_1 .

Let us assume the result for k=q-1. Let d be a regular value of f such that W(c,d) contains all the q critical points of index 1. We claim that M(d) is homeomorphic to U_q . Choose a regular value d_1 of f such that $W(c,d_1)$ contains q-1 critical points of f of index 1. Then by induction hypothesis $M(d_1)$ is homeomorphic to U_{q-1} . Now M(d) is homeomorphic to U_q follows from Exercise 1. Hence M homeomorphic to P_q .

Case 2. Let all critical points of index 1 in W(c, d) be of type I. Since the number of critical points of index 1 of type I in W(c, d) is even, q = 2p for some integer p. (Corollary 24.) In this case we claim that M is homeomorphic to T_p .

We modify the function f in W(c,d) to an ordered Morse function having same number of critical points as f with the same indices and such that every critical point of index 1 of type I of second kind is followed by a critical point of index 1 of type I of the first kind. (Theorem 16 and Corollary 24.) We assume that this is done.

Since M'(d) is a disc, it is enough to prove that M(d) is homeomorphic to V(p). Let us proceed by induction on the number of critical points in W(c, a) where a > c is a regular value of f. We claim that if W(c, a) contains n pairs of critical points of index 1 then M(a) is homeomorphic to V_n .

Let a be a regular value of f such that W(c,a) contains a pair of critical points of index 1 of type I. Then, by Proposition 23, M(a) is homeomorphic to V_1 . Let us assume the result for k = p - 1. Let d be a regular value of f such that W(c,d) contains all p pairs of critical points of index 1 of type I. We show that M(d) is homeomorphic to V_p . Let d_1 be a regular value of f such that $W(c,d_1)$ contains p-1 pairs of critical points of index 1 of type I of f. By induction hypothesis $M(d_1)$ is homeomorphic to V_{p-1} . Hence M(d) is homeomorphic to V_p . (Exercise 1.) Thus M is homeomorphic to T_p .

Case 3. Let f have critical points of index 1 of type I as well as of type II in W(c,d). In this case we claim that M is homeomorphic P_q .

There exists an integer k, 2k < q such that f has 2k critical points of index 1 of type I in W(c,d). (Corollary 24.) The remaining critical points of f of index 1 are of type II. We modify f in W(c,d) to an ordered Morse function g such that if $Y,Y' \in W(c,d)$ are critical points of index 1 of type I and type II respectively, then g(Y) < g(Y'). Further assume that every critical point of g of index 1 of type I of the second kind is followed by a critical point of type I of the first kind in W(c,d). Let g be a regular value of g which separates critical points of type I and type II in g in g case 1, g is homeomorphic g by g. Similarly, by case 2, g is homeomorphic to g by g in the first kind in g is homeomorphic to g by g in the first kind in g in the first kind in g is homeomorphic to g by g in the first kind in g in the

Theorem 26. Let M be a compact surface with boundary. Let k be the number of boundary components of M. Then M is homeomorphic to either T_n with k holes or P_m with k holes for some integers m and n.

Proof. Since the boundary of a surface is a closed 1-manifold, each boundary component is a circle. Attach k discs along the boundary components. The resulting surface is a closed surface. By Theorem 25, the resulting surface is either T_n or P_m .

Acknowledgement

The first author thanks the Department of Mathematics, University of Mumbai for the hospitality during the period in which this work was undertaken.

The second author thanks the National Board for Higher Mathematics for its research fellowship.

We thank Professor Akhil Ranjan for going through this article.

References

- [1] Coddington & Levinson, Ordinary Differential Equations
- [2] A. Gramain, Topologie des Surfaces, Presses Universitarires de France (1971)
- [3] Guillemin & Pollack, Differential Topology Prentice-Hall (1974)
- [4] Hirsch, Differential Topology, GTM Series (33), Springer-Verlag
- [5] S. Kumaresan, A Basic Course in Differential Geometry and Lie Groups, To be published
- [6] W. S. Massey, Basic Course in Algebraic Topology, GTM Series (127), Springer-Verlag
- [7] J. Milnor, Morse Theory, Annals of Math. Studies (51), 1963

Abhijit Champanerkar

Research Fellow, Department of Mathematics, Columbia University, New York abhijit@math.columbia.edu

Ajit Kumar, Research Fellow

Department of Mathematics, University of Mumbai, Mumbai - 400 098, India ajitkumar@math.mu.ac.in

S. Kumaresan

Department of Mathematics, University of Mumbai, Mumbai - 400 098, India kumaresa@math.mu.ac.in