O Teorema de Sard
e suas Aplicações

Publicações Matemáticas

O Teorema de Sard e suas Aplicações

Edson Durão Júdice
PUC Minas

impa

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

Publicações Matemáticas

- Introdução à Topologia Diferencial – Elon Lages Lima
- Criptografia, Números Primos e Algoritmos – Manoel Lemos
- Introdução à Economia Dinâmica e Mercados Incompletos – Aloísio Araújo
- Conjuntos de Cantor, Dinâmica e Aritmética – Carlos Gustavo Moreira
- Geometria Hiperbólica – João Lucas Marques Barbosa
- Introdução à Economia Matemática – Aloísio Araújo
- Superfícies Mínimas – Manfredo Perdigão do Carmo
- The Index Formula for Dirac Operators: an Introduction – Levi Lopes de Lima
- Introduction to Symplectic and Hamiltonian Geometry – Ana Cannas da Silva
- Primos de Mersenne (e outros primos muito grandes) – Carlos Gustavo T. A. Moreira e Nicolau Saldanha
- The Contact Process on Graphs – Márcia Salzano
- Canonical Metrics on Compact almost Complex Manifolds – Santiago R. Simanca
- Introduction to Toric Varieties – Jean-Paul Brasselet
- Birational Geometry of Foliations – Marco Brunella
- Introdução à Teoria das Probabilidades – Pedro J. Fernandez
- Teoria dos Corpos – Otto Endler
- Introdução à Dinâmica de Aplicações do Tipo Twist – Clodoaldo G. Ragazzo, Mário J. Dias Carneiro e Salvador Addas Zanata
- Elementos de Estatística Computacional usando Plataformas de Software Livre/Gratuito – Alejandro C. Frery e Francisco Cribari-Neto
- Uma Introdução a Soluções de Viscosidade para Equações de Hamilton-Jacobi – Helena J. Nussenzveig Lopes, Milton C. Lopes Filho
- Elements of Analytic Hypoellipticity – Nicholas Hanges
- Métodos Clássicos em Teoria do Potencial – Augusto Ponce
- Variedades Diferenciáveis – Elon Lages Lima
- O Método do Referencial Móvel – Manfredo do Carmo
- A Student's Guide to Symplectic Spaces, Grassmannians and Maslov Index – Paolo Piccione e Daniel Victor Tausk
- Métodos Topológicos en el Análisis no Lineal – Pablo Amster
- Tópicos em Combinatória Contemporânea – Carlos Gustavo Moreira e Yoshiharu Kohayakawa
- Uma Iniciação aos Sistemas Dinâmicos Estocásticos – Paulo Ruffino
- Compressive Sensing – Adriana Schulz, Eduardo A.B.. da Silva e Luiz Velho
- O Teorema de Poncelet – Marcos Sebastiani
- Cálculo Tensorial – Elon Lages Lima
- Aspectos Ergódicos da Teoria dos Números – Alexander Arbieto, Carlos Matheus e C. G. Moreira
- A Survey on Hiperbolicity of Projective Hypersurfaces – Simone Diverio e Erwan Rousseau
- Algebraic Stacks and Moduli of Vector Bundles – Frank Neumann
- O Teorema de Sard e suas Aplicações – Edson Durão Júdice
- Tópicos de Mecânica Clássica – Artur Lopes

IMPA - ddic@impa.br - http://www.impa.br - ISBN: 978-85-244-0333-0

Apresentação

O presente trabalho é a dissertação apresentada pelo Professor Edson Durão Júdice, em 1961, para concorrer a uma cátedra de Matemática na Escola de Engenharia da Universidade de Minas Gerais. Trata-se de uma exposição cuidadosa e elegante do Teorema de Sard, acompanhada de resumo histórico, comentários esclarecedores e aplicações interessantes. Seu estilo é pausado e a clareza do tratamento faz jus à reputação do autor por suas qualidades de professor. Ao publicá-lo, agradecemos ao Professor Edson pela permissão dada. Estamos certos de que esta edição, que vem à luz tantos anos após a primeira, vem prestar um valioso serviço à nova geração de matemáticos do país.

Rio de Janeiro, setembro de 2011.

Elon Lages Lima

Prefácio

O extraordinário desenvolvimento da Matemática em nossos dias impõe a quem cultiva essa ciência a necessidade de fazer uma escolha entre as diversas partes em que ela se ramifica, a fim de restringir suas atividades, em certa medida, a um dos muitos e vastos campos abertos à investigação científica. No Brasil, com respeito à Matemática Pura, já se distingue nitidamente a formação de três correntes: a da Álgebra, a da Análise e a da Geometria. Tendo em vista, porém, que as doutrinas matemáticas são estreitamente interligadas, é claro que ninguém pode aspirar ao ingresso em uma dessas correntes, sem antes se submeter a adequada preparação. Nessa fase preliminar, são estudadas as noções fundamentais de Análise Matemática e Geometria Analítica clássicas, de Álgebra Abstrata, de Álgebra Linear, de Topologia dos Espaços Métricos, de Topologia Geral e de Teoria da Medida.

Quanto a nós, depois que nos familiarizamos com os conhecimentos básicos indispensáveis, iniciamos nossa marcha no caminho da Geometria. Essa escolha não só satisfaz a uma tendência natural do nosso espírito, mas concorre para incrementar a eficiência da colaboração que modestamente vimos prestando à Escola em que nos formamos, onde servimos, há alguns anos, em uma cadeira de Geometria.

Atente-se, contudo, em que a Geometria tem, para os matemáticos modernos, um sentido muito amplo, pois abrange, além das teorias geométricas clássicas, estudos sobre Variedades Diferenciáveis, Topologia Diferencial, Geometria Riemanniana, Teoria Geométrica das Equações Diferenciais Ordinárias, etc.

Este trabalho, com o qual nos apresentamos à egrégia Congregação da Escola de Engenharia da Universidade de Minas Gerais como candidato à cátedra de Geometria Analítica e Projetiva, é fruto dos estudos a que nos dedicamos, em 1960, no Instituto de Matemática Pura e Aplicada, onde, na qualidade de professor agregado, tivemos ocasião de frequentar cursos e seminários orientados pelos ilustres professores Elon Lages Lima e Maurício Matos Peixoto.

No período em que realizamos tal estágio, contamos com o apoio da Escola de Engenharia da U.M.G., da Campanha Nacional de Aperfeiçoamento do Pessoal de Nível Superior, do Conselho Nacional de Pesquisas, do Instituto de Matemática Pura e Aplicada e da Força Aérea Norte-Americana, que, mediante o pagamento de vencimentos, bolsas ou auxílios, nos permitiram viver no Rio de Janeiro, em regime de total dedicação ao estudo. Deixamos aqui consignado o nosso sentimento de gratidão para com essas organizações.

Agradecemos ao Prof. Maurício Matos Peixoto, que nos sugeriu o assunto para este escrito, e nos auxiliou com observações úteis e palavras de incentivo.

De modo especial, agradecemos ao caro amigo Prof. Elon Lages Lima, que foi o nosso orientador, a quem várias vezes recorremos para tirar dúvidas, pedir opiniões, discutir situações ou solicitar indicações bibliográficas.

Queremos, também, deixar gravada a nossa palavra de homenagem a duas figuras de relevo da ciência brasileira, que, por ensinamentos ministrados, e principalmente pelo exemplo vivo que nos deram, contribuiram eficazmente para a nosssa formação cultural: ao Prof. Christovam Colombo dos Santos, que tanto soube dignificar a cátedra a que agora nos candidatamos, fazendo dela um foco irradiante de sabedoria, e que nos concedeu a mais subida honra, tornando-nos depositário de sua irrestrita confiança durante todo o tempo em que fomos seu assistente; ao Prof. Leopoldo Nachbin, justo orgulho da nova geração de matemáticos brasileiros, que, iluminado pela chama de um grande idealismo, tem trabalhado infatigavelmente pelo desenvolvimento das ciências matemáticas em nossa Pátria.

Finalmente, nossos agradecimentos ao pessoal do Departamento de Publicações do Diretório dos Estudantes de Engenharia, pelo carinho e zelo com que se encarregou da impressão deste trabalho, e ao Sr. Henrique V. Corrêa, que com admirável competência o datilografou.

E. D. Judice

Belo Horizonte
Setembro, 1961

Sumário

Capítulo I

Noções sobre Variedades Diferenciáveis

1 Introdução

Pode afirmar-se, sem receio de incorrer em exagero, que a concepção cartesiana da geometria foi das idéias que mais marcada influência tiveram no progresso das ciências matemáticas. Na verdade, a semente plantada por Descartes desenvolveu-se em uma grande árvore, que até hoje vem lançando novos ramos e produzindo belos frutos. Parece-nos que o mais novo rebento dessa árvore é a teoria das variedades diferenciáveis, na qual estão trabalhando notáveis matemáticos da atualidade. Descrever minuciosamente a evolução das idéias matemáticas, desde a obra de Descartes até as suas consequências mais recentes, é uma sugestão tentadora, que podia suscitar interessante estudo histórico-crítico, mas não é isso, certamente, o que pretendemos fazer aqui. Vamos apenas apresentar, a título de introdução ao nosso trabalho, e da maneira mais resumida, um esboço desse desenvolvimento de idéias.

A Geometria Analítica Clássica, fundada por Descartes, e o Cálculo Infinitesimal, instituído por Newton e Leibniz, desenvolveram-se de modo notável, auxiliando-se e completando-se mutuamente. A Geometria foi (e continua a ser) grande inspiradora de problemas e teorias que iam sendo examinadas, criticadas, enfim, filtradas no crivo da Análise, sob cuja linguagem austera, mas exata, revestiam-se do rigor lógico que caracteriza as doutrinas matemáticas. Surgiu a Geometria Diferencial, na qual aparece como figura de primeira plana o vulto genial de Gauss. Nessa teoria, as curvas e superfícies do espaço euclidiano ordinário foram estudadas, originalmente, do ponto de vista local, e, mais tarde, também em seus aspectos globais. Em incontida ânsia de generalizar cada vez mais os resultados obtidos, lançaram-se os matemáticos ao estudo dos espaços euclidianos n-dimensionais, e conseguiram estender a estes muitas noções já antes estudadas na geometria do espaço R^3.

A teoria dos conjuntos, na qual colaborou de modo decisivo o espírito profundamente penetrante de Cantor, veio abrir novos horizontes em matemática. Os conjuntos abstratos passaram a constituir o mais importante objeto de estudo. Construíram-se variadas estruturas matemáticas, por meio de diversas maneiras de introduzir relações entre os elementos de um conjunto. Apareceram as estruturas algébricas, fundadas no conceito de operação (ou lei de composição), e nasceram as estruturas topológicas, baseadas na idéia de proximidade entre os pontos de um conjunto. Essa proximidade foi definida, no princípio, em termos da idéia de distância, e assim surgiu a teoria dos espaços métricos, estudada minuciosamente por Fréchet em sua tese apresentada ao mundo científico em 1906. Descobriu-se, porém, que o conceito de distância não é indispensável para a definição da vizinhança entre elementos de um conjunto, e tiveram então origem espaços topológicos mais gerais, que não são necessariamente métricos. Dentre as muitas maneiras de definir espaço topológico, verificou-se que a mais vantajosa, por motivo de simplicidade, é a que faz apelo à noção de conjunto aberto. Não precisamos repetir aqui essa definição, naturalmente bem conhecida do leitor.

Alguns espaços topológicos, por serem excessivamente gerais, não apresentam grande interesse, salvo para certos efeitos de crítica. Os espaços que desde cedo se revelaram da maior importância para o desenvolvimento de muitas teorias, são os de Hausdorff. Um espaço de Hausdorff é caracterizado pelo fato de que nele, dois pontos distintos quaisquer admitem vizinhanças disjuntas. Em virtude de ulteriores vantagens e simplificações, tornaram-se particularmente importantes os espaços de Hausdorff que têm base enumerável de conjuntos abertos; tais espaços possuem a conhecida propriedade de Lindelöf.

Os espaços de Hausdorff com base enumerável de abertos, que, do ponto de vista topológico, se comportam localmente como espaços euclidianos de dimensão n, constituem as variedades topológicas n-dimensionais. Nestas, é possível introduzir, localmente, da maneira que adiante descreveremos, sistemas de coordenadas, graças aos quais se atinge a possibilidade de realizar um estudo analítico das propriedades desses espaços. Mediante a introdução de convenientes hipóteses de diferenciabilidade, chega-se, finalmente, ao conceito de variedade diferenciável, entidade abstrata cujas propriedades podem ser estabelecidas pela aplicação dos métodos da geometria analítica e do cálculo infinitesimal.

O nosso trabalho tratará especialmente do teorema de Sard, que é um dos resultados relevantes da teoria das variedades diferenciáveis. Deverá o leitor possuir alguns conhecimentos dessa teoria, que repousa sobre noções de Geometria Analítica, Cálculo Infinitesimal, Topologia Geral e Álgebra Linear. Recomendamos, especialmente, a leitura do magnífico trabalho do Prof. Elon Lages Lima, intitulado "Introdução às Variedades Diferenciáveis", recentemente dado à estampa pelo Instituto de Matemática da Universidade do Rio Grande do Sul.

Para o esclarecimento de dúvidas relativas à Topologia, poderá o leitor consultar Bourbaki (Topologie Générale, livre III, chapitre I), ou MacLane (Curso de Topologia Geral, tradução de Joviano C. Valadares), ou ainda, Elon Lages Lima (Topologia dos Espaços Métricos).

Quanto à Álgebra Linear, as melhores fontes para consulta são: Halmos (Finite Dimensional Vector Spaces), Birkhoff & MacLane (A Survey of Modern Algebra) e Bourbaki (Algèbre, livre II, chapitre II). Em nível mais elementar que as três obras mencionadas, existe um trabalho de nossa autoria (Introdução à Álgebra Linear) que também pode contribuir para elucidar algumas questões.

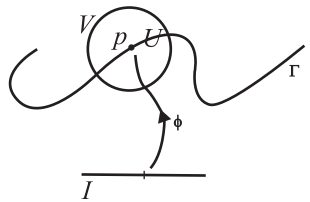
Sobre Geometria Analítica e Cálculo Infinitesimal, grande é o número de livros e apostilas disponíveis e bem conhecidos, e não necessitamos dar indicações bibliográficas especiais a esse respeito.

É claro que um trabalho da natureza deste, contido em tão poucas páginas, não pode ser auto-suficiente. Encontrar-se-ão, a cada passo, referências à bibliografia indicada no fim do volume. Animados pela preocupação de tornar suave a leitura, na medida do possível, resolvemos antepor aos dois últimos capítulos, que encerram a parte essencial do trabalho, algumas noções básicas, indispensáveis à boa assimilação do assunto sobre o qual discorreremos. Essas questões preliminares formam o presente capítulo; através dele, iremos também fixando algumas notações e apresentando a terminologia de que nos serviremos.

2 Curvas e superfícies regulares no espaço euclidiano R^3

Sem entrar em considerações sobre a evolução da idéia de curva, recordemos a maneira como se pode definir essa importante entidade no espaço euclidiano R^3, que é o ambiente onde se realiza a maior parte dos estudos da geometria clássica.

Uma *curva* (simples) é um subespaço de R^3 localmente homeomorfo a um intervalo real aberto. Em termos mais precisos: uma curva é um subconjunto $\Gamma \subset R^3$, tal que para cada ponto $p \in \Gamma$ existe uma vizinhança V de p em R^3, e um homeomorfismo $\phi\colon I \to \Gamma \cap V$, onde $I \subset R$ é um intervalo aberto. Observe-se que $\Gamma \cap V = U$ é uma vizinhança de p no subespaço $\Gamma \subset R^3$.



O homeomorfismo $\phi\colon I \to U$ é uma *parametrização local* da curva Γ, se $q = (x, y, z) \in U$, existe $t \in I$ tal que $\phi(t) = q$. As coordenadas de q são, pois, funções do *parâmetro* t. As equações

$$x = x(t), \qquad y = y(t), \qquad z = z(t),$$

que exprimem essa dependência, dizem-se *equações paramétricas* de Γ (válidas na vizinhança U). Pode adotar-se a representação vetorial $r = r(t)$, onde $r(t)$ é o vetor $(x(t), y(t), z(t))$.

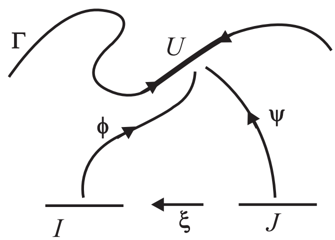
A parametrização local $\phi\colon I \to U$ diz-se *diferenciável de classe* C^k $(1 \leq k \leq \infty)$ se as funções x, y, z, acima consideradas, admitem derivadas sucessivas contínuas de ordens $1, \ldots, k$ (o caso $k = \infty$ corresponde à situação em que existem as derivadas de todas as ordens).

Uma parametrização diferenciável $\phi\colon I \to U$, válida numa vizinhança U do ponto $p \in \Gamma$, é *regular* se para todo $t \in I$ as derivadas $x'(t)$, $y'(t)$, $z'(t)$ não se anulam simultaneamente. Em outros termos: ϕ é regular se o vetor $\frac{dr}{dt} = (x'(t), y'(t), z'(t))$ é $\neq 0$ para todo $t \in I$.

Dizemos que Γ é uma *curva regular de classe* C^k $(1 \leq k \leq \infty)$ se para cada $p \in \Gamma$ existe uma parametrização regular de classe C^k, válida numa vizinhança de p. Referir-nos-emos mais brevemente a tais parametrizações chamando-lhes *admissíveis* em relação a Γ.

Não vemos necessidade de recapitular a noção de tangente a uma curva, mas é interessante acentuarmos que as curvas regulares são justamente as que possuem tangente bem definida em cada ponto.

Consideremos uma curva regular Γ, de classe C^k, e seja $\phi\colon I \to U$ uma parametrização local admissível, válida num aberto $U \subset \Gamma$. Seja $J \subset R$ um intervalo aberto, e $\xi\colon J \to I$ um homeomorfismo de J sobre I, diferenciável de classe C^k, tal que se tenha $\xi'(s) \neq 0$ para todo $s \in J$.

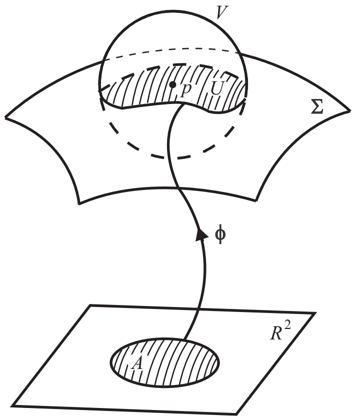


Nessas condições, o homeomorfismo inverso $\xi^{-1}\colon I \to J$ é também diferenciável de classe C^k, e dizemos que ξ é um *difeomorfismo* de classe C^k. O homeomorfismo composto $\phi \circ \xi = \psi\colon J \to U$ é, evidentemente, uma parametrização local admissível em relação a Γ. É importante observarmos que na situação que acabamos de descrever, o difeomorfismo ξ é justamente o homeomorfismo composto $\phi^{-1} \circ \psi$. Demonstra-se que essa é a situação geral, isto é, com palavras precisas: se $\phi\colon I \to U$ e $\psi\colon J \to V$ são duas parametrizações locais admissíveis em relação à curva regular Γ, tais que $U \cap V \neq \varnothing$, então $\xi = \phi^{-1} \circ \psi$ é um difeomorfismo de $\psi^{-1}(U \cap V)$ sobre $\phi^{-1}(U \cap V)$.

As superfícies do espaço euclidiano R^3 podem ser definidas de modo análogo ao que adotamos para as curvas.

Uma *superfície* (simples) é um subespaço de R^3 localmente homeomorfo a um disco aberto do plano. (Em vez de um disco, pode usar-se um aberto qualquer de R^2, que quase sempre se escolhe conexo.) Com maior precisão, definiremos: uma superfície é um subconjunto $\Sigma \subset R^3$, tal que para cada ponto $p \in \Sigma$ existe uma vizinhança V de p em R^3 e um homeomorfismo $\phi\colon A \to \Sigma \cap V$, onde $A \subset R^2$ é um conjunto aberto. O conjunto $\Sigma \cap V = U$ é uma vizinhança de p no subespaço $\Sigma \subset R^3$.

Da mesma maneira que para as curvas, o homeomorfismo $\phi\colon A \to U$ diz-se uma *parametrização local* da superfície Σ. Se $q = (x, y, z) \in U$, existe $(u, v) \in A$ tal que $\phi(u, v) = q$.



As coordenadas de q são, portanto, funções dos parâmetros u, v. As equações

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v),$$

por meio das quais se expressa essa dependência, são *equações paramétricas* de Σ (válidas na vizinhança U). É muito empregada a representação vetorial $r = f(u, v)$, onde $r(u, v)$ é o vetor $(x(u, v), y(u, v), z(u, v))$.

A parametrização local $\phi\colon A \to U$ é dita *diferenciável de classe* C^k $(1 \leq k \leq \infty)$ se as funções x, y, z admitem todas as derivadas parciais de ordens $1, \ldots, k$ contínuas (no caso $k = \infty$, existem as derivadas de todas as ordens das ditas funções).

Uma parametrização diferenciável $\phi\colon A \to U$, válida num aberto $U \subset \Sigma$, é *regular* se a matriz

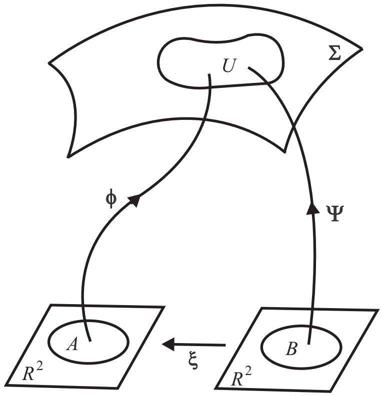
$$\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{bmatrix}$$

tem característica 2 em todo ponto $(u, v) \in A$. Em outras palavras: ϕ é regular se os vetores $\frac{\partial r}{\partial u}$ e $\frac{\partial r}{\partial v}$ são linearmente independentes, qualquer que seja $(u, v) \in A$.

A superfície Σ diz-se *regular de classe* C^k, onde $1 \leq k \leq \infty$, se para todo $p \in \Sigma$ existe uma parametrização regular de classe C^k, válida numa vizinhança de p. Diremos que tais parametrizações são *admissíveis* em relação a Σ.

As superfícies regulares são precisamente as que admitem um plano tangente bem definido em cada ponto.

Seja $\phi\colon A \to U$ uma parametrização local admissível em uma superfície regular Σ, de classe C^k. Seja $B \subset R^2$ um conjunto aberto e $\xi\colon B \to A$ um homeomorfismo de B sobre A, diferenciável de classe C^k, com determinante jacobiano $\neq 0$ em todo ponto de B. Resulta que o homeomorfismo inverso $\xi^{-1}\colon A \to B$ também é diferenciável de classe C^k. Para sintetizar tudo isso em uma palavra, diremos que ϕ é um *difeomorfismo* (de classe C^k). O homeomorfismo



$\phi \circ \xi = \psi\colon B \to U$ é uma parametrização admissível em relação a Σ, e é claro que $\xi = \phi^{-1} \circ \psi$. Se $\phi\colon A \to U$ e $\psi\colon B \to V$ são duas parametrizações locais admissíveis relativamente à superfície regular Σ, tais que $U \cap V \neq \varnothing$, demonstra-se que $\phi^{-1} \circ \psi = \xi$ é um difeomorfismo de $\psi^{-1}(U \cap V)$ sobre $\phi^{-1}(U \cap V)$.

O estudo introdutório que acabamos de apresentar para as curvas e superfícies regulares do espaço euclidiano R^3 pode ser estendido imediatamente ao espaço n-dimensional R^n, no qual podemos definir, de maneira óbvia, *superfícies regulares de dimensão* r $(r < n)$. A $r = n-1$ correspondem as chamadas *hipersuperfícies* de R^n. O leitor interessado pode encontrar o estudo dessas superfícies em [6], Cap. II.

A noção de variedade diferenciável deriva, por via direta, da idéia de superfície regular, a que nos referimos. Trata-se apenas de recorrer a um processo de abstração, e definir uma entidade matemática que não se supõe imersa em um espaço euclidiano, mas que possui algumas propriedades sugeridas pelas superfícies regulares de R^n. Estas passarão a ser particulares variedades diferenciáveis.

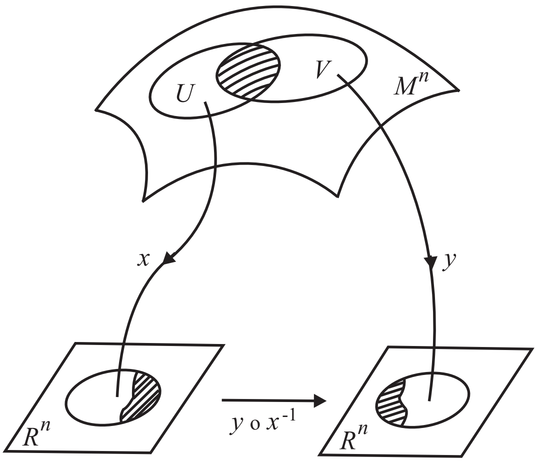
3 Conceito de variedade diferenciável

Conforme dissemos na introdução, uma *variedade topológica* de dimensão n é um espaço de Hausdorff M, com base enumerável de conjuntos abertos, localmente homeomorfo ao espaço euclidiano R^n. Essa última propriedade significa que cada ponto $p \in M$ tem uma vizinhança (aberta) U homeomorfa ao espaço R^n (ou, de modo mais geral, a um conjunto aberto $A \subset R^n$). Indicaremos a variedade pela mesma letra M que representa o espaço topológico subjacente, e, quando houver conveniência em dar ênfase à dimensão n, usaremos o símbolo M^n. O homeomorfismo local $x\colon U \to R^n$ permite-nos a introdução de coordenadas na variedade M. Para cada $q \in U$, tem-se $x(q) = (x^1, \dots, x^n) \in R^n$; os números reais $x^1, \dots, x^n$ são, por definição, as coordenadas do ponto p no *sistema local* $x\colon U \to R^n$.

Para poder atribuir coordenadas a todo ponto da variedade M^n, devemos imaginar uma cobertura de M formada de conjuntos abertos U_α, cada um dos quais seja o domínio de um sistema local $x_\alpha\colon U_\alpha \to R^n$. Uma tal coleção de sistemas de coordenadas locais em M é o que chamaremos um *atlas de dimensão* n sobre a variedade M.

Em uma variedade M^n, consideremos dois sistemas locais $x\colon U \to R^n$ e $y\colon V \to R^n$, tais que $U \cap V \neq \varnothing$. Se $p \in U \cap V$, deve ser $x(p) = (x^1, \dots, x^n)$ e $y(p) = (y^1, \dots, y^n)$. A aplicação $(x^1, \dots, x^n) \to (y^1, \dots, y^n)$ é precisamente o homeomorfismo composto $y \circ x^{-1}\colon x(U \cap V) \to y(U \cap V)$, que se diz uma *transformação de coordenadas locais* na variedade M. O homeomorfismo inverso de $y \circ x^{-1}$ é evidentemente $x \circ y^{-1}$, e define a *transformação inversa* da primeira.

Para passar do conceito de variedade topológica à idéia mais particular de variedade diferenciável, é mister supor que essas transformações satisfaçam a certas hipóteses de diferenciabilidade, que passamos a descrever de modo preciso. A transformação $y \circ x^{-1}$, acima considerada, associa



ao ponto $(x^1, \ldots, x^n) \in x(U \cap V)$ o ponto $(y^1, \ldots, y^n) \in y(U \cap V)$, e pode, pois, representar-se por um sistema de equações

$$y^i = y^i\,(x^1, \ldots, x^n),$$
$$i = 1, \ldots, n.$$

Diz-se que essa transformação é *diferenciável de classe* C^k $(1 \leq k \leq \infty)$ se as funções reais $y^1, \ldots, y^n$ admitem derivadas parciais contínuas de ordens $1, \ldots, k$ (ou de todas as ordens, no caso $k = \infty$). Quando a transformação inversa de $y \circ x^{-1}$ (isto é, $x \circ y^{-1}$) é também diferenciável de classe C^k, dizemos que $y \circ x^{-1}$ é um *difeomorfismo de classe* C^k, e, neste caso, o jabobiano $\det\left(\frac{\partial y^i}{\partial x^j}\right)$ é $\neq 0$ em todo ponto onde é definido.

Um atlas $\mathcal{A}$ sobre uma variedade topológica M diz-se *diferenciável de classe* C^k se as transformações de coordenadas locais $y \circ x^{-1}$, correspondentes a todo par de sistemas $x, y \in \mathcal{A}$, são diferenciáveis de classe C^k (e então é claro que tais transformações são difeomorfismos de classe C^k).

Seja $\mathcal{A}$ um atlas diferenciável de classe C^k sobre a variedade topológica M^n. Um sistema de coordenadas locais $z\colon V \to R^n$ diz-se *admissível* em relação a esse atlas, se a transformação de coordenadas $z \circ x^{-1}$ é um difeomorfismo de classe C^k, qualquer que seja o sistema local $x\colon U \to R^n$, pertencente a $\mathcal{A}$, tal que $V \cap U \neq \varnothing$. É claro que são admissíveis relativamente a $\mathcal{A}$ todos os sistemas locais $x \in \mathcal{A}$. Um atlas $\mathcal{A}$ é *máximo* quando a ele pertence todo sistema local admissível em relação a $\mathcal{A}$. Todo atlas diferenciável sobre uma variedade M está evidentemente contido em um único atlas máximo de mesma classe.

Podemos agora definir: uma *variedade diferenciável* de dimensão n e de classe C^k $(1 \leq k \leq \infty)$ é uma variedade topológica M^n dotada de um atlas diferenciável máximo de classe C^k.

Citaremos, agora, alguns exemplos de variedades diferenciáveis, mas não nos deteremos em discutí-los, para não alongar este capítulo introdutório. O leitor interessado poderá encontrar em [6] o estudo minucioso desses exemplos, e de outros muito interessantes. São variedades diferenciáveis:

1) O espaço euclidiano R^n, munido do atlas máximo de classe C^∞ ao qual pertence o sistema $x\colon R^n \to R^n$, tal que $x(p) = p$ para todo $p \in R^n$.

2) Todo subconjunto aberto de uma variedade diferenciável.

3) As curvas e superfícies regulares de R^3, e, de modo mais geral, as superfícies regulares de dimensão r do espaço euclidiano R^n $(r < n)$.

4) A esfera unitária $S^n \subset R^{n+1}$. Essa esfera é o conjunto $S^n = \{p \in R^{n+1}\,;\, |p| = 1\}$; trata-se de uma hipersuperfície de R^{n+1}. Este exemplo é um caso particular importante do exemplo precedente.

4 Espaço vetorial tangente a uma variedade diferenciável

A idéia de variedade diferenciável é uma generalização abstrata da idéia de superfície regular de um espaço euclidiano. É possível definir objetos matemáticos que exercem, em relação a uma variedade M, um papel análogo ao que desempenham na geometria diferencial os vetores tangentes a uma superfície r-dimensional do espaço R^n $(r < n)$. A tais objetos chamaremos vetores tangentes à variedade M, mas, ao contrário do que ocorre no caso das superfícies, eles não são vetores usuais de um espaço euclidiano, pois a variedade M não está necessariamente contida em um R^n. Existem várias maneiras de definir vetor tangente a uma variedade diferenciável. Descreveremos apenas uma delas, a qual, aliás, não é a mais intuitiva, mas a que nos parece mais simples do ponto de vista formal. A idéia básica é a de que as entidades abstratas a que chamaremos vetores tangentes a uma variedade diferenciável num ponto dela, constituam um espaço vetorial, e possam ser descritas, de modo preciso, em cada sistema de coordenadas locais, válido numa vizinhança do dito ponto.

Seja M^n a variedade considerada. Tomemos um ponto $p \in M$, e seja $\mathcal{M}_p$ a coleção dos sistemas locais $x\colon U \to R^n$, tais que $p \in U$. Um *vetor tangente* à variedade M no ponto p é uma função $v\colon \mathcal{M}_p \to R^n$ que cumpre a seguinte condição: se $x, y \in \mathcal{M}_p$, se $v(x) = (\alpha^1, \ldots, \alpha^n)$ e se $v(y) = (\beta^1, \ldots, \beta^n)$, então

$$\beta^i = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j}(p) \cdot \alpha^j, \qquad i = 1, \ldots, n,$$

onde $\left(\frac{\partial y^i}{\partial x^j}(p)\right)$ é a matriz jacobiana da transformação de coordenadas $y \circ x^{-1}$, calculada no ponto $x(p) \in R^n$. Os números reais $\alpha^1, \ldots, \alpha^n$ são as *componentes* do vetor tangente v no sistema x. A fórmula acima exprime a lei de transformação das componentes de v, quando se passa do sistema x ao sistema y, e mostra que o vetor v fica completamente determinado quando se conhecem as suas componentes em um sistema (*).

Representemos por M_p o conjunto de todos os vetores tangentes à variedade M no ponto p. Podemos definir a soma de dois vetores $u, v \in M_p$ como sendo o vetor $u + v \in M_p$ tal que $(u + v)(x) = u(x) + v(x)$, qualquer que seja o sistema $x \in \mathcal{M}_p$. Se $\lambda \in R$, podemos também definir

---

(*) É a chamada lei de transformação por contravariância.

o produto de λ pelo vetor $u \in M_p$ como sendo o vetor $\lambda u \in M_p$ tal que $(\lambda u)(x) = \lambda u(x)$, para todo $x \in \mathcal{M}_p$. Relativamente a essas duas leis de composição, o conjunto M_p apresenta uma estrutura algébrica de espaço vetorial sobre R; trata-se do *espaço vetorial tangente à variedade M no ponto p*, o qual representaremos ainda pelo símbolo M_p. Observe-se que o vetor zero desse espaço tangente é o vetor $0 \in M_p$ tal que $0(x) = (0, \ldots, 0)$ para todo sistema $x \in \mathcal{M}_p$.

Consideremos um sistema de coordenadas locais $x\colon U \to R^n$, tal que $p \in U$. Representemos por $X_1, X_2, \ldots, X_n$ os vetores tangentes a M em p, definidos assim:

$$\begin{aligned} &X_1(x) = (1, 0, \ldots, 0),\\ &X_2(x) = (0, 1, \ldots, 0),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &X_n(x) = (0, 0, \ldots, 1). \end{aligned}$$

Se $y\colon V \to R^n$ é outro sistema local tal que $p \in V$, verifica-se trivialmente que

$$X_i(y) = \left(\frac{\partial y^1}{\partial x^i}(p), \ldots, \frac{\partial y^n}{\partial x^i}(p)\right), \qquad i = 1, \ldots, n.$$

Os vetores $X_1, \ldots, X_n$ são linearmente independentes; com efeito, se $\lambda^1, \ldots, \lambda^n \in R$ e $\lambda^1 X_1 + \cdots + \lambda^n X_n = 0$, devemos ter:

$$(\lambda^1 X_1 + \cdots + \lambda^n X_n)(z) = (0, 0, \ldots, 0)$$

para todo sistema local $z \in \mathcal{M}_p$; em particular, para o sistema x resulta

$$(\lambda^1 X_1 + \cdots + \lambda^n X_n)(x) = (\lambda^1, \lambda^2, \ldots, \lambda^n) = (0, 0, \ldots, 0);$$

logo $\lambda^1 = \lambda^2 = \cdots \lambda^n = 0$, como queríamos provar.

Os vetores tangentes $X_1, \ldots, X_n$ geram o espaço vetorial M_p. De fato, se $v \in M_p$ e $v(x) = (\xi^1, \ldots, \xi^n) \in R^n$, podemos escrever:

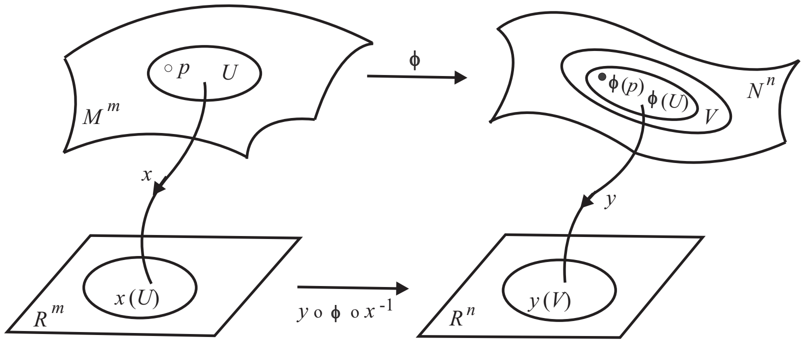
$$v(y) = \left(\sum_j \frac{\partial y^1}{\partial x^j}(p)\, \xi^j, \ldots, \sum_j \frac{\partial y^n}{\partial x^j}(p)\, \xi^j \right),$$
$$v(y) = \sum_j \left(\frac{\partial y^1}{\partial x^j}(p)\, \xi^j, \ldots, \frac{\partial y^n}{\partial x^j}(p)\, \xi^j \right),$$
$$v(y) = \sum_j \xi^j \left(\frac{\partial y^1}{\partial x^j}(p), \ldots, \frac{\partial y^n}{\partial x^j}(p) \right) = \sum_j \xi^j X_j(y),$$
$$v(y) = \left(\sum_j \xi^j X_j \right)(y).$$

Como $y \in \mathcal{M}_p$ é genérico, segue-se que $v = \sum_j \xi^j X_j$.

Podemos concluir que o conjunto $\{X_1, \ldots, X_n\}$ é uma base do espaço vetorial M_p, e que este é n-dimensional. Diremos que $\{X_1, \ldots, X_n\}$ é a *base de M_p associada ao sistema x.*

5 Aplicações diferenciáveis

Consideremos duas variedades M^m, N^n, diferenciáveis de classe C^k.



Uma aplicação ϕ: $M \to N$ diz-se *diferenciável de classe* C^h $(h \leq k)$ se satisfaz às seguintes condições: 1) para todo $p \in M$ e para todo sistema local $y\colon V \to R^n$ tal que $\phi(p) \in V$, existe um sistema local $x\colon U \to R^m$ tal que $p \in U$ e $\phi(U) \subset V$; 2) a aplicação composta $y \circ \phi \circ x^{-1}\colon x(U) \to y(V)$ é diferenciável de classe C^h.

Da condição 1) decorre a continuidade de ϕ. A condição 2) tem o seguinte significado: se $q \in U$, se $x(q) = (x^1, \ldots, x^m)$ e se $y(\phi(q)) = (y^1, \ldots, y^n)$, então a aplicação $y \circ \phi \circ x^{-1}$ é representada por

$$y^i = y^i(x^1, \ldots, x^m), \qquad i = 1, \ldots, n,$$

onde $y^1, \ldots, y^n$ são funções que admitem derivadas parciais contínuas de ordens $1, \ldots, h$ em todo ponto do seu domínio $x(U)$. Da hipótese $h \leq k$ resulta que se a condição 2) é satisfeita para um par x, y de sistemas locais, ela é também verificada para qualquer outro par (supõe-se, evidentemente, que esses pares de sistemas satisfazem à condição 1)). É claro que só se pode ter $h = \infty$ se é também $k = \infty$.

Incluem-se como casos particulares das aplicações diferenciáveis as duas importantes categorias seguintes:

1) *As funções reais diferenciáveis* definidas sobre a variedade M^n: são as funções $f\colon M \to R$ tais que para todo sistema de coordenadas locais $x\colon U \to R^n$, a função real $f \circ x^{-1}\colon x(U) \to R$ é diferenciável.

2) As *curvas diferenciáveis*: são as aplicações diferenciáveis $C\colon I \to M^n$, onde I é um intervalo real aberto. Seja $x\colon U \to R^n$ um sistema local em M, para o qual se tenha $U \cap C(I) \neq \varnothing$; seja $J \subset I$ um intervalo aberto tal que $C(J) \subset U$. A hipótese de diferenciabilidade de C significa que a aplicação $x \circ C\colon J \to x(U)$ se representa por n funções diferenciáveis $x^i = x^i(t)$, $i = 1, \ldots, n$.

6 Diferencial de uma aplicação diferenciável

Consideremos duas variedades M^m e N^n, e uma aplicação diferenciável $\phi\colon M \to N$. Vamos mostrar que ϕ induz, em cada ponto $p \in M$, uma aplicação linear $\phi_p\colon M_p \to N_{\phi(p)}$ (cada vetor tangente a M em p é aplicado num vetor tangente a N em $\phi(p)$). Ponhamos $\phi(p) = q$, e consideremos sistemas de coordenadas locais $x\colon U \to R^m$, em M, e $y\colon V \to R^n$, em N, tais que $p \in U$ e $q \in V$. Podemos admitir que $\phi(U) \subset V$. A aplicação ϕ representa-se, em termos dos ditos sistemas locais, por equações:

$$y^i = y^i(x^1, \ldots, x^m), \qquad i = 1, \ldots, n.$$

Seja $u \in M_p$ um vetor cujas componentes na base associada ao sistema x são $(\alpha^1, \ldots, \alpha^m)$. Por definição, a aplicação linear ϕ_p associa ao vetor u um vetor $v = \phi_p(u) \in N_q$, cujas componentes $(\beta^1, \ldots, \beta^n)$ na base associada ao sistema y são

$$\beta^i = \sum_{j=1}^{m} \alpha^j \frac{\partial y^i}{\partial x^j}(p), \qquad i = 1, \ldots, n.$$

Como se vê, a matriz da aplicação linear ϕ_p, relativamente às bases associadas aos sistemas de coordenadas x e y, é a matriz jacobiana $\left(\frac{\partial y^i}{\partial x^j}(p)\right)$ da aplicação diferenciável $y \circ \phi \circ x^{-1} \colon x(U) \to y(V)$, calculada no ponto $x(p)$. Diremos, para abreviar, que essa é a matriz jacobiana da aplicação ϕ, relativamente aos sistemas x, y.

Como a definição de ϕ_p, acima dada, foi estabelecida mediante o emprego de sistemas de coordenadas locais nas variedades M e N, precisamos provar que essa definição é legítima, isto é, que ϕ_p não depende da escolha dos sistemas locais x, y. Tomemos novos sistemas de coordenadas $\overline{x} \colon \overline{U} \to R^m$, em M, e $\overline{y} \colon \overline{V} \to R^n$, em N, tais que $p \in \overline{U}$, $q \in \overline{V}$ e $\phi(\overline{U}) \subset \overline{V}$. A aplicação ϕ representar-se-á por novas equações

$$\overline{y}^r = \overline{y}^r(\overline{x}^1, \ldots, \overline{x}^m), \qquad r = 1, \ldots, n.$$

Sejam $(\overline{\alpha}^1, \ldots, \overline{\alpha}^m)$ as componentes do vetor $u \in M_p$ na base associada ao sistema $\overline{x}$, e $(\overline{\beta}^1, \ldots, \overline{\beta}^n)$ as componentes de $v = \phi_p(u) \in N_q$ na base associada ao sistema $\overline{y}$. Precisamos provar que

$$\overline{\beta}^r = \sum_{s=1}^{m} \overline{\alpha}^s \frac{\partial \overline{y}^r}{\partial \overline{x}^s}(p), \quad r = 1, \ldots, n.$$

Com efeito, levando em conta a maneira, já conhecida, como se transformam as componentes de um vetor tangente, quando se muda de sistema de coordenadas, podemos escrever:

$$\overline{\beta}^r = \sum_{i=1}^{n} \beta^i \frac{\partial \overline{y}^r}{\partial y^i} = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \alpha^j \frac{\partial y^i}{\partial x^j} \right) \frac{\partial \overline{y}^r}{\partial y^i} =$$

$$= \sum_{j=1}^{m} \alpha^j \left(\sum_{i=1}^{n} \frac{\partial \overline{y}^r}{\partial y^i} \frac{\partial y^i}{\partial x^j} \right) = \sum_{j=1}^{m} \alpha^j \frac{\partial \overline{y}^r}{\partial x^j} =$$

$$= \sum_{j=1}^{m} \left(\sum_{s=1}^{m} \overline{\alpha}^s \frac{\partial x^j}{\partial \overline{x}^s} \right) \frac{\partial \overline{y}^r}{\partial x^j} = \sum_{s=1}^{m} \overline{\alpha^s} \left(\sum_{j=1}^{m} \frac{\partial \overline{y}^r}{\partial x^j} \frac{\partial x^j}{\partial \overline{x}^s} \right) =$$

$$= \sum_{s=1}^{m} \overline{\alpha}^s \frac{\partial \overline{y}^r}{\partial \overline{x}^s} \,.$$

A definição acima dada mostra claramente que a aplicação $\phi_p \colon M_p \to N_q$ é linear. Essa aplicação diz-se *diferencial* de $\phi \colon M \to N$, no ponto $p \in M$. Para representá-la, pode empregar-se o símbolo $(d\phi)_p$, ou simplesmente $d\phi$. Também é usado o símbolo ϕ_*, quando o ponto p pode ficar subentendido, sem prejuízo da clareza. No presente trabalho, referir-nos-emos, de modo geral, à aplicação $\phi_p \colon M_p \to N_q$ como sendo a *aplicação linear induzida em* p *pela aplicação* $\phi \colon M \to N$.

Sejam M, N, P três variedades e $\phi \colon M \to N$, $\psi \colon N \to P$ duas aplicações diferenciáveis. Sejam $p \in M$, $q = \phi(p) \in N$ e $r = \psi(q) = (\psi \circ \phi)(p) \in P$. Verifica-se, sem dificuldade, que $(\psi \circ \phi)_p = \psi_q \circ \phi_p$.

Observemos também que se $i \colon M \to M$ é a apliação idêntica, e se $p \in M$, então $i_p \colon M_p \to M_p$ é a aplicação idêntica.

7 Regularidade de uma aplicação diferenciável. Imersões. Difeomorfismos

Sejam M^m, N^n duas variedades e $\phi\colon M \to N$ uma aplicação diferenciável. Dizemos que ϕ é *regular no ponto* $p \in M$ quando a aplicação linear induzida $\phi_p\colon M_p \to N_{\phi(p)}$ é *biunívoca.* Neste caso, a imagem $\phi_p(M_p)$ é um subespaço de dimensão m do espaço vetorial $N_{\phi(p)}$, e é claro que isso só pode ocorrer quando $m \leq n$ (no caso $m = n$, ϕ_p é um isomorfismo entre os espaços vetoriais M_p e $N_{\phi(p)}$). Uma *aplicação regular* $\phi\colon M \to N$ é uma aplicação que é regular em todo ponto $p \in M$.

Vamos exprimir essa idéia de regularidade em termos de coordenadas locais. Seja $y\colon V \to R^n$ um sistema local em N, tal que $\phi(p) \in V$, e seja $x\colon U \to R^m$ um sistema em M, tal que $p \in U$ e $\phi(U) \subset V$. Relativamente às bases de M_p e $N_{\phi(p)}$ associadas aos sistemas x e y, respectivamente, a matriz da aplicação linear $\phi_p\colon M_p \to N_{\phi(p)}$ é, segundo já assinalamos, a matriz jacobiana $\left(\frac{\partial y^i}{\partial x^j}(p)\right)$ da aplicação $y \circ \phi \circ x^{-1}\colon x(U) \to y(V)$. De acordo com a definição acima, ϕ é regular em p se e somente se a característica dessa matriz é m. Verifica-se, trivialmente, que essa condição independe da escolha dos sistemas de cordenadas x e y.

Um fato importante, que devemos mencionar aqui, é o seguinte: se a aplicação $\phi\colon M \to N$ é regular no ponto $p \in M$, existe uma vizinhança $U \ni p$ em M, tal que a restrição $\phi|U$ é biunívoca; se ϕ é regular, então é localmente biunívoca (isto é, cada ponto $p \in M$ tem uma vizinhança U na condição descrita). A demonstração desse resultado pode ser vista em [6], capítulo III.

A $\phi\colon M \to N$ é uma *imersão* quando é ao mesmo tempo uma aplicação regular e um homeomorfismo de M em N.

A aplicação $\phi\colon M \to N$ é um *difeomorfismo* quando é um homeomorfismo diferenciável de M sobre N, cujo inverso $\phi^{-1}\colon N \to M$ é também diferenciável. Nessas condições, se $p \in M$ e $q = \phi(p) \in N$, podemos escrever: $(\phi^{-1})_q \circ \phi_p = (\phi^{-1} \circ \phi)_p = i_p\colon M_p \to M_p$ (identidade) e $\phi_p \circ (\phi^{-1})_q = (\phi \circ \phi^{-1})_q = i_q\colon N_q \to N_q$ (identidade). Isso mostra que ϕ_p aplica M_p biunivocamente sobre N_q (e é, pois, um isomorfismo de espaços vetoriais) e que $(\phi_p)^{-1} = (\phi^{-1})_q$; em particular, podemos concluir que as variedades M e N têm a mesma dimensão.

O conceito de difeomorfismo é o que substitui o de homeomorfismo, quando passamos da Topologia Geral à Topologia Diferencial.

8 Valores regulares e valores críticos de uma aplicação diferenciável

Seja $\phi\colon M^m \to N^n$ uma aplicação diferenciável. Um ponto $q \in N$ diz-se *valor regular* de ϕ se para todo $p \in \phi^{-1}(q)$ a aplicação linear ϕ_p aplica o espaço tangente M_p *sobre* o espaço tangente N_q (quer dizer, todo vetor $v \in N_q$ é a imagem $\phi_p(u)$ de algum vetor $u \in M_p$, ou, em símbolos: $\phi_p(M_p) = N_q$). De acordo com essa definição, todo ponto $q \in N$ tal que $\phi^{-1}(q) = \varnothing$ é um valor regular de ϕ.

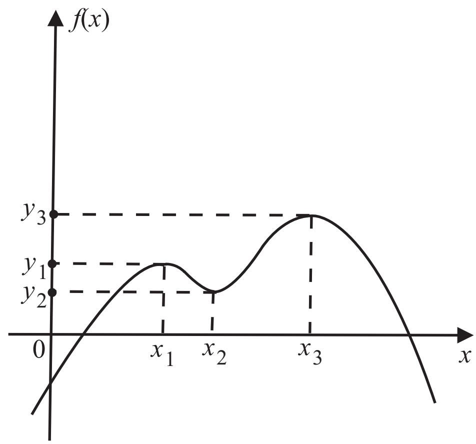
Um ponto $q \in N$ é um *valor crítico* de ϕ quando não é um valor regular. Portanto, se $q \in N$ é um valor crítico de ϕ, existe pelo menos um ponto $p \in \phi^{-1}(q) \subset M$, tal que $\phi_p(M_p)$ é um subespaço próprio de N_q (neste caso, ϕ_p não aplica M_p sobre N_q); o ponto p diz-se um *ponto crítico* de ϕ. É claro que o conjunto dos valores críticos de ϕ é a imagem do conjunto dos seus pontos críticos. Os valores regulares e os valores críticos de ϕ constituem conjuntos complementares na variedade N.

A definição dada acima permite concluirmos imediatamente que se $q \in N$ é um valor regular de $\phi\colon M^m \to N^n$, tal que $\phi^{-1}(q) \neq \varnothing$, então deve ser $m \geq n$. Segue-se que se $m < n$, todo $q \in \phi(M)$ é um valor crítico de ϕ; neste caso, $\phi(M)$ é justamente o conjunto dos valores críticos de ϕ.

Vejamos como se exprimem essas idéias em termos de coordenadas locais. Seja $q \in N$ um valor crítico de $\phi\colon M^m \to N^n$. Tomemos um sistema de coordenadas locais em N, $y\colon V \to R^n$, tal que $q \in V$. Existe algum ponto $p \in \phi^{-1}(q)$, tal que a aplicação $\phi_p\colon M_p \to N_q$ não é sobre. Seja $x\colon U \to R^m$ um sistema local em M, tal que $p \in U$ e $\phi(U) \subset V$. Consideremos a matriz jacobiana $\left(\frac{\partial y^i}{\partial x^j}(p)\right)$ da ϕ, relativamente aos ditos sistemas; ela é a matriz da aplicação linear ϕ_p, com respeito às bases de M_p e N_q associadas àqueles sistemas. Dizer que $\phi_p(M_p)$ é subespaço próprio de N_q é dizer que $\dim \phi_p(M)_p < \dim N_q$, ou seja, que a matriz jacobiana acima tem característica $< n$. É claro que essa condição não depende da escolha dos sistemas de coordenadas x e y.

Quando $q \in N$ é valor regular de ϕ, a matriz jacobiana de ϕ, em todo $p \in \phi^{-1}(q)$, tem característica n, quaisquer que sejam os sistemas de coordenadas locais usados.

Os conceitos de valor crítico e de valor regular de uma aplicação diferenciável são fundamentais no presente trabalho. Por essa razão, vamos ilustrá-los por meio de exemplos. Seja $f\colon R \to R$ uma função diferenciável. Os pontos críticos de f são os pontos $x \in R$ tais que $f'(x) = 0$. Um valor $y \in R$ é crítico para f se existe pelo menos um ponto $x \in R$ tal que $f(x) = y$ e $f'(x) = 0$. No gráfico da função $f\colon R \to R$, os pontos críticos são aqueles em que a função passa por máximo ou mínimo, ou a curva apresenta inflexão com tangente paralela ao eixo dos xx. Na figura abaixo, x_1 , x_2 , x_3 são pontos críticos de f, e y_1 , y_2 , y_3 são valores

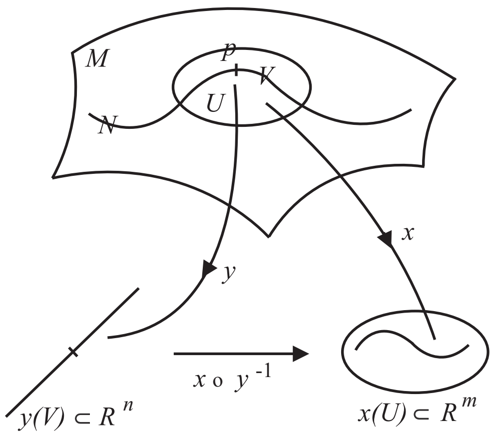


críticos. A função seno ($y = \text{sen}\ x$), por exemplo, tem uma infinidade (enumerável) de pontos críticos, e os seus únicos valores críticos são $+1$ e -1.

Consideremos, agora, uma função diferenciável $f\colon R^2 \to R$. Ponhamos $f(x, y) = z$. Um ponto $p = (x_0, y_0) \in R^2$ é crítico para f se $\frac{\partial f}{\partial x}(p) = 0$ e $\frac{\partial f}{\partial y}(p) = 0$. Um valor $z_0 \in R$ é crítico para f se existe pelo menos um ponto de R^2, $p = (x_0, y_0)$, tal que $f(p) = z_0$, $\frac{\partial f}{\partial x}(p) = 0$ e $\frac{\partial f}{\partial y}(p) = 0$.

9 Subvariedades

Consideremos duas variedades diferenciáveis M^m e N^n, tais que $N \subset M$. Diremos que N é uma *subvariedade* de M se a aplicação de inclusão $F\colon N \to M$ é uma imersão. O significado preciso dessa definição é o seguinte: 1) como conjunto de pontos, a subvariedade N é um subconjunto de M; 2) como espaço topológico, a subvariedade possui a topologia induzida pela da variedade; 3) a estrutura diferenciável da subvariedade N satisfaz à seguinte condição: para todo ponto $p \in N$, e para todo sistema de coordenadas locais $x\colon U \to R^m$, válido num aberto $U \subset M$, tal que $p \in U$, existe um sistema local $y\colon V \to R^n$, válido num aberto $V \subset N$, tal que $p \in V \subset U$,



e a aplicação $x \circ y^{-1}\colon y(V) \to x(U)$ é diferenciável, e tem matriz jacobiana com característica n em todo ponto do seu domínio.

A última condição mostra claramente que se N^n é subvariedade de M^m, deve ser $n \leq m$. Simples considerações, que omitiremos por brevidade, permitem concluirmos que a igualdade $n = m$ ocorre se e somente se N é um conjunto aberto de M.

Observemos também que se M, N são duas variedades diferenciáveis, e se $f\colon M \to N$ é uma imersão, então a imagem $f(M)$ é uma subvariedade de N, e a f aplica M difeomorficamente sobre $f(M)$.

Por várias vezes, no texto subsequente, teremos necessidade de considerar a imagem inversa de um valor regular relativamente a uma aplicação diferenciável. Vale, a propósito, o seguinte

Teorema. *Sejam M^m, N^n duas variedades e $f\colon M \to N$ uma aplicação diferenciável. Seja $q \in N$ um valor regular de f.*

1) *Se $m < n$, o conjunto $f^{-1}(q)$ é vazio.*

2) *Se $m = n$, $f^{-1}(q)$ é vazio ou é um conjunto de pontos isolados.*

3) *Se $m > n$, $f^{-1}(q)$ é vazio ou é uma subvariedade de M, de dimensão $m - n$.*

Demonstração. 1) Se $m < n$, todo $p \in M$ é ponto crítico de f, e então é claro que $f^{-1}(q) = \varnothing$.

2) Seja $m = n$. Se $f^{-1}(q)$ não é vazio, a f é regular em cada $p \in f^{-1}(q)$; existe então uma vizinhança $U \ni p$ em M, tal que a restrição $f|U$ é biunívoca; resulta daí que p é o único ponto de U que se aplica em q. Não existe, pois, outro ponto de $f^{-1}(q)$ arbitrariamente próximo de p; em outros termos: os pontos de $f^{-1}(q)$ são isolados.

3) Seja agora $m > n$, e admitamos $f^{-1}(q)$ não vazio. Tomemos um sistema de coordenadas locais $y\colon V \to R^n$, em N, válido numa vizinhança $V \ni q$, e suponhamos, por

simplicidade, que $y(q) = 0 \in R^n$. Consideremos um ponto qualquer $p \in f^{-1}(q)$, e seja $x\colon U \to R^m$ um sistema local arbitrário em M, tal que $p \in U$ e $f(U) \subset V$. Ponhamos, para fixar idéias, $x(p) = (x_0^1, \ldots, x_0^m)$. Com referência aos ditos sistemas de coordenadas, podemos representar a aplicação f por equações

$$y^i = y^i(x^1, \ldots, x^m), \qquad i = 1, \ldots, n$$

(note-se que o domínio das funções $y^1, \ldots, y^n$ é o conjunto aberto $x(U) \subset R^m$).

Representemos por $\left(\frac{\partial y^i}{\partial x^j}(p)\right)$ a matriz jacobiana desse sistema de equações, calculada no ponto $x(p)$. Essa matriz tem característica n, por hipótese. Logo, podemos supor, mediante eventual reenumeração das coordenadas de p, que

$$\det\left(\frac{\partial y^i}{\partial x^j}(p)\right) \neq 0, \qquad i, j = 1, \ldots, n.$$

É claro que as coordenadas $x^1, \ldots, x^m$ de todo ponto do conjunto $U \cap f^{-1}(q)$ satisfazem as equações

$$y^i(x^1, \ldots, x^m) = 0, \qquad i = 1, \ldots, n.$$

De acordo com o teorema das funções implícitas, podemos concluir que essas equações definem, localmente, $x^1, \ldots, x^n$ como funções de $x^{n+1}, \ldots, x^m$. Em termos precisos: existe uma vizinhança W do ponto $(x_0^{n+1}, \ldots, x_0^m)$ em R^{m-n}, e uma vizinhança $Z \ni p$ em M, $Z \subset U$, de maneira que a cada ponto $(x^{n+1}, \ldots, x^m) \in W$ corresponde um e um só ponto $r \in Z$, com $x(r) = (x^1, \ldots, x^n, x^{n+1}, \ldots, x^m)$, tal que as equações

$$y^i(x^1, \ldots, x^n, x^{n+1}, \ldots, x^m) = 0, \qquad i = 1, \ldots, n,$$

são satisfeitas simultaneamente. Segue-se que $f^{-1}(q)$ é uma subvariedade de dimensão $m - n$ da variedade M (a aplicação $r \to (x^{n+1}, \ldots, x^m)$, definida em $Z \cap f^{-1}(q)$, com valores em W, é um sistema de coordenadas locais em $f^{-1}(q)$, válido numa vizinhança de p).

10 Variedades com bordo

De acordo com a definição de variedade topológica (e, em particular, diferenciável), todo ponto de uma variedade M^n tem uma vizinhança (em M) homeomorfa a um conjunto aberto do espaço euclidiano R^n. Assim, por exemplo, pode a reta real, ou um intervalo real aberto, ou um disco aberto do plano, receber uma estrutura de variedade diferenciável, mas não acontece o mesmo com um intervalo real fechado, ou com um disco fechado do plano. Com efeito, se I é o intervalo fechado $[a, b] \subset R$, é claro que nenhuma vizinhança do ponto a em I pode ser homeomorfa a um aberto de R, e, no caso do disco fechado $B = \{(x, y) \in R^2;\ x^2 + y^2 \leq 1\}$, uma impossibilidade do mesmo tipo ocorre com todo ponto da circunferência $S = \{(x, y) \in R^2;\ x^2 + y^2 = 1\}$.

Entretanto, é possível (e muito conveniente) ampliar, em certo sentido, o conceito de variedade, mediante modificação da definição dada, de maneira que o novo conceito venha a englobar alguns casos importantes (como os acima citados) que se não enquadram como variedades segundo a primeira definição. A alteração de que se trata consiste essencialmente em alargar, de certo modo, o conceito de sistema de coordenadas locais. Temos considerado um tal sistema, até agora, como um homeomorfismo $x \colon U \to R^n$, onde U é um conjunto aberto da variedade M^n e $x(U)$ é um aberto do espaço euclidiano R^n. Consideremos em R^n o *semi-espaço*

$$R^n_0 = \{(x^1, \ldots, x^n) \in R^n;\, x^n \geq 0\}.$$

O hiperplano $H^{n-1} = \{(x^1, \ldots, x^n) \in R^n;\, x^n = 0\}$ diz-se bordo do semi-espaço R^n_0. A topologia que consideramos sempre em R^n_0 é a induzida pelo espaço R^n (isto é, pensamos em R^n_0 como subespaço de R^n); é claro que um conjunto $A \subset R^n_0$ pode ser aberto neste semi-espaço sem o ser no espaço R^n (e isso ocorre precisamente quando A contém pontos do bordo de R^n_0).

Já dissemos que uma variedade topológica de dimensão n é um espaço de Hausdorff M, com base enumerável de conjuntos abertos, no qual cada ponto tem uma vizinhança homeomorfa a um aberto do espaço euclidiano R^n. Modifiquemos essa definição em sua parte final, e passemos a admitir que cada ponto $p \in M$ deva possuir uma vizinhança homeomorfa a um aberto do espaço R^n *ou do semi-espaço* R^n_0. O conceito de variedade topológica a que chegamos graças a essa alteração, inclui, evidentemente, o conceito ao qual corresponde a definição antiga, e é, pois, mais abrangente.

Seja M uma variedade no sentido da nova definição. Consideremos o conjunto $N \subset M$, constituído pelos pontos $p \in M$, para cada um dos quais existe em M uma vizinhança $V \ni p$ homeomorfa a um aberto do espaço R^n (o homeomorfismo $x\colon V \to R^n$ é um sistema de coordenadas válido em V). Esse conjunto N é aberto em M, pois é claro que só possui pontos interiores; o seu complemento $M - N$ é um conjunto fechado em M, ao qual chamamos *bordo* da variedade M. O bordo de M, que indicamos com o símbolo ∂M, pode ser um conjunto vazio, ou não. No primeiro caso, M é uma variedade no sentido da antiga definição, e, quando quisermos dar realce à circunstância de que $\partial M = \varnothing$, diremos que M é uma *variedade sem bordo.* No segundo caso, M é uma *variedade com bordo.* A uma variedade compacta e sem bordo, chamaremos simplesmente *variedade fechada.*

Em uma variedade com bordo M^n, consideremos um ponto $p \in \partial M$. Se $x\colon U \to R^n$ é qualquer sistema de coordenadas locais em M, tal que $p \in U$, é claro que $x(U)$ é um aberto do semi-espaço R^n_0, mas não é um aberto de R^n, e então, $x(U)$ contém necessariamente pontos do hiperplano H^{n-1}. Pode demonstrar-se, sem dificuldade, que os pontos do conjunto $x(U) \cap H^{n-1}$ são as imagens (por x) dos pontos de $U \cap \partial M$. Portanto, todo ponto $q \in U \cap \partial M$ (e em particular o ponto p) tem coordenadas da forma $(x^1, \ldots, x^{n-1}, 0)$ no sistema x, e vice-versa, todo ponto $q \in U$ cujas coordenadas no sistema x sejam dessa forma, pertence ao bordo ∂M.

Os conceitos de atlas diferenciável e de atlas máximo estendem-se imediatamente às variedades com bordo. Podemos, pois, daqui em diante, considerar também variedades diferenciáveis de dimensão n e classe C^k, dotadas de bordo. Para estas podemos definir, da mesma maneira como o fizemos para as variedades sem bordo, os conceitos de espaço vetorial tangente, de aplicação diferenciável, de diferencial, de valor regular, de imersão, de subvariedade, etc.

Observações. Entre as propriedades interessantes relacionadas com as variedades dotadas de bordo, citamos as seguintes:

1ª) A estrutura diferenciável de uma variedade com bordo M^n induz sobre o bordo ∂M uma estrutura de variedade diferenciável sem bordo, de mesma classe e de dimensão $n-1$. (Em termos menos refinados: o bordo de uma variedade diferenciável de dimensão n é uma variedade de mesma classe, de dimensão $n-1$, sem bordo.)

2ª) Seja M^n uma variedade diferenciável com bordo, e seja $p \in \partial M$. O espaço vetorial tangente ao bordo em p, $(\partial M)_p$, pode ser identificado, de modo natural, com um subespaço de dimensão $n-1$ do espaço vetorial M_p, tangente à variedade M em p.

3ª) O teorema estudado no final do n. 9 é válido igualmente para variedades com bordo, conforme atesta um mero exame da demonstração apresentada. Uma situação interessante, que merece registro aqui, é a que passamos a descrever. Seja $f\colon M^n \to N^{n-r}$ uma aplicação diferenciável da variedade M (com bordo) na variedade N (com ou sem bordo). Estamos supondo $1 \leq r < n$. Se $q \in N$ é valor regular de f, a imagem inversa $f^{-1}(q) = P \subset M$ é uma subvariedade de M, de dimensão r, de acordo com o dito teorema. Nessas condições, dois fatos devem ser assinalados:

1) o bordo da subvariedade P é a interseção dela com o bordo da variedade M; em símbolos: $\partial P = P \cap \partial M$;

2) em cada ponto $p \in \partial P$, os espaços vetoriais tangentes $(\partial P)_p$, P_p e $(\partial M)_p$ são tais que

$$(\partial P)_p = P_p \cap (\partial M)_p\,.$$

O leitor interessado pode encontrar as demonstrações desses resultados em [12], pag. 7; observe-se, a propósito, que, para Pontrjagin, no trabalho que acabamos de citar, as únicas subvariedades de M são aquelas que podem ser definidas como o foi P nas considerações acima (isto é, como imagem inversa de um valor regular em uma aplicação diferenciável). Aliás, também para as necessidades deste nosso trabalho, tais subvariedades são as que realmente têm interesse, mas, num capítulo cuja finalidade era apresentar um resumo das noções básicas da teoria das variedades diferenciáveis, não podíamos deixar de definir as subvariedades da maneira mais geral que adotamos.

Capítulo II

O Teorema de Sard

Tem este capítulo por principal objetivo a demonstração do teorema de Sard. Antes de iniciarmos a marcha na direção dessa meta, queremos dar sobre esse teorema algumas explicações, por meio das quais possa o leitor adquirir um conhecimento perfeito do substrato da proposição que é o tema central do nosso trabalho. Para começar, nada melhor do que focalizar alguns resultados familiares (e que, na verdade, são filiados à ordem de estudos com que nos ocupamos) e, ao mesmo tempo, fazer um pouco de história, e descrever a evolução das idéias matemáticas que culminaram com o estabelecimento do dito teorema.

1 Introdução. Notícia histórica

Comecemos com observações de caráter intuitivo, aparentemente desprovidas de sentido mais profundo. Adiante, veremos que essas observações se referem a um estado de coisas muito mais geral, o qual é governado pela proposição de Sard. A Geometria Diferencial Clássica estuda as curvas e superfícies regulares do espaço ordinário R^3. Tais entidades são definidas da maneira já recordada no Capítulo I, por meio de aplicações diferenciáveis (de classe C^k, $1 \leq k \leq \infty$). As curvas e as superfícies do espaço euclidiano ordinário são, como vimos, subespaços de R^3 localmente homeomorfos a conjuntos abertos da reta R e do plano R^2, respectivamente. O aspecto para o qual desejamos chamar a atenção do leitor neste momento é o seguinte: tanto as curvas quanto as superfícies, consideradas como partes de R^3, são conjuntos de volume nulo (são conjuntos de medida nula em R^3, e, portanto, de interior vazio em R^3). Observemos que todos os pontos de uma curva ou superfície de R^3 são obviamente valores críticos da correspondente aplicação diferenciável. Desse modo, quando consideramos curvas e superfícies (de classe C^k) do espaço ordinário, estamos lidando com aplicações diferenciáveis para as quais tem medida nula o conjunto dos valores críticos. A demonstração rigorosa desse fato será apresentada adiante; contentemo-nos, por equanto, com obsrvações intuitivas.

Recordemos outro resultado interessante, muito conhecido da Análise Clássica. Seja $f(x)$ uma função de classe C^1, definida num aberto Ω da reta. Se $f'(x) = 0$ para todo $x \in I$, onde $I \subset \Omega$ é um conjunto conexo (um intervalo), então a função f é constante sobre I. Esse fato é consequência imediata do teorema do valor médio.

Surge naturalmente a idéia de estender o mesmo resultado às funções reais de várias variáveis. Consideremos, por exemplo, uma função $f(x, y)$, de classe C^1, definida num aberto Ω do plano, e suponhamos que $\frac{\partial f}{\partial x} = 0$, $\frac{\partial f}{\partial y} = 0$ sobre um conjunto conexo $A \subset \Omega$. Não é fora de propósito conjecturar que a função f seja constante sobre A.

H. Whitney, em 1935, demonstrou que essa conjectura é falsa. Conseguiu Whitney construir uma função real de classe C^1, $f(x,y)$, para a qual são críticos todos os pontos de um arco do plano, assumindo, sobre esse arco, todos os valores do intervalo $[0,1]$. Os valores críticos dessa função formam, como se vê, um conjunto que *não é* de medida nula. Poderá o leitor interessado estudar esse notável e-xemplo no próprio artigo em que Whitney o apresentou (v. [16]).

A literatura existente sobre a matéria atribui a Marston Morse o pioneirismo da idéia de considerar a medida do conjunto dos valores críticos de uma função real de várias variáveis. Não se pode afirmar, sem maior cuidado, que tal conjunto é sempre de medida nula, pois aí está o exemplo de Whitney para invalidar essa asserção. Estudos feitos sobre essa questão mostraram, porém, que a dita afirmação é verdadeira desde que a classe de diferenciabilidade da função seja suficientemente elevada. Para pôr as idéias em termos precisos, seja $f\colon \Omega \to R$ uma função de classe C^k, definida num aberto $\Omega \subset R^n$. Estamos supondo $n \geq 1$ e $k \geq 1$. Marston Morse e Arthur Sard, segundo consta da literatura disponível, provaram que os valores críticos de f constituem um conjunto de medida nula em R, se $k \geq N$, onde N é o maior inteiro que não excede o número $n + \frac{1}{16}(n-3)^2$. Esse resultado mostra, em particular, que para os valores $n = 1, \ldots, 6$, basta que se tenha $k = n$ para que seja válida a afirmação acima.

Prosseguindo nessa ordem de estudos, Anthony P. Morse demonstrou, em 1938, que a hipótese $k = n$ é sempre suficiente para o fim visado, isto é: se $f\colon \Omega \to R$ é uma função de classe C^n, definida em um aberto $\Omega \subset R^n$ $(n \geq 1)$, então o conjunto dos valores críticos de f é de medida nula em R. O trabalho de A.P. Morse intitula-se "The behavior of a function on its critical set" (v. [11]).

Em artigo publicado em 1942, sob o título "The measure of the critical values of differentiable maps" (v. [15]), Arthur Sard estendeu o resultado supra às aplicações diferenciáveis definidas num aberto de R^m e tomando valores em R^n. A proposição de A. Sard pode enunciar-se nestes termos: seja $F\colon \Omega \to R^n$ uma aplicação diferenciável de classe C^k $(k \geq 1)$, cujo domínio Ω é um aberto de R^m. Se $m \leq n$, o conjunto dos valores críticos de F é de medida nula, qualquer que seja k; se $m > n$, o conjunto dos valores críticos de F é de medida nula, desde que seja $k \geq m-n+1$.

Leon S. Pontrjagin, em seu trabalho "Smooth manifolds and their applications in homotopy theory" (v. [12]), atribui a A. Dubovitsky a autoria de um teorema que nos traz à lembrança o teorema de Sard. A proposição de Dubovitsky não emprega a idéia de medida nula, mas, em lugar desta, usa o conceito de categoria de conjuntos. Esse teorema foi demonstrado, pela primeira vez, em 1953, portanto 11 anos depois da publicação de Arthur Sard. No mencionado trabalho, que data de 1955, Pontrjagin demonstra esse teorema com a hipótese de que seja $k \geq 2 + \frac{1}{2}(m-n)(m-n+1)$, graças à qual diz que obtém certas simplificações no decurso da prova.

No presente escrito, mostraremos que o teorema de Dubovitsky é um corolário do teorema de Sard, e daremos deste último uma demonstração à qual chegamos através do estudo da prova que Pontrjagin apresenta para o primeiro.

A demonstração que desenvolveremos será um tanto simplificada, porque vamos supor que as aplicações com as quais lidaremos sejam de classe C^∞ (e, portanto, que também sejam de classe C^∞ as variedades consideradas). Procedendo assim, estaremos descansados quanto à possibilidade de derivar as nossas funções quantas vezes quisermos, e não necessitaremos contar, em cada fase da demonstração, o número de vezes que já usamos a diferenciação. Tal atitude é mais consentânea à tendência que hoje predomina na teoria das variedades diferenciáveis, e que pode ser observada nos trabalhos de alguns dos matemáticos mais notáveis da atualidade, entre os quais podemos citar: R, Thom, J. Milnor, W. Ambrose, G. de Rham, S. Smale. Tal ponto de vista, não obstante apresentar o inconveniente de ser um tanto restritivo, justifica-se cabalmente pelas seguintes razões:

1ª) a classe C^∞ é a única invariante relativamente à diferenciação;

2ª) se k é finito, nunca se pode garantir que a consideração da classe C^k seja suficiente para uma teoria completa que envolva variedades e aplicações diferenciáveis;

3ª) a adoção da classe C^∞ numa teoria tem a vantagem de impedir que o verdadeiro conteúdo das proposições estudadas nessa doutrina possa ser mascarado ou obscurecido pela enfadonha enumeração de hipóteses de diferenciabilidade, que de outro modo seriam inevitáveis.

Após essas explicações, deixemos convencionado, de uma vez por todas, que o adjetivo "diferenciável" significará sempre, no que segue, "diferenciável de classe C^∞". Somente faremos referência explícita à classe C^∞ quando tivermos a intenção de dar mais ênfase a essa condição.

2 Teorema de Sard e Teorema de Dubovitsky

Através da introdução que acabamos de apresentar, pode o leitor observar que o teorema de Sard, em sua forma original, trata de aplicações diferenciáveis do tipo $f\colon \Omega \to R^n$, onde Ω é um aberto do espaço euclidiano R^m. O teorema afirma que os valores críticos de f constituem um conjunto de medida nula em R^n.

Recordemos que um conjunto $C \subset R^n$ tem medida nula se é possível cobrir C por meio de uma coleção enumerável(*) de cubos (ou bolas, ou paralelepípedos) de R^n, cujos volumes tenham soma arbitrariamente pequena.

Se C é de medida nula em R^n, C é um conjunto de interior vazio (em R^n), e então o complementar $R^n - C$ é denso em R^n. Sabe-se também que toda parte de um conjunto de medida nula tem medida nula, e que toda reunião finita ou enumerável de conjuntos de medida nula é um conjunto de medida nula. Do ponto de vista intuitivo, a afirmação de que C é de medida nula em R^n sugere-nos a idéia de que C é pequeno em relação a R^n, ou, em outras palavras, que o complementar $R^n - C$ é quase todo o espaço R^n.

O teorema de Sard, como veremos, estende-se a aplicações diferenciáveis do tipo $F\colon M^m \to N^n$, onde M e N são variedades diferenciáveis. Embora a demonstração desse fato possa reduzir-se facilmente ao caso euclidiano, como adiante veremos, a possibilidade dessa extensão vem aumentar consideravelmente o alcance do teorema de Sard, e

(*) Aqui, e em muitas outras situações no texto seguinte, enumerável significa finito ou infinito enumerável.

dilatar o campo de suas aplicações. Surge aqui a questão preliminar de definir o conceito de medida nula para subconjuntos de uma variedade diferenciável. Para fazer isso do modo mais satisfatório, comecemos demonstrando a seguinte

Proposição. *Se $\phi\colon R^n \to R^n$ é diferenciável, e se $C \subset R^n$ é um conjunto de medida nula, então $\phi(C)$ é também de medida nula.*

Demonstração. Suponhamos, inicialmente, que o conjunto C seja limitado. Seja $A \subset R^n$ uma bola aberta tal que $C \subset A$. Por ser diferenciável, a aplicação ϕ é lipschitziana sobre o compacto $\overline{A}$, isto é, existe um número real $K > 0$ tal que $|\phi(q)-\phi(p)| < K|p-q|$, quaisquer que sejam os pontos $p, q \in \overline{A}$. Nessas condições, se $B \subset \overline{A}$ é uma bola de raio δ, a imagem $\phi(B)$ há de estar contida numa bola B' de raio $K\delta$, e então é claro que vol $B' = K^n$vol B. Como C é de medida nula, podemos cobrir C com uma coleção de bolas $\{B_1, B_2, \dots\}$ tal que $\sum\limits_i \text{vol } B_i < \frac{\varepsilon}{K^n}$, onde $\varepsilon > 0$ é arbitrário. Podemos, evidentemente, admitir que $B_i \subset A$, qualquer que seja i. Resulta, pois, que

$$\phi(C) \subset \phi\left(\bigcup_i B_i\right) = \bigcup_i \phi(B_i) \subset \bigcup_i B_i'$$

e

$$\text{vol}\left(\bigcup_i B_i'\right) \leq \sum_i \text{vol } B_i' = \sum_i K^n \text{vol } B_i = \\ = K^n \sum_i \text{vol } B_i < K^n \, \frac{\varepsilon}{K^n} = \varepsilon.$$

Isso mostra que $\phi(C)$ é de medida nula.

Se o conjunto C não é limitado, podemos considerar uma cobertura enumerável de C por meio de bolas abertas $B_1, B_2, \ldots$. Segue-se que $C = \bigcup_i (C \cap B_i)$, e, para cada i, $C \cap B_i$ é um conjunto limitado de medida nula. Logo, $\phi(C) = \bigcup_i \phi(C \cap B_i)$ é de medida nula.

Do resultado que acabamos de provar decorre, em particular, que a propriedade da medida nula é preservada pelos difeomorfismos. Em outras palavras: a classe constituída pelos subconjuntos de R^n que possuem medida nula é invariante através dos difeomorfismos.

Consideremos, agora, uma variedade diferenciável M, de dimensão n, e seja $C \subset M$. Diremos que *C é de medida nula em M* se existir uma coleção finita ou enumerável de sistemas de coordenadas locais em M, $\{x_i : U_i \to R^n\}$, tal que: 1) $C \subset \bigcup_i U_i$; 2) para todo i, o conjunto $x_i(U_i \cap C)$ é de medida nula em R^n.

Resulta do estudo acima feito que se C é de medida nula em M, $y(V \cap C)$ é de medida nula em R^n, qualquer que seja o sistema de coordenadas locais $y : V \to R^n$, pertencente ao atlas máximo de M.

Se $C \subset M$ é de medida nula, é claro que C é um conjunto de interior vazio (em M), e então $M - C$ é denso em M.

Estamos, agora, em condições de apresentar o teorema de Sard em sua forma geral.

Teorema (de Sard). *Sejam M^m, N^n duas variedades, e $\phi : M \to N$ uma aplicação diferenciável. Os valores críticos de ϕ constituem um conjunto de medida nula em*

N. Em consequência, o conjunto dos valores regulares de ϕ é denso em N.

Antes de passsar à demonstração do teorema de Sard, queremos apresentar o enunciado do teorema de Dubovitsky, a que já nos referimos. Vamos recordar alguns conceitos, estudados em Topologia Geral, em termos dos quais é formulado o enunciado do dito teorema.

Consideremos um espaço topológico X, e seja $C \subset X$. Dizemos que C é um *conjunto de primeira categoria em* X se C é uma reunião finita ou enumerável de partes de X, cujos fechos são conjuntos de interior vazio em X. Simbolicamente, $C = \bigcup_{i=1}^{\infty} C_i$, onde $C_i \subset X$ e int $\overline{C}_i = \varnothing$, para todo indice i. (O caso finito enquadra-se nessa mesma representação, quando se supõe que os conjuntos C_i são vazios para $i > p$, onde p é um inteiro natural.)

Se C é de primeira categoria em X, o mesmo ocorre com toda parte de C. É também óbvio que toda reunião finita ou enumerável de conjuntos de primeira categoria em X é um conjunto de primeira categoria em X.

Verifica-se, trivialmente, que C é de primeira categoria em X se e somente se $C \subset \bigcup_{i=1}^{\infty} F_i$, onde cada F_i é um conjunto fechado de interior vazio em X.

O espaço topológico X diz-se um *espaço de Baire* se todo conjunto de primeira categoria em X tem interior vazio. Dessa definição resulta que num espaço de Baire, é denso o complementar de todo conjunto de primeira categoria.

Para o estudo que estamos fazendo desempenha importante papel o conhecido

Teorema (de Baire). *Os espaços topológicos localmente compactos e os espaços métricos completos são espaços de Baire.*

A demonstração desse teorema pode ser encontrada em [2], cap. IX, e em [5], cap. III.

As variedades diferenciáveis são espaços de Baire; com efeito, por serem espaços localmente euclidianos, elas são também espaços localmente compactos.

Observe-se que num espaço de Baire não é verdade que um conjunto de interior vazio tenha que ser de primeira categoria. Por exemplo, na reta real R, que é um espaço de Baire, o conjunto dos pontos irracionais tem interior vazio, mas não é de primeira categoria.

A idéia intuitiva encerrada na definição dos espaços de Baire é a seguinte: em um espaço desses, os conjuntos de primeira categoria são pequenos em comparação com o espaço inteiro. Mas, atente-se bem em que essa pequenez não deve ser interpretada do ponto de vista da potência dos conjuntos considerados. Um conjunto C de primeira categoria num espaço de Baire X pode ser denso em X (é o que ocorre, por exemplo, com o conjunto dos números racionais na reta real). Há casos em que C pode até ter a mesma potência que X. O aspecto realmente importante, que dá o sentido exato da assinalada pequenez, é o que já registramos acima: num espaço de Baire, é denso o complemento de um conjunto de primeira categoria; esse complemento nunca pode ser vazio.

A idéia de pequenez a que nos referimos é a mesma que nos sugere a propriedade da medida nula (em espaços onde ela tenha sentido), mas não se deve supor que haja equivalência entre o conceito de primeira categoria e o de medida nula. É possível apresentar exemplos de conjuntos de primeira categoria que não têm medida nula, e também de conjuntos de medida nula que não são de primeira categoria.

Tendo em vista que o teorema de Sard é baseado sobre esses dois conceitos, é essencial, para a utilidade do presente trabalho, que não deixemos no espírito do leitor nem a mais leve dúvida sobre o significado exato de tais idéias. Vamos, pois, apresentar alguns exemplos esclarecedores. Consideremos, inicialmente, o conjunto de Cantor C. Esse conjunto é construído da seguinte maneira: tomemos o intervalo real fechado $[0,1]$ e suprimamos o seu terço médio aberto $\left(\frac{1}{3}, \frac{2}{3}\right)$; de cada um dos dois intervalos fechados restantes, suprimamos o respectivo terço médio aberto; de cada um dos quatro intervalos fechados remanescentes, tiremos o correspondente terço médio aberto; e assim sucessivamente. O conjunto de Cantor é o conjunto C que resta do intervalo $[0,1]$ após a supressão de todos os terços médios. A reunião de todos os intervalos abertos removidos é um conjunto aberto cujo complementar em relação a $[0,1]$ é C. Concluímos que o conjunto de Cantor é fechado. Observemos que a soma dos comprimentos de todos os intervalos suprimidos é $\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots=1$. Segue-se que C não pode conter nenhum intervalo, e é, pois, um conjunto de interior vazio. Fica assim provado que o conjunto de Cantor é de primeira categoria. Esse conjunto é também de medida nula. Com efeito, dado $\varepsilon>0$ arbitrário (suporemos $\varepsilon<1$), na série

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots+\frac{2^{n-1}}{3^n}+\cdots,$$

que representa a soma dos comprimentos dos terços médios removidos de $[0,1]$, podemos escolher n bastante grande para que seja

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots+\frac{2^{n-1}}{3^n}>1-\frac{\varepsilon}{2}\cdot$$

Suprimidos os terços médios correspondentes aos n termos da série acima considerados, resta em $[0,1]$ um conjunto A, constituído por 2^n intervalos fechados, cujos comprimentos têm uma soma $< \frac{\varepsilon}{2}\cdot$ É claro que $C \subset A$. Cobrindo cada um desses 2^n intervalos com um intervalo de comprimento duplo, obtemos uma cobertura finita de A (e também de C) para a qual a soma dos comprimentos dos intervalos componentes é $< \varepsilon$. Isso prova que C tem medida nula. O conjunto de Cantor é exemplo de conjunto de medida nula e, ao mesmo tempo, de primeira categoria. Pode demonstrar-se que tal conjunto tem a potência do contínuo (a prova dessa asserção, que não tem interesse para este trabalho, pode ser encontrada em [1], cap. I, sec. 6).

Vamos apresentar um exemplo de conjunto de primeira categoria que não tem medida nula. Consideremos, novamente o intervalo fechado $[0,1]$, e tomemos uma série numérica de termos positivos (convergente), cuja soma seja < 1, por exemplo:

$$\frac{1}{3}+\frac{1}{6}+\frac{1}{12}+\cdots$$

(observe-se que a soma desta série é $\frac{2}{3}$). Suprimamos de $[0,1]$ o terço central aberto; de cada um dos dois intervalos fechados restantes suprimamos um intervalo aberto concêntrico de comprimento $\frac{1}{2}\cdot\frac{1}{6}$; de cada um dos quatro intervalos fechados remanescentes, removamos um intervalo aberto concêntrico de comprimento $\frac{1}{4}\cdot\frac{1}{12}$; e assim prossigamos indefinidamente. O conjunto K que resta de $[0,1]$ após a remoção de todos os intervalos abertos, da maneira descrita, é o exemplo procurado. Notemos que a soma dos comprimentos de todos os intervalos suprimidos é igual a $\frac{2}{3}$ (é a soma da série acima). O conjunto K é fechado, pois é o complementar de uma reunião de abertos, em relação a $[0,1]$. Representando por $\mathcal{C}K$ o complementar de K em relação ao espaço $[0,1]$, é fácil concluir que $\overline{\mathcal{C}K} = [0,1]$, pois todo ponto de K é ponto de acumulação do conjunto $\mathcal{C}K$; segue-se que K é um conjunto de interior vazio. Concluímos que K é um conjunto de primeira categoria em $[0,1]$. Provemos, agora, que K não tem medida nula. Se fosse K de medida nula, poderíamos cobrí-lo com uma coleção de intervalos cujos comprimentos tivessem uma soma $< \frac{1}{3} \cdot$ Essa coleção reunida com a de todos os intervalos abertos removidos de $[0,1]$, formaria uma cobertura de $[0,1]$ por intervalos cujos comprimentos teriam uma soma menor que a unidade, o que é absurdo.

Descreveremos, a seguir, um exemplo de conjunto de medida nula, mas que não é de primeira categoria. Para construí-lo, comecemos considerando o conjunto K do exemplo anterior. O complemento de K em relação a $[0,1]$ é uma coleção enumerável de intervalos abertos disjuntos, cujos comprimentos somam $\frac{2}{3}$ (lembrar que são justamente os intervalos removidos de $[0,1]$ na construção de K). Podemos pensar nesses intervalos considerando-os como lacunas deixadas pelo conjunto K no espaço $[0,1]$. Com cada um desses intervalos (suposto fechado para efeito da construção a ser executada), procedamos da mesma maneira que fizemos com $[0,1]$ no exemplo precedente, isto é, em outras palavras, substituamos cada uma das lacunas acima mencionadas por um conjunto do mesmo tipo de K. Seja K_1 a reunião dessa coleção enumerável de conjuntos análogos a K. O conjunto $K \cup K_1$ deixa em $[0,1]$ novas lacunas. Substituamos cada uma destas por um conjunto de tipo de K, construído sobre a mesma, e designemos por K_2 a reunião de tais conjuntos. Procedamos da mesma maneira com as lacunas deixadas em $[0,1]$ pelo conjunto $K \cup K_1 \cup K_2$, e assim por diante, indefinidamente. O conjunto $H = \mathcal{C}(K \cup K_1 \cup K_2 \cup \dots)$, onde $\mathcal{C}$ indica o complemento em relação a $[0,1]$, é um exemplo de conjunto de medida nula, mas que não é de primeira categoria. Provemos essa afirmação. Já vimos que a soma dos comprimentos dos intervalos que compõem $\mathcal{C}K$ é $\frac{2}{3}\cdot$ Os intervalos que compõem $\mathcal{C}(K \cup K_1)$ têm comprimentos cuja soma é evidentemente igual a $\frac{2}{3} \cdot \frac{2}{3} = \left(\frac{2}{3}\right)^2$. De modo geral, os intervalos que compõem o conjunto $\mathcal{C}(K \cup K_1 \cup K_2 \cup \dots \cup K_{n-1})$ têm por soma de seus comprimentos $\left(\frac{2}{3}\right)^n$. Dado $\varepsilon > 0$, arbitrariamente pequeno, podemos achar n_0 tal que $\left(\frac{2}{3}\right)^{n_0} < \varepsilon$. Os intervalos que compõem $\mathcal{C}(K \cup K_1 \cup \dots \cup K_{n_0-1})$, cuja coleção é enumerável, cobrem H, e a soma dos seus comprimentos é $< \varepsilon$. Logo, H é de medida nula. Por outro lado, como K é de primeira categoria, e como $K_1, K_2, \dots$ são reuniões enumeráveis de conjuntos do tipo de K, segue-se que $K \cup K_1 \cup K_2 \cup \dots$ é um conjunto de primeira categoria. Resulta, então, do teorema de Baire, que H, como complemento de um conjunto de primeira categoria no espaço métrico completo $[0,1]$, não pode ser de primeira categoria.

Agora, que julgamos ter esclarecido completamente os conceitos de medida nula e de categoria, vamos tratar de uma situação particularmente importante para o nosso trabalho, na qual a propriedade da medida nula implica a de primeira categoria. Sabemos, da Topologia Geral, que um espaço X se diz σ-compacto, quando é uma reunião enumerável de partes compactas. Um conjunto $C \subset X$ diz-se σ-compacto quando o subespaço C é σ-compacto. Todo espaço localmente compacto, com base enumerável, é σ-compacto, e é claro que todo conjunto fechado em tal espaço é σ-compacto. Segue-se que toda variedade topológica, e, em particular, toda variedade diferenciável, é σ-compacta, e todo conjunto fechado em uma variedade é σ-compacto. Um resultado que será de capital importância para o que vai seguir é o que demonstraremos sob forma da seguinte

Proposição. *Em uma variedade topológica M, todo conjunto σ-compacto de medida nula em M é também de primeira categoria em M.*

Demonstração. Seja $C \subset M$ um conjunto σ-compacto de medida nula em M. Podemos escrever $C = \bigcup_{i=1}^{\infty} K_i$, onde cada K_i é um compacto. Para concluir que C é de primeira categoria em M, basta provar que K_i tem interior vazio, qualquer que seja o índice i. Suponhamos que algum desses conjuntos, por exemplo K_j, tenha um ponto interior p. Existe, então, uma vizinhança $W \ni p$, tal que $W \subset K_j \subset C$. Segue-se que p é interior a C, o que é absurdo, porque C é de medida nula. Logo, C é de primeira categoria em M.

Voltemos, agora, ao nosso tema principal. O teorema de Dubovitsky enuncia-se da maneira que segue.

Teorema (de Dubovitsky). *Sejam M^m, N^n duas variedades, e $\phi\colon M \to N$ uma aplicação diferenciável. Os* *valores críticos de ϕ constituem um conjunto de primeira categoria em N. Consequentemente, o conjunto dos valores regulares de ϕ é denso em N.*

Demonstremos que, uma vez provado o teorema de Sard, o de Dubovitsky resulta como corolário. Se $S \subset M$ é o conjunto dos pontos críticos de ϕ, então $C = \phi(S) \subset N$ é o conjunto dos valores críticos de ϕ. Como S é fechado em M (pois o seu complementar $M - S$ é evidentemente aberto), S é σ-compacto, e podemos, por isso, representá-lo na forma $S = \bigcup_{i=1}^{\infty} K_i$, onde cada K_i é um compacto. Segue-se daí que $C = \phi(S) = \bigcup_{i=1}^{\infty} \phi(K_i)$ é também σ-compacto, pois os conjuntos $\phi(K_i)$, $i = 1, 2, \ldots$, são compactos. Em virtude da proposição que acima demonstramos, se C for de medida nula, será também de primeira categoria, e isso é precisamente o que desejávamos provar.

Daqui em diante, podemos considerar o teorema de Dubovitsky como uma segunda formulação do teorema de Sard, um tanto mais fraca que a primeira. Para muitas aplicações, podem ser empregadas indiferentemente as duas formulações. Isso ocorre, por exemplo, toda vez que a questão essencial é o fato de ser denso na variedade N o conjunto dos valores regulares da aplicação $\phi\colon M \to N$. Existem, todavia, aplicações para as quais o segundo enunciado do teorema de Sard é insuficiente, e nesses casos tem-se que apelar para o primeiro.

Tratemos, agora, de provar esse notável teorema. Por ser longa a demonstração, dividi-la-emos em partes, e, para que o leitor bem conheça o significado de cada parte, e a sua posição relativamente ao todo, passamos a explicar, em poucas palavras, o que adiante será feito com a devida minúcia.

Admitiremos, inicialmente, que M é uma variedade sem bordo, e mostraremos que é então possível reduzir a demonstração ao caso euclidiano, quer dizer, ao caso de aplicações diferenciáveis do tipo $\phi\colon \Omega \to R^n$, onde Ω é um conjunto aberto de R^m. (Note-se que este é justamente o caso original, estudado por A. Sard.)

A seguir, demonstraremos que é verdadeiro o teorema no caso euclidiano; a prova será dada em duas fases: na primeira, suporemos $m < n$; na segunda, $m \geq n$. A primeira fase, em razão de sua relativa simplicidade, corresponde ao que chamaremos "o caso elementar do teorema de Sard".

Provaremos, finalmente, que o teorema é verdadeiro também se M é uma variedade com bordo.

3 Redução ao caso euclidiano

Consideremos as variedades diferenciáveis M^m e N^n. Suponhamos M sem bordo, e seja $\phi\colon M \to N$ uma aplicação diferenciável. Para cada ponto $a \in M$, seja U_a uma vizinhança de a em M. As vizinhanças U_a cobrem M; como as variedades possuem a propriedade de Lindelöf, a cobertura $\{U_a\}$ admite uma subcobertura enumerável, que indicaremos mais simplesmente por $\{U_1, U_2, \dots\}$. Seja $C \subset M$ o conjunto dos pontos críticos de $\phi\colon M \to N$. É claro que $C \cap U_i = C_i$ $(i = 1, 2, \dots)$ é o conjunto dos pontos críticos da restrição $\phi\colon U_i \to N$. Mas,

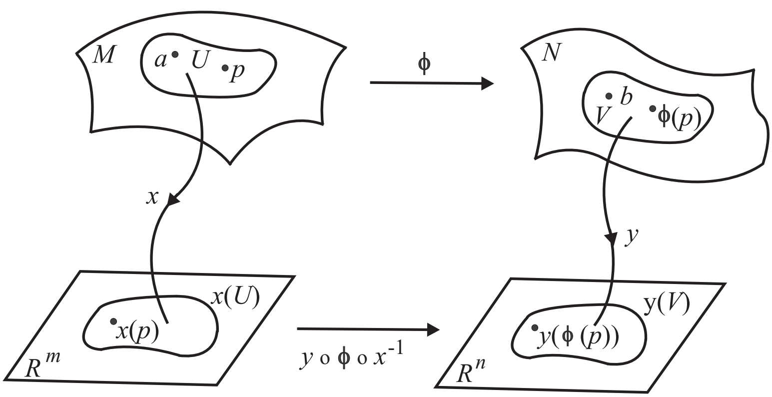
$$C = C \cap \left(\bigcup_{i=1}^{\infty} U_i\right) = \bigcup_{i=1}^{\infty}(C \cap U_i) = \bigcup_{i=1}^{\infty} C_i\,.$$

Logo:

$$\phi(C) = \bigcup_{i=1}^{\infty} \phi(C_i).$$

A expressão que acabamos de obter para o conjunto dos valores críticos de $\phi\colon M \to N$ mostra que se $\phi(C_i)$, $i = 1, 2, \ldots,$ for de medida nula em N, outro tanto ocorrerá com $\phi(C)$. Essa observação reduz a demonstração do teorema de Sard (global) a um caso local: basta provarmos que cada ponto $a \in M$ tem uma vizinhança U_a em M, tal que o teorema de Sard seja verdadeiro para a restrição $\phi\colon U_a \to N$.

Se $a \in M$ e se $\phi(a) = b$, podemos escolher sistemas de coordenadas locais $y\colon V \to R^n$ e $x\colon U \to R^m$, tais que $b \in V$, $a \in U$ e $\phi(U) \subset V$.



Por definição, $p \in U$ é ponto crítico de $\phi\colon U \to N$ se e só se $x(p)$ é ponto crítico da aplicação

$$y \circ \phi \circ x^{-1}\colon x(U) \to y(V).$$

Seja $Q \subset U$ o conjunto dos pontos críticos de $\phi\colon U \to N$. Já sabemos que o conjunto dos valores críticos dessa aplicação, $\phi(Q) \subset V$, é de medida nula em N se e somente se a sua imagem $y(\phi(Q)) \subset y(V)$ é de medida nula em R^n. Tudo se reduz, pois, a provar que o teorema de Sard é verdadeiro para a aplicação $y \circ \phi \circ x^{-1}\colon x(U) \to R^n$, onde $x(U)$ é um conjunto aberto de R^m (porque estamos supondo que M é variedade sem bordo).

Daqui em diante, para efeito de demonstração do teorema de Sard, poderemos restringir nossas considerações às aplicações diferenciáveis do tipo $\phi\colon \Omega \to R^n$, onde Ω é um aberto de R^m.

4 O caso elementar do teorema de Sard (caso $m < n$)

Sejam M^m, N^n duas variedades diferenciáveis, e suponhamos que $m < n$. Se $\phi\colon M \to N$ é uma aplicação diferenciável, é claro que todo $a \in M$ é ponto crítico de ϕ. O teorema de Sard corresponde, neste caso, à afirmação de que o conjunto $\phi(M)$ é de medida nula em N.

Para fins de demonstração, podemos restringir-nos, conforme já foi visto, ao caso de uma aplicação diferenciável $\phi\colon \Omega \to R^n$, onde Ω é um conjunto aberto de R^m. Supondo $m < n$, vamos provar que $\phi(\Omega)$ é de medida nula em R^n. Considerações triviais, nas quais se faz uso da propriedade de Lindelöf, reduzem a questão a provar que cada ponto $p \in \Omega$ tem uma vizinhança U em Ω, tal que $\phi(U)$ é de medida nula em R^n.

A nossa tarefa será facilitada pela possibilidade de escolher vizinhanças cúbicas. Seja U um cubo aberto de centro p e arestas paralelas aos eixos coordenados de R^m, tal que $\overline{U} \subset \Omega$. Basta demonstrarmos que $\phi(\overline{U})$ é de medida nula em R^n, pois daí, e da inclusão $\phi(U) \subset \phi(\overline{U})$, resultará que $\phi(U)$ é também de medida nula em R^n.

Tendo em vista que a função ϕ é diferenciável, e que o conjunto $\overline{U}$ é compacto, existe, como sabemos, um número $k > 0$, tal que

$$|\phi(q) - \phi(q')| < k|q - q'|$$

para todo par de pontos $q, q' \in \overline{U}$. Em outros termos, a aplicação ϕ, por ser diferenciável, é lipschitziana sobre o compacto $\overline{U}$; a demonstração desse fato pode ser vista em [6], cap. I.

Dividamos cada aresta do cubo $\overline{U}$ em r partes iguais, e consideremos os hiperplanos que passam pelos pontos de divisão, paralelos aos hiperplanos coordenados de R^m; eles decompõem $\overline{U}$ em r^m cubos iguais. Seja K_i um qualquer desses cubos (fechados). Se h é o comprimento da aresta de $\overline{U}$, o diâmetro de K_i é $\frac{h}{r}\sqrt{m}$. Por força da desigualdade acima, o diâmetro do conjunto $\phi(K_i)$ não pode exceder $\frac{hk}{r}\sqrt{m}$, e então é claro que $\phi(K_i) \subset H_i$, onde $H_i \subset R^n$ é um cubo (fechado) de arestas paralelas aos eixos coordenados de R^n, tendo $\frac{2hk}{r}\sqrt{m}$ como comprimento de sua aresta. O volume de H_i é $\left(\frac{2hk}{r}\sqrt{m}\right)^n = \frac{c}{r^n}$, onde pusemos $c = \left(2hk\sqrt{m}\right)^n$. Como $\overline{U}$ é a reunião dos r^m cubos K_i, $\phi(\overline{U})$ há de estar contido na reunião dos r^m cubos H_i, cujos volumes somam $r^m \cdot \frac{c}{r^n} = c\,r^{m-n}$. Em virtude da hipótese $m < n$, dado $\varepsilon > 0$ arbitrário, é possível escolher r bastante grande para que resulte $c\,r^{m-n} < \varepsilon$. Logo, $\phi(\overline{U})$ é um conjunto de medida nula em R^n, como queríamos demonstrar.

5 O caso $m \geq n$

Seja $\phi\colon \Omega \to R^n$ uma aplicação diferenciável cujo domínio Ω é um conjunto aberto de R^m, e suponhamos que $m \geq n$. Seja K um cubo de arestas paralelas aos eixos coordenados de R^m, aberto, tal que $\overline{K} \subset \Omega$. Todo ponto $p \in \Omega$ é centro de um cubo tal como K, e então, conforme vimos no n. 3, basta demonstrarmos o teorema de Sard para a aplicação $\phi\colon K \to R^n$ (restrição de ϕ ao cubo K).

Sejam $x^1, \ldots, x^m$ as coordenadas do ponto genérico $q \in K$, e $y^1, \ldots, y^n$ as coordenadas de $\phi(q) \in R^n$. A aplicação ϕ é descrita analiticamente por equações

$$y^i = \phi^i(x^1, \ldots, x^m), \qquad i = 1, \ldots, n.$$

Representemos a matriz jacobiana de ϕ, no ponto $p \in K$, assim:

$$D\phi(p) = \left(\frac{\partial \phi^i}{\partial x^j}(p)\right), \qquad i = 1, \ldots, n;\ j = 1, \ldots, m.$$

Seja C o conjunto dos pontos críticos de $\phi\colon K \to R^n$, isto é:

$$C = \{p \in K;\ D\phi(p) \text{ tem característica } < n\}.$$

Precisamos demonstrar que $\phi(C)$ é um conjunto de medida nula em R^n, e isso é tudo o que devemos fazer.

Seja $s = m - n + 1$; a hipótese $m \geq n$ implica $s > 0$. Fixemos, de uma vez por todas, uma das n funções reais ϕ^i, por exemplo a última, ϕ^n, e, para cada inteiro natural r, tal que $1 \leq r \leq s$, seja C_r o conjunto dos pontos de K nos quais se anulam todas as derivadas parciais de ϕ^n de ordens $1, \ldots, r$. Simbolicamente, podemos indicar esse conjunto assim:

$$C_r = \{p \in K;\, d\phi^n(p) = 0,\, d^2\phi^n(p) = 0, \ldots, d^r\phi^n(p) = 0\}.$$

Observemos que $C, C_1, \ldots, C_s$ são conjuntos limitados e fechados, e, portanto, compactos. É evidente que

$$C \supset C_1 \supset \cdots \supset C_s\,,$$
$$C = (C - C_1) \cup (C_1 - C_2) \cup \cdots \cup (C_{s-1} - C_s) \cup C_s\,,$$
$$\phi(C) = \phi(C - C_1) \cup \phi(C_1 - C_2) \cup \cdots \cup \phi(C_{s-1} - C_s) \cup \phi(C_s).$$

O nosso trabalho fica reduzido a provar que são de medida nula em R^n os conjuntos $\phi(C - C_1), \ldots, \phi(C_{s-1} - C_s), \phi(C_s)$. Daqui em diante, a demonstração será feita em três partes. Na primeira, mostraremos que $\phi(C_s)$ é de medida nula em R^n. Na segunda, usando o método de indução, provaremos que o mesmo se pode dizer de $\phi(C_1 - C_2), \ldots, \phi(C_{s-1} - C_s)$. Na terceira etapa, estudaremos em particular o conjunto $\phi(C - C_1)$; empregaremos de novo a indução, mas o argumento é diferente do que se aplica na segunda etapa.

1a. parte. Mostremos que $\phi(C_s)$ é de medida nula em R^n. Se $p \in C_s$, deve ser $d\phi^n(p) = 0$, $d^2\phi^n(p) = 0, \ldots, d^s\phi^n(p) = 0$, e então, no desenvolvimento tayloriano da função ϕ^n, em torno de p, não comparecem os termos de graus $1, \ldots, s$. Como $\overline{K}$ é compacto, podemos afirmar que existe um número $k > 0$, tal que

$$|\phi^n(q) - \phi^n(p)| < k|q-p|^{s+1},$$

qualquer que seja $q \in \overline{K}$.

Por outro lado, tendo ainda em vista a compacidade de $\overline{K}$, subsistem as $n-1$ seguintes desigualdades (lipschitzianas):

$$|\phi^j(q) - \phi^j(q')| < k|q-q'|, \qquad j = 1, \ldots, n-1,$$

para todo par de pontos $q, q' \in \overline{K}$. Na verdade, a cada índice j corresponde uma constante $k_j > 0$, porém é mais simples adotarmos uma só constante k que sirva para as n desigualdades, e é óbvio que isso se pode fazer sem mais comentários.

Dividamos cada uma das arestas do cubo $\overline{K}$ em r partes iguais, e, por meio de hiperplanos paralelos às faces, decomponhamos $\overline{K}$ em r^m cubos iguais (fechados). Se h é o comprimento da aresta de $\overline{K}$, o diâmetro de cada um desses cubos é $\frac{h}{r}\sqrt{m}$. Sejam $L_1, \ldots, L_t$ os cubos da decomposiçao que encontram o conjunto C_s (quer dizer que $L_j \cap C_s \neq \varnothing$, $j = 1, \ldots, t$); é claro que $C_s \subset \bigcup_{j=1}^{t} L_j$ e que $t \leq r^m$. Cada cubo L_j contém algum ponto de C_s, e então, em face das n desigualdades acima estabelecidas, podemos escrever: $\phi(L_j) \subset H_j$, onde $H_j \subset R^n$ é um paralelepípedo de arestas paralelas aos eixos coordenados, no qual uma das arestas mede $2k\left(\frac{h}{r}\sqrt{m}\right)^{s+1}$ e cada uma das $n-1$ arestas restantes mede $2k\,\frac{h}{r}\sqrt{m}$. Calculemos o volume desse paralelepípedo:

$$\begin{aligned}\text{vol } H_j &= \left(2k\frac{h}{r}\sqrt{m}\right)^{n-1} \cdot 2k\left(\frac{h}{r}\sqrt{m}\right)^{s+1} = \\ &= 2^n k^n \left(\frac{h}{r}\sqrt{m}\right)^{n+s} = \frac{c}{r^{n+s}}\,,\end{aligned}$$

onde $c = 2^n k^n \big(h\sqrt{m}\big)^{n+s}$. De tudo isso resulta:

$$\phi(C_s) \subset \phi\left(\bigcup_{j=1}^{t} L_j\right) = \bigcup_{j=1}^{t} \phi(L_j) \subset \bigcup_{j=1}^{t} H_j$$

e

$$\text{vol}\left(\bigcup_{j=1}^{t} H_j\right) \leq \frac{tc}{r^{n+s}} \leq r^m \cdot \frac{c}{r^{n+s}} = \frac{c}{r^{n+s-m}} = \frac{c}{r}\,\cdot$$

Ora, dado $\varepsilon > 0$ arbitrário, podemos escolher r bastante grande para que seja $\frac{c}{r} < \varepsilon$. Isso prova que $\phi(C_s)$ é de medida nula, como queríamos demonstrar.

Decorre imediatamente do resultado ao qual chegamos, que o teorema de Sard é verdadeiro no caso $m = 1$. Com efeito, se $m = 1$, tem-se também $n = 1$, já que $m \geq n$; segue-se que $s = m - n + 1 = 1$, e então $C_s = C_1$. Mas, no presente caso, $C_1 = C$, pois a aplicação ϕ, em termos de coordenadas, exprime-se por $y = \phi(x)$, e a matriz jacobiana de ϕ reduz-se a $\frac{\partial\phi}{\partial x}$; um ponto $p \in K$ é crítico se e só se $\frac{\partial\phi}{\partial x}(p) = 0$.

Essa observação nos dá a base para o emprego do método de indução, sobre o qual será fundada a parte seguinte da demonstração. Vamos supor verdadeiro o teorema de

Sard para aplicações definidas em variedades de dimensão $< m$, e provar que ele é verdadeiro para a aplicação $\phi\colon K \to R^n$ que vem sendo objeto do nosso estudo (recorde-se que K é um aberto de R^m).

2a. parte. Provemos que são de medida nula em R^n os conjuntos $\phi(C_1 - C_2), \ldots, \phi(C_{s-1} - C_s)$. Tomemos um qualquer deles, $\phi(C_r - C_{r+1})$, onde $1 \leq r \leq s-1$. Consideremos um ponto $p \in C_r - C_{r+1}$. Como $p \notin C_{r+1}$, pelo menos uma derivada parcial de ordem $r+1$ da função ϕ^n não se anula em p. Existe, pois, uma vizinhança U_p do ponto p, $U_p \subset K$, na qual a mesma derivada se conserva $\neq 0$. Quando p descreve $C_r - C_{r+1}$, a coleção de vizinhanças $\{U_p\}$ constitui uma cobertura do conjunto $C_r - C_{r+1}$, a qual admite uma subcobertura enumerável, que preferiremos indicar por $\{U_1, U_2, \ldots\}$. É claro que

$$C_r - C_{r+1} = \bigcup_{i=1}^{\infty} [U_i \cap (C_r - C_{r+1})],$$

e daí resulta:

$$\phi(C_r - C_{r+1}) = \bigcup_{i=1}^{\infty} \phi[U_i \cap (C_r - C_{r+1})].$$

Basta demonstrarmos que, para cada índice i, o conjunto $\phi[U_i \cap (C_r - C_{r+1})]$ é de medida nula em R^n.

Tomemos, pois, uma qualquer das vizinhanças U_i, que passamos a indicar simplesmente por U. Em U, como vimos, uma bem determinada derivada de ordem $r+1$ da função ϕ^n é $\neq 0$; suponhamos que se trate de $\frac{\partial \omega}{\partial x^i}$, onde ω é uma derivada de ordem r de ϕ^n, que não precisamos especificar. Observemos que ω é uma função real diferenciável, definida no aberto $U \subset R^m$, tal que $\frac{\partial\omega}{\partial x^i} \neq 0$ em todo ponto de U. Sabemos que o conjunto

$$\omega^{-1}(0) = \{p \in U; \omega(p) = 0\}$$

é uma superfície regular de dimensão $m-1$ contida em U (e, pois, uma variedade diferenciável de dimensão $m - 1$). Notemos que $U \cap (C_r - C_{r+1}) \subset \omega^{-1}(0)$. Com um raciocínio trivial, vemos, ainda, que todo $p \in U \cap (C_r - C_{r+1})$, por ser ponto crítico da aplicação $\phi\colon K \to R^n$, é também ponto crítico da restrição $\phi\colon \omega^{-1}(0) \to R^n$. Logo, $\phi[U \cap (C_r - C_{r+1})]$ está contido no conjunto dos valores críticos de $\phi\colon \omega^{-1}(0) \to R^n$, o qual é de medida nula em R^n, em virtude da hipótese de indução.

3a. parte. Demonstremos, finalmente, que o conjunto $\phi(C - C_1)$ é de medida nula em R^n. Se $p \in C - C_1$, pelo menos uma das derivadas parciais primeiras da função ϕ^n não se anula em p; suponhamos, para fixar idéias, que seja $\frac{\partial\phi^n}{\partial x^m}(p) \neq 0$. Consideremos as equações

$$z^1 = x^1, \ldots, z^{m-1} = x^{m-1},\ z^m = \phi^n(x^1, \ldots, x^m). \qquad (*)$$

O jacobiano do sistema que elas compõem, calculado no ponto p, é:

$$\det\left(\frac{\partial z^i}{\partial x^j}(p)\right) = \frac{\partial\phi^n}{\partial x^m}(p) \neq 0.$$

significa isso que as equações (*) estabelecem uma mudança de coordenadas locais $(x^1, \ldots, x^m) \to (z^1, \ldots, z^m)$, válida numa vizinhança V_p de p, $V_p \subset K$. Como C_1 é um conjunto fechado, e $p \notin C_1$, podemos supor V_p suficientemente pequena para que seja $V_p \cap C_1 = \varnothing$. Seja W_p uma vizinhança de p, tal que $\overline{W}_p \subset V_p$.

Se construirmos, para cada ponto $p \in C - C_1$, a vizinhança W_p nas condições que acabamos de descrever, obteremos uma cobertura $\{W_p\}$ do conjunto $C - C_1$, a qual admite uma subcobertura enumerável, que indicaremos por $\{W_1, W_2, \dots\}$. Podemos escrever:

$$C - C_1 = (C - C_1) \cap \left(\bigcup_{i=1}^{\infty} \overline{W}_i\right) = \bigcup_{i=1}^{\infty} [(C - C_1) \cap \overline{W}_i] = \\ = \bigcup_{i=1}^{\infty} (C \cap \overline{W}_i).$$

Resulta daí que

$$\phi(C - C_1) = \bigcup_{i=1}^{\infty} \phi(C \cap \overline{W}_i).$$

Basta, agora, provarmos que $\phi(C \cap \overline{W}_i)$ é de medida nula em R^n, qualquer que seja i.

A partir deste momento, podemos limitar nossas considerações a um (qualquer) dos conjuntos $C \cap \overline{W}_i$; por isso, deixemos de lado o índice i, e escrevamos W em lugar de W_i. Suponhamos que $p \in C - C_1$ seja o ponto ao qual corresponde a dita vizinhança W, e recordemos que V é uma vizinhança de p, tal que $\overline{W} \subset V$, na qual valem as coordenadas $z^1, \dots, z^m$. Em vista das equações (*), a aplicação

ϕ, restrita a V, pode expressar-se assim:

$$
\begin{aligned}
&y^1 = \phi^1(x^1, \ldots, x^m) = \psi^1(z^1, \ldots, z^m),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&y^{n-1} = \phi^{n-1}(x^1, \ldots, x^m) = \psi^{n-1}(z^1, \ldots, z^m),\\
&y^n = \phi^n(x^1, \ldots, x^m) = z^m.
\end{aligned}
\tag{**}
$$

O conjunto V, como aberto de R^m, possui estrutura de variedade diferenciável de dimensão m. Seja H o conjunto dos pontos críticos da aplicação $\phi\colon V \to R^n$ (restrição de ϕ a V), representada pelas equações (**). A propriedade de ser crítico um ponto, independe do sistema de coordenadas; logo:

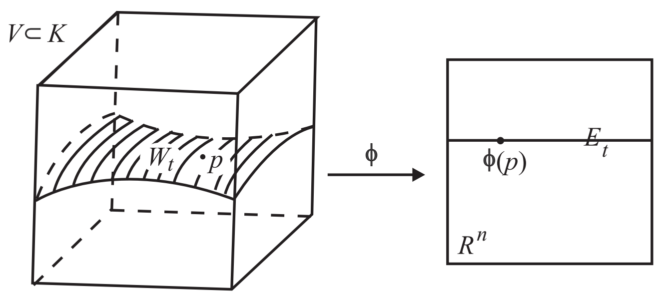
$$H = C \cap V.$$

Definamos os conjuntos W_t da seguinte maneira:

$$W_t = \{q = (z^1, \ldots, z^m) \in V;\ z^m = t\}.$$

Vemos imediatamente que W_t (suposto não vazio) é uma superfície regular de dimensão $m-1$ contida em V; basta observar que W_t é a imagem inversa do valor $t \in R$ pela função $F\colon V \to R$, tal que $F(z^1, \ldots, z^m) = z^m$, e notar que todo valor dessa função é regular.

Seja E_t o hiperplano de R^n cuja equação é $y^n = t$.



Em vista das equações (**) e da definição de W_t, podemos escrever:

$$\phi(W_t) \subset E_t\,,$$

quer dizer: a restrição de ϕ à superfície W_t é uma função que toma valores no hiperplano E_t, e pode, pois, ser indicada por $\phi\colon W_t \to E_t$ (W_t e E_t são variedades diferenciáveis). Se $q \in W_t$, q tem coordenadas $z^1, \ldots, z^{m-1}, t$ em V, e $\phi(q)$ tem coordenadas $y^1, \ldots, y^{n-1}, t$ em R^n; podemos adotar $z^1, \ldots, z^{m-1}$ como coordenadas de q em W_t e, analogamente, $y^1, \ldots, y^{n-1}$ como coordenadas de $\phi(q)$ em E_t. Nessas condições, se H_t é o conjunto dos pontos críticos de $\phi\colon W_t \to E_t$, é fácil ver que

$$H_t = H \cap W_t\,;$$

para se chegar à evidência desse resultado, basta examinar as matrizes jacobianas das aplicações $\phi\colon V \to R^n$ e $\phi\colon W_t \to E_t$, que são respectivamente:

$$\begin{bmatrix} \dfrac{\partial\psi^1}{\partial z^1} \cdots \dfrac{\partial\psi^1}{\partial z^{m-1}} & \dfrac{\partial\psi^1}{\partial z^m} \\ \cdots\cdots\cdots\cdots\cdots & \\ \dfrac{\partial\psi^{n-1}}{\partial z^1} \cdots \dfrac{\partial\psi^{n-1}}{\partial z^{m-1}} & \dfrac{\partial\psi^{n-1}}{\partial z^m} \\ 0 \quad \ldots \quad 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} \dfrac{\partial\psi^1}{\partial z^1} \cdots \dfrac{\partial\psi^1}{\partial z^{m-1}} \\ \cdots\cdots\cdots\cdots \\ \dfrac{\partial\psi^{n-1}}{\partial z^1} \cdots \dfrac{\partial\psi^{n-1}}{\partial z^{m-1}} \end{bmatrix}.$$

A segunda matriz é a submatriz da primeira, formada pelas $n-1$ primeiras linhas e pelas $m-1$ primeiras colunas. É claro que a primeira matriz tem característica $< n$, num ponto $q \in W_t$, se e somente se a segunda matriz tem característica $< n - 1$ nesse mesmo ponto.

Como W_t é uma variedade de dimensão $m-1$, $\phi(H_t)$, conjunto dos valores críticos de $\phi\colon W_t \to E_t$, é de medida nula em E_t, em virtude da hipótese de indução.

Mostremos, enfim, que $\phi(C \cap \overline{W})$ é um conjunto de medida nula em R^n. Para esse fim, provemos, primeiramente, que

$$E_t \cap \phi(C \cap \overline{W}) \subset \phi(H_t).$$

De fato, observemos que sendo $C \cap \overline{W} \subset C \cap V = H$, resulta:

$$E_t \cap \phi(C \cap \overline{W}) \subset E_t \cap \phi(H).$$

Por outro lado, é fácil ver que

$$E_t \cap \phi(H) = \phi(H \cap W_t) = \phi(H_t).$$

Realmente, se $q \in E_t \cap \phi(H)$, existe $p \in H \subset V$, tal que $\phi(p) = q$; como $q \in E_t$, a última coordenada de q em R^n é $y^n = t$; mas $y^n = z^m$, quer dizer, a última coordenada de p em V é $z^m = t$; logo, $p \in W_t$. Concluímos que $p \in H \cap W_t = H_t$, e então, $q \in \phi(H_t)$. Vice-versa, se $q \in \phi(H \cap W_t)$, então $q = \phi(p)$, onde $p \in H$ e $p \in W_t$. Logo, $q \in \phi(H)$ e $q \in \phi(W_t) \subset E_t$, e, finalmente, $q \in E_t \cap \phi(H)$. É, pois, verdadeira a inclusão $E_t \cap \phi(C \cap \overline{W}) \subset \phi(H_t)$.

Como parte de um conjunto de medida nula em E_t, o conjunto $E_t \cap \phi(C \cap \overline{W})$ é também de medida nula em E_t, e isto, qualquer que seja t. A conclusão final de que $\phi(C \cap \overline{W})$ é de medida nula em R^n, é agora uma consequência imediata do conhecido teorema de Fubini, pois $\phi(C \cap \overline{W})$ é compacto, e, portanto, mensurável.

6 Extensão ao caso em que M é uma variedade com bordo

Acabamos de demonstrar que o teorema de Sard é verdadeiro para toda aplicação diferenciável $\phi\colon M^m \to N^n$, onde as dimensões m, n são quaisquer (≥ 1), N é uma variedade diferenciável arbitrária, e M é uma variedade diferenciável sem bordo. Esta última restrição foi imposta, como vimos, a fim de que pudéssemos reduzir a demonstração ao caso das aplicações do tipo $\phi\colon \Omega \to R^n$, onde Ω é um conjunto aberto de R^m. É fácil provar, agora, que tal restrição sobre M pode ser eliminada.

Se M^m é uma variedade diferenciável com bordo, o seu bordo ∂M é uma variedade diferenciável de dimensão $m-1$, sem bordo. Seja $C \subset M$ o conjunto dos pontos críticos de $\phi\colon M \to N$, e seja $C' \subset \partial M$ o conjunto dos pontos críticos da restrição $\phi\colon \partial M \to N$. Verifica-se imediatamente que

$$C \cap \partial M \subset C'.$$

Por outro lado, é óbvio que

$$C = (C \cap \partial M) \cup (C - \partial M).$$

Podemos então escrever:

$$C \subset C' \cup (C - \partial M),$$

$$\phi(C) \subset \phi(C') \cup \phi(C - \partial M).$$

Observemos que $\phi(C')$ é o conjunto dos valores críticos de $\phi\colon \partial M \to N$, e $\phi(C - \partial M)$ é o conjunto dos valores críticos de $\phi\colon M - \partial M \to N$. Ora, ∂M e $M - \partial M$ são variedades diferenciáveis sem bordo; de acordo com a parte já demonstrada, podemos afirmar que $\phi(C')$ e $\phi(C - \partial M)$ são conjuntos de medida nula em N; o mesmo podemos dizer da reunião desses dois conjuntos, e também de $\phi(C)$, em face da última inclusão.

Capítulo III

Aplicações do Teorema de Sard

O teorema de Sard é um dos resultados importantes da teoria das variedades diferenciáveis, e tem aplicações notáveis, algumas das quais constituem o objeto deste capítulo. Não temos a intenção de descrever muitas dessas aplicações, nem a pretensão de afirmar que escolhemos as mais interessantes. Observe-se, a propósito, que um dos resultados de máximo relevo naquela teoria é o conhecido teorema de imersão de Whitney: toda variedade diferenciável compacta[(*)] de dimensão n é difeomorfa a uma superfície regular do espaço euclidiano R^{2n+1}. A demonstração desse fato baseia-se no teorema de Sard, usado em seu caso elementar. Tal aplicação, embora tão fundamental, não é descrita neste trabalho, pois já aparece em certas obras ao

[(*)] Na realidade, o teorema de Whitney estende-se a variedades não compactas, mas o caso compacto é o único que precisa ser mencionado no presente trabalho.

alcance do leitor, inclusive na literatura matemática brasileira (v. [6], cap. III, onde se encontra uma demonstração bem clara do teorema de Whitney). Outras aplicações importantes do teorema de Sard relacionam-se com o estudo da posição geral de duas variedades imersas num espaço euclidiano, e também com o das auto-interseções a que dá origem uma aplicação regular $\phi\colon M^n \to R^{2n}$. Esses e outros problemas podem ser vistos em [12], cap. I.

As aplicações às quais dedicamos as páginas seguintes são aquelas que ainda não vimos descritas, e que, pela utilidade e interesse das questões envolvidas, justificam plenamente, a nosso ver, a elaboração do presente trabalho, no qual encontram a merecida divulgação. Tais aplicações ocorrem em Análise, Topologia e Geometria, e são aqui tratadas sob forma de estudos.

1 Estudo da dependência funcional

Apresentamos, nas linhas que seguem, uma conceituação precisa da dependência funcional de funções reais diferenciáveis, definidas num conjunto aberto $\Omega \subset R^n$. Trata-se de assunto interessante, mas que nem sempre vem exposto de modo satisfatório nos livros de Análise Matemática. Em alguns escritores, as demonstrações falham por omissões nas hipóteses (não descrevem claramente as condições a que deve satisfazer a função que estabelece a dependência). Em outros, aparecem somente resultados de caráter local. Há também autores que fazem um estudo irrepreensível da questão, como é o caso de Rudin (v. [14], pag.

182), mas estes quase sempre obtêm uma relação funcional que é apenas contínua.

O estudo que desenvolvemos apresenta as seguintes vantagens: 1) tem caráter global; 2) a dependência funcional é estabelecida por uma função diferenciável (de classe C^∞); 3) a demonstração do teorema final é simplificada, graças ao emprego do teorema de Sard. Aliás, quanto a esta última vantagem, cumpre observar que o teorema de Sard é essencial no presente estudo, e que os autores que discorrem sobre essa matéria sem mencioná-lo explicitamente, outra coisa não fazem, no fundo, senão intercalar no texto a demonstração desse teorema (em situações particulares), ou de asserções que dele decorrem (v., p. ex., [14], loc. cit.).

Comecemos nossas considerações com a recordação de um resultado trivial: qualquer fechado $A \subset R^n$ é, como se sabe, o conjunto dos zeros de uma função real contínua $\phi\colon R^n \to R$; basta definir ϕ como a função distância $\phi(x) = \operatorname{dist}(x, A)$; é claro que ϕ é contínua, e que $\phi(x) = 0$ se e só se $x \in A$.

Como no presente trabalho estamos lidando com funções diferenciáveis (de classe C^∞), e vamos exigir que a dependência funcional seja estabelecida por uma função também de classe C^∞, necessitaremos, adiante, de um resultado algo mais forte do que esse que acabamos de recordar. Teremos necessidade de considerar um compacto qualquer $K \subset R^n$ como conjunto dos zeros de uma função diferenciável $\phi\colon R^n \to R$. Esse fato, que pode ser de utilidade também em outras questões, vai ser aqui demonstrado sob forma do seguinte

Lema. *Se $K \subset R^n$ é um conjunto compacto, existe uma função diferenciável $\phi\colon R^n \to R$, tal que $\phi(x) = 0$ se e somente se $x \in K$.*

Demonstração. Vamos provar, inicialmente, que se V é um aberto de R^n tal que $K \subset V$, existe uma função diferenciável $f\colon R^n \to R$ com a seguinte propriedade: $f(x) = 0$ se $x \in K$, e $f(x) = 1$ se $x \in R^n - V$.

Para cada $x \in K$, podemos considerar cubos abertos U_x, W_x, de centro x e arestas paralelas aos eixos coordenados de R^n, tais que $\overline{W}_x \subset U_x \subset V$. Por causa da compacidade de K, um número finito de cubos da família $\{W_x\}$ cobrem K; indiquemos tais cubos com $W_1, \ldots, W_s$, e sejam $U_1, \ldots, U_s$ os cubos da família $\{U_x\}$ que lhes correspondem. Para cada índice $i \in \{1, \ldots, s\}$, existe uma função diferenciável $g_i\colon R^n \to R$, tomando valores no intervalo $[0, 1]$, tal que $g_i = 0$ sobre $\overline{W}_i$, $0 < g_i < 1$ sobre $U_i - \overline{W}_i$, $g_i = 1$ sobre $R^n - U_i$. A demonstração da existência das funções g_i pode ser vista, com ligeira modificação, em [6], cap. III.

A função $f\colon R^n \to R$, que queríamos obter, pode ser definida assim:

$$f(x) = g_1(x) \ldots g_s(x).$$

Verifica-se, imediatamente, que essa função f satisfaz às condições acima impostas.

Observemos que f é uma função limitada, pois qualquer que seja $x \in R^n$, tem-se $0 \le f(x) \le 1$. Se supusermos que o aberto V seja limitado, o que é possível, porque K, sendo compacto, é limitado, resultará que as derivadas de f, de todas as ordens, serão também limitadas. Com efeito, por ser f constante (igual a 1) no exterior do compacto $\overline{V}$, essas derivadas são nulas no exterior de $\overline{V}$, e, sendo elas funções contínuas, são limitadas no compacto $\overline{V}$.

O conjunto compacto $K \subset R^n$ é uma interseção enumerável de abertos de R^n. Podemos, por exemplo, considerar, para cada inteiro positivo p, o conjunto

$$V_p = \left\{ x \in R^n ; \, \mathrm{dist}(x, K) < \frac{1}{p} \right\};$$

é claro que $K = \bigcap\limits_{p=1}^{\infty} V_p$.

De acordo com o que acabamos de ver, a cada aberto V_p podemos associar uma função $f_p \colon R^n \to R$, de classe C^∞, tal que $f_p = 0$ sobre K, e $f_p = 1$ sobre $R^n - V_p$. Pode ser que a função f_p se anule em pontos situados fora de K (pelo menos, nada nos garante o contrário), mas se $y \notin K$, existe algum inteiro positivo r tal que $y \notin V_r$, e então $f_r(y) = 1$.

Tomemos constantes positivas $c_1, c_2, \ldots,$ e formemos a série de funções

$$\phi = \sum_{p=1}^{\infty} c_p \, f_p \, ,$$

onde ϕ representa a função para a qual estamos supondo, momentaneamente, que essa série converge. Se $x \in K$, temos $f_p(x) = 0$ para todo p, e então é claro que $\phi(x) = 0$. Se $x \notin K$, já vimos que existe algum índice p tal que $f_p(x) = 1$, e como a série só tem termos positivos, é claro que $\phi(x) \neq 0$.

O nosso lema estará, pois, completamente demonstrado, se conseguirmos provar que a função ϕ, soma da série considerada, é definida em R^n e diferenciável (classe C^∞).

Recordemo-nos de que para todo p, f_p é uma função limitada em R^n, bem como as suas derivadas de todas as ordens.

Seja $M_{op} > 0$ um número tal que $|f_p| < M_{op}$, sobre R^n. Escolhamos, para cada valor de p, uma constante c_{op} que satisfaça à condição

$$0 < c_{op} < \frac{1}{2^p\, M_{op}}\,.$$

Segue-se que $|c_{op}\, f_p| < \frac{1}{2^p}$, e daí resulta que a série $\sum\limits_{p=1}^{\infty} c_{op}\, f_p$ converge uniformemente, em R^n, para uma função $\psi_0 \colon R^n \to R$.

Seja, agora, $M_{1p} > 0$ um número tal que sobre R^n se tenha

$$\max\left\{\left|\frac{\partial f_p}{\partial x^1}\right|, \ldots, \left|\frac{\partial f_p}{\partial x^n}\right|\right\} < M_{1p}\,.$$

Para condensar um pouco a escrita, ponhamos

$$df_p = \left(\frac{\partial f_p}{\partial x^1}, \ldots, \frac{\partial f_p}{\partial x^n}\right),$$

isto é, suponhamos que o símbolo df_p represente o vetor de R^n cujas componentes são as derivadas primeiras da função f_p. Escolhamos, para cada p, uma constante c_{1p} tal que

$$0 < c_{1p} < \frac{1}{2^p\, M_{1p}}\,,$$

e consideremos a série

$$\sum_{p=1}^{\infty} c_{1p}\, df_p = \left(\sum_p c_{1p}\,\frac{\partial f_p}{\partial x^1}, \ldots, \sum_p c_{1p}\,\frac{\partial f_p}{\partial x^n}\right).$$

Como $\left|c_{1p}\,\frac{\partial f_p}{\partial x^i}\right| < \frac{1}{2^p}$, $i = 1, \dots, n$, segue-se que a série $\sum\limits_{p=1}^{\infty} c_{1p}\, df_p$ converge uniformemente, em R^n, para um vetor $(\psi_1^1, \dots, \psi_1^n)$, onde ψ_1^i é a função para a qual converge a série $\sum\limits_{p} c_{1p}\,\frac{\partial f_p}{\partial x^i}$.

Vamos supor, também, que as constantes c_{1p} satisfaçam à desigualdade

$$c_{1p} < c_{op}\,;$$

a conveniência dessa suposição aparecerá brevemente.

Para trabalhar com as derivadas segundas das funções f_p, ponhamos

$$d^2 f_p = \left(\frac{\partial^2 f_p}{\partial x^1 \partial x^1}, \frac{\partial^2 f_p}{\partial x^1 \partial x^2}, \dots, \frac{\partial^2 f_p}{\partial x^1 \partial x^n}, \frac{\partial^2 f_p}{\partial x^2 \partial x^1}, \dots, \frac{\partial^2 f_p}{\partial x^n \partial x^n}\right),$$

quer dizer, $d^2 f_p$ indica um vetor de R^{n^2} cujas componentes são as derivadas segundas da função f_p. Seja $M_{2p} > 0$ um número tal que

$$\max\left\{\left|\frac{\partial^2 f_p}{\partial x^1 \partial x^1}\right|, \left|\frac{\partial^2 f_p}{\partial x^1 \partial x^2}\right|, \dots, \left|\frac{\partial^2 f_p}{\partial x^n \partial x^n}\right|\right\} < M_{2p}\,.$$

Para cada valor de p, escolhamos uma constante c_{2p} que satisfaça à desigualdade

$$0 < c_{2p} < \frac{1}{2^p}\, M_{2p}\,.$$

Nessas condições, é claro que a série $\sum\limits_{p=1}^{\infty} c_{2p}\, d^2 f_p$ converge uniformemente, em R^n, para um vetor $(\psi_2^{11}, \psi_2^{12}, \dots, \psi_2^{nn})$, onde ψ_2^{ij} é a função para a qual converge a série

$$\sum_p c_{2p} \frac{\partial^2 f_p}{\partial x^i \partial x^j} .$$

Suporemos, ainda, que as constantes c_{2p} verificam a desigualdade

$$c_{2p} < c_{1p} .$$

De modo geral, construamos a série $\sum\limits_{p=1}^{\infty} c_{kp}\, d^k f_p$, onde as constantes c_{kp} são tais que

$$0 < c_{kp} < \frac{1}{2^p M_{kp}} ,$$

M_{kp} e $d^k f_p$ tendo significado óbvio. Essa série converge uniformemente, em R^n, para um vetor de R^{n^k} cujas componentes são n^k funções.

Admitiremos que as constantes c_{kp} satisfazem à desigualdade

$$c_{kp} < c_{(k-1)p} .$$

Uma vez definidas as constantes c_{kp}, da maneira que acabamos de ver, ponhamos

$$c_p = c_{pp} , \qquad p = 1, 2, \dots$$

e construamos a série $\sum\limits_{p=1}^{\infty} c_p\, f_p$. Podemos afirmar que ela converge uniformemente, em R^n, para uma função $\phi \colon R^n \to R$; isso resulta imediatamente das desigualdades $c_{pp} < c_{op}$ $(p = 1, 2, \dots)$, aliadas ao fato, já estabelecido, de que a série $\sum_p c_{op} f_p$ é uniformemente convergente em R^n. Escrevamos, pois:

$$\phi = \sum_{p=1}^{\infty} c_p f_p .$$

Por outro lado, examinemos a série $\sum_{p=1}^{\infty} c_p d^k f_p$. Qualquer que seja o valor de k ($k = 1, 2, \dots$), ela converge uniformemente em R^n. Com efeito, $c_p = c_{pp} < c_{kp}$, desde que seja $p > k$; quer isso dizer que a partir do termo de ordem $k+1$, a série acima é majorada pela série $\sum_{p=1}^{\infty} c_{kp} d^k f_p$, a qual converge uniformemente em R^n, como já vimos.

Ora, sabemos que uma série uniformemente convergente pode ser derivada termo a termo justamente quando a série formada pelas derivadas dos termos da primeira é também uniformemente convergente. Encontramo-nos aqui exatamente nessa situação, e podemos, portanto, afirmar que

$$d^k \phi = \sum_{p=1}^{\infty} c_p d^k f_p .$$

Isso mostra que a função $\phi \colon R^n \to R$ é de classe C^∞, o que conclui a demonstração do lema.

De posse desse lema, estamos em condições de passar propriamente ao estudo da dependência funcional.

Sejam $f^1, \dots, f^n \colon \Omega \to R$ funções reais diferenciáveis, definidas num aberto $\Omega \subset R^n$. Diremos que essas n funções são *funcionalmente dependentes* sobre um conjunto $K \subset \Omega$, se e somente se existe uma função diferenciável $\phi \colon R^n \to R$, que não se anula identicamente em nenhum aberto de R^n, e tal que $\phi(f^1(x), \ldots, f^n(x)) = 0$ para todo $x \in K$.

Seja $F \colon \Omega \to R^n$ a aplicação cujas componentes são as funções $f^1, \ldots, f^n$, isto é:

$$F(x) = (f^1(x), \ldots, f^n(x)).$$

Podemos dizer: as funções $f^1, \ldots, f^n$ são funcionalmente dependentes sobre $K \subset \Omega$, se e somente se existe $\phi \colon R^n \to R$, diferenciável, a qual não se anula identicamente em nenhum aberto de R^n, e tal que $\phi(F(x)) = 0$ para todo $x \in K$.

De acordo com essa definição, se $f^1, \ldots, f^n$ são funcionalmente dependentes sobre K, existe a função ϕ nas condições descritas, e então $F(K) \subset \phi^{-1}(0)$. Como $\phi^{-1}(0)$ não pode conter nenhum aberto de R^n, concluímos: uma condição necessária para a dependência funcional das funções $f^1, \ldots, f^n$ sobre K, é que $F(K)$ seja um conjunto de interior vazio em R^n.

Não podemos garantir que essa condição seja suficiente, salvo se $F(K)$ é um conjunto compacto, o que ocorre certamente quando K é compacto. Com efeito, é verdadeiro o seguinte teorema, que encerra tudo o que existe de essencial no estudo da dependência funcional.

Teorema. *Seja $F \colon \Omega \to R^n$ uma aplicação diferenciável, definida num aberto $\Omega \subset R^n$, e seja $K \subset \Omega$ um conjunto compacto. As funções $f^1, \ldots, f^n \colon \Omega \to R$, componentes de F, são funcionalmente dependentes sobre K, se e somente se $F(K)$ é um conjunto de interior vazio em R^n.*

Demonstração. A necessidade da condição já foi estabelecida. Provemos a suficiência. Da compacidade de K

e da continuidade de F, resulta que $F(K)$ é compacto. De acordo com o lema previamente estudado, existe uma função $\phi\colon R^n \to R$, diferenciável, a qual se anula exclusivamente sobre $F(K)$. Se $F(K)$ tem interior vazio, a ϕ não pode anular-se em nenhum aberto de R^n. Segue-se daí que as funções $f^1, \ldots, f^n$, componentes de F, são funcionalmente dependentes sobre K.

O teorema que acabamos de demonstrar dá um critério de dependência funcional que, embora seja teoricamente muito simples, não é o resultado final aonde queremos chegar. O teorema final é o que dá a condição de dependência funcional das funções componentes de $F\colon \Omega \to R^n$ em termos do determinante jacobiano de F. A demonstração torna-se extraordinariamente elegante, graças ao emprego do teorema de Sard, como passamos a mostrar.

Teorema. *Seja $F\colon \Omega \to R^n$ uma aplicação diferenciável cujo domínio Ω é um aberto de R^n. Para que as funções $f^1, \ldots, f^n\colon \Omega \to R$, componentes de F, sejam funcionalmente dependentes sobre todo conjunto compacto $K \subset \Omega$, é necessário e suficiente que o determinante jacobiano de F se anule identicamente sobre Ω.*

Demonstração. Representemos por $J(x)$ o jacobiano de F no ponto $x \in \Omega$. A necessidade da condição é um fato quase trivial, pois se $J(x) \neq 0$ para algum $x \in \Omega$, o teorema da função inversa garante que existe uma vizinhança U de x, $U \subset \Omega$, a qual é aplicada pela F difeomorficamente sobre uma vizinhança $F(U)$ de $F(x)$. Se V é uma vizinhança compacta de x contida em U, $F(V)$ é uma vizinhança (compacta) de $F(x)$; logo $F(V)$ contém um aberto de R^n, e então, por força do teorema precedente, as funções $f^1, \ldots, f^n$ não podem ser funcionalmente dependentes sobre V.

Para demonstrar a suficiência, suponhamos que $J(x) = 0$ para todo $x \in \Omega$, e seja K qualquer conjunto compacto contido em Ω. O anulamento do jacobiano sobre K significa que todo $x \in K$ é ponto crítico de F. Quer dizer que $F(K)$ está contido no conjunto dos valores críticos de F, o qual tem interior vazio em R^n, pelo teorema de Sard. Segue-se que $F(K)$ também tem interior vazio em R^n, e então, de acordo com o teorema anterior, as funções $f^1, \ldots, f^n$ são funcionalmente dependentes sobre K.

Estudamos, até aqui, a dependência funcional de uma coleção de n funções de n variáveis. Tal caso é o único a que usualmente se referem vários livros de Análise, e é o mais interessante, porque o teorema final se exprime elegantemente em termos do jacobiano da aplicação F, da qual as ditas funções são as componentes. É natural que procuremos discorrer também sobre o caso mais geral, onde ocorrem m funções de n variáveis. Aliás, nenhuma dificuldade aparecerá nesse estudo, pois dispomos do teorema de Sard, que poderá ser empregado em sua máxima força (isto é, quaisquer que sejam as dimensões em jogo).

Consideremos, pois, uma coleção de m funções $f^1, \ldots,$ $f^m \colon \Omega \to R$, diferenciáveis, definidas no conjunto aberto $\Omega \subset R^n$. Seja $F \colon \Omega \to R^m$ a aplicação definida assim:

$$F(x) = (f^1(x), \ldots, f^m(x)),$$

onde $x = (x^1, \ldots, x^n) \in \Omega$. As funções $f^1, \ldots, f^m$ dizem-se *funcionalmente dependentes* sobre o conjunto $K \subset \Omega$, se e somente se existe uma função $\phi\colon R^m \to R$, diferenciável, a qual não se anula identicamente em nenhum aberto de R^m, e tal que $\phi(F(x)) = 0$ para todo $x \in K$.

Dessa definição se deduz, imediatamente, que se $f^1, \ldots, f^m$ são funcionalmente dependentes sobre o conjunto $K \subset \Omega$, então $F(K)$ é um conjunto de interior vazio em R^m. Vice-versa, se $F(K)$ tem interior vazio em R^m, *e se K é compacto*, podemos provar que $f^1, \ldots, f^m$ são funcionalmente dependentes sobre K (o argumento é idêntico ao que empregamos há pouco, no caso $m = n$).

Observemos, agora, que se $m > n$, as funções $f^1, \ldots, f^m$ são funcionalmente dependentes sobre todo compacto $K \subset \Omega$. De fato, nessas condições o teorema de Sard afirma que $F(\Omega)$ é um conjunto de interior vazio em R^m, e então $F(K)$ também tem interior vazio, qualquer que seja o compacto $K \subset \Omega$.

O único caso que ainda temos a estudar é aquele em que $m < n$; a propósito, vale o seguinte

Teorema. *Seja $F\colon \Omega \to R^m$ uma aplicação diferenciável cujo domínio Ω é um aberto de R^n, e admitamos que $m < n$. Para que as funções $f^1, \ldots, f^m\colon \Omega \to R$, componentes de F, sejam funcionalmente dependentes sobre todo conjunto compacto $K \subset \Omega$, é necessário e suficiente que a matriz jacobiana de F tenha característica $< m$ em todo ponto de Ω.*

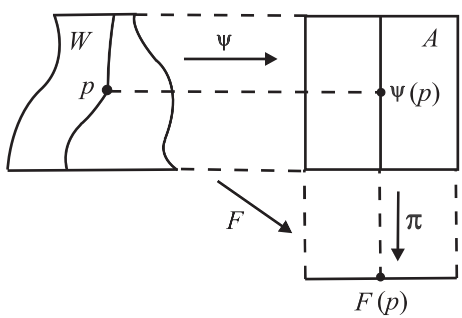
Demonstração. A suficiência da condição é consequência do teorema de Sard. Com efeito, qualquer que seja o compacto $K \subset \Omega$, a dita condição assegura-nos que $F(K)$ está contido no conjunto dos valores críticos de F; logo $F(K)$ tem interior vazio em R^m.

Provemos a necessidade. Suponhamos que as funções $f^1, \dots, f^m$ sejam funcionalmente dependentes sobre todo compacto $K \subset \Omega$. Admitamos que exista um ponto $p \in \Omega$, tal que a matriz jacobiana $\left(\frac{\partial f^i}{\partial x^j}(p)\right)$ tenha característica m, e mostemos que isso conduz a um absurdo. Podemos supor, para fixar idéias, que $D = \det\left(\frac{\partial f^i}{\partial x^j}(p)\right) \neq 0$, onde $i, j = 1, \dots, m$.

Consideremos a aplicação $\psi\colon \Omega \to R^n$ definida assim:

$$\psi(q) = \big(f^1(q), \dots, f^m(q), x^{m+1}, \dots, x^n\big),$$

para todo $q = (x^1, \dots, x^n) \in \Omega$. Verifica-se, sem dificuldade, que o determinante jacobiano da aplicação ψ, no ponto p, é justamente igual a D, e é, pois, $\neq 0$. Então, de acordo com o teorema da função inversa, ψ aplica uma vizinhança U de p difeomorficamente sobre uma vizinhança V de $\psi(p)$. Seja A um cubo de R^n, de centro $\psi(p)$ e de arestas paralelas aos eixos coordenados de R^n, tal que $A \subset V$, e seja $U \cap \psi^{-1}(A) = W$. A aplicação ψ é um difeomorfismo de W sobre A. Seja $\pi\colon R^n \to R^m$ a projeção $(x^1, \dots, x^n) \to (x^1, \dots, x^m)$, que corresponde a abandonar as $n - m$ últimas coordenadas. Tem-se, evidentemente: $F = \pi \circ \psi$.



Essa decomposição de F no difeomorfismo ψ e na projeção π, mostra claramente que toda vizinhança Z de p, tal que $Z \subset W$, é transformada pela F em uma vizinhança de $F(p)$. Podemos escolher a vizinhança Z compacta, e então $F(Z)$ há de ser uma vizinhança compacta de $F(p)$; como o interior de $F(Z)$ não é vazio, as funções componentes $f^1, \ldots, f^m$ não podem ser funcionalmente dependentes sobre o compacto Z, o que é uma contradição. Concluímos que em todo ponto $p \in \Omega$, a matriz jacobiana de F deve ter característica $< m$.

2 Estudo do teorema do ponto fixo de Brouwer

Representemos por B^n uma bola fechada do espaço euclidiano R^n, e por S^{n-1} a esfera que limita B^n. Suponhamos, por motivo de simplicidade, que B^n seja a bola de centro na origem e raio unitário; podemos escrever:

$$B^n = \{x \in R^n;\ |x| \leq 1\},$$
$$S^{n-1} = \{x \in R^n;\ |x| = 1\}.$$

Um famoso teorema de Brouwer afirma que toda transformação contínua da bola B^n admite pelo menos um ponto fixo. Apresentamos, a seguir, uma prova desse fato, baseada no teorema de Sard. Para maior clareza, dividimos o nosso estudo em três partes: na primeira, introduzimos uma formulação equivalente ao teorema de Brouwer, a qual serve melhor ao nosso fim; na segunda parte, mostramos que é possível restringirmo-nos ao caso diferenciável, o que é um passo essencial para a aplicação do teorema de Sard;

na parte final, concluímos a demonstração do teorema de Brouwer.

1ª **parte**. Inicialmente enunciemos, de modo preciso, o teorema de Brouwer.

I) Se $f\colon B^n \to B^n$ é uma aplicação contínua, existe pelo menos um ponto $x \in B^n$ tal que $f(x) = x$.

Recordemos, em seguida, o conceito de retração, introduzido na Topologia pelo matemático Karol Borsuk. Se E é um espaço topológico, e se $F \subset E$, uma *retração* de E sobre F é qualquer aplicação contínua $r\colon E \to F$ tal que $r(x) = x$ para todo $x \in F$ (quer dizer: r deixa fixos os pontos de F). O subespaço F diz-se um *retrato* de E, se existe uma retração de E sobre F.

Pois bem, o teorema de Brouwer equivale à afirmação de que a esfera não é um retrato da bola, ou melhor:

II) Não existe nenhuma retração $r\colon B^n \to S^{n-1}$.

Demonstremos que as proposição I) e II) são equivalentes.

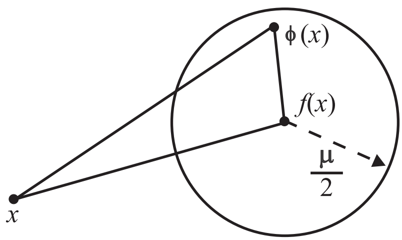
I) $\Rightarrow$ II). Suponhamos que exista uma retração $r\colon B^n \to S^{n-1}$. Definamos $f\colon B^n \to B^n$ pondo $f(x) = -r(x)$, para todo $x \in B^n$. A aplicação f é evidentemente contínua, e é claro que $f(x) \neq x$, qualquer que seja $x \in B^n$, contrariamente ao que afirma I).

II) $\Rightarrow$ I). Admitamos que exista uma aplicação contínua $f\colon B^n \to B^n$, sem pontos fixos. Tem-se, então, $f(x) \neq x$, para cada $x \in B^n$; a semirreta de origem $f(x)$ que passa por x corta a esfera S^{n-1} num ponto que designaremos por $r(x)$. A aplicação $x \to r(x)$, que é contínua, como se verifica facilmente, é uma retração $r\colon B^n \to S^{n-1}$, em contradição com II).

$2^{\underline{a}}$ **parte**. O teorema de Brouwer refere-se a transformações contínuas da bola B^n. Vamos ver que é suficiente demonstrá-lo para as transformações diferenciáveis (de classe C^∞). Suponhamos que exista uma aplicação $f\colon B^n \to B^n$, contínua, sem pontos fixos. Definamos a função $g\colon B^n \to R$, pondo $g(x) = |f(x) - x|$. Qualquer que seja $x \in B^n$, temos $f(x) \neq x$, e, portanto, $g(x) > 0$. No conjunto compacto B^n, a função g, evidentemente contínua, atinge um mínimo μ, e é óbvio que $\mu > 0$. Usando o teorema de aproximação de Weierstrass, podemos afirmar que existe uma função $\phi\colon B^n \to B^n$, de classe C^∞, tal que $|\phi(x) - f(x)| < \frac{\mu}{2}$ para todo $x \in B^n$. Segue-se que

$$|x-\phi(x)|+\frac{\mu}{2} > |x-\phi(x)|+|\phi(x)-f(x)| \geq |x-f(x)| \geq \mu,$$

donde, finalmente: $|x - \phi(x)| > \frac{\mu}{2}\cdot$



A última desigualdade mostra que $\phi\colon B^n \to B^n$ não admite nenhum ponto fixo. Portanto, se provarmos que toda aplicação diferenciável $\phi\colon B^n \to B^n$ admite algum ponto fixo, o mesmo estará demonstrado para toda aplicação $f :$ $B^n \to B^n$ contínua.

Observemos que a proposição de Brouwer no caso diferenciável equivale a afirmar que não existe nenhuma retração diferenciável $r\colon B^n \to S^{n-1}$; a demonstração dessa equivalência é a mesma que apresentamos na 1ª parte, trocando-se simplesmente a condição de continuidade pela de diferenciabilidade.

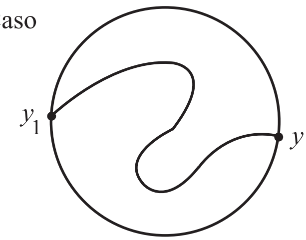
3ª **parte**. Provemos o teorema de Brouwer. De acordo com as duas partes anteriores, basta demonstrarmos que não existe nenhuma retração diferenciável $F\colon B^n \to S^{n-1}$. Procedamos pela redução ao absurdo. Suponhamos que exista uma retração diferenciável $F\colon B^n \to S^{n-1}$. Notemos que B^n é uma variedade diferenciável que tem S^{n-1} por bordo. O teorema de Sard garante a existência de valores regulares de F; na verdade, ele nos assegura que os valores regulares de F constituem um conjunto denso em S^{n-1}. Seja, pois, $y \in S^{n-1}$ um valor regular de F. Por ser F uma retração, $F(y) = y$; logo $y \in F^{-1}(y)$. Sabemos que $F^{-1}(y)$ é uma subvariedade unidimensional de B^n. Seja A a componente conexa de $F^{-1}(y)$ que contém y; como conjunto fechado contido no compacto B^n, A é também compacto. Portanto, A é uma variedade diferenciável compacta e conexa de dimensão 1 e, como tal, é difeomorfa a um segmento de reta (fechado) ou a uma circunferência.(*) Examinemos separadamente os dois casos.

No primeiro, A é uma subvariedade de B^n dotada de bordo, e então o seu bordo, constituído pelos dois pontos extremos, deve estar contido em S^{n-1}, que é o bordo de

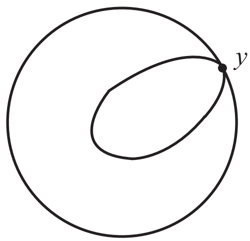
(*) Como se sabe, existem essencialmente apenas duas variedades diferenciáveis unidimensionais conexas: a reta R^1 e a circunferência S^1; no caso compacto, a variedade é um segmento (fechado) ou uma circunferência (a menos de difeomorfismos).

B^n. Seja y_1 um desses extremos, distinto de y. Temos, ao mesmo tempo: $y_1 \in S^{n-1}$ e $F(y_1) = y \neq y_1$, em contradição com a hipótese de que F é retração.

1º Caso



2º Caso



Consideremos agora o segundo caso. Seja f a restrição de F à subvariedade A de B^n. Como $f(x) = y$ para todo $x \in A$, a aplicação $f\colon A \to S^{n-1}$ é constante, e daí resulta que a aplicação linear $f_x\colon R \to R^{n-1}$ é tal que para todo vetor v, tangente a A em x, $f_x(v) = 0$, e isso qualquer que seja o ponto $x \in A$. Como $y \in A$, tem-se, em particular, $f_y(v) = 0$. Por outro lado, a restrição g de F a S^{n-1} é a aplicação idêntica, porque F é uma retração. Logo, para cada $x \in S^{n-1}$, a aplicação linear g_x, induzida por g, é a identidade, isto é, $g_x(v) = v$. Por ser $y \in S^{n-1}$, tem-se $g_y(v) = v$, para todo vetor v tangente a S^{n-1} em y. Tomando, pois, um vetor $v \neq 0$ que seja tangente a A em y, v há de ser também tangente a S^{n-1} em y, e então será, ao mesmo tempo, $f_y(v) = 0$ e $g_y(v) = v$. Mas isso é impossível, porque sendo f e g restrições da mesma aplicação F, tem-se: $f_y(v) = g_y(v) = F_y(v)$, para todo vetor $v \in A_y \subset S_y^{n-1}$.

Concluímos que não existe nenhuma retração diferenciável $F\colon B^n \to S^{n-1}$. Fica assim demonstrado o teorema do ponto fixo de Brouwer.

3 Estudo do grau de uma aplicação diferenciável

Como importante aplicação do teorema de Sard, estudamos, nas páginas que seguem, a noção de grau de uma aplicação diferenciável $f\colon M \to N$, onde M e N são variedades diferenciáveis que satisfazem às condições que adiante descreveremos. Trata-se de um conceito cuja introdução se revelou útil não só na Topologia Algébrica, mas também na Geometria Diferencial. As variedades com as quais trabalhamos são supostas orientadas; comecemos, pois, com alguns esclarecimentos sobre a orientabilidade.

3.1 Variedades orientadas

Um atlas diferenciável $\mathcal{A}$ sobre a variedade M^n diz-se *coerente* quando cumpre a seguinte condição: quaisquer que sejam os sistemas de coordenadas locais $x\colon U \to R^n$ e $y\colon V \to R^n$, pertencentes a $\mathcal{A}$, ou $U \cap V = \varnothing$ ou, em caso contrário, a aplicação $y \circ x^{-1}\colon x(U \cap V) \to y(U \cap V)$ tem jacobiano positivo em todos os pontos do seu domínio.

Se $\mathcal{A}$ é um atlas diferenciável coerente sobre M^n, um sistema local $z\colon W \to R^n$ é *admissível* em relação a $\mathcal{A}$ se para todo sistema $x\colon U \to R^n$ pertencente a $\mathcal{A}$, tal que $U \cap W \neq \varnothing$, a aplicação $z \circ x^{-1}$ tem jacobiano positivo em todo ponto do seu domínio.

Se todos os sistemas admissíveis em relação a $\mathcal{A}$ pertencem a $\mathcal{A}$, o atlas coerente $\mathcal{A}$ diz-se *máximo*. É claro que todo atlas coerente $\mathcal{A}$ está contido em um único atlas coerente máximo.

Uma variedade diferenciável M é *orientável* se admite um atlas coerente. Se M é orientável, podemos introduzir em M uma orientação, e M passa a ser, então, uma variedade *orientada*.

Para orientar uma variedade (orientável) M^n, basta distinguir sobre M um atlas coerente $\mathcal{A}$. Uma vez escolhido esse atlas, diremos, para abreviar, que um sistema local $x\colon U \to R^n$ é *positivo*, se é admissível em relação a $\mathcal{A}$ (nesse caso, x pertence ao atlas coerente máximo que contém $\mathcal{A}$).

Existe uma segunda maneira de orientar M, de caráter talvez mais intuitivo que a primeira: consiste em orientar cada espaço vetorial tangente a M, de sorte que as orientações em dois espaços tangentes vizinhos sejam coerentes, num sentido que precisaremos brevemente.

Recordemos, a propósito, a idéia de orientação de um espaço vetorial E. Se $\mathcal{E}$ e $\mathcal{F}$ são bases de E, representemos por $\det(\mathcal{E},\mathcal{F})$ o determinante da matriz de passagem de $\mathcal{E}$ a $\mathcal{F}$. A condição $\det(\mathcal{E},\mathcal{F}) > 0$ estabelece uma relação de equivalência na coleção das bases de E. A totalidade dessas bases fica repartida em duas classes de equivalência: dentro de cada classe, a passagem de uma base a outra faz-se com determinante positivo. Orientar E é escolher uma dessas duas classes; uma base dir-se-á *compatível* com a orientação de E, se e somente se pertencer à classe escolhida.

Voltemos à variedade M^n. Para cada $p \in M$, orientemos o espaço vetorial tangente M_p, porém, em vez de fazer isso de maneira arbitrária em cada ponto, suponhamos que seja possível escolher as orientações em pontos vizinhos, de modo que sejam coerentes na seguinte acepção: se p, q pertencem à mesma vizinhança coordenada U, onde é válido o sistema $x\colon U \to R^n$, as bases de M_p e M_q associadas ao sistema x são ambas compatíveis (ou ambas incompatíveis) com as orientações introduzidas em M_p e M_q. Da possibilidade de escolher orientações em todos os espaços vetoriais tangentes a M, de maneira que seja satisfeita essa condição de coerência, depende a orientabilidade de M. A variedade orientável M torna-se *orientada* quando cada espaço vetorial M_p é efetivamente orientado, sob a condição acima descrita.

Não é difícil provar a equivalência entre os dois mencionados processos de orientar a variedade M.

Passemos, agora, ao estudo de alguns resultados fundamentais para o que vai seguir.

Proposição 1. *Sejam M, N duas variedades de mesma dimensão, e $f\colon M \to N$ uma aplicação diferenciável. Se M é compacta e se $a \in N$ é valor regular de f, então $f^{-1}(a)$ é um conjunto finito.*

Demonstração. Já provamos no Capítulo I, n. 9, que $f^{-1}(a)$ é vazio ou é um conjunto de pontos isolados. Por ser um fechado contido na variedade compacta M, $f^{-1}(a)$ é compacto, e, nessas condições, só pode ser finito.

Proposição 2. *Sejam M, N duas variedades de mesma dimensão, e $f\colon M \to N$ uma aplicação diferenciável. Se*

M é compacta, e se $a \in N$ é um valor regular de f, existe uma vizinhança V de a, em N, cuja imagem inversa $f^{-1}(V)$ é a reunião de um número finito de abertos de M, disjuntos, cada um dos quais se aplica pela f difeomorficamente sobre V.

Demonstração. Em face da Proposição 1, seja $f^{-1}(a) = \{a_1, \dots, a_s\}$. Por terem M e N a mesma dimensão, e por ser f regular em a_i $(i = 1, \dots, s)$, existe uma vizinhança U_i de a_i que se aplica pela f, difeomorficamente, sobre uma vizinhança Z_i de a, isto é, as restrições $f|U_i = f_i$ são difeomorfismos $f_i \colon U_i \to Z_i$. Podemos supor que as vizinhanças $U_1, \dots, U_s$ são disjuntas, bastando para isso tomá-las suficientemente pequenas. Seja $Z = \bigcap_{i=1}^{s} Z_i$. Se $W_i = f_i^{-1}(Z)$, então $W_1, \dots, W_s$ são vizinhanças disjuntas de $a_1, \dots, a_s$, respectivamente, as quais se aplicam, pela f, difeomorficamente sobre Z. Basta agora demonstrarmos que existe uma vizinhança V de a, $V \subset Z$, bastante pequena para que seja $f^{-1}(V) \subset \bigcup_{i=1}^{s} W_i$. Provemos isso pela redução ao absurdo.

Suponhamos que para toda vizinhança V de a, $V \subset Z$, exista algum ponto $q \in V$ que se possa escrever na forma $q = f(p)$, com $p \notin \bigcup_{i=1}^{s} W_i$. Podemos, então, tomar uma base enumerável de vizinhanças de a, $V_1 \supset V_2 \supset \dots$, e construir para cada V_k os correspondentes pontos $q_k \in V_k$ e $p_k \in M$, tais que $q_k = f(p_k)$ e $p_k \notin \bigcup_{i=1}^{s} W_i$. Em virtude da compacidade de M, a sucessão (p_k) contém uma subsucessão (p'_k) que converge para um ponto $b \in M$. Como a

f é contínua, a sucessão correspondente (q'_k) converge para $f(b)$. Mas, $q_k \to a$, por construção. Logo, $f(b) = a$. Segue-se que $b \in \{a_1, \ldots, a_s\}$. Suponhamos, para fixar idéias, que $b = a_j$ $(1 \leq j \leq s)$. Aqui é que aparece o absurdo, porque sendo a_j o limite da sucessão (p'_k), existem valores de k suficientemente grandes para que seja $p_k \in W_j \subset \bigcup_{i=1}^{s} W_i$, contrariamente à hipótese.

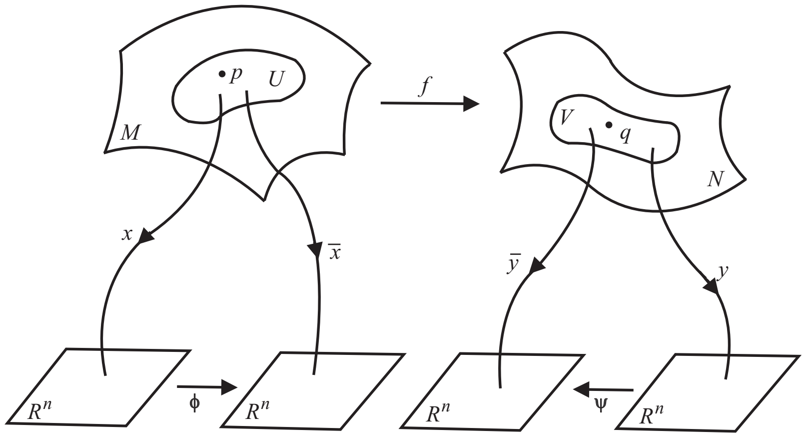
Proposição 3. *Sejam M, N duas variedades de mesma dimensão n, orientadas, e $f\colon M \to N$ uma aplicação diferenciável. Se f é regular no ponto $p \in M$, o sinal do determinante jacobiano de f, em p, relativo a sistemas positivos de coordenadas locais nas variedades M e N, não depende da escolha desses sistemas.*

Demonstração. Seja $f(p) = q$. Tomemos sistemas positivos de coordenadas locais $x\colon U \to R^n$, $y\colon V \to R^n$, nas variedades M, N, respectivamente, tais que $p \in U$, $q \in V$, $f(U) \subset V$. Seja $J(p)$ o determinante jacobiano de f, no ponto p, calculado em termos dos sistemas x, y. Mais precisamente, $J(p)$ é o determinante jacobiano da aplicação $y \circ f \circ x^{-1}\colon x(U) \to y(V)$, calculado no ponto $x(p)$, ou seja:

$$J(p) = \det\left(\frac{\partial y^j}{\partial x^k}(p)\right), \quad j, k = 1, \ldots, n.$$

Por ser f regular em p, temos $J(p) \neq 0$.

Consideremos novos sistemas positivos de coordenadas locais, $\overline{x}\colon U \to R^n$, $\overline{y}\colon V \to R^n$, que, por simplicidade, supomos definidos nos mesmos abertos $U \subset M$ e $V \subset N$.



Seja ϕ o difeomorfismo que estabelece a passagem de x a $\overline{x}$, ψ o que permite passar de y a $\overline{y}$ (a figura serve para facilitar a visualização de tudo isso). Seja $\overline{J}(p)$ o jacobiano de f, em p, calculado nas novas coordenadas. Podemos escrever:

$$\begin{gathered}\overline{x} = \phi \circ x, \qquad\qquad \overline{y} = \psi \circ y,\\ \overline{y} \circ f \circ \overline{x}^{-1} = (\psi \circ y) \circ f \circ (\phi \circ x)^{-1} = \\ = \psi \circ (y \circ f \circ x^{-1}) \circ \phi^{-1}.\end{gathered}$$

Resulta daí que

$$\overline{J}(p) = \det \psi \cdot J(p) \cdot \det \phi^{-1}.$$

Como as variedades M e N são orientadas, $\det \psi > 0$ e $\det \phi^{-1} > 0$, e então é claro que $\overline{J}(p)$ e $J(p)$ têm o mesmo sinal, como queríamos provar.

Daqui em diante, sempre que nos referirmos ao jacobiano $J(p)$ da aplicação $f\colon M \to N$, no ponto $p \in M$, estaremos supondo que esse determinante é calculado em relação a sistemas positivos de coordenadas locais nas variedades M e N.

Definição. Relativamente à aplicação $f\colon M \to N$ (M, N variedades orientadas de mesma dimensão), diremos, doravante, para simplificar a linguagem, que $p \in M$ *é ponto positivo* ou *ponto negativo*, conforme seja $J(p) > 0$ ou $J(p) < 0$. A Proposição 3 confere a essa terminologia um sentido preciso.

Observação. Existe um modo aparentemente mais geral de definir ponto positivo e ponto negativo, relativamente a uma aplicação $f\colon M \to N$ (variedades orientadas de mesma dimensão n). Suponhamos que f seja regular no ponto $p \in M$, e que $\{e_1, \ldots, e_n\}$ seja uma base do espaço vetorial tangente M_p compatível com a orientação de M. Se $f(p) = q$, a aplicação linear f_p é um isomorfismo de M_p sobre N_q, e podemos afirmar que o conjunto $\{f_p(e_1), \ldots, f_p(e_n)\}$ é uma base de N_q. O ponto p é positivo ou negativo, com respeito a f, conforme tal base seja ou não seja compatível com a orientação de N. Expliquemos o assunto em termos mais explícitos. Seja $y\colon V \to R^n$ um sistema positivo de coordenadas locais em N, tal que $q \in V$. Os vetores $f_p(e_1), \ldots, f_p(e_n)$ têm componentes bem determinadas em relação à base de N_q associada ao sistema y, e podemos considerar a matriz quadrada

$$A = \big(f_p(e_1), \ldots, f_p(e_n)\big),$$

na qual a i-ésima coluna é constituída pelas componentes de $f_p(e_i)$. O ponto p é positivo ou negativo consoante seja

$\det A > 0$ ou $\det A < 0$. Verifica-se, sem dificuldade, que o sinal de $\det A$ independe da escolha do sistema y.

A definição dada antes (por meio do jacobiano $J(p)$) corresponde ao caso em que a base $\{e_1, \ldots, e_n\}$ de M_p é a base associada a um sistema positivo de coordenadas locais $x\colon U \to R^n$, válido num aberto $U \subset M$, tal que $p \in U$. Ora, dada qualquer base de M_p, compatível com a orientação de M, sabemos que é possível considerá-la como sendo a base associada a um conveniente sistema de coordenadas locais $x\colon U \to R^n$, pertencente ao atlas coerente de M, e tal que $p \in U$. Essa observação justifica a asserção, há pouco feita, de que a nova definição de ponto positivo (ou negativo) era apenas na aparência mais geral que a primeira.

3.2 O conceito de grau

Consideremos duas variedades M e N, de mesma dimensão n, orientadas e fechadas, e uma aplicação diferenciável $f : M \to N$. Seja $a \in N$ um valor regular de f; já mostramos que $f^{-1}(a)$ é um conjunto finito. Nessas condições, definamos: o *grau da aplicação f no ponto a* é a diferença entre o número de pontos positivos e o número de pontos negativos em $f^{-1}(a)$. Representemos esse grau por $\mathrm{gr}_a(f)$.

O grau de f é definido para cada $a \in N$ que seja valor regular de f e é de se supor que o número $\mathrm{gr}_a(f)$ dependa do ponto a considerado (ou, pelo menos, nada nos autoriza a pensar que não seja assim). A definição supra representará, para nós, apenas um meio de passagem a uma idéia mais útil. Queremos chegar a um conceito de grau que tenha caráter global. Conseguiremos isso mediante a hipótese adicional de que a variedade N seja conexa; se for satisfeita essa condição, veremos que $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$, quaisquer que sejam $a, b \in N$, valores regulares de f. Pontrjagin demonstra esse resultado fazendo uso da idéia de homotopia (v. [12]). Apresentaremos aqui uma prova baseada no teorema de Sard. Dividiremos a demonstração em três partes. Supondo sempre que $a, b \in N$ sejam valores regulares de f, provaremos, primeiramente, que $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$ se a e b são pontos suficientemente próximos. A seguir, mostraremos que essa igualdade subsiste se a e b pertencem ao domínio de um sistema de coordenadas locais em N (para tal fim, esse domínio deve ser suposto homeomorfo ao espaço R^n). Finalmente, demonstraremos que a dita igualdade continua válida, quaisquer que sejam $a, b \in N$, desde que N seja uma variedade conexa.

Das três referidas partes, somente a segunda nos dará algum trabalho, especialmente se quisermos entrar imediatamente na demonstração, pois que seremos então obrigados a inserir no texto várias considerações laterais, com a desvantagem de fazer que o leitor, nesses desvios, perca o fio da meada. Será preferível, pois, começarmos com uma adequada preparação.

Façamos, inicialmente, uma observação importante. Sejam $a, b \in N$ valores regulares de $f\colon M \to N$. Se pudéssemos ligar a e b por um arco de Jordan em N cujos pontos fossem todos valores regulares de f seria mais fácil provar que $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$. Acontece, porém, que nem sempre existe um arco nessas condições. Para contornar a dificuldade daí resultante, lançaremos mão do conceito de regularidade transversa, introduzido por R. Thom em seu trabalho "Quelques propriétés globales des variétés différentiables", publicado em Comm. Math. Helv., vol. 28 (1954), pp. 17-86. Abramos um parêntese para explicar essa idéia.

Sejam M^m, N^n duas variedades e $\phi\colon M \to N$ uma aplicação diferenciável. Seja P^p uma subvariedade de N. Consideremos um ponto $x \in M$ tal que $\phi(x) = y \in P$. Em linguagem geométrica, dizemos que π é *transversalmente regular* em relação a P, no ponto $x \in M$, se os espaços vetoriais tangentes M_x, N_y, P_y satisfazem à condição

$$N_y = P_y + \phi_x(M_x).$$

(A soma acima indicada não é necessariamente direta.) Se a dita condição subsiste para todo $y \in P$ e todo $x \in \phi^{-1}(y)$, dizemos que ϕ é *transversalmente regular sobre P.*

Consideremos as aplicações

$$\phi_x\colon M_x \to N_y\,, \qquad \alpha\colon N_y \to N_{y/P_y}\,,$$

(supomos que α é a aplicação canônica de N_y sobre o espaço vetorial quociente N_{y/P_y}). Façamos a composição

$$\alpha \circ \phi_x\colon M_x \to N_{y/P_y}\,.$$

Outro modo de definir a regularidade transversa de ϕ sobre P consiste em afirmar simplesmente que $\alpha \circ \phi_x$ aplica M_x *sobre* N_{y/P_y}. Verifica-se trivialmente que as duas definições se equivalem.

Continuando na fase preparatória do nosso estudo, estabeleçamos, a seguir, um resultado do qual necessitaremos mais adiante.

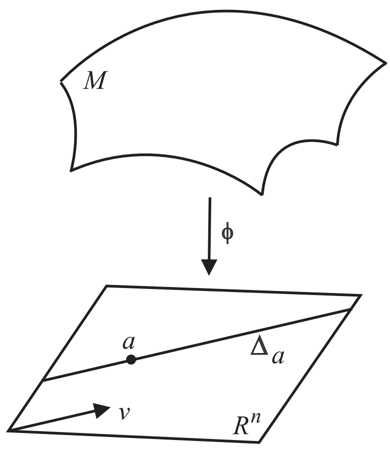
Lema. *Seja $\phi\colon M^n \to R^n$ uma aplicação diferenciável da variedade M no espaço euclidiano de mesma dimensão.*

Consideremos um ponto $a \in R^n$, um vetor $v \neq 0$ em R^n, e a reta $\Delta_a = \{a + tv;\ t \in R\}$, que é uma subvariedade de R^n. Definamos a aplicação $\widetilde{\phi}\colon M \times R \to R^n$ assim:

$$\widetilde{\phi}(p,t) = \phi(p) + tv.$$

Nessas condições, ϕ é transversalmente regular sobre Δ_a se e somente se $a \in R^n$ é valor regular de $\widetilde{\phi}$.

Demonstração. Seja $x\colon U \to R^n$ qualquer sistema de coordenadas locais em M. Se $p \in U$, representemos por $D\phi(p)$ a matriz jacobiana de ϕ, calculada em p. As colunas de $D\phi(p)$ são os vetores imagens, pela aplicação linear ϕ_p, dos vetores da base de M_p associada ao sistema x. Acrescentemos à direita da matriz $D\phi(p)$ a coluna formada pelas componentes do vetor v, de R^n, e indiquemos a matriz assim obtida com o símbolo $(D\phi(p), v)$. Observemos que $(D\phi(p), v)$ é exatamente a matriz jacobiana da aplicação $\widetilde{\phi}$, calculada no ponto $(p,t) \in M \times R$.



De acordo com a definição de regularidade transversa, podemos escrever:

1) ϕ é transversalmente regular sobre Δ_a se e somente se para todo $p \in M$, tal que $\phi(p) \in \Delta_a$, a matriz $(D\phi(p), v)$ tem característica n.

Por outro lado, tendo em vista a definição de valor regular, resulta:

2) O ponto $a \in R^n$ é valor regular de $\widetilde{\phi}$ se e somente se para todo $p \in M$, tal que $\widetilde{\phi}(p,t) = a$, a matriz $(D\phi(p), v)$ tem característica n.

Ora, se $p \in M$ e $\widetilde{\phi}(p,t) = a$, é claro que $\phi(p) \in \Delta_a$, porque $\widetilde{\phi}(p,t) = \phi(p) + tv$. Vice-versa, se $p \in M$ e $\phi(p) \in \Delta_a$, podemos escrever $\phi(p) = a - t'v$, onde $t' \in R$, e resulta $\widetilde{\phi}(p,t') = \phi(p) + t'v = a$. Segue-se que os primeiros membros das equivalências proposicionais 1) e 2) se equivalem, como queríamos demonstrar.

Passemos, agora, a provar o importante resultado já anunciado, que transcrevemos de novo sob forma do seguinte:

Teorema. *Sejam M, N duas variedades de mesma dimensão n, fechadas, orientadas, e seja $f\colon M \to N$ uma aplicação diferenciável. Se N é conexa e se $a, b \in N$ são dois quaisquer valores regulares de f, tem-se $gr_a(f) = gr_b(f)$.*

Demonstração. Tratemos sucessivamente as três partes esboçadas no esquema antes feito.

1ª **parte**. Provemos que o teorema vale se a, b são suficientemente vizinhos. Seja $f^{-1}(a) = \{a_1, \ldots, a_s\}$, de acordo com a Proposição 1. Tomemos uma vizinhança conexa $V \ni a$, em N, nas condições a que se refere a Proposição 2. Podemos escrever: $f^{-1}(V) = \bigcup_{i=1}^{s} W_i$, onde W_i é uma vizinhança de a_i em M, a qual se aplica pela f difeomorficamente sobre V, e $W_j \cap W_k = \varnothing$, se $j \neq k$. Suponhamos que $b \in V$. Então é claro que $f^{-1}(b) = \{b_1, \dots, b_s\}$, onde $b_i \in W_i$ $(i = 1, \dots, s)$. Mostremos que b_i é ponto positivo ou negativo, conforme seja a_i positivo ou negativo. Sendo W_i conexa (por ser difeomorfa a V), é possível ligar a_i e b_i por um arco de Jordan em W_i. Se $J(a_i)$ e $J(b_i)$ tivessem sinais contrários, teria que ser $J(p) = 0$ em algum ponto p do dito arco (em virtude da continuidade do jacobiano). Ora, isso é impossível, porque a restrição $f|W_i$ é um difeomorfismo. Logo, $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$.

2ª **parte**. O teorema é verdadeiro se a e b pertencem à mesma vizinhança coordenada, homeomorfa ao R^n. Seja $y \colon U \to R^n$ um sistema de coordenadas locais em N, tal que $a, b \in U$ e $y(U) = R^n$. Admitamos que $y(a) = 0$, origem de coordenadas em R^n (isso se pode sempre conseguir, com uma eventual mudança de coordenadas em R^n). Seja $y(b) = b'$. Designemos por v o vetor $0b' = b' - 0$. A reta $0b'$ é o conjunto

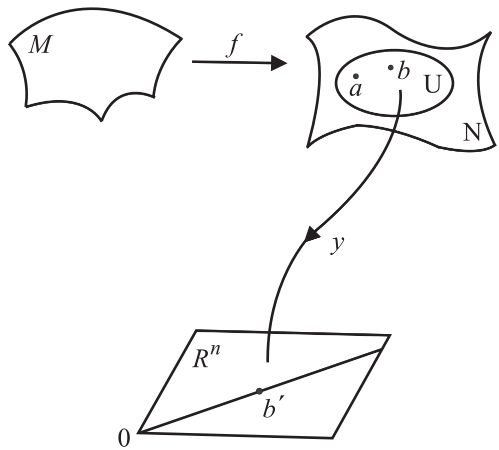
$$\Delta = \{tv;\ t \in R\}.$$

Seja

$$\phi = y \circ f \colon M \to R^n.$$

Definamos $\widetilde{\phi} \colon M \times R \to R^n$ pondo

$$\widetilde{\phi}(p, t) = \phi(p) + tv.$$



De acordo com o lema há pouco demonstrado, ϕ é transversalmente regular sobre Δ se e só se 0 é valor regular de $\widetilde{\phi}$. Ora, não podemos supor, a priori, que 0 seja valor regular de $\widetilde{\phi}$, pois não dispomos de nenhum argumento que nos garanta isso. Mas, graças ao teorema de Sard, podemos conseguir, arbitrariamente próximo de 0, um ponto que seja valor regular de $\widetilde{\phi}$, e que sirva igualmente ao nosso fim. Expliquemos esse passo mais minuciosamente. Suponhamos que 0 não seja valor regular de $\widetilde{\phi}$. Podemos tomar uma vizinhança $V \ni a$ em N, $V \subset U$, na qual o grau de f seja constante (a existência de tal vizinhança foi provada na 1ª parte). Em $y(V)$, que é vizinhança de 0 em R^n, tomemos um ponto c' que seja valor regular de $\widetilde{\phi}$ (o teorema de Sard assegura-nos que tal ponto existe). Seja $c = y^{-1}(c') \in V$. Nessas condições, $\mathrm{gr}_c(f) = \mathrm{gr}_a(f)$. Daqui em diante, basta provarmos que $\mathrm{gr}_b(f) = \mathrm{gr}_c(f)$.

As considerações que acabamos de fazer justificam a atitude mais simplista que vamos tomar, de supor que 0 seja valor regular de $\widetilde{\phi}$. Então, de acordo com o lema estudado, a aplicação ϕ é transversalmente regular sobre a reta Δ.

Mostremos, a seguir, que $\phi^{-1}(\Delta)$ é uma subvariedade de M, de dimensão 1. (Na realidade, existe uma proposição que diz: se $F\colon M^m \to N^n$ é transversalmente regular sobre a subvariedade $P^p \subset N$, a imagem inversa $F^{-1}(P)$ é uma subvariedade de M, de dimensão $m-n+p$. A demonstração desse resultado, que é uma consequência do teorema das funções implícitas, pode ser encontrada em [10]. Em vez de usar essa proposição, vamos dar uma prova direta para o caso que nos interessa.)

Seja $\pi\colon R^n \to R^n$ a projeção ortogonal de R^n sobre o hiperplano perpendicular à reta Δ na origem. Se $u = (\alpha^1, \dots, \alpha^n)$ é o vetor unitário de $v = 0b'$ (isto é, da reta Δ, orientada positivamente de 0 para b'), é fácil ver que

$$\pi(x^1, \dots, x^n) = (x^1, \dots, x^n) - \left(\sum_{i=1}^{n} \alpha^i x^i\right)(\alpha^1, \dots, \alpha^n).$$

A imagem $\pi(R^n)$ é o conjunto

$$\left\{(x^1, \dots, x^n) \in R^n;\ \sum_i \alpha^i x^i = 0\right\},$$

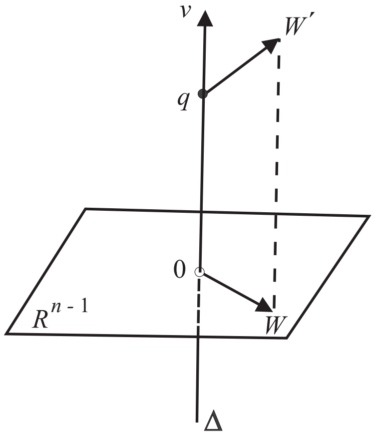
que é justamente o hiperplano acima referido. Escolhendo uma base nesse hiperplano, podemos identificá-lo com R^{n-1}. Ponhamos $\psi = \pi \circ \phi\colon M \to R^{n-1}$.

Observemos que $\pi^{-1}(0) = \Delta$, isto é, os pontos de R^n que se projetam na origem são todos os pontos da reta Δ, e somente estes. Notemos também que

$$(\pi \circ \phi)^{-1}(0) = \phi^{-1}(\pi^{-1}(0)) = \phi^{-1}(\Delta),$$

e então basta provarmos que 0 é valor regular da aplicação $\pi \circ \phi$ (v. teorema estudado no Capítulo I, §9). A projeção π é uma transformação linear, e coincide, pois, com a transformação linear por ela induzida em cada ponto de R^n. Se $p \in M$ e $\phi(p) = q \in R^n$, podemos escrever: $(\pi \circ \phi)_p = \pi_q \circ \phi_p = \pi \circ \phi_p$.

Para qualquer $p \in (\pi \circ \phi)^{-1}(0) = \phi^{-1}(\Delta)$, precisamos mostrar que $(\pi \circ \phi)_p$ aplica M_p sobre R^{n-1}. Ora, dado o



vetor w, de origem 0 em R^{n-1}, existe um vetor w', de origem q em R^n, tal que $\pi_q(w') = \pi(w') = w$ (realmente, há uma infinidade de vetores tais como w'). Como ϕ é transversalmente regular sobre Δ, existe algum vetor $u' \in M_p$, tal que seja

$$w' = \phi_p(u') + \lambda v,$$

onde $\lambda \in R$ (lembrar que $v = 0b'$ é vetor tangente a Δ). Segue-se que

$$\pi_q(w') = \pi_q(\phi_p(u')) + \pi_q(\lambda v) = (\pi_q \circ \phi_p)(u') + 0 = (\pi \circ \phi)_p(u'),$$

como queríamos.

Está deste modo provado que $\phi^{-1}(\Delta)$ é uma subvariedade unidimensional de M.

A fim de dar continuação a este estudo, vamos aqui abrir um parêntese e fazer algumas considerações que serão importantes mais adiante. Para que tenham sentido preciso as afirmações que seguem, suponhamos que a variedade M^n esteja imersa num espaço euclidiano (o teorema de Whitney assegura-nos que é sempre possível imergir a variedade M^n num espaço euclidiano R^k, desde que seja $k \geq 2n+1$). Nessas condições, para cada $p \in M$ podemos pensar no espaço tangente M_p como sendo um espaço vetorial euclidiano, e podemos falar no produto escalar de dois vetores de M_p; também adquire sentido a idéia de ortogonalidade de vetores. Após essas observações, vamos mostrar que sobre a subvariedade $\phi^{-1}(\Delta) \subset M$ é possível definir $n-1$ campos diferenciáveis de vetores linearmente independentes, ao mesmo tempo tangentes à variedade M e normais à subvariedade $\phi^{-1}(\Delta)$. Seja $p \in \phi^{-1}(\Delta) \subset M$. Reconsideremos a aplicação

$$\psi = \pi \circ \phi \colon M \to R^{n-1}$$

que em termos de um sistema de coordenadas locais $x \colon U \to R^n$, tal que $p \in U$, se exprime por meio de $n-1$ funções reais

$$y^i = \psi^i(x^1, \ldots, x^n), \qquad i = 1, \ldots, n-1.$$

Consideremos as diferenciais $d\psi^1, \ldots, d\psi^{n-1}$, que são formas lineares sobre o espaço vetorial tangente M_p. Por ser euclidiano, o espaço vetorial M_p é canonicamente isomorfo ao seu dual M_p^*; esse isomorfismo é, como se sabe, $J\colon M_p \to M_p^*$ assim definido: para cada $v \in M_p$, $J(v) = g \in M_p^*$ tal que $g(w) = v \cdot w$, onde $w \in M_p$ e o ponto indica o produto escalar. Sejam, então, $\nabla\psi^1, \ldots, \nabla\psi^{n-1}$ os vetores de M_p correspondentes, nesse isomorfismo, às formas $d\psi^1, \ldots, d\psi^{n-1} \in M_p^*$. Esses vetores são normais à subvariedade $\phi^{-1}(\Delta) \subset M$ no ponto p; com efeito, se u é um vetor tangente a essa subvariedade em p, resulta:

$$u \cdot \nabla\psi^i = d\psi^i(u) = \frac{\partial\psi^i}{\partial u}(p) = 0,$$

porque ψ^i é constante sobre $\phi^{-1}(\Delta)$. Os $n-1$ vetores $\nabla\psi^1, \ldots, \nabla\psi^{n-1}$ são linearmente independentes, porque são independentes as formas lineares $d\psi^1, \ldots, d\psi^{n-1}$ que lhes correspondem no dito isomorfismo. De fato,

$$(d\psi^i)_p = \sum_{j=1}^{n} \frac{\partial\psi^i}{\partial x^j}(p)\, dx^j, \quad i = 1, \ldots, n-1,$$

e como $p \in \psi^{-1}(0) = \phi^{-1}(\Delta)$ e 0 é valor regular de ψ, segue-se que a matriz jacobiana $\left(\frac{\partial\psi^i}{\partial x^j}(p)\right)$, $i = 1, \ldots, n-1$; $j = 1, \ldots, n$, tem característica $n-1$, e daí resulta a independência linear das formas $d\psi^1, \ldots, d\psi^{n-1}$ (no ponto p).

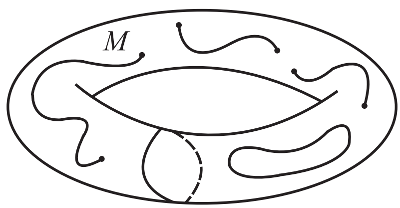
O vetor $\nabla\psi^i$ diz-se *gradiente* da função ψ^i.

Fechemos o parêntese e voltemos ao ponto onde nos achávamos.

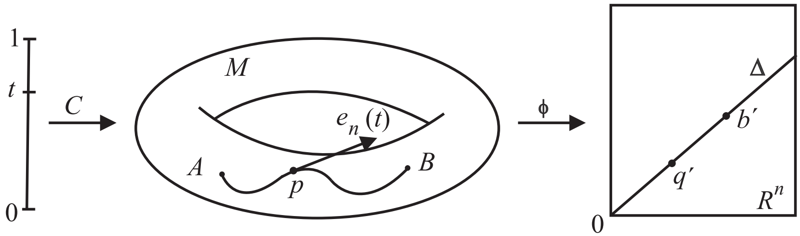
Por ser uma subvariedade de dimensão 1 da variedade fechada M, o conjunto $\phi^{-1}(\Delta)$ é uma reunião de retas e circunferências (topológicas). (Recorde o leitor que toda variedade diferenciável conexa unidimensional, sem bordo, é difeomorfa à reta R^1 ou à circunferência S^1.) A subvariedade $\phi^{-1}(\Delta)$ contém os conjuntos

$$\phi^{-1}(0) = f^{-1}(a) = \{a_1, \ldots, a_r\},$$
$$\phi^{-1}(b') = f^{-1}(b) = \{b_1, \ldots, b_s\}.$$

Vamos agora restringir as nossas considerações ao segmento $0b' \subset \Delta$. É claro que $\phi^{-1}(0b')$ é uma reunião de circunferências e de segmentos (topológicos), ou seja, uma reunião de arcos e curvas fechadas de Jordan, na variedade M (v. figura).



É intuitivo (e pode demonstrar-se sem dificuldade) que os pontos $a_1, \ldots, a_r$, $b_1, \ldots, b_s$ são necessariamente as extremidades desses arcos de Jordan; daí resulta que tais arcos são em número finito. Portanto, para efeito de avaliação de $\mathrm{gr}_a(f)$ e de $\mathrm{gr}_b(f)$, as curvas fechadas eventualmente contidas em $\phi^{-1}(0b')$ podem ser transcuradas, já que não contêm pontos de $f^{-1}(a)$, nem de $f^{-1}(b)$. Resta fazermos um estudo cuidadoso relativo aos ditos arcos de Jordan. Tomemos um qualquer deles, cujas extremidades representaremos por A, B, e ao qual nos referiremos como sendo o arco AB.



Seja $C\colon I \to M$, onde $I = [0,1]$, uma parametrização diferenciável do arco AB, tal que $C(0) = A$, $C(1) = B$, $C(t) = p$.

Chamemos $e_1, \ldots, e_{n-1}$ aos $n-1$ campos diferenciáveis de vetores cuja existência há pouco demonstramos (esses campos são bem definidos sobre o arco AB). No ponto $p = C(t)$, os vetores $e_1(t), \ldots, e_{n-1}(t)$ são, como vimos, tangentes à variedade M e normais ao arco AB. Seja $e_n(t)$ o vetor $C'(t)$, tangente ao arco AB no ponto p. É claro, então, que o conjunto

$$\{e_1(t), \ldots, e_{n-1}(t), e_n(t)\}$$

constitui uma base do espaço vetorial tangente M_p. Podemos admitir que essa base seja compatível com a orientação existente em M_p como espaço tangente à variedade orientada M (se o não fosse, substituiríamos um dos $n-1$ primeiros vetores pelo seu simétrico, e obteríamos uma base na dita condição).

Consideremos o determinante

$$D(t) = \det\big(\phi_p(e_1(t)), \ldots, \phi_p(e_{n-1}(t)), \phi_p(e_n(t))\big),$$

onde $\phi_p(e_i(t))$ indica a coluna formada pelas n componentes do vetor $\phi_p(e_i(t)) \in R^n$. Se u é o vetor unitário de

$v = b' - 0$, podemos escrever:

$$\phi_p(e_n(t)) = \phi_p(C'(t)) = \lambda(t)u,$$

onde $\lambda(t)$ é uma função real diferenciável. Consideremos também o determinante

$$A(t) = \det\big(\phi_p(e_1(t)), \ldots, \phi_p(e_{n-1}(t)), u\big).$$

Em virtude da regularidade transversa de ϕ sobre a reta Δ, segue-se que $A(t) \neq 0$ para todo $t \in I$. Como $A(t)$ depende continuamente de t, podemos afirmar que $A(t)$ não muda de sinal quando t descreve I. Observemos que entre $D(t)$ e $A(t)$ existe a seguinte relação

$$D(t) = \lambda(t) \cdot A(t). \tag{*}$$

Recordemos, agora, que os pontos A, B são elementos do conjunto

$$f^{-1}(a) \cup f^{-1}(b) = \{a_1, \ldots, a_r\} \cup \{b_1, \ldots, b_s\}.$$

Pode acontecer que A pertença a um dos conjuntos $f^{-1}(a)$, $f^{-1}(b)$ e B pertença ao outro, mas pode também suceder que A e B estejam ambos num só desses conjuntos. É essencial que estudemos as duas possibilidades.

1º **caso**. Para fixar idéias, suponhamos que seja $A \in f^{-1}(a)$ e $B \in f^{-1}(b)$. Então $A = a_i$, $B = b_j$ e temos $\phi(A) = 0$, $\phi(B) = b'$. Como a aplicação ϕ é regular nos pontos a_i e b_j, existem vizinhanças de a_i e b_j em M, as quais se aplicam pela ϕ difeomorficamente sobre vizinhanças de 0 e de b' em R^n. Seja $q' = \phi(p) = \phi(C(t))$.

Podemos afirmar que para valores não nulos de t suficientemente próximos de 0, os correspondentes pontos q' são vizinhos de 0, e distintos de 0. De modo análogo se comportam os pontos q' que correspondem a valores de t bastante próximos de 1, e diferentes de 1; tais pontos são vizinhos de b', e distintos de b'. Vamos exprimir tudo isso em linguagem precisa. Considerando $0b'$ como segmento orientado de origem 0, seja $g(t)$ a abscissa de q'; é claro que

$$g(t) = (\phi(C(t)) - 0) \cdot u,$$

onde u é o vetor unitário de $b' - 0$ e o ponto indica o produto escalar. Por derivação, obtemos:

$$g'(t) = \big((\phi \circ C)'(t)\big) \cdot u = \big(\phi_p(C'(t))\big) \cdot u = \lambda(t).$$

As observações acima feitas equivalem justamente a dizer que a função real $g(t)$ é crescente à direita de 0 e também à esquerda de 1; quer dizer que $g'(0) = \lambda(0) > 0$ e $g'(1) = \lambda(1) > 0$ (derivadas unilaterais, evidentemente). Segue-se daí, e da relação (*), que $D(0)$ tem o mesmo sinal que $A(0)$ e $D(1)$ tem o mesmo sinal que $A(1)$. Mas, já vimos que $A(0)$ e $A(1)$ têm o mesmo sinal; portanto, o mesmo ocorre com $D(0)$ e $D(1)$, e isso significa que os pontos $a_i, b_j \in M$ são ambos positivos ou ambos negativos.

$2^{\underline{o}}$ **caso**. Admitamos agora, para fixar idéias, que $A, B \in f^{-1}(a)$; seja então $A = a_i$, $B = a_j$, onde $i \neq j$. Nesse caso, $\phi(A) = \phi(B) = 0$. Empregando a mesma função real $g(t)$ definida no $1^{\underline{o}}$ caso, verifica-se imediatamente que $g(t)$ é crescente à direita de 0 e decrescente à esquerda de 1; portanto, $g'(0) = \lambda(0) > 0$ e $g'(1) = \lambda(1) < 0$. Em face da relação (*), $D(0)$ e $A(0)$ têm o mesmo sinal, ao passo que $D(1)$ e $A(1)$ têm sinais contrários. Como $A(0)$ e $A(1)$ têm o mesmo sinal, segue-se que $D(0)$ e $D(1)$ têm sinais contrários. Indica isso que se a_i é ponto positivo, a_j é negativo, e vice-versa.

Chegamos, pois, ao fim desta 2ª parte de nossa demonstração. Se o arco de Jordan AB está no 2º caso, ele não pesa no cálculo de $\mathrm{gr}_a(f)$ (ou de $\mathrm{gr}_b(f)$), porque as suas extremidades são pontos de $f^{-1}(a)$ (ou de $f^{-1}(b)$) de sinais contrários. Os únicos arcos que influem no cálculo do dito grau são os que estão no 1º caso. Para estes, as extremidades são ambas positivas ou ambas negativas, e, pela f, uma delas se aplica em a e a outra em b, e então é óbvio que $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$, como queríamos demonstrar.

3ª **parte**. Se a variedade N é conexa, $\mathrm{gr}_a(f) = \mathrm{gr}_b(f)$, quaisquer que sejam $a, b \in N$, valores regulares de f.

Podemos ligar a e b por um arco em N, definido por uma função contínua $C\colon I \to N$, tal que $C(0) = a$, $C(1) = b$. Argumentando com a compacidade de I e com a continuidade de C, é fácil ver que se pode dividir o intervalo I por meio de pontos $0 = t_0, t_1, \ldots, t_{p-1}, t_p = 1$, de maneira que pontos consecutivos, tais como $q_{i-1} = C(t_{i-1})$ e $q_i = C(t_i)$ pertençam a uma mesma vizinhança coordenada (conexa). Nessas condições, tendo em vista a 2ª parte da demonstração, resulta:

$$\mathrm{gr}_a(f) = \mathrm{gr}_{q_1}(f) = \cdots = \mathrm{gr}_{q_{p-1}}(f) = \mathrm{gr}_b(f),$$

e isso conclui a demonstração.

O teorema precedente confere sentido preciso à seguinte

Definição. O *grau da aplicação diferenciável* $f\colon M \to N$, onde M e N são variedades de mesma dimensão, fechadas e orientadas, N sendo conexa, é o número inteiro $\mathrm{gr}_a(f)$, calculado em qualquer ponto $a \in N$ que seja valor regular de f. Esse grau será indicado por $\mathrm{gr}(f)$.

Antes de prosseguir no estudo do grau, ilustremos esse conceito por meio de um exemplo. A variedade com que vamos lidar nesse exemplo é a esfera de Riemann S^2, que equivale topologicamente ao plano complexo compactificado pelo acrescentamento do ponto ∞. A homeomorfia entre esses dois espaços pode ser estabelecida, como sabemos, por meio da projeção estereográfica, que faz corresponder a 0 o polo sul da esfera, e a ∞ o polo norte. Como variedade orientável imersa em R^3, S^2 pode ser orientada, de modo natural, assim: um sistema de coordenadas locais $x\colon U \to R^2$, válido num aberto $U \subset S^2$, é considerado positivo quando para cada $z \in U$, a base do plano tangente S_z associada ao sistema x forma com a normal exterior à esfera um triedro positivo em R^3.

Consideremos a aplicação $f\colon S^2 \to S^2$ definida por $f(z) = z^n$, onde n é um número inteiro > 0. A f é diferenciável (e até analítica), e deixa fixos os pontos 0 e ∞ (polos da esfera), os quais são os seus únicos valores críticos. É fácil verificar que $\mathrm{gr}(f) = n$. Com efeito, se $w \in S^2$ é valor regular de f (isto é, distinto de 0 e de ∞), $f^{-1}(w)$ se compõe de n pontos $z_1, \ldots, z_n$, os quais são todos positivos, como vamos mostrar. Seja M um meridiano da esfera, que não passe por nenhum dos $n+1$ ditos pontos (meridiano aqui significa uma semicircunferência contendo os polos da esfera como extremidades). No conjunto aberto $U = S^2 - M$, podemos adotar o usual sistema de coordenadas polares $x\colon U \to R^2$ tal que, se $z \in U$ e $z = \rho(\cos\theta + i \operatorname{sen}\theta)$, se tenha $x(z) = (\rho, \theta)$. Para cada $z \in U$, tal que $f(z) \in U$, podemos escrever:

$$f(z) = \rho'(\cos\theta' + i \text{ sen } \theta') = z^n = \rho^n(\cos n\theta + i \text{ sen } n\theta),$$

ou seja:

$$\rho' = \rho^n, \qquad \theta' = n\theta.$$

O jacobiano de f no ponto z é

$$J(z) = \begin{bmatrix} n\rho^{n-1} & 0 \\ 0 & n \end{bmatrix} = n^2\,\rho^{n-1} > 0.$$

Como os $n+1$ pontos $z_1, \ldots, z_n$, w pertencem a U, podemos concluir que $J(z_i) > 0$ para $i = 1, \ldots, n$, como queríamos provar.

Para valores negativos do inteiro n, a aplicação $f\colon S^2 \to S^2$ definida por $f(z) = z^n$ é ainda diferenciável, mas aplica 0 em ∞ e ∞ em 0; estes são os únicos valores críticos de f. Procedendo como acima, pode concluir-se que o grau de f é o número positivo $-n$.

Se $n = 0$, a aplicação $f(z) = z^n$, de S^2 em S^2, reduz-se à aplicação constante $z \to 1$. Então, 1 é o único valor crítico de f. Todo ponto $w \in S^2$, tal que $w \neq 1$, é um valor regular de f, para o qual $f^{-1}(w) = \varnothing$, e isso mostra que $\operatorname{gr}(f) = 0$.

No exemplo acima, o grau é um inteiro positivo ou nulo. O grau de uma aplicação diferenciável pode, porém, ser um inteiro negativo, e não é difícil dar exemplo de aplicação para a qual o grau é um inteiro negativo arbitrário. Não nos deteremos em tais exemplos; o leitor interessado poderá encontrá-los em [8], cap. II.

3.3 Duas aplicações da teoria do grau

Vamos descrever dois interessantes e úteis resultados aos quais nos conduz o estudo que estamos desenvolvendo. Em toda a teoria que apresentamos a respeito do grau de uma aplicação diferenciável $f\colon M^n \to N^n$, supusemos que M era uma variedade compacta, e, para definir o grau (global) de f, foi necessário supor N conexa. A hipótese da compacidade de M pode ser substituída por outra um tanto mais fraca, sem que a teoria do grau deixe de ser válida. Em vez de admitir que M é compacta, podemos supor, somente, que a f é uma *aplicação própria*. Com isso queremos dizer que é compacta a imagem inversa $f^{-1}(K)$ de toda parte compacta $K \subset N$. Evidentemente, se M é compacta, qualquer aplicação $f\colon M \to N$, que seja contínua, é também própria, mas a recíproca não é verdadeira. É fácil ver que a teoria do grau continua de pé se substituirmos a hipótese de que a variedade M é compacta, pela hipótese de que a aplicação f é própria. A razão disso é que a Proposição 1 e a Proposição 2 resistem a essa modificação, isto é, são verdadeiras as duas proposições que seguem.

Proposição 1′ – *Sejam M, N duas variedades de mesma dimensão, e $f\colon M \to N$ uma aplicação diferenciável própria. Se $a \in N$ é valor regular de f, o conjunto $f^{-1}(a)$ é finito.*

Proposição 2′ – *Se M, N são variedades de mesma dimensão, e se $f\colon M \to N$ é uma aplicação diferenciável própria, cada ponto $a \in N$, que seja valor regular de f, possui uma vizinhança V, em N, cuja imagem inversa $f^{-1}(V)$ é reunião de um número finito de abertos de M, disjuntos,* *cada um dos quais se aplica pela f difeomorficamente sobre V.*

As demonstrações são análogas às que apresentamos para as Proposições 1 e 2, com ligeiras adaptações à nova hipótese.

Passemos às prometidas aplicações do estudo do grau. Descrevê-las-emos sob forma de teoremas.

Teorema 2. *Sejam M, N variedades de mesma dimensão, orientadas, N conexa, e $f\colon M \to N$ uma aplicação diferenciável própria. Se f é biunívoca, o seu jacobiano $J(p)$ não pode mudar de sinal.*

Demonstração. Sejam $x, y \in M$ dois pontos distintos. Suponhamos que $J(x)$ e $J(y)$ tivessem sinais contrários, e que fosse, por exemplo, $J(x) > 0$ e $J(y) < 0$. Tendo em vista a biunivocidade de f, concluímos que o grau de f seria $+1$ em $f(x)$ e -1 em $f(y)$, o que é impossível. Logo, $J(p)$ não pode mudar de sinal quando p descreve M.

Observações. 1) Édouard Goursat, em seu Cours d'Analyse Matématique, vol. I, no capítulo dedicado às integrais duplas, dá uma demonstração intuitiva do resultado a que se refere o teorema acima, em uma situação particular. Goursat considera uma aplicação $f\colon A_1 \to A$, onde A_1 e A são, na realidade, variedades bidimensionais orientadas (com bordo), e a f é suposta de classe C^1 e biunívoca. Querendo provar que o jacobiano J de f não muda de sinal em A_1, Goursat admite, por absurdo, que J se anule sobre uma curva γ_1, que separe a porção de A_1 onde $J > 0$ da porção onde $J < 0$. Ora, a hipótese de que J muda de sinal em A_1 implica o anulamento de J em pontos de A_1, mas não se pode garantir, sem maior cuidado, que seja uma curva o conjunto de tais pontos. Descrevamos uma situação em que é válido o argumento apresentado por Goursat. Definamos uma função $\phi\colon A_1 \to R$ da seguinte maneira: $\phi(p) = J(p)$ para todo $p \in A_1$. O conjunto dos pontos de A_1 onde $J = 0$ é $\phi^{-1}(0)$. Suponhamos que a f seja de classe C^2, donde resulta que a ϕ é de classe C^1. Nessas condições, se 0 é valor regular de ϕ, o conjunto $\phi^{-1}(0)$ é uma subvariedade unidimensional de A_1 (e, portanto, uma curva diferenciável, não necessariamente conexa).

A falha acima apontada, aliada à circunstância inconveniente de que o final da prova é baseado em uma observação puramente intuitiva, não deixa de ser uma imperfeição na demonstração dada por aquele eminente matemático.

2) Com as mesmas hipóteses do Teorema 2, se $f\colon M \to N$ é biunívoca mas não é sobre N, podemos garantir que o jacobiano $J(p)$ é identicamente nulo em M. De fato, existe então algum ponto $a \in N$ tal que $f^{-1}(a) = \varnothing$, e resulta $\mathrm{gr}(f) = \mathrm{gr}_a(f) = 0$. Qualquer que seja $p \in M$, se $J(p) \neq 0$, $f(p)$ é valor regular de f, e o grau de f em $f(p)$ é $+1$ ou -1, conforme seja $J(p) > 0$ ou $J(p) < 0$. Em face da contradição achada, concluímos que $J(p) = 0$.

Teorema 3. *Sejam M, N variedades de mesma dimensão, orientadas, N conexa, e $f\colon M \to N$ uma aplicação diferenciável própria. Se o jacobiano de f, $J(p)$, não muda de sinal em M, e nem é identicamente nulo, a f aplica M sobre N.*

Demonstração. Suponhamos, para fixar idéias, que $J(p) \geq 0$ em M. Como $J(p)$ não é identicamente nulo, seja $q \in M$ tal que $J(q) > 0$. Existe uma vizinhança V de

q, em M, tal que $J(p) > 0$ para todo $p \in V$. O conjunto $f(V) = W$ é uma vizinhança de $f(q) = a$ em N. De acordo com o teorema de Sard, existe um ponto $a' \in W$ que é valor regular de f. O grau de f no ponto a' é > 0, pois $f^{-1}(a')$ contém um ponto q' no qual $J > 0$, e em todos os demais pontos de $f^{-1}(a')$ (se houver algum) tem-se $J > 0$, já que $J \geq 0$ em M. Resulta daí que o grau de f é > 0. Segue-se que f é necessariamente sobre N, pois se existisse um ponto $b \in N$ tal que $f^{-1}(b) = \varnothing$, seria $\mathrm{gr}_b(f) = 0$.

Observações. 1) Dentro das condições em que vale o Teorema 3, se acrescentarmos a hipótese de que a variedade M seja compacta, poderemos concluir, como corolário, que N é também compacta.

2) O Teorema 3, no caso de variedades de dimensão 2, foi obtido por S.S. Chern, como corolário de um teorema que ele demonstra no artigo "Complex Analytic Mappings of Riemann Surfaces, I", American Journal of Mathematics, vol. 82, n.2 (1960), pp. 323-337.

3) O Teorema 3, no caso geral acima demonstrado, foi estabelecido por S. Sternberg e R.G. Swan, no artigo "On Maps with Nonnegative Jacobian", The Michigan Mathematical Journal, vol. 6 (1959), pp. 339-342. A demonstração que apresentam é essencialmente a mesma que acabamos de fazer, baseada na consideração do grau, mas este é estudado de modo diferente (nesse trabalho, os autores empregam técnicas da Topologia Algébrica).

4 Estudo do grau à luz da teoria da integração

Em muitas das considerações que seguem, as variedades são supostas riemannianas. A propósito, recordemos que em toda variedade diferenciável compacta é possível introduzir uma métrica riemanniana; esse fato é uma consequência imediata do teorema de imersão de Whitney (v. [6], cap. III, pag. 170). Com base na teoria da integração em uma variedade riemanniana, o grau de uma aplicação diferenciável pode ser interpretado de um modo interessante, que também serve como nova definição desse conceito (de caráter global). Para maior clareza da exposição, e para tornar mais amena a leitura do texto seguinte, vamos apresentar um resumo dos conhecimentos que serão necessários.

Comecemos com a recordação de algumas noções da Álgebra Exterior. Seja E um espaço vetorial n-dimensional sobre R. Um funcional n-linear $\omega\colon E\times\cdots\times E\to R$ diz-se *alternado* quando

$$\omega(v_1,\ldots,v_i,\ldots,v_k,\ldots,v_n)=-\omega(v_1,\ldots,v_k,\ldots,v_i,\ldots,v_n),$$

quaisquer que sejam os índices distintos $i,k\in\{1,\ldots,n\}$. Podemos definir a soma $\omega_1+\omega_2$ e o produto $\lambda\omega$, onde $\lambda\in R$, da maneira usual:

$$(\omega_1+\omega_2)(v_1,\ldots,v_n)=\omega_1(v_1,\ldots,v_n)+\omega_2(v_1,\ldots,v_n),$$

$$(\lambda\omega)(v_1,\ldots,v_n)=\lambda\omega(v_1,\ldots,v_n).$$

Em relação a essas operações, os funcionais n-lineares alternados formam um espaço vetorial unidimensional sobre

R. O vetor zero desse espaço vetorial é o funcional 0 tal que $0(v_1, \ldots, v_n) = 0$, quaisquer que sejam os vetores $v_1, \ldots, v_n$ pertencentes a E.

Se $f_1, \ldots, f_n$ são funcionais lineares sobre E, isto é, elementos do espaço vetorial dual E^*, a aplicação $\omega\colon E \times \cdots \times E \to R$ definida por

$$\omega(v_1, \ldots, v_n) = \det \begin{bmatrix} f_1(v_1) \ldots f_1(v_n) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(v_1) \ldots f_n(v_n) \end{bmatrix}$$

é evidentemente um funcional n-linear alternado. Esse funcional é, por definição o *produto exterior* dos n funcionais $f_1, \ldots, f_n$, e representa-se assim: $\omega = f_1 \wedge \cdots \wedge f_n$. A definição acima conduz-nos às seguintes conclusões:

1) $\omega = f_1 \wedge \cdots \wedge f_n = 0$ se e só se os funcionais $f_1, \ldots, f_n$ são linearmente dependentes;

2) se $\omega \neq 0$, $\omega(v_1, \ldots, v_n) = 0$ se e somente se os vetores $v_1, \ldots, v_n \in E$ são linearmente dependentes. Enquanto esta última conclusão nos oferece um critério útil para a verificação da dependência linear de n vetores de E, a primeira nos assegura que se $f_1, \ldots, f_n$ são linearmente independentes, $\omega = f_1 \wedge \cdots \wedge f_n$ é uma base do espaço vetorial dos funcionais n-lineares alternados definidos sobre $E \times \cdots \times E$ (n fatores). Se $\overline{\omega}$ é qualquer desses funcionais, deve ser $\overline{\omega} = \lambda\omega$, onde $\lambda \in R$.

4.1 Formas diferenciais exteriores sobre uma variedade

Admitamos, daqui por diante, que a variedade M^n seja orientada. Vamos ver como é que se podem definir sobre M

formas diferenciais exteriores de grau n, ou simplesmente formas de grau n. (Na realidade, uma tal forma é uma seção de um espaço fibrado, mas, para o fim que temos em mira, podemos evitar a menção explícita desse espaço, e contentar-nos com a descrição sumária que passamos a fazer.) Uma *forma* ω*, de grau* n, sobre a variedade M, é uma aplicação que associa a cada ponto $p \in M$ um funcional n-linear alternado

$$\omega_p \colon M_p \times \cdots \times M_p \to R$$

(recorde-se que M_p é o espaço vetorial tangente a M em p).

Seja $x \colon U \to R^n$ um sistema positivo de coordenadas locais em M, tal que $p \in U$. Seja $\{e_1, \ldots, e_n\}$ a base de M_p associada ao sistema x. As diferenciais $dx^1, \ldots, dx^n$ constituem, como sabemos, uma base do espaço M_p^*, dual de M_p (trata-se justamente da base dual da primeira, pois é claro que $dx^i(e_j) = \delta_j^i$). De acordo com uma observação anterior, podemos escrever:

$$\omega_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n,$$

onde a é uma função real definida em $x(U) \subset R^n$. O valor que a forma ω_p assume na n-pla de vetores $v_1, \ldots, v_n \in M_p$ é:

$$\omega_p(v_1, \ldots, v_n) = a(x(p)) \det(dx^i(v_j)).$$

Se $y \colon V \to R^n$ é outro sistema positivo de coordenadas locais em M, tal que $p \in V$, deve ser

$$\omega_p = b(y(p))dy^1 \wedge \cdots \wedge dy^n,$$

e é fácil verificar que

$$b(y(p)) = a(x(p)) \det\left(\frac{\partial x^i}{\partial y^j}(p)\right) = a(x(p)) \cdot J(p).$$

Como o jacobiano $J(p)$ é > 0, podemos concluir que $a(x(p))$ e $b(y(p))$ têm o mesmo sinal.

Reconsideremos a expressão da forma ω no ponto p:

$$\omega_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n.$$

A forma ω diz-se *positiva, negativa* ou *nula* no ponto $p \in M$, conforme seja $a(x(p)) > 0$, $a(x(p)) < 0$ ou $a(x(p)) = 0$; esse sinal, como acabamos de ver, não depende da escolha do sistema (positivo) x.

A forma ω é dita *contínua* se para todo $p \in M$ existe um sistema de coordenadas locais $x\colon U \to R^n$, com $p \in U$, tal que $\omega_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n$, onde $a\colon x(U) \to R$ é uma função contínua.

Em nosso estudo, consideraremos inicialmente as formas contínuas, porém teremos interesse, mais adiante, em lidar também com certas formas que não satisfazem necessariamente a essa condição.

4.2 Suporte de uma função ou de uma forma. Partição da unidade

Se M é uma variedade e $F\colon M \to R$ uma função, chamaremos *suporte* dessa função ao fecho do conjunto dos pontos de M nos quais F assume valores diferentes de zero. Define-se de modo análogo o suporte de uma forma ω, de grau n, sobre M^n: é o fecho do conjunto dos pontos $p \in M$, tais que $\omega_p \neq 0$. Quando a variedade M é compacta, o suporte de uma função ou de uma forma é sempre compacto.

Outra idéia importante, que convém relembrarmos, é a de partição da unidade. Tal conceito, que pode ser estudado em [10], em [13] ou em [4], tem um desempenho muito útil em várias questões nas quais o objetivo é a passagem da situação local à global. Seja M uma variedade diferenciável, $\{U_\alpha\}$ uma cobertura de M, aberta e localmente finita. Uma *partição da unidade sobre M, subordinada à cobertura* $\{U_\alpha\}$, é uma coleção $\{\phi_\alpha\}$ de funções $\phi_\alpha \colon M \to R$, contínuas, $\phi_\alpha(p) \geq 0$ sobre M, com as duas seguintes propriedades: 1) para todo α, o suporte de ϕ_α está contido em U_α; 2) para todo $p \in M$, $\sum_\alpha \phi_\alpha(p) = 1$ (observe-se que nesta soma há apenas um número finito de parcelas não nulas).

Demonstra-se que se M é uma variedade arbitrária, e se $\{U_\alpha\}$ é qualquer cobertura aberta de M, existe uma cobertura aberta $\{V_\alpha\}$ de M, localmente finita, tal que $V_\alpha \subset U_\alpha$, qualquer que seja α. Prova-se, também, que para toda cobertura aberta e localmente finita de M, existe uma partição da unidade, subordinada a essa cobertura. Mais ainda: é possível tomar as funções ϕ_α dessa partição diferenciáveis. As demonstrações desses resultados podem ser encontradas em [10]; não as reproduziremos aqui, para não alongar esta exposição com questões que pertencem mais propriamente à Topologia Geral, e que estão um tanto afastadas da linha de assuntos que vimos seguindo neste momento.

4.3 Integração de formas contínuas sobre uma variedade compacta

Vamos mostrar, a seguir, como é que se pode integrar uma forma contínua, de grau n, sobre uma variedade M^n, compacta e orientada.

Seja $x\colon U \to R^n$ um sistema (positivo) de coordenadas locais em M. Conforme observamos, em cada ponto $p \in U$ a forma ω assume um valor

$$\omega_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n,$$

onde $a\colon x(U) \to R$ é uma função contínua. Também podemos escrever:

$$\omega_p = a(x^1, \ldots, x^n)dx^1 \wedge \cdots \wedge dx^n,$$

onde $x^1, \ldots, x^n$ são as coordenadas de p no sistema x. Suponhamos que o suporte S de ω esteja contido em U; nessas condições, a função a se anula no exterior do conjunto compacto $x(S) \subset x(U) \subset R^n$, e podemos definir a integral da forma ω sobre M da seguinte maneira:

$$\int_M \omega = \int_{x(U)} a(x^1, \ldots, x^n)dx^1 \ldots dx^n,$$

onde o segundo membro é uma integral múltipla clássica, no sentido de Riemann.

Para que essa definição seja legítima, é necessário que o valor que ela atribui à integral da forma ω não dependa da escolha das coordenadas. É fácil verificar que essa condição se cumpre. Seja $y\colon V \to R^n$ outro sistema (positivo) de coordenadas locais em M, tal que o suporte de ω esteja contido em V. Sabemos que

$$\omega_p = b(y(p))dy^1 \wedge \cdots \wedge dy^n = b(y^1, \ldots, y^n)dy^1 \wedge \cdots \wedge dy^n,$$

e já provamos que

$$b(y(p)) = a(x(p)) \cdot J(p).$$

Usando o teorema clássico da mudança de variáveis em uma integral múltipla, podemos escrever:

$$\begin{aligned}\int_M \omega &= \int_{x(U)} a(x^1, \ldots, x^n)dx^1 \ldots dx^n = \\ &= \int_{y(V)} a(x^1, \ldots, x^n)J(p)dy^1 \ldots dy^n = \\ &= \int_{y(V)} b(y^1, \ldots, y^n)dy^1 \ldots dy^n.\end{aligned}$$

Passemos à situação mais geral em que o suporte de ω não está necessariamente contido numa vizinhança coordenada. Seja $\{U_\alpha\}$ uma cobertura qualquer de M por vizinhanças coordenadas. Como M é compacta, $\{U_\alpha\}$ admite uma subcobertura finita, que indicaremos simplesmente por $\{U_1, \ldots, U_r\}$. Seja $\{\phi_1, \ldots, \phi_r\}$ uma partição da unidade, subordinada a essa cobertura finita. Qualquer que seja $p \in M$, podemos escrever:

$$\omega_p = \left(\sum_{i=1}^r \phi_i(p)\right)\omega_p = \sum_{i=1}^r \phi_i(p) \cdot \omega_p\,,$$

ou, abreviadamente:

$$\omega = \sum_{i=1}^r \phi_i\,\omega.$$

Definiremos a integral de ω sobre M por meio da fórmula:

$$\int_M \omega = \sum_{i=1}^{r} \int_M \phi_i\, \omega.$$

Para cada $i \in \{1, \ldots, r\}$, $\phi_i\, \omega$ é uma forma cujo suporte está contido na vizinhança coordenada U_i, e já sabemos como se calcula $\displaystyle\int_M \phi_i \omega$ nessas condições.

Para mostrar que a definição acima tem sentido, precisamos provar que a integral ali definida não depende da cobertura $\{U_1, \ldots, U_r\}$ escolhida, nem da partição da unidade, subordinada a essa cobertura. Seja $\{V_1, \ldots, V_s\}$ outra cobertura de M por vizinhanças coordenadas, e $\{\psi_1, \ldots, \psi_s\}$ qualquer partição da unidade, a ela subordinada. Observemos que os abertos $U_i \cap V_j$ $(i = 1, \ldots, r;$ $j = 1, \ldots, s)$ constituem uma cobertura de M, e que as rs funções θ_{ij}, definidas por $\theta_{ij}(p) = \phi_i(p) \cdot \psi_j(p)$, compõem uma partição da unidade, subordinada à mesma cobertura. Como $\sum\limits_i \phi_i = 1$ e $\sum\limits_j \psi_j = 1$, resulta:

$$\begin{aligned}\sum_i \int_M \phi_i \omega &= \sum_i \int_M \phi_i \left(\sum_j \psi_j \right) \omega = \sum_{i,j} \int_M \phi_i \psi_j \omega = \\ &= \sum_{i,j} \int_M \theta_{ij} \omega,\end{aligned}$$

$$\begin{aligned}\sum_j \int_M \psi_j \omega &= \sum_j \int_M \psi_j \left(\sum_i \phi_i \right) \omega = \sum_{i,j} \int_M \psi_j \phi_i \omega = \\ &= \sum_{i,j} \int_M \theta_{ij} \omega.\end{aligned}$$

Fica assim concluída a prova de que a definição da integral acima dada tem caráter intrínseco.

Observação. A definição que demos de $\int_M \omega$ pode estender-se ao caso em que a variedade M não é compacta, desde que seja compacto o suporte da forma ω. Se $\{U_\alpha\}$ é uma cobertura de M por vizinhanças coordenadas, não podemos, em geral, extrair dela uma subcobertura finita, mas uma subcobertura enumerável $\{V_1, V_2, \dots\}$ sempre existe (propriedade de Lindelöf). Podemos obter, em seguida, uma cobertura $\{W_1, W_2, \dots\}$ de M, localmente finita, tal que $W_i \subset V_i$ para todo índice i. Seja $\{\phi_1, \phi_2, \dots\}$ uma partição da unidade, subordinada a $\{W_i\}$; podemos admitir que o suporte de cada ϕ_i seja compacto (para isso, basta supor que cada W_i é o domínio de um sistema de coordenadas locais $x_i\colon W_i \to R^n$, tal que $x_i(W_i)$ seja um aberto limitado de R^n). Ponhamos, então:

$$\int_M \omega = \sum_{i=1}^{\infty} \int_M \phi_i\, \omega.$$

Por ser compacto, o suporte de ω só encontra um número finito dos abertos W_i, de sorte que a soma acima indicada apresenta somente um número finito de parcelas que podem ser $\neq 0$.

4.4 Integração sobre uma variedade riemanniana

Na definição de uma variedade riemanniana, desempenha preponderante papel o conceito de produto escalar. Para maior clareza do texto seguinte, talvez seja útil que apresentemos uma ligeira recapitulação desse assunto. Os espaços vetoriais a que nos referiremos serão sempre supostos de dimensão finita.

Um *produto escalar* num espaço vetorial real V é uma forma bilinear simétrica sobre V, definida positiva. Em termos mais explícitos, um produto escalar em V é uma função $g\colon V \times V \to R$, com as seguintes propriedades:

1) se $u, v, w \in V$ e se $\lambda \in R$, tem-se:

$$\begin{aligned} g(u+v, w) &= g(u,w) + g(v,w), \\ g(u, v+w) &= g(u,v) + g(u,w), \\ g(\lambda u, v) &= \lambda g(u,v), \\ g(u, \lambda v) &= \lambda g(u,v), \end{aligned}$$

2) quaisquer que sejam $u, v \in V$, tem-se:

$$g(u,v) = g(v,u),$$

3) para todo $u \in V$, $g(u,u) \geq 0$, e $g(u,u) = 0$ se e somente se $u = 0$.

Preferiremos usar a notação $u \cdot v$ em lugar de $g(u,v)$. Um exemplo bem conhecido do produto escalar é o clássico produto escalar de dois vetores do espaço R^n: se $u = (x^1, \ldots, x^n)$ e $v = (y^1, \ldots, y^n)$, então

$$u \cdot v = x^1 y^1 + \cdots + x^n y^n.$$

Um espaço vetorial V munido de um produto escalar diz-se *euclidiano*, e em tal espaço se define o *comprimento* de um vetor u como sendo o número real não negativo $\sqrt{u \cdot u}$. Também podemos definir a *distância* do vetor u ao vetor v como sendo o comprimento do vetor $v - u$, isto é: $d(u, v) = \sqrt{(v - u) \cdot (v - u)}$. Não é difícil verificar que a função $d\colon V \times V \to R$, assim definida, satisfaz aos axiomas de uma função distância. Nessas condições, V torna-se um espaço métrico. Num espaço vetorial euclidiano V, podemos também definir a noção de ortogonalidade. Dois vetores $u, v \in V$ dizem-se *ortogonais* se e somente se $u \cdot v = 0$. Se V tem dimensão n, uma base $\{e_1, \ldots, e_n\}$ de V diz-se *ortonormal* quando os vetores $e_1, \ldots, e_n$ são unitários (isto é, de comprimento 1) e ortogonais entre si (quer dizer: $e_i \cdot e_j = 0$, se $i \neq j$). É claro, então, que se e_i, e_j são dois vetores quaisquer de uma base ortonormal de V, distintos ou não, vale a relação: $e_i \cdot e_j = \delta_{ij}$ (símbolo de Kronecker). Num espaço vetorial euclidiano V, de dimensão finita, existem sempre bases ortonormais; a partir de uma base arbitrária $\{u_1, \ldots, u_n\}$ de V, é possível construir uma base ortonormal $\{e_1, \ldots, e_n\}$, tal que cada e_i seja combinação linear de $u_1, \ldots, u_i$ (processo de ortonormalização de Gram-Schmidt).

Suponhamos, agora, que o espaço vetorial euclidiano V, de dimensão n, seja orientado. Se $v_1, \ldots, v_n$ são n vetores de V, linearmente independentes, definiremos da seguinte maneira o *volume* (orientado) do paralelepípedo por eles formado:

$$\mathrm{vol}(v_1, \ldots, v_n) = \pm\sqrt{\det(v_i \cdot v_j)},$$

onde se toma o sinal $+$ ou o sinal $-$, conforme a base $\{v_1, \ldots, v_n\}$ seja ou não seja compatível com a orientação de V. Mostremos que esse volume é sempre um número real bem definido. Para isso, consideremos em V uma base ortonormal $\{e_1, \ldots, e_n\}$, compatível com a orientação de V, e representemos os vetores $v_1, \ldots, v_n$ em termos dessa base:

$$v_k = \sum_r \alpha_k^r \, e_r \,.$$

Podemos escrever:

$$v_i \cdot v_j = \sum_{r,s} \alpha_i^r \alpha_j^s (e_r \cdot e_s) = \sum_{r,s} \alpha_i^r \alpha_j^s \delta_{rs} = \sum_h \alpha_i^h \alpha_j^h \,.$$

Seja α a matriz (α_i^k), e α' a sua transposta. O produto escalar $v_i \cdot v_j$ é evidentemene o elemento da i-ésima linha e da j-ésima coluna da matriz produto $\alpha\alpha'$. Então

$$\det(v_i \cdot v_j) = \det(\alpha\alpha') = \det\,\alpha \cdot \det\,\alpha' = (\det\,\alpha)^2,$$

e resulta que

$$\mathrm{vol}(v_1, \ldots, v_n) = \pm\sqrt{(\det\,\alpha)^2},$$

ou, com as convenções de sinal há pouco admitidas:

$$\mathrm{vol}(v_1, \ldots, v_n) = \det\,\alpha.$$

Esta última fórmula também tem sentido quando os vetores $v_1, \ldots, v_n$ são linearmente dependentes, e, neste caso, é claro que $\mathrm{vol}(v_1, \ldots, v_n) = 0$. Daqui em diante, usaremos o símbolo $\mathrm{vol}(v_1, \ldots, v_n)$ para indicar o volume associado aos n vetores $v_1, \ldots, v_n$, quer sejam eles independentes, quer sejam dependentes.

Voltemos ao estudo das variedades. Uma variedade diferenciável M^n diz-se *riemanniana* quando em cada ponto

$p \in M$ o espaço vetorial tangente M_p é dotado de um produto escalar, o qual varia diferenciavelmente, na seguinte acepção: se $x\colon U \to R^n$ é qualquer sistema de coordenadas locais em M, tal que $p \in U$, e se $\{X_1, \ldots, X_n\}$ é a base de M_p associada a esse sistema, as funções reais $g_{ij} = X_i \cdot X_j$, definidas em U, são diferenciáveis (de classe C^∞).

Os produtos escalares definidos nos espaços vetoriais tangentes a uma variedade riemanniana constituem o que se chama a *métrica riemanniana* da variedade. Cada vetor u, tangente a M em p, tem um comprimento bem determinado, definido por $\sqrt{u \cdot u}$, onde o ponto indica o produto escalar em M_p.

Se a variedade riemanniana M^n é orientada, podemos falar no volume (orientado) do paralelepípedo formado por n vetores linearmente independentes $v_1, \ldots, v_n \in M_p$.

Seja M^n uma variedade compacta, orientada, riemanniana. Mostraremos que é possível integrar não só as formas contínuas de grau n sobre M, mas também as funções reais contínuas definidas sobre M. Essa possibilidade resulta da existência de uma forma diferencial canônica de grau n sobre a variedade M. Tal forma, que designaremos por σ, e à qual daremos o nome de *elemento de volume*, assume em cada ponto $p \in M$ um valor σ_p definido assim:

$$\sigma_p(v_1, \ldots, v_n) = \mathrm{vol}(v_1, \ldots, v_n),$$

onde $v_1, \ldots, v_n \in M_p$.

Tomemos em M_p uma base ortonormal $\{e_1, \ldots, e_n\}$, compatível com a orientação de M_p. Em relação a essa base, podemos escrever:

$$v_k = \sum_r \alpha_k^r \, e_r \, ,$$

e já provamos que

$$\mathrm{vol}(v_1, \dots, v_n) = \det\, \alpha,$$

onde α é a matriz (α^i_j). Segue-se que

$$\sigma_p(v_1, \dots, v_n) = \det\, \alpha,$$

e torna-se agora evidente que σ, como foi acima definida, é, de fato, uma forma diferencial exterior de grau n sobre a variedade M (basta pensar nas propriedades dos determinantes).

Se $f\colon M \to R$ é uma função contínua, definamos a integral de f sobre a variedade M assim:

$$\int_M f = \int_M f\sigma,$$

onde $f\sigma$ é a forma que em cada $p \in M$ assume o valor $f(p)\sigma_p$.

Procuremos expressar a forma σ em termos de coordenadas locais. Seja $x\colon U \to R^n$ um sistema positivo de coordenadas em M. Se $p \in U$, já sabemos que

$$\sigma_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n.$$

Precisamos determinar a função $a\colon x(U) \to R$. Designando por $\{X_1, \dots, X_n\}$ a base de M_p associada ao sistema x, podemos escrever:

$$\sigma_p(X_1, \dots, X_n) = a(x(p)) \det(dx^i(X_j)) = a(x(p)).$$

Por outro lado,

$$\sigma_p(X_1, \dots, X_n) = \mathrm{vol}(X_1, \dots, X_n) = +\sqrt{\det(X_i \cdot X_j)}.$$

Observe-se que o radical acima aparece precedido do sinal $+$, porque a base $\{X_1, \dots, X_n\}$ de M_p é compatível com a orientação desse espaço vetorial, já que x é, por hipótese, um sistema positivo. Adotemos a notação antes introduzida: $X_i \cdot X_j = g_{ij}$ (não esquecer que g_{ij} são funções do ponto $p \in U$), e façamos $\det(g_{ij}) = g$. Nessas condições,

$$\sigma_p(X_1, \dots, X_n) = \sqrt{g},$$

e podemos concluir, finalmente, que

$$a(x(p)) = \sqrt{g},$$

e que

$$\sigma_p = \sqrt{g}\, dx^1 \wedge \cdots \wedge dx^n.$$

Convém assinalar que $g = \det(g_{ij}) > 0$ em todo ponto $p \in U$, pois g é o volume (orientado) do paralelepípedo formado em M_p pelos vetores $X_1, \dots, X_n$, os quais constituem uma base de M_p compatível com a orientação deste espaço vetorial.

4.5 Formas localmente somáveis. Conjuntos mensuráveis

Seja M^n uma variedade compacta, orientada, riemanniana, e ω uma forma de grau n sobre M (não se supõe ω contínua). Dizemos que ω é *localmente somável* se para cada ponto $p \in M$ existe um sistema (positivo) de coordenadas locais $x\colon U \to R^n$, com $p \in U$, no qual a forma ω se exprime como $\omega_p = a(x(p))dx^1 \wedge \cdots \wedge dx^n$, onde a é uma função integrável segundo Lebesgue no conjunto aberto $x(U) \subset R^n$. É claro que toda forma contínua é localmente somável.

Se ω é localmente somável, e se o suporte de ω está contido no domínio U do sistema de coordenadas x, podemos definir a integral de ω sobre M assim:

$$\int_M \omega = \int_{x(U)} a(x^1, \ldots, x^n)\, dx^1 \ldots dx^n.$$

Da mesma maneira que no caso das formas contínuas, demonstra-se que essa integral não depende do sistema x. Se o suporte de ω não está contido numa vizinhança coordenada U nas condições descritas, podemos proceder como no caso das formas contínuas, e definir

$$\int_M \omega = \sum_{i=1}^{r} \int_M \phi_i\, \omega,$$

onde $\{\phi_1, \ldots, \phi_r\}$ é uma partição da unidade, subordinada a uma cobertura finita $\{U_1, \ldots, U_r\}$ de M, tal que cada U_i satisfaça às condições acima indicadas.

Consideremos um conjunto qualquer $C \subset M$. Seja $\phi\colon M \to R$ a função característica de C (quer dizer: $\phi(p) = 1$, se $p \in C$, e $\phi(p) = 0$, se $p \notin C$). Representemos por σ a forma que define o elemento de volume em M. Diremos que o conjunto C é *mensurável*, se a forma $\phi\sigma$ for localmente somável. Como M é compacta, a forma $\phi\sigma$, correspondente ao conjunto mensurável C, é integrável, e podemos definir o *volume* de C por meio da integral.

$$\operatorname{vol} C = \int_M \phi\sigma.$$

Todo conjunto fechado $F \subset M$ é mensurável. Com efeito, qualquer que seja o sistema (positivo) de coordenadas locais $x\colon U \to R^n$, o conjunto $F \cap U$ é fechado em U; logo, $x(F \cap U)$ é fechado em $x(U) \subset R^n$, e, portanto, mensurável em R^n. Se ϕ é a função característica de F, a forma $\phi\sigma$ se exprime, no sistema x, para cada $p \in U$, da seguinte maneira:

$$(\phi\sigma)_p = \phi(p) \cdot \sigma_p = \psi(x(p)) \cdot \sqrt{g}\, dx^1 \wedge \cdots \wedge dx^n,$$

onde indicamos por ψ a função característica de $x(F \cap U)$ em R^n. Concluímos que a função $\psi(x(p))\sqrt{g}$ é integrável segundo Lebesgue no conjunto $x(U)$, e daí segue que a forma $\phi\sigma$ é localmente somável e, finalmente, que F é mensurável.

É fácil provar que todo conjunto A de medida nula em M tem volume nulo. Com efeito, qualquer que seja o sistema de coordenadas locais $x\colon U \to R^n$, o conjunto $x(A \cap U)$ tem medida nula em R^n. Se ϕ é a função característica de A, podemos escrever:

$$(\phi\sigma)_p = \phi(p) \cdot \sigma_p = \psi(x(p)) \cdot \sqrt{g}\, dx^1 \wedge \cdots \wedge dx^n,$$

onde $p \in U$ e ψ é a função característica de $x(A \cap U)$. Como ψ se anula sobre $x(U)$, exceto num conjunto de medida nula, resulta:

$$\int_{x(U)} \psi(x(p)) \cdot \sqrt{g}\, dx^1 \ldots dx^n = 0.$$

Ao calcular o volume de A, por meio de uma partição da unidade, encontraremos vol $A = 0$, porque são nulas todas as parcelas da soma que representa esse volume.

4.6 Integração sobre variedades não compactas

Consideremos uma variedade diferenciável M^n, orientada, mas que não supomos compacta. A mesma definição que demos no caso compacto pode ser usada para definir formas localmente somáveis sobre M.

Seja ω uma forma localmente somável sobre M. Consideremos uma cobertura aberta enumerável $\{U_1, U_2, \dots\}$ de M, localmente finita, tal que cada U_r seja o domínio de um sistema (positivo) de coordenadas $x_r \colon U_r \to R^n$, no qual a forma ω, em cada $p \in U_r$, se exprime por

$$\omega_p = a_r(x_r(p))dx_r^1 \wedge \cdots \wedge dx_r^n\,,$$

onde a_r é uma função integrável segundo Lebesgue no aberto $x_r(U_r) \subset R^n$. Seja $\{\phi_1, \phi_2, \dots\}$ uma partição da unidade, subordinada à cobertura $\{U_r\}$. Para cada r, a forma $\phi_r\omega$ exprime-se, no sistema x_r, da seguinte maneira:

$$\begin{aligned}(\phi_r\omega)_p &= \phi_r(p) \cdot a_r(x_r(p))dx_r^1 \wedge \cdots \wedge dx_r^n = \\ &= A_r(x_r^1, \dots, x_r^n)dx_r^1 \wedge \cdots \wedge dx_r^n\,.\end{aligned}$$

Diremos que ω é integrável se a série

$$\sum_{r=1}^{\infty} \int_M \phi_r\omega = \sum_{r=1}^{\infty} \int_{x_r(U_r)} A_r(x_r^1, \dots, x_r^n)dx_r^1 \dots dx_r^n$$

é convergente, qualquer que seja a escolha da cobertura $\{U_r\}$ e da partição $\{\phi_r\}$, nas condições acima descritas. Se tal convergência se verifica para toda partição $\{\phi_r\}$, resulta, em particular, que a série converge qualquer que seja a ordem dos seus termos (observe-se que mudar a ordem desses termos corresponde a reordenar as funções ϕ_r). Podemos, pois, afirmar que a série é absolutamente convergente. Decorre daí que todas as manipulações que usamos, no caso finito, para provar que a integral, definida por partição da unidade, independe dessa partição, são ainda válidas no presente caso. Podemos, então, com legitimidade, definir a integral $\displaystyle\int_M \omega$ como sendo a soma da série acima.

Quando a forma ω, localmente somável, tem suporte compacto, a série considerada tem todos os termos nulos, exceto um número finito deles, e então ω é integrável. A esse respeito, reportamo-nos à observação que fizemos no final do §3.

4.7 Novas considerações sobre o grau de uma aplicação diferenciável

Chegamos agora à parte final do nosso trabalho. Vamos dar uma interpretação do grau que também serve como nova definição desse conceito.

Sejam M e N duas variedades de dimensão n, fechadas, orientadas, riemannianas; admitamos, ainda, que N seja conexa. Consideremos uma aplicação diferenciável $f\colon M \to N$. Se $p \in M$ e $q = f(p)$, representemos, como sempre, por $f_p\colon M_p \to N_q$ a aplicação linear induzida por f. Se ω é uma forma sobre N, a aplicação f induz sobre M uma forma $f^*\omega$, definida da seguinte maneira:

$$(f^*\omega)_p(v_1, \ldots, v_n) = \omega_q(f_p v_1, \ldots, f_p v_n),$$

onde $v_1, \ldots, v_n \in M_p$, e $f_p v_i$ significa $f_p(v_i)$.

Verifica-se imediatamente que $f^*(\lambda\omega) = \lambda(f^*\omega)$, para todo $\lambda \in R$.

Estamos particularmente interessados na consideração da forma σ (elemento de volume) sobre N. Se $y\colon V \to R^n$ é um sistema (positivo) de coordenadas locais em N, e se $q \in V$, já vimos que

$$\sigma_q = \sqrt{g}\, dy^1 \wedge \cdots \wedge dy^n,$$

onde $g = \det(g_{ij}(q))$.

Seja, agora, $x\colon U \to R^n$ um sistema (positivo) de coordenadas locais em M, tal que $p \in U$. Podemos escrever:

$$(f^*\omega)_p = a(x(p))\, dx^1 \wedge \cdots \wedge dx^n.$$

Indiquemos por $X_1, \ldots, X_n$ os vetores da base de M_p associada ao sistema x. Por um lado, temos:

$$(f^*\sigma)_p(X_1, \ldots, X_n) = a(x(p)) \det(dx^i(X_j)) = a(x(p)),$$

e, por outra parte:

$$(f^*\sigma)_p(X_1, \ldots, X_n) = \sigma_q(f_p X_1, \ldots, f_p X_n) =$$
$$= \sqrt{g}\, dy^1 \wedge \ldots \wedge dy^n(f_p X_1, \ldots, f_p X_n) = \sqrt{g}\, \det(dy^i(f_p X_j)).$$

Mas,

$$f_p X_j = \sum_{k=1}^{n} \frac{\partial y^k}{\partial x^j}(p)\, Y_k\,,$$

onde $Y_1, \ldots, Y_n$ são os vetores da base de N_q associada ao sistema y.

$$dy^i(f_p X_j) = dy^i\left(\sum_{k=1}^{n} \frac{\partial y^k}{\partial x^j}(p) Y_k\right) = \sum_{k=1}^{n} \frac{\partial y^k}{\partial x^j}(p)\, dy^i(Y_k) =$$
$$= \sum_{k=1}^{n} \frac{\partial y^k}{\partial x^j}(p)\, \delta_k^i = \frac{\partial y^i}{\partial x^j}(p).$$

Segue-se daí que

$$(f^*\sigma)_p(X_1, \ldots, X_n) = \sqrt{g}\, \det\left(\frac{\partial y^i}{\partial x^j}(p)\right) = \sqrt{g} \cdot J(p),$$

onde $J(p)$ é o jacobiano da aplicação f no ponto p (calculado nos sistemas x e y considerados).

De tudo o que precede, podemos concluir:

$$a(x(p)) = \sqrt{g} \cdot J(p),$$
$$(f^*\sigma)_p = \sqrt{g} \cdot J(p)\, dx^1 \wedge \cdots \wedge dx^n.$$

É bom recordar que g deve ser calculado no ponto $q = f(p)$ (pertencente à variedade N).

Reconsideremos, agora, a aplicação diferenciável $f \colon M \to N$ acima mencionada, e seja γ o grau de f, no sentido já estudado. Vamos demonstrar uma igualdade notável relacionada com esse grau, a saber:

$$\int_M f^*\sigma = \gamma \int_N \sigma.$$

Comecemos a demonstração com um estudo de natureza local. Se $q \in N$ é um valor regular de f (lembrar que o conjunto de tais valores nunca é vazio, em virtude do teorema de Sard), existe uma vizinhança V de q, em N, tal que $f^{-1}(V)$ é uma reunião finita de partes de M, $W_1, \dots, W_r$, abertas, disjuntas, cada uma das quais é aplicada pela f difeomorficamente sobre V (v. Proposição 2, pag. 84). Para abreviar, diremos que uma vizinhança V (em N), nessas condições, é uma *vizinhança regular*. Se $p_i \in W_i$ é tal que $f(p_i) = q$, é claro que o sinal do jacobiano da aplicação f em todos os pontos de W_i é positivo ou negativo, conforme seja p_i um ponto positivo ou negativo em relação a f. Podemos admitir, evidentemente, que os abertos $V \subset N$ e $W_i \subset M$ $(i = 1, \dots, r)$ sejam domínios de sistemas de coordenadas locais $y\colon V \to R^n$ e $x_i\colon W_i \to R^n$. De acordo com as considerações há pouco feitas, escreveremos:

$$\int_V \sigma = \int_{y(V)} \sqrt{g}\, dy^1 \dots dy^n,$$
$$\int_{W_i} f^*\sigma = \int_{x_i(W_i)} \sqrt{g}\, J(p)\, dx_i^1 \dots dx_i^n\,.$$

Esta última integral é > 0 ou < 0, consoante seja $J(p) > 0$ ou $J(p) < 0$ sobre W_i, isto é, conforme seja p_i ponto positivo ou negativo. Seja ε_i um símbolo que suporemos igual a $+1$ se p_i é ponto positivo, e igual a -1 se p_i é ponto negativo. Nessas condições, podemos escrever:

$$\int_{W_i} f^*\sigma = \varepsilon_i \int_{x_i(W_i)} \sqrt{g}\, |J(p)|\, dx_i^1 \dots dx_i^n\,.$$

Usando o teorema da mudança de variáveis numa integral múltipla, resulta:

$$\int_{W_i} f^*\sigma = \varepsilon_i \int_{y(V)} \sqrt{g}\, dy^1 \dots dy^n\,,$$

ou seja:

$$\int_{W_i} f^*\sigma = \varepsilon_i \int_V \sigma.$$

Resulta daí que

$$\int_{f^{-1}(V)} f^*\sigma = \sum_{i=1}^{r} \varepsilon_i \int_V \sigma = \left(\sum_{i=1}^{r} \varepsilon_i\right) \int_V \sigma,$$

ou, finalmente:

$$\int_{f^{-1}(V)} f^*\sigma = \gamma \int_V \sigma.$$

O que nos resta agora a fazer, é estender esta igualdade do caso local ao global. Para esse fim, necessitaremos empregar os dois lemas que seguem.

Lema 1. *Se a forma ω é integrável sobre a variedade orientada M, e se ω se anula sobre um conjunto fechado $F \subset M$, então $\int_M \omega = \int_{M-F} \omega$.*

Demonstração. Seja $\{U_r\}$ uma cobertura aberta de M, enumerável e localmente finita, cada U_r sendo o domínio de um sistema (positivo) de coordenadas locais $x_r \colon U_r \to R^n$, no qual a forma ω se expressa por meio de uma função a_r, integrável segundo Lebesgue em $x_r(U_r)$. Tomemos uma partição qualquer da unidade, $\{\phi_r\}$, subordinada a essa cobertura. A integral de ω sobre M é, por definição:

$$\int_M \omega = \sum_{r=1}^{\infty} \int_{x_r(U_r)} A_r(x_r^1, \ldots, x_r^n)\, dx_r^1 \ldots dx_r^n\,,$$

onde pusemos $\phi_r(p) \cdot a_r(x_r(p)) = A_r(x_r^1, \dots, x_r^n)$. Façamos $U_r' = U_r \cap (M-F)$. É claro que a coleção $\{U_r'\}$ é uma cobertura aberta de $M - F$, localmente finita. Como $\phi_r(p) = 0$ se $p \notin U_r$, e como por hipótese $a_r(x_r(p)) = 0$ se $p \in F$, concluímos que a função A_r é nula fora de $x_r(U_r')$. Usando uma propriedade conhecida das integrais em R^n, podemos escrever:

$$\int_M \omega = \sum_{r=1}^{\infty} \int_{x_r(U_r')} A_r(x_r^1, \dots, x_r^n) dx_r^1 \dots dx_r^n = \int_{M-F} \omega,$$

como queríamos demonstrar.

Lema 2. *Se a forma ω é integrável sobre a variedade orientada M, e se $F \subset M$ é um conjunto fechado de medida nula em M, tem-se* $\int_M \omega = \int_{M-F} \omega.$

Demonstração. Procedendo exatamente como na demonstração do lema anterior, chegamos à expressão

$$\int_M \omega = \sum_{r=1}^{\infty} \int_{x_r(U_r)} A_r(x_r^1, \dots, x_r^n)\, dx_r^1 \dots dx_r^n\,,$$

onde $A_r(x_r^1, \dots, x_r^n) = \phi_r(p) \cdot a_r(x_r(p))$. Ainda aqui, ponhamos $U_r' = U_r \cap (M - F) = U_r - (U_r \cap F)$. A coleção $\{U_r'\}$ é uma cobertura aberta e localmente finita da variedade $M - F$. Por hipótese, $x_r(U_r \cap F)$ é de medida nula em R^n, qualquer que seja r. Como x_r é um homeomorfismo, é claro que

$$x_r(U_r') = x_r(U_r) - x_r(U_r \cap F).$$

Nessas condições, uma bem conhecida propriedade das integrais em R^n permite que escrevamos

$$\int_{x_r(U'_r)} = \int_{x_r(U_r)} .$$

Segue-se daí que

$$\int_M \omega = \sum_{r=1}^{\infty} \int_{x_r(U_r)} A_r(x_r^1, \ldots, x_r^n)\, dx_r^1 \ldots dx_r^n = \int_{M-F} \omega,$$

conforme queríamos provar.

Retomemos as variedades M^n, N^n, orientadas, fechadas e riemannianas, N sendo conexa, e a aplicação diferenciável $f\colon M \to N$. Seja $C \subset N$, o conjunto dos valores críticos de f. A respeito da imagem inversa de C, podemos escrever: $f^{-1}(C) = C_0 \cup A_0$, onde

$$C_0 = \{p \in M; J(p) = 0\},\ A_0 = \{p \in M; J(p) \neq 0 \text{ e } f(p) \in C\}.$$

Com $J(p)$ indicamos o jacobiano de f em p, calculado em termos de sistemas genéricos de coordenadas locais nas duas variedades. O conjunto C_0 é evidentemente fechado em M, e, portanto, compacto. Segue-se que $C = f(C_0)$ é compacto, donde também fechado em N. Resulta daí que $f^{-1}(C)$ é fechado em M. Como $A_0 = (M - C_0) \cap f^{-1}(C)$, concluímos que A_0 é fechado em $M - C_0$.

O conjunto A_0 é de medida nula em M. Para prová-lo, basta mostrar que cada ponto $p \in A_0$ tem uma vizinhança U tal que $A_0 \cap U$ é de medida nula em M. Como a aplicação f é regular em p (lembrar que $J(p) \neq 0$), existe uma vizinhança $U \ni p$ que é aplicada pela f difeomorficamente sobre uma vizinhança V de $f(p)$ em N. É evidente que $f(A_0) \cap U = C \cap V$. Como C tem medida nula (teorema de Sard), o mesmo ocorre com $C \cap V$, e, portanto, também com $A_0 \cap U$ (porque a restrição $f|U$ é um difeomorfismo).

Consideremos, agora, a forma $f^*\sigma$ sobre M, antes definida. A expressão

$$(f^*\sigma)_p = \sqrt{g}\, J(p)\, dx^1 \wedge \cdots \wedge dx^n,$$

que já deduzimos, mostra claramente que $f^*\sigma$ se anula sobre o conjunto fechado C_0. Podemos aplicar o Lema 1 e escrever:

$$\int_M f^*\sigma = \int_{M-C_0} f^*\sigma.$$

O conjunto A_0 é fechado em $M - C_0$, e é de medida nula em $M - C_0$ (pois o é em M). De acordo com o Lema 2, resulta:

$$\int_{M-C_0} f^*\sigma = \int_{(M-C_0)-A_0} f^*\sigma.$$

Tendo em vista que $(M - C_0) - A_0 = M - (C_0 \cup A_0)$, concluímos que

$$\int_M f^*\sigma = \int_{M-(C_0\cup A_0)} f^*\sigma. \tag{1}$$

Consideremos agora uma cobertura $\{V_r\}$ da variedade $N - C$, enumerável e localmente finita, formada por vizinhanças regulares. É claro que $\{f^{-1}(V_r)\}$ é uma cobertura enumerável e localmente finita da variedade $M - f^{-1}(C) = M - (C_0 \cup A_0)$. Seja $\{\phi_r\}$ uma partição da unidade, subordinada à cobertura $\{V_r\}$. Então $\{\phi_r \circ f\}$ é uma partição da unidade, subordinada à cobertura $\{f^{-1}(V_r)\}$. De conformidade com a definição da integral por meio de uma partição da unidade, temos:

$$\int_{M-(C_0\cup A_0)} f^*\sigma = \sum_{r=1}^{\infty} \int_{M-(C_0\cup A_0)} (\phi_r \circ f) f^*\sigma. \qquad (2)$$

Observemos que a função real $\phi_r \circ f$ se anula fora de $f^{-1}(V_r)$ e, para cada $p \in f^{-1}(V_r)$, tem-se $(\phi_r \circ f)(p) = \phi_r(q)$, onde $q = f(p) \in V_r$. Notemos também que $\lambda f^*\sigma = f^*(\lambda\sigma)$, qualquer que seja $\lambda \in R$. Em vista desses fatos, podemos escrever:

$$\sum_{r=1}^{\infty} \int_{M-(C_0\cup A_0)} (\phi_r \circ f) f^*\sigma = \sum_{r=1}^{\infty} \int_{f^{-1}(V_r)} f^*(\phi_r\sigma). \qquad (3)$$

Ora, já mostramos, no estudo local previamente feito, que

$$\int_{f^{-1}(V_r)} f^*\sigma = \gamma \int_{V_r} \sigma.$$

Um mero exame na demonstração desse resultado evidencia que

$$\int_{f^{-1}(V_r)} f^*(\phi_r\sigma) = \gamma \int_{V_r} \phi_r\, \sigma.$$

Segue-se daí que

$$\sum_{r=1}^{\infty} \int_{f^{-1}(V_r)} f^*(\phi_r\sigma) = \sum_{r=1}^{\infty} \gamma \int_{V_r} \phi_r\, \sigma. \qquad (4)$$

Considerando que a função real ϕ_r se anula fora de V_r, e levando em conta, outra vez, a definição da integral, resulta:

$$\sum_{r=1}^{\infty} \gamma \int_{V_r} \phi_r \sigma = \sum_{r=1}^{\infty} \gamma \int_{N-C} \phi_r \, \sigma =$$
$$= \gamma \sum_{r=1}^{\infty} \int_{N-C} \phi_r \sigma = \gamma \int_{N-C} \sigma. \tag{5}$$

Por ser C um conjunto de medida nula em N (teorema de Sard), podemos empregar novamente o Lema 2, e concluir:

$$\int_{N-C} \sigma = \int_{N} \sigma. \tag{6}$$

Examinando, finalmente, as igualdades $(1), \ldots, (6)$, chegamos à conclusão que buscávamos:

$$\int_M f^* \sigma = \gamma \int_N \sigma.$$

Não vamos descrever aqui as aplicações da teoria do grau, pois estaríamos fugindo do assunto ao qual dedicamos este trabalho. Registremos, porém, que tais aplicações existem, e algumas são muito úteis e belas. Entre estas podemos citar: 1) a classificação homotópica das aplicações contínuas $f\colon M^n \to S^n$, onde M^n é uma variedade diferenciável orientada, compacta e conexa, e S^n é a esfera n-dimensional; 2) o teorema da curvatura integral de Hopf.

O leitor interessado encontrará em [8] uma bela descrição dessas duas aplicações da teoria do grau.

BIBLIOGRAFIA

[1] Boas (Jr.), Ralph, P. - A primer of real functions. The Carus Mathematical Monographs, The Mathematical Association of America. 1960.

[2] Bourbaki, N. Topologie Générale, Chapitres I, II, deuxième édition, Hermann, Paris, 1951. Chapitre IX, deuxième édition, Hermann, Paris, 1958.

[3] Bourbaki, N. Algèbre, Chapitre II (Algèbre Linéaire) Hermann, Paris, 1947. Chapitre III (Algèbre Multilinéaire), Hermann, Paris, 1948.

[4] Chern, S.S. Differentiable Manifolds. Textos de Matemática, n. 5 - Instituto de Física e Matemática da Universidade do Recife.

[5] Hönig, Chaim S. Aplicações da Topologia à Análise. Textos de Matemática, n. 8 - Instituto de Física e Matemática da Universidade do Recife, 1961.

[6] Lima, Elon Lages Introdução às Variedades Diferenciáveis. Instituto de Matemática da Universidade do Rio Grande do Sul, 1960.

[7] Lima, Elon Lages Topologia dos Espaços Métricos. Notas de Matemática, n. 10 - Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1954.

[8] Lima, Elon Lages Introdução à Topologia Diferencial. Notas de Matemática, n. 22 - Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1961.

[9] MacLane, Saunders Curso de Topologia Geral. Tradução de Joviano C. Valadares - Notas de Matemática, n. 11 - Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1954.

[10] Milnor, John Differential Topology - apostilas redigidas por J. Munkres - Princeton University, 1958.

[11] Morse, Anthony P. The behavior of a function on its critical set. Annals of Mathematics, vol. 40 (1939), pp. 62-70.

[12] Pontrjagin, Leon S. Smooth manifolds and their applications in homotopy theory. Amer. Math. Soc. Translations, Series 2, vol. 11, 1959.

[13] de Rham, Georges Variétés Différentiables. Hermann, Paris, 1955.

[14] Rudin, Walter Principles of Mathematical Analysis. McGraw-Hill Book Company, 1953.

[15] Sard, Arthur The measure of the critical values of differentiable maps. Bulletin of the American Mathematical Society, vol. 48 (1942), pp. 883-890.

[16] Whitney, Hassler A function not constant on a connected set of critical points. Duke Mathematical Journal, vol. 1 (1935), pp. 514-517.