Modern Birkhäuser Classics

V.I. Arnold
S.M. Gusein-Zade
A.N. Varchenko

Singularities of Differentiable Maps, Volume 2

Monodromy and Asymptotics of Integrals

Birkhäuser

Modern Birkhäuser Classics

Many of the original research and survey monographs, as well as textbooks, in pure and applied mathematics published by Birkhäuser in recent decades have been groundbreaking and have come to be regarded as foundational to the subject. Through the MBC Series, a select number of these modern classics, entirely uncorrected, are being re-released in paperback (and as eBooks) to ensure that these treasures remain accessible to new generations of students, scholars, and researchers.

Singularities of Differentiable Maps

Volume 2

Monodromy and Asymptotics of Integrals

V.I. Arnold
S.M. Gusein-Zade
A.N. Varchenko

Reprint of the 1988 Edition

Birkhäuser

V.I. Arnold (deceased)

A.N. Varchenko
Department of Mathematics
The University of North Carolina
at Chapel Hill
Chapel Hill, NC, USA

S.M. Gusein-Zade
Department of Mechanics and Mathematics
Moscow State University
Moscow, Russia

Originally published as Vol. 83 in the series *Monographs in Mathematics*

ISBN 978-0-8176-8342-9 ISBN 978-0-8176-8343-6 (eBook)
DOI 10.1007/978-0-8176-8343-6
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012938547

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

V. I. Arnold
S. M. Gusein-Zade
A. N. Varchenko

Singularities of Differentiable Maps

Volume II

Monodromy and Asymptotics of Integrals

Under the Editorship of V. I. Arnold

1988

Birkhäuser
Boston · Basel · Berlin

Originally published as
Osobennosti differentsiruemykh otobrazhenii
by Nauka, Moscow, 1984

Translated by Hugh Porteous

Translation revised by the authors
and James Montaldi

Library of Congress Cataloging in Publication Data
(Revised for vol. 2)

Arnol'd, V. I. (Vladimir Igorevich), 1937.
Singularities of differentiable maps.

(Monographs in matematics ; vol. 82 –)
Translation of: Osobennosti differentsiruemykh
otobrazheniĭ.
Includes bibliographies and indexes.
Contents: v. 1. The classification of critical
points, caustics and wave fronts -- v. 2. Monodromy
and asymptotics of integrals.
1. Differentiable mappings. 2. Singularities (Mathematics)
I. Guseĭn-Zade. S. M. (Sabir Medzhidovich)
II. Varchenko, A. N. (Aleksandr Nikolaevich) III. Title.

IV. Series.
QA614.58.A7513 1985 514'.72 84-12134
ISBN 0-8176-3187-9

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Arnol'd, Vladimir I.:
Singularities of differentiable maps / V. I.
Arnold ; S. M. Guseĭn-Zade ; A. N. Varchenko. –
Boston ; Basel ; Stuttgart : Birkhäuser
Einheitssacht.: Osobennosti differenciruemych
otobraženij ⟨engl.⟩

NE: Gusejn-Zade, Sabir M.:; Varčenko, Aleksandr N.:

Vol. 2. Monodromy and asymptotics of integrals / under
the editorship of V. I. Arnold. Transl. by Hugh
Porteous. – 1987.

(Monographs in mathematics ; Vol. 83)
ISBN 3-7643-3185-2 (Basel...)
ISBN 0-8176-3185-2 (Boston)

NE: GT

Printed in Germany
ISBN 0-8176-3185-2
ISBN 3-7643-3185-2

Preface

The present volume is the second volume of the book "Singularities of Differentiable Maps" by V. I. Arnold, A. N. Varchenko and S. M. Gusein-Zade. The first volume, subtitled "Classification of critical points, caustics and wave fronts", was published by Moscow, "Nauka", in 1982. It will be referred to in this text simply as "Volume 1".

Whilst the first volume contained the zoology of differentiable maps, that is it was devoted to a description of what, where and how singularities could be encountered, this volume contains the elements of the anatomy and physiology of singularities of differentiable functions. This means that the questions considered in it are about the structure of singularities and how they function.

Another distinctive feature of the present volume is that we take a hard look at questions for which it is important to work in the complex domain, where the first volume was devoted to themes for which, on the whole, it was not important which field (real or complex) we were considering. Such topics as, for example, decomposition of singularities, the connection between singularities and Lie algebras and the asymptotic behaviour of different integrals depending on parameters become clearer in the complex domain.

The book consists of three parts. In the first part we consider the topological structure of isolated critical points of holomorphic functions. We describe the fundamental topological characteristics of such critical points: vanishing cycles, distinguished bases, intersection matrices, monodromy groups, the variation operator and their interconnections and method of calculation.

The second part is devoted to the study of the asymptotic behaviour of integrals of the method of stationary phase, which is widely met with in applications. We give an account of the methods of calculating asymptotics, we discuss the connection between asymptotics and various characteristics of critical points of the phases of integrals (resolution of singularities, Newton polyhedra), we give tables of the orders of asymptotics for critical points of the phase which were classified in Volume 1 of this book (in particular for simple, unimodal and bimodal singularities).

The third part is devoted to integrals evaluated over level manifolds in a neighbourhood of the critical point of a holomorphic function. In it we shall consider integrals of holomorphic forms, given in a neighbourhood of a critical point, over cycles, lying on level hypersurfaces of the function. Integral of a holomorphic form over a cycle changes holomorphically under continuous deformation of the cycle from one level hypersurface to another. In this way there arise many-valued holomorphic functions, given on the complex line in a neighbourhood of a critical value of the function. We show that the asymptotic behaviour of these functions (that is the asymptotic behaviour of the integrals) as the level tends to the critical one is connected with a variety of characteristics of the initial critical point of the holomorphic function.

The theory of singularities is a vast and rapidly developing area of mathematics, and we have not sought to touch on all aspects of it.

The bibliography contains works which are directly connected with the text (although not always cited in it) and also works connected with volume 1 but for some or other reason not contained in its bibliography.

References in the text to volume 1 refer to the above-mentioned book "Singularities of Differentiable Maps".

The authors offer their thanks to the participants in the seminar on singularity theory at Moscow State University, in particular A. M. Gabrielov, A. B. Givental, A. G. Kushnirenko, D. B. Fuks, A. G. Khovanski and S. V. Chmutov. The authors also wish to thank V. S. Varchenko and T. V. Ogorodnikova for rendering inestimable help in preparing the manuscript for publication.

The authors.

Contents

Part I

The topological structure of isolated critical points of functions

Introduction

In the topological investigation of isolated critical points of complex-analytic functions the problem arises of describing the topology of its level sets. The topology of the level sets or infra-level sets of smooth real-valued functions on manifolds may be investigated with the help of Morse theory (see [255]). The idea there is to study the change of structure of infra-level sets and level sets of functions upon passing critical values. In the complex case passing through a critical value does not give rise to an interesting structure, since all the non-singular level sets near one critical point are not only homeomorphic but even diffeomorphic. The complex analogue of Morse theory, describing the topology of level sets of complex analytic functions, is the theory of Picard-Lefschetz (which historically precedes Morse theory). In Picard-Lefschetz theory the fundamental principle is not passing through a critical point but going round it in the complex plane.

Let us fix a circle, going round the critical value. Each point of the circle is a value of the function. The level sets, corresponding to these values, give a fibration over the circle. Going round the circle defines a mapping of the level set above the initial point of the circle into itself. This mapping is called the (classical) *monodromy* of the critical point.

The simplest interesting example in which one can observe all this clearly and carry through the calculations to the end is the function of two variables given by

$$f(z,w)=z^2+w^2, \qquad (z,w)\in\mathbb{C}^2.$$

It has a unique critical point $z=w=0$. The critical value is $f=0$. The critical level set $V_0=\{(z,w): z^2+w^2=0\}$ consists of two complex lines intersecting in the point 0. All the other level sets

$$V_\lambda = \{(z, w) : z^2 + w^2 = \lambda\} \qquad (\lambda \neq 0)$$

are topologically the same; they are diffeomorphic to a cylinder $S^1 \times \mathbb{R}^1$ (figure 1).

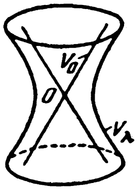

Fig. 1.

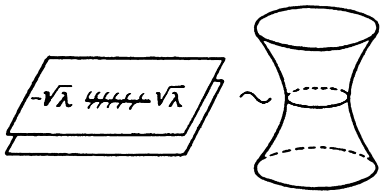

Fig. 2.

To show this, we consider the Riemann surface of the function $w = \sqrt{(\lambda - z^2)}$ (figure 2). This surface is glued together from two copies of the complex z-plane, joined along the cut $(-\sqrt{\lambda}, \sqrt{\lambda})$. Each copy of the cut plane is homeomorphic to a half cylinder; the line of the cut corresponds to a circumference of the cylinder. In this way, the whole (four-real-dimensional) space $\mathbb{C}^2$ decomposes into the singular fibre V_0 and the non-singular fibres V_λ, diffeomorphic to cylinders, mapping to the critical value 0 and the non-critical values $\lambda \neq 0$ by the mapping

$$f : (\mathbb{C}^2, 0) \to (\mathbb{C}, 0).$$

Let us proceed to the construction of the monodromy. We consider on the target plane a path going round the critical value 0 in the positive direction (anticlockwise):

$$\lambda(t) = \exp(2\pi i t)\alpha, \quad 0 \leqslant t \leqslant 1, \quad \alpha > 0 \text{ (figure 3)}$$

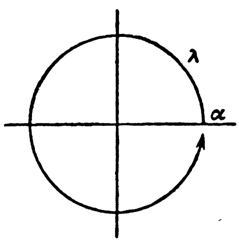

Fig. 3.

Let us observe how the fibre $V_{\lambda(t)}$ changes as t varies from 0 to 1. For this we consider the Riemann surfaces of the functions

$$w=\sqrt{(\lambda(t)-z^2)}.$$

As the parameter t increases, both the branch points $z=\pm\sqrt{\lambda(t)}=\exp(\pi it)\times(\pm\sqrt{\alpha})$ move around the point $z=0$ in the positive direction. As t varies from 0 to 1, each of these points performs a half turn and arrives at the other's place. In this way, as $\lambda(t)$ goes round the critical value 0, it corresponds to a sequence of Riemann surfaces, depicted in figure 4, beginning and ending with the same surface V_α.

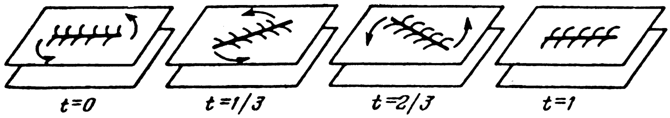

Fig. 4.

Now it is easy to construct a family continuous in t, of diffeomorphisms from the initial fibre $V_{\lambda(0)}=V_\alpha$ to the fibre $V_{\lambda(t)}$ over the point $\lambda(t)$

$$\Gamma_t: V_{\lambda(0)}\to V_{\lambda(t)}$$

beginning with the identity map, Γ_0, and ending with the *monodromy* $\Gamma_1=h$. For example one may define Γ_t in the following fashion. Choose a smooth "bump function", $\chi(\tau)$, such that

$$\chi(\tau)=1 \text{ for } 0\leqslant\tau\leqslant 2\sqrt{\alpha},$$

$$\chi(\tau)=0 \text{ for } \tau\geqslant 3\sqrt{\alpha}.$$

We let

$$g_t(z)=\exp\{\pi it\cdot\chi(|z|)\}\cdot z.$$

The family of diffeomorphisms g_t from the complex z-plane into itself defines the desired family of diffeomorphisms Γ_t. The diffeomorphism $h=\Gamma_1: V_\alpha\to V_\alpha$ of the cylinder is the identity outside a sufficiently large compact set (for $|z|>3\sqrt{\alpha}$).

We consider now the action of the monodromy h on the homology of a non-singular fibre V_α. The first homology group $H_1(V_\alpha; \mathbb{Z}) \approx \mathbb{Z}$ of the cylinder V_α is generated by the homology class of the "gutteral" circle Δ (figure 5). As $\alpha \to 0$ the circle Δ tends to the point 0. Therefore it is called the *vanishing cycle of Picard-Lefschetz.*

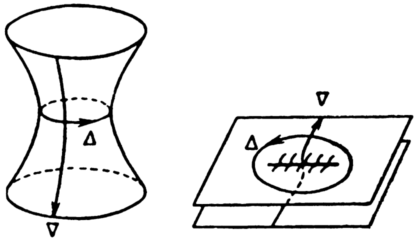

Fig. 5.

We consider further the first homology group $H_1^{cl}(V_\alpha; \mathbb{Z})$ of the fibre V_α with closed support. According to Poincaré duality, this group is also isomorphic to the group $\mathbb{Z}$ of integers. It is generated by the homology class of the "covanishing cycle" ∇ – a line on the cylinder going from infinity to infinity and intersecting the vanishing cycle, Δ, once transversely (see figure 5). We shall suppose that the cycle ∇ is oriented in such a way that its intersection number $(\nabla \circ \Delta)$ with the vanishing cycle Δ, determined by the complex orientation of the fibre V_α, is equal to $+1$.

Figure 4 allows us to observe the action of the diffeomorphisms Γ_t on the vanishing and covanishing cycles (figure 6).

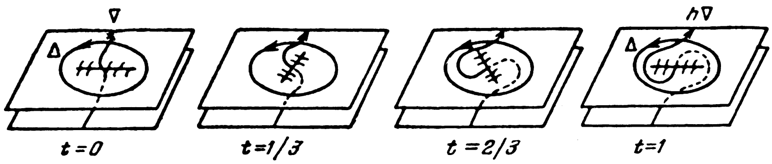

Fig. 6.

We notice that the diffeomorphism $h = \Gamma_1$ of the cylinder $V_\alpha \approx S^1 \times \mathbb{R}^1$ can be described as follows: it is fixed outside a certain annulus, the circles forming the annulus rotate through various angles varying from 0 at one edge to 2π at the other. In this way, under the action of the monodromy mapping h. the vanishing cycle Δ is mapped into itself, the covanishing cycle winds once around the cylinder (figure 7).

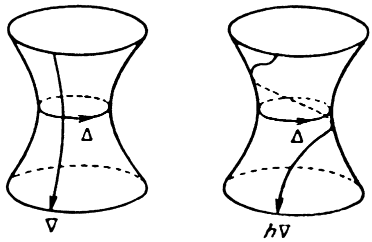

Fig. 7.

The diffeomorphism h is the identity outside some compact set. Outside this compact set the cycles ∇ and $h\nabla$ coincide. Therefore the cycle $h\nabla - \nabla$ is concentrated in a compact part of the cylinder. From figure 7 (or from figure 6) it is clear that

$$h\nabla - \nabla = -\Delta.$$

In this way any cycle δ with closed support gives rise to a cycle $h\delta - \delta$ with compact support. This defines a mapping from the homology of the fibre V_α with closed support into its homology with compact support. It is called the *variation* and is denoted by

$$\mathrm{Var}: H_1^{cl}(V_\alpha; \mathbb{Z}) \to H_1(V_\alpha; \mathbb{Z}).$$

From figure 7 or figure 6 it can be seen that we have

$$\mathrm{Var}\ \delta = (\Delta \circ \delta)\, \Delta$$

for every cycle

$$\delta \in H_1^{cl}(V_\alpha; \mathbb{Z}).$$

Here $(\Delta \circ \delta)$ is the intersection number of the cylces Δ and δ, defined by the complex orientation of the fibre V_α. This relationship is called the *formula of Picard-Lefschetz*.

We notice that, generally speaking, the diffeomorphisms Γ_t are defined only up to homotopy and that there is no a priori reason why the mapping Γ_1 should be fixed outside a compact set. For example, the family of diffeomorphisms

$$\Gamma_t' : (z, w) \to (z \exp(\pi i t), w \exp(\pi i t))$$

defines the mapping

$$\Gamma_1' = h' : (z, w) \to (-z, -w),$$

which is not fixed outside a compact set and is not, therefore, suitable for defining the variation (though it is suitable for defining the action of the monodromy on the compact homology). Thus we have considered the fundamental concepts of the theory of Picard-Lefschetz: vanishing cycles, monodromy and variation for the simplest example of the function $f(z, w) = z^2 + w^2$.

In the general case of an arbitrary function of any number of variables, the topology of the fibre V_λ will not be as simple as in the example we analysed. The investigation of the topology of the fibre V_λ, the monodromy and the variation in the general case is a difficult problem, solved completely only for a few special cases. In this part we shall recount several methods and results which have been obtained along these lines.

The fundamental method which we shall make use of is the method of deformation (or perturbation). Under a small perturbation, a complicated critical point of a function of n variables breaks up into simple ones. These simple critical points look like the critical point 0 of the function

$$f(z_1, \ldots, z_n) = z_1^2 + \ldots + z_n^2$$

and can be investigated completely in the same way as we analysed the case $n = 2$ above. In place of the cylinders V_λ, which occured in the case $n = 2$, the non-singular fibre in the general case are smooth manifolds

$$V_\lambda = \{(z_1, \ldots, z_n) : z_1^2 + \ldots + z_n^2 = \lambda\}, \quad \lambda \neq 0,$$

which are diffeomorphic to the space TS^{n-1}, the tangent bundle of the $(n-1)$-dimensional sphere (giving a cylinder for $n = 2$). The vanishing cycle in V_1 is the real sphere

$$S^{n-1} = \{z \in \mathbb{R}^n \subset \mathbb{C}^n : z_1^2 + \ldots + z_n^2 = 1\}.$$

If complicated critical points break down under deformation into μ simple ones, then the perturbed function will have, in general, μ critical values (figure 8).

In this case it is possible in the target plane of the perturbed function to go round each of the μ critical values. In this way we get, not one monodromy diffeomorphism h, but a whole *monodromy group* $\{h_\gamma\}$, where γ runs through the fundamental group of the set of non-critical values.

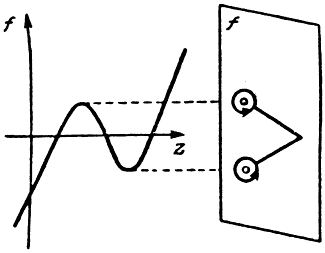

Fig. 8.

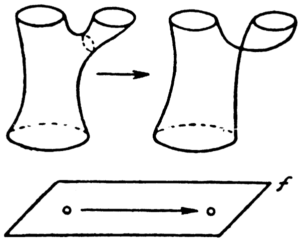

Fig. 9.

The non-singular fibre V_λ of the perturbed function have the same structure (inside some ball surrounding the critical point of the initial function) as the non-singular fibre of the initial function. When the value of λ tends to one of the critical values of the perturbed function, a certain cycle on the non-singular fibre vanishes. This cycle is a sphere whose dimension is a half of the (real) dimension of the fibre V_λ (figure 9). Tending in this manner to all μ critical values, we define in the non-singular fibre μ vanishing cycles, each a sphere of the middle dimension. It happens that the non-singular fibre is homotopy equivalent to a bouquet of these spheres.

In the case when the real dimension of the non-singular fibre is divisible by 4 (that is when the number n of variables is odd), the intersection number gives a symmetric bilinear form in the homology group $H_{n-1}(V_\lambda; \mathbb{Z})$ of the non-singular level manifold. The self-intersection number of each of the vanishing cycles is equal to 2 or -2, depending on the number of variables n. The action of going round the critical value corresponding to a vanishing cycle is equivalent to reflection in a mirror which is orthogonal to this cycle, where orthogonality is defined by the scalar product given by the intersection numbers.

For example, for the function of three variables

$$f(z_1, z_2, z_3) = z_1^{k+1} + z_2^2 + z_3^2$$

a suitable perturbation is

$$\tilde{f}(z_1, z_2, z_3) = f(z_1, z_2, z_3) - \varepsilon z_1.$$

This function has $\mu=k$ critical points $(\sqrt[k]{\varepsilon/(k+1)}\,\xi_m, 0, 0)$, where ξ_m $(m=1,\ldots,k)$ are the k-th roots of unity. The corresponding vanishing cycles $\Delta_1,\ldots,\Delta_k$ can be chosen so that they have the following intersection numbers

$$(\Delta_i \circ \Delta_i) = -2,$$

$$(\Delta_1 \circ \Delta_2) = (\Delta_2 \circ \Delta_3) = \ldots = (\Delta_{k-1} \circ \Delta_k) = 1,$$

and all other intersection numbers are zero. The monodromy group is generated by the resulting reflections in the orthogonal complements of the cycles Δ_m, and coincides with the Weyl group A_k (see [53]), that is with the group $S(k+1)$ of permutations of $(k+1)$ elements.

In this Part we shall generally (except in Chapter 5) be concerned with isolated singularities of functions. Therefore by the term "singularity" we shall understand the germ of a holomorphic function $f:(\mathbb{C}^n, 0)\to(\mathbb{C}, 0)$, having at the origin an isolated critical point (that is, a point at which all the partial derivatives of the function f are equal to zero).

Let $G:(\mathbb{C}^n, 0)\to(\mathbb{C}^p, 0)$ be the germ at the origin of an analytic function, let U be a neighbourhood of the origin in the space $\mathbb{C}^n$, in which a representative of the germ G is defined, and let G_λ be a family of functions from U to $\mathbb{C}^p$, analytic in λ in a neighbourhood of 0 in $\mathbb{C}$, such that $G_0=G$. We shall refer to the function G_λ, for sufficiently small λ, as a *small perturbation* $\tilde{G}$ of the function G, without spelling out each time the dependence on the parameter λ.

Throughout, the absolute homology groups will be considered reduced modulo a point; the relative groups of a pair "manifold-boundary" will be modulo a fundamental cycle. (For this reason the tilde over the letter H, which usually indicates a reduced homology group, will be omitted.) All the homology will be considered with coefficients in the group $\mathbb{Z}$ of integers, unless we specifically indicate otherwise.

Let $\mathbb{C}^n$ be the n-dimensional complex vector space with coordinates

$$x_j = u_j + iv_j, \qquad (j=1,\ldots,n;\ u_j \text{ and } v_j \text{ real}).$$

The space $\mathbb{C}^n \cong \mathbb{R}^{2n}$, considered as a real $2n$-dimensional vector space, has a preferred orientation, which we shall call the complex one. This orientation is defined so that the system of coordinates in the space $\mathbb{R}^{2n}$ given by $u_1, v_1, u_2, v_2, \ldots, u_n, v_n$ has positive orientation. Complex manifolds will be considered to have this complex orientation unless we specifically indicate otherwise. With this choice of orientation the intersection numbers of complex submanifolds will always be non-negative.

Chapter 1

Elements of the theory of Picard-Lefschetz

In this chapter we shall define concepts of Picard-Lefschetz theory such as vanishing cycles, the monodromy and variation operators, the Picard-Lefschetz operators, etc. As we have already said, they are used to investigate the topology of critical points of holomorphic functions.

1.1 The monodromy and variation operators

Let $f: M^n \to \mathbb{C}$ be a holomorphic function on an n-dimensional complex manifold M^n, with a smooth boundary ∂M^n (in the real sense). Let U be a contractible compact region in the complex plane with smooth boundary ∂U. We shall suppose that the following conditions are satisfied:

(i) For some neighbourhood U' of the region U, the restriction of f to the preimage of U' is a proper mapping $f^{-1}(U') \to U'$, that is a mapping for which the preimage of any compact set is compact.

(ii) The restriction of f to $\partial M^n \cap f^{-1}(U')$ is a regular mapping into U', that is a mapping, the differential of which is an epimorphism.

(iii) The function f has in the preimage, $f^{-1}(U')$, of the region U' a finite number of critical points p_i $(i = 1, \ldots, \mu)$ with critical values $z_i = f(p_i)$ lying inside the region U, that is in $U \setminus \partial U$.

From condition (ii) it follows that the restriction of the function f to $\partial M^n \cap f^{-1}(U)$ defines a locally trivial, and consequently (since the region U is assumed contractible) also a trivial fibration $\partial M^n \cap f^{-1}(U) \to U$. The direct product structure in the space of this fibration is unique up to homotopy. In addition, the restriction of the function f to the preimage $f^{-1}(U \setminus \{z_i\})$ of the set of non-critical values is a locally trivial fibration.

We will denote by F_z $(z \in U)$ the level set of the function f $(F_z = f^{-1}(z))$. If $z \in U$ is a non-critical value of the function f, then the corresponding level set F_z is a compact $(n-1)$-dimensional complex manifold with smooth boundary $\partial F_z = F_z \cap \partial M^n$. Let us fix a non-critical value z_0 lying on the boundary ∂U of the region U. Let γ be a loop in the complement of the set of critical values

V.I. Arnold et al., *Singularities of Differentiable Maps, Volume 2*, Modern Birkhäuser Classics,
DOI 10.1007/978-0-8176-8343-6_1,

$U\setminus\{z_i|i=1,\ldots,\mu\}$ with initial and end points at z_0 ($\gamma:[0,1]\to U\setminus\{z_i\}, \gamma(0)=\gamma(1)=z_0$). (We can suppose, without loss of generality, that all the loops and paths we encounter are piecewise smooth.) Going round the loop γ generates a continuous family of mappings $\Gamma_t: F_{z_0}\to M^n$ (lifting homotopy), for which Γ_0 is the identity map from the level manifold F_{z_0} into itself, $f(\Gamma_t(x))=\gamma(t)$, that is Γ_t maps the level manifold F_{z_0} into the level manifold $F_{\gamma(t)}$. The homotopy Γ_t can and will be chosen to be consistent with the direct product structure on $\partial M^n \cap f^{-1}(U)$. Indeed we can choose as $\{\Gamma_t\}$ a family of diffeomorphisms $F_{z_0}\to F_{\gamma(t)}$, but we shall not need this in the sequel. Thus the map

$$h_\gamma=\Gamma_1: F_{z_0}\to F_{z_0}$$

is the identity map on the boundary ∂F_{z_0} of the level manifold F_{z_0}. It is defined uniquely up to homotopy (fixed on the boundary ∂F_{z_0}) by the class of the loop γ in the fundamental group $\pi_1(U\setminus\{z_i\}, z_0)$ of the complement of the set of critical values.

Definitions. The transformation h_γ of the non-singular level set F_{z_0} into itself is called the *monodromy* of the loop γ. The action $h_{\gamma *}$ of the transformation h_γ on the homology of the non-singular level set $H_*(F_{z_0})$ is called the *monodromy operator* of the loop γ.

The monodromy operator is uniquely defined by the class of the loop γ in the fundamental group of the complement of the set of critical values.

We shall discuss also the automorphism $h_{\gamma *}^{(r)}$ induced by the transformation h_γ in the relative homology group $H_*(F_{z_0}, \partial F_{z_0})$ of the non-singular level set modulo its boundary. In the introduction to this Part we used, instead of the relative homology group $H_*(F_{z_0}, \partial F_{z_0})$, the homology group $H_1^{cl}(V_\alpha)$ with closed support (using the isomorphism

$$H_*(F_{z_0}, \partial F_{z_0})\cong H_*^{cl}(F_{z_0}\setminus \partial F_{z_0})).$$

Let δ be a relative cycle in the pair $(F_{z_0}, \partial F_{z_0})$. Since the transformation h_γ is the identity on the boundary ∂F_{z_0} of the level manifold F_{z_0}, the boundary of the cycle $h_\gamma\delta$ coincides with the boundary of the cycle δ. Therefore the difference $h_\gamma\delta-\delta$ is an absolute cycle in the manifold F_{z_0}. It is not hard to see that the mapping $\delta\mapsto h_\gamma\delta-\delta$ gives the correct definition of the homomorphism

$$\operatorname{var}_\gamma: H_*(F_{z_0}, \partial F_{z_0})\to H_*(F_{z_0}).$$

Definition. The homomorphism

$$\mathrm{var}_\gamma : H_*(F_{z_0}, \partial F_{z_0}) \to H_*(F_{z_0})$$

is called the *variation operator* of the loop γ.

It is not difficult to see that the automorphisms $h_{\gamma *}$ and $h^{(r)}_{\gamma *}$ are connected with the variation operator by the relations

$$h_{\gamma *} = id + \mathrm{var}_\gamma \cdot i_*,$$

$$h^{(r)}_{\gamma *} = id + i_* \cdot \mathrm{var}_\gamma,$$

where

$$i_* : H_*(F_{z_0}) \to H_*(F_{z_0}, \partial F_{z_0})$$

is the natural homomorphism induced by the inclusion

$$F_{z_0} \subset (F_{z_0}, \partial F_{z_0}).$$

If the class of the loop γ in the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values is equal to the product $\gamma_1 \cdot \gamma_2$ of the classes γ_1 and γ_2, then

$$\mathrm{var}_\gamma = \mathrm{var}_{\gamma_1} + \mathrm{var}_{\gamma_2} + \mathrm{var}_{\gamma_2} \cdot i_* \cdot \mathrm{var}_{\gamma_1},$$

$$h_{\gamma *} = h_{\gamma_2 *} \cdot h_{\gamma_1 *},$$

$$h^{(r)}_{\gamma *} = h^{(r)}_{\gamma_2 *} \cdot h^{(r)}_{\gamma_1 *}.$$

Therefore the mapping $\gamma \mapsto h_{\gamma *}$ is an (anti)homomorphism of the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values into the group Aut $H_*(F_{z_0})$ of automorphisms of the homology group $H_*(F_{z_0})$ of the non-singular level set. We shall denote by $(a \circ b)$ the intersection number of the cycles (or homology classes) a and b. This notation will be used both in the case when both the cycles a and b are absolute and in the case when one of them is relative. Remember that the level manifold F_{z_0} is a complex manifold and therefore possesses the preferred orientation which defines the intersection number of the cycles on it.

Lemma 1.1. Let

$$a, b \in H_*(F_{z_0}, \partial F_{z_0})$$

be relative homology classes,

$$\dim a + \dim b = 2n - 2,$$

$$\gamma \in \pi_1(U \setminus \{z_i\}, z_0).$$

Then

$$(\mathrm{var}_\gamma a \circ \mathrm{var}_\gamma b) + (a \circ \mathrm{var}_\gamma b) + (\mathrm{var}_\gamma a \circ b) = 0.$$

Proof. Choose relative cycles which are representatives of the homology classes a and b so that their boundaries (lying in the boundary ∂F_{z_0} of the level manifold F_{z_0}) do not intersect. This can be done using the dimensional relationships

$$\begin{aligned} \dim \partial a + \dim \partial b &= 2n - 4 \\ &< 2n - 3 \\ &= \dim \partial F_{z_0}. \end{aligned}$$

The chosen cycles we shall also denote by a and b. For such cycles the intersection number makes sense, though, of course, it is not an invariant of the classes in the homology group $H_*(F_{z_0}, \partial F_{z_0})$. We have

$$\mathrm{var}_\gamma a = h_\gamma a - a, \qquad \mathrm{var}_\gamma b = h_\gamma b - b,$$

$$\begin{aligned} &(\mathrm{var}_\gamma a \circ \mathrm{var}_\gamma b) + (\mathrm{var}_\gamma a \circ b) + (a \circ \mathrm{var}_\gamma b) = \\ &\qquad = (h_\gamma a \circ h_\gamma b) - (a \circ h_\gamma b) - (h_\gamma a \circ b) + (a \circ b) + \\ &\qquad + (h_\gamma a \circ b) - (a \circ b) + (a \circ h_\gamma b) - (a \circ b) = 0, \end{aligned}$$

since $(h_\gamma a \circ h_\gamma b) = (a \circ b)$.

Corollary. For $a, b \in H_*(F_{z_0}, \partial F_{z_0})$

$$(h_{\gamma *}^{(r)} a \circ \mathrm{var}_\gamma b) + (\mathrm{var}_\gamma a \circ b) = 0.$$

Proof.

$$\begin{aligned} &(h_{\gamma *}^{(r)} a \circ \mathrm{var}_\gamma b) + (\mathrm{var}_\gamma a \circ b) \\ &= (i_* \cdot \mathrm{var}_\gamma a \circ \mathrm{var}_\gamma b) + (a \circ \mathrm{var}_\gamma b) + (\mathrm{var}_\gamma a \circ b) \\ &= 0 \end{aligned}$$

since

$$h^{(r)}_{\gamma *} = id + i_* \cdot \mathrm{var}_\gamma,$$
$$(i_* \cdot \mathrm{var}_\gamma a \circ \mathrm{var}_\gamma b) = (\mathrm{var}_\gamma a \circ \mathrm{var}_\gamma b).$$

1.2 Vanishing cycles and the monodromy group

Let us suppose now that all critical points p_i of the function f are non-degenerate (that is that $\det(\partial^2 f/\partial x_j \partial x_k) \neq 0$), and all critical values $z_i = f(p_i)$ are different ($i = 1, \ldots, \mu$). Remember that in this case the function f is said to be *Morse*.

Definition. The *monodromy group* of the (Morse) function f is the image of the homomorphism of the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values in the group Aut $H_*(F_{z_0})$ of automorphisms of the homology group $H_*(F_{z_0})$ of the non-singular level set F_{z_0} which is obtained by mapping the loop γ into the monodromy operator

$$h_{\gamma *} : H_*(F_{z_0}) \to H_*(F_{z_0}).$$

Let us be given in the region U a path $u : [0, 1] \to U$, joining some critical value z_i with the non-critical value z_0 ($u(0) = z_i$, $u(1) = z_0$) and not passing through critical values of the function f for $t \neq 0$. By the Morse lemma, there exists a local coordinate system $x_1, \ldots, x_n$ in a neighbourhood of the non-degenerate critical point p_i on the manifold M^n, in which the function f can be written in the form $f(x_1, \ldots, x_n) = z_i + \Sigma_{j=1}^n x_j^2$. For values of the parameter t near zero, we fix in the level manifold $F_{u(t)}$ the sphere $S(t) = \sqrt{(u(t) - z_i)}\, S^{n-1}$, where

$$S^{n-1} = \{(x_1, \ldots, x_n) : \Sigma_j x_j^2 = 1, \mathrm{Im}\, x_j = 0\}$$

is the standard unit $(n-1)$-dimensional sphere.

Lifting the homotopy t from zero to one defines a family of $(n-1)$-dimensional spheres $S(t) \subset F_{u(t)}$ in the level manifolds $F_{u(t)}$ for all $t \in (0, 1]$. Note that for $t = 0$ the sphere $S(t)$ reduces to the critical point p_i.

Definition. The homology class $\Delta \in H_{n-1}(F_{z_0})$, represented by the $(n-1)$-dimensional sphere $S(1)$ in the chosen non-singular level manifold F_{z_0} is called a *vanishing* (along the path u) *cycle of Picard-Lefschetz*.

It is easy to see that the homotopy class of the path u in the set of all paths in the region U joining the critical value z_i with the non-critical value z_0 and not passing through critical values of the function f for $t \neq 0$, defines the homology class of the vanishing cycle Δ modulo orientation.

Definition. The set of cycles $\Delta_1, \ldots, \Delta_\mu$ from the $(n-1)$st homology group $H_{n-1}(F_{z_0})$ of the non-singular level set F_{z_0} is called *distinguished* if:

(i) the cycles $\Delta_i (i=1, \ldots, \mu)$ are vanishing along non-self-intersecting paths u_i, joining the critical value z_i with the non-critical value z_0;

(ii) the paths u_i and u_j have, for $i \neq j$, a unique common point $u_i(1) = u_j(1) = z_0$;

(iii) the paths $u_1, \ldots, u_\mu$ are numbered in the same order in which they enter the point z_0, counting clockwise, beginning at the boundary ∂U of the region U (see figure 10).

Remark. The need to choose a non-critical value z_0 on the boundary ∂U of the region U was dictated by the need to number the elements of the distinguished set of vanishing cycles according to condition (iii).

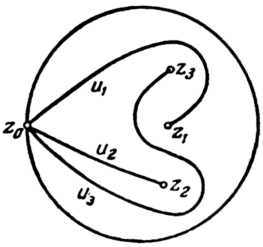

Fig. 10.

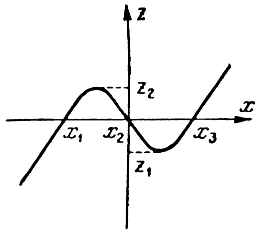

Fig. 11.

Examples. 1. Let us consider the Morse function $f(x) = x^3 - 3\lambda x$, where λ is a small positive number. This function is a perturbation of the function $f_0(x) = x^3$ (having the singularity type A_2 in the sense of volume 1), but we do not need that fact just now. The function f has two critical points ($x = \sqrt{\lambda}$ and $x = -\sqrt{\lambda}$) with critical values $z_1 = -2\lambda\sqrt{\lambda}$ and $z_2 = 2\lambda\sqrt{\lambda}$ respectively. As the non-critical value of the function f we take $z_0 = 0$. Let us join the critical values z_i $(i = 1, 2)$ with the non-critical value z_0 by line segments u_1 and u_2. The level manifold $\{f = 0\}$ consists of three points $x_1 = -\sqrt{3\lambda}$, $x_2 = 0$ and $x_3 = \sqrt{3\lambda}$ (see figure 11). It is easy to see that the cycles, vanishing along the described paths u_1 and u_2 joining the critical values z_1 and z_2 with the non-critical value 0, are the differences

$\Delta_1=\{x_3\}-\{x_2\}$ and $\Delta_2=\{x_2\}-\{x_1\}$ of zeroth homology class represented by the points x_1, x_2 and x_3. Note that the orientation of the cycles was chosen by us arbitrarily: any of them can be multiplied by -1.

For greater clarity we chose the non-critical value $z_0=0$. It presupposes, certainly, a special choice for the region U. On this occasion it is not very important, but later, for the definition of a distinguished basis of vanishing cycles in the homology of a non-singular level manifold of a degenerate singularity, we shall consider the region U to be a disk of sufficiently large radius, in comparison with the critical values of the perturbed function. The need to choose a non-critical value on the boundary of the sufficiently large disk is dictated as, otherwise, firstly, the identification of the homology group of the non-singular level manifold of a singularity and its perturbation would not be unique and, secondly, the order in which the vanishing cycles must enter the distinguished basis would not be unique. In order to "correct" the example we considered above, we can choose a non-critical value z_0^* sufficiently large in absolute value ($|z_0^*|\gg 2\lambda\sqrt{\lambda}$), joining it with the non-critical value $z_0=0$ by a path which does not pass through the critical values of the function f, and observe the change of the non-singular level manifolds $f=z$ as z moves along this path from $z_0=0$ to z_0^*. We shall consider later an analogous construction in a more general case (§2.9). Here for simplicity we modify our example somewhat.

1*. We consider the Morse function $f(x)=x^3+3\lambda x$, where λ is a positive number. The critical points of the function f are $x=-\sqrt{\lambda}i$ and $x=\sqrt{\lambda}i$, the critical values are $z_1=-2\lambda\sqrt{\lambda}i$ and $z_2=2\lambda\sqrt{\lambda}i$. We choose as the region U a disk of sufficiently large radius r with centre at zero ($r\gg 2\lambda\sqrt{\lambda}$). We consider two non-critical values of the function f: $z_0=0$ and $z_0^*=r$. The critical values $z_{1,2}$ are joined to $z_0=0$ by segments, going along the imaginary axis, $z_0=0$ is joined to $z_0^*=r$ by a segment of the positive real half-axis. In this way we get paths, u_1 and u_2, joining the critical values $z_{1,2}$ with the non-critical value z_0^*. As before the zero level manifold of the function f consists of three points

$$x_1=-\sqrt{3\lambda}i,\ x_2=0 \text{ and } x_3=+\sqrt{3\lambda}i.$$

The level manifold $\{f=z_0^*\}$ is near to the level manifold $\{f_0=z_0^*\}$ of the function $f_0(x)=x^3$ (since $|z_0^*|=r\gg 2\lambda\sqrt{\lambda}$). Therefore it consists of three points

$$x_1^*\approx\exp(-2\pi i/3)\sqrt[3]{z_0^*},$$
$$x_2^*\approx\sqrt[3]{z_0^*},$$
$$x_3^*\approx\exp(2\pi i/3)\sqrt[3]{z_0^*}.$$

It is not difficult to see that, along the line segment joining the critical value $z_1 = -2\lambda\sqrt{\lambda}i$ (respectively $z_2 = 2\lambda\sqrt{\lambda}i$) with the non-critical value $z_0 = 0$, the cycle $\{x_2\} - \{x_1\}$ (respectively $\{x_3\} - \{x_2\}$) vanishes. Further, it is clear that as the non-critical value z moves along the segment of the positive half-axis from $z_0 = 0$ to $z_0^* = r$ the points of the manifold $\{f = z\}$ change in such a manner that the point x_2 remains on the real axis, the point x_1 is in the lower and the point x_3 is in the upper half-plane. Therefore as z moves from z_0 to z_0^*, the points x_1, x_2 and x_3 go to the points x_1^*, x_2^* and x_3^* respectively. Consequently, along the paths u_1 and u_2 which we described joining the critical values z_1 and z_2 of the function f with the non-critical value z_0^* the cycles

$$\Delta_1 = \{x_2^*\} - \{x_1^*\}$$

and

$$\Delta_2 = \{x_3^*\} - \{x_2^*\}$$

respectively vanish. It is easy to see that the vanishing cycles Δ_1 and Δ_2 form a distinguished set.

2. As another example we consider the function of two variables $f(x, y) = x^3 - 3\lambda x + y^2$ (λ is a small positive number). This function is a perturbation of the function $f_0(x, y) = x^3 + y^2$, which also has singularity type A_2 in the sense of volume 1. The function has the same critical values, $z_1 = -2\lambda\sqrt{\lambda}$ and $z_2 = 2\lambda\sqrt{\lambda}$, as the function in the first example. These values are taken at the points $(\sqrt{\lambda}, 0)$ and $(-\sqrt{\lambda}, 0)$ respectively. We join the critical values z_1 and z_2 with the non-critical value $z_0 = 0$ by segments u_1 and u_2 of the real axis. The zero level manifold of the function f (the complex curve $\{f = 0\}$) is the graph of the two-valued function $y = \pm\sqrt{(-x^3 + 3\lambda x)}$ and therefore is a double covering of the plane of the complex variable x, branching at the points $x_1 = -\sqrt{3\lambda}$, $x_2 = 0$ and $x_3 = \sqrt{3\lambda}$. It can be obtained from two copies of the plane of the complex variable x with cuts from the point x_1 to the point x_2 and from the point x_3 to infinity (see figure 12), glued together criss-cross along these cuts.

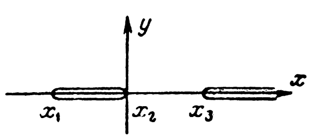

Fig. 12.

As z moves along the real axis from $z_1 = -2\lambda\sqrt{\lambda}$ to $z_2 = 2\lambda\sqrt{\lambda}$ the manifold $\{f = z\}$ is deformed. The movement of the branch points $\tilde{x}_1 = \tilde{x}_1(z)$, $\tilde{x}_2 = \tilde{x}_2(z)$, and $\tilde{x}_3 = \tilde{x}_3(z)$ as a double covering of the plane of the complex variable x is illustrated in figure 13.

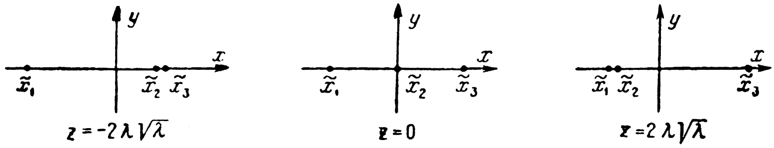

Fig. 13.

From this it is clear that the vanishing cycles corresponding to the critical values $z_1 = -2\lambda\sqrt{\lambda}$ and $z_2 = 2\lambda\sqrt{\lambda}$ and the paths u_1 and u_2 which we described joining them to the non-critical value 0 are the one-dimensional cycles Δ_1 and Δ_2 portrayed in figure 14 (we have indicated by dashes the part of the cycle lying on the second sheet of the surface; the orientation of the vanishing cycles again can be chosen arbitrarily).

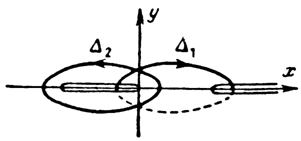

Fig. 14.

Once again let u be a path joining some critical value z_i with a non-critical value z_0.

Definition. A *simple loop* corresponding to the path u is an element of the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values represented by the loop going along the path u from the point z_0 to the point z_i, going round the point z_i in the positive direction (anticlockwise) and returning along the path u to the point z_0.

The region U, with the μ critical values $\{z_i | i = 1, \ldots, \mu\}$ of the function f removed from it is homotopically equivalent to a bouquet of μ circles. Therefore the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values of the function f is a free group on μ generators. If $\{u_i | i = 1, \ldots, \mu\}$ is a system of loops, defining a distinguished set of vanishing cycles $\{\Delta_i\}$, then the group $\pi_1(U \setminus \{z_i\}, z_0)$ is generated by the simple loops $\tau_1, \ldots, \tau_\mu$ corresponding to the paths $u_1, \ldots, u_\mu$.

Definition. The set of vanishing cycles $\Delta_1, \ldots, \Delta_\mu$, defined by the set of paths $\{u_i\}$, is called *weakly distinguished* if the fundamental group $\pi_1(U \setminus \{z_i\}, z_0)$ of the complement of the set of critical values is the free group on the generators $\tau_1, \ldots, \tau_\mu$, corresponding to the paths $u_1, \ldots, u_\mu$.

We note that permutation of the elements preserves weak distinguishment of a set, but does not preserve its distinguishment.

If the set of paths $\{u_i | i=1, \ldots, \mu\}$ defines a weakly distinguished set of vanishing cycles $\{\Delta_i\}$ in the $(n-1)$st homology group of the non-singular level manifold, then the monodromy group of the function f is generated by the monodromy operators $h_{\tau_i *}$ of the simple loops τ_i $(i=1, \ldots, \mu)$, corresponding to the paths u_i. Therefore the monodromy group of the (Morse) function f is always a group generated by μ generators.

Definition. The monodromy operator

$$h_i = h_{\tau_i *}: H_*(F_{z_0}) \to H_*(F_{z_0})$$

of the simple loop τ_i is called the *Picard-Lefschetz operator* corresponding to the path u_i (or the vanishing cycle Δ_i).

Examples. 1. We consider the Morse function $f(x) = x^3 + 3\lambda x$ of example 1* following the definition of distinguished sets of vanishing cycles. Let τ_i be the simple loop (with initial and final points at the point z_0^*) corresponding to the path u_i. As the non-critical value z moves along the loop τ_1, the level manifold $\{f = z\}$ changes in the following manner: The points x_1^* and x_2^* approach each other, make a half-turn about a common centre, changing places, then move apart to the other's former place; the point x_3^* returns to its own place. Therefore the monodromy h_{τ_1} of the loop τ_1 exchanges the points x_1^* and x_2^* and fixes the points x_3^*. From this it follows that

$$h_1 \Delta_1 = h_{\tau_1 *} \Delta_1 = h_{\tau_1 *}(\{x_2^*\} - \{x_1^*\}) = \{x_1^*\} - \{x_2^*\} = -\Delta_1,$$

$$h_1 \Delta_2 = h_{\tau_1 *} \Delta_2 = h_{\tau_1 *}(\{x_3^*\} - \{x_2^*\}) = \{x_3^*\} - \{x_1^*\} = \Delta_2 + \Delta_1.$$

Similarly

$$h_2 \Delta_2 = -\Delta_2,$$

$$h_2 \Delta_1 = \Delta_2 + \Delta_1.$$

The homology group $H_{n-1}(F_z, \partial F_z)$ of the non-singular level manifold modulo its boundary is the dual group to the group $H_{n-1}(F_z)$ (reduced modulo a point for $n=1$). In the given case its rôle is filled by the ordinary zeroth homology group of the level manifold $\{f=z_0^*\}$ (consisting of the three points x_1^*, x_2^* and x_3^*), factored by the subgroup generated by the "maximal cycle"

$$\{x_1^*\}+\{x_2^*\}+\{x_3^*\}.$$

It is generated by two cycles ∇_1 and ∇_2 such that

$$(\nabla_i \circ \Delta_j)=\delta_{ij}.$$

We can take as these cycles

$$\nabla_1=-\{x_1^*\},\ \nabla_2=\{x_3^*\}.$$

From the description of the monodromy transformation h_{τ_1} it follows that

$$\operatorname{var}_{\tau_1}\nabla_1=-\{x_2^*\}+\{x_1^*\}=-\Delta_1,$$
$$\operatorname{var}_{\tau_1}\nabla_2=0.$$

For the loop τ_2 we have

$$\operatorname{var}_{\tau_2}\nabla_1=0,$$
$$\operatorname{var}_{\tau_2}\nabla_2=-\Delta_2.$$

We consider now the loop τ, defined by the formula $\tau(t)=z_0^*\exp(2\pi it)$. The loop τ goes once round the critical values of the function f in the positive direction (anticlockwise) along a circle of large radius. From the fact that for large $|z|$ the level set $\{f=z\}$ is close to the level set $\{x^3=z\}$, it follows that the monodromy transformation h_τ, of the loop τ cyclically permutes the points x_1^*, x_2^* and x_3^*

$$(x_1^*\to x_2^*\to x_3^*\to x_1^*).$$

From this it follows that

$$h_{\tau*}\Delta_1=h_{\tau*}(\{x_2^*\}-\{x_1^*\})=\{x_3^*\}-\{x_2^*\}=\Delta_2,$$
$$h_{\tau*}\Delta_2=h_{\tau*}(\{x_3^*\}-\{x_2^*\})=\{x_1^*\}-\{x_3^*\}=-\Delta_1-\Delta_2.$$