publicações matemáticas

Cálculo Tensorial

Elon Lages Lima

Cálculo Tensorial

Publicações Matemáticas

Cálculo Tensorial

Elon Lages Lima
IMPA

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 - E-mail: ddic@impa.br - http://www.impa.br ISBN: 978-85-244-0313-2

Prefácio da primeira edição

Com exceção do capítulo final, estas são notas de aula de um curso que lecionei duas vezes, em 1960 e em 1962. Os três capítulos iniciais são baseados numa redação do primeiro curso, feita por J. Ubyrajara Alves. O quarto capítulo, dado somente na segunda vez, foi redigido por Alciléa Augusto. J.B. Pitombeira colaborou na redação do primeiro capítulo. Esses amigos, a quem agradeço agora, são responsáveis pelo aparecimento das notas, mas certamente não pelos erros e defeitos nelas contidos, dos quais sou o único autor.

A idéia aqui é a de apresentar uma introdução, moderna mas sem "modernismos", ao Cálculo e à Análise Tensoriais. No que tange ao primeiro (onde nos restringimos aos espaços vetoriais reais de dimensão finita e assim tiramos partido das várias simplificações que estas hipóteses acarretam) a exposição é bastante para efeitos da Geometria Diferencial e da Análise. De propósito, não foram mencionados os produtos tensoriais de módulos, e assim a introdução aqui apresentada não serve senão de motivação às teorias gerais da Álgebra Homológica. Quanto à Análise

Tensorial, mal foi arranhada a superfície. Foi feito um começo de introdução, no Capítulo 4, levando ao Teorema de Stokes como resultado final e, no Capítulo 5, foi demonstrado o Teorema de Frobenius sob o ponto de vista do colchete de Lie de dois campos vetoriais. As bases foram lançadas solidamente, com discussões dos conceitos fundamentais, tais como variedades diferenciáveis, orientabilidade, partições da unidade, integração de formas diferenciais, diferencial exterior e Teorema de Stokes. Mas, em obediência ao caráter essencialmente introdutório das notas, nenhum desenvolvimento mais profundo é tentado. No fim do trabalho, uma lista de indicações bibliográficas é apresentada, como desencargo de consciência.

O Cálculo Tensorial Clássico, em sua parte estritamente algébrica, é pouco mais do que um repertório de trivialidades. Isto se reflete na natureza do Capítulo 2, onde é feita uma apresentação intrínseca e conceitual dos tensores. Nota-se ali uma ausência conspícua de teoremas "astutos". É interessante contrastar este fato com o Capítulo 3, onde é estudada a Álgebra de Grassmann e onde surgem aplicações interessantes à Teoria dos Determinantes e à Geometria. No Capítulo 4, que relaciona com a Geometria Diferencial os fundamentos algébricos lançados nos três primeiros capítulos, os únicos tensores usados são os anti-simétricos do Capítulo 3.

Nestas notas nunca aparece o famoso "delta de Kronecker" nem é adotada a chamada "convenção de Einstein". O primeiro poderia trazer alguma vantagenzinha ao nos poupar de escrever duas ou três palavras a mais. A segunda não só é desnecessária mas é francamente absurda. As atrapalhações que causa não compensam o trabalho de se escrever

um $\sum$ aqui e alí. Ao deixar de lado essas notações, prestamos nossa singela homenagem a Kronecker e Einstein, que devem ser lembrados por motivos mais sérios. Por outro lado, adotamos o uso clássico de, sempre que possível, pôr em diferentes alturas os índices repetidos. Assim procedendo, poupam-se erros quando se escrevem os somatórios: primeiro põem-se as letras principais e depois preenchem-se os índices, de modo que "dê certo".

Brasília, 6 de novembro de 1964

Elon Lages Lima

Prefácio da segunda edição

Para esta edição, publicada tantos anos após a primeira, o texto mimeografado original foi digitado eletronicamente e algumas correções tipográficas foram feitas. A notação e a terminologia anteriores foram mantidas.

Agradeço a Arturo Ulises Fernández Pérez e a Renan Finder pela revisão final do texto.

Rio de Janeiro, outubro de 2010

Elon Lages Lima

Prefácio da terceira edição

Este livro apareceu inicialmente, sob forma mimeografada, na coleção "Notas de Matemática", hoje extinta. Muitos anos depois, algumas pessoas me convenceram que valeria a pena uma republicação. O texto foi então digitado eletronicamente e incluído na série "Publicações Matemáticas". Esse processo fez ocorrerem diversos erros tipográficos, os quais foram cuidadosamente apontados pelo Professor Carlos Matheus, a quem agradeço. As devidas correções foram incorporadas nesta terceira edição.

Rio de Janeiro, abril de 2012

Elon Lages Lima

Sumário

Capítulo 1

Espaços Vetoriais

Faremos, neste capítulo, uma revisão sucinta dos conceitos básicos da teoria dos espaços vetoriais, com a finalidade de fixar terminologia e notações que usaremos no decorrer deste curso. Tencionamos ser breve nesta recapitulação, ao mesmo tempo em que nos esforçamos para dar ao leitor os conhecimentos de Álgebra Linear necessários a uma leitura proveitosa destas notas. Para uma exposição mais detalhada, o leitor poderá consultar o texto de Álgebra Linear que será publicado pelo autor, daqui a trinta anos, na coleção "Matemática Universitária" do IMPA.

1.1 Noção de espaço vetorial

Chamaremos de *escalares* os elementos do corpo T dos números reais; assim faremos porque nos interessam apenas os espaços vetoriais sobre o corpo dos reais.

Um *espaço vetorial* V é um conjunto de elementos, denominados *vetores*, satisfazendo aos seguintes axiomas:

a) A todo par u, v de vetores de V corresponde um vetor $u+v$, chamado *soma* de u e v de tal maneira que:

1) $u+v=v+u$ (comutatividade da adição);

2) $u+(v+w)=(u+v)+w$ (associatividade da adição);

3) existe em V um vetor 0 (denominado vetor nulo, zero ou origem) tal que $u+0=u$ para todo u em V;

4) a cada vetor u em V corresponde um vetor $-u$ tal que $u+(-u)=0$.

b) A todo par λ, u, onde λ é um escalar e u um vetor em V corresponde um vetor λu, chamado *produto* de λ e u, de tal maneira que:

1) $\alpha(\beta u)=(\alpha\beta)u$ (a multiplicação por escalar é associativa);

2) $1\cdot u=u$, para todo u em V;

3) $\lambda(u+v)=\lambda u+\lambda v$ (distributividade da multiplicação por escalar em relação à soma de vetores);

4) $(\alpha+\beta)u=\alpha u+\beta u$.

Decorrem dos axiomas acima algumas regras formais de cálculo algébrico nos espaços vetoriais, que são inteiramente análogas às regras de operações com números. A única diferença é que não se postula a possibilidade de multiplicar vetores. Daremos alguns exemplos. Outras regras adicionais serão livremente usadas a seguir.

Usamos o mesmo símbolo 0 para indicar o escalar nulo e o vetor nulo em V. Tal prática nunca originará confusão.

(i) *Se* $u + w = v + w$, *então* $u = v$.

Com efeito, somando $-w$ a ambos os membros, vem: $(u+w)+(-w) = (v+w)+(-w)$ ou seja, $u+(w+(-w)) = v + (w + (-w))$, donde $u + 0 = v + 0$ e, finalmente, $u = v$.

Para abolir o pedantismo, nunca usaremos parênteses em expressões do tipo $(u + v) + w$, já que o significado de $u + v + w$ é unívoco, em virtude da associatividade. Analogamente, escreveremos $u - v$, em vez de $u + (-v)$.

(ii) *Para todo* $v \in V$, *e todo escalar* α, *tem-se* $0 \cdot v = 0$ *e* $\alpha \cdot 0 = 0$.

Em virtude de (i), basta mostrar que $0 \cdot v + v = v$ e $\alpha \cdot 0 + \alpha \cdot 0 = \alpha \cdot 0$. Em primeiro lugar, temos $0 \cdot v + v = \cdot v + 1 \cdot v = (0 + 1)v = 1 \cdot v = v$. No segundo caso, $\alpha \cdot 0 + \alpha \cdot 0 = \alpha(0 + 0) = \alpha \cdot 0$, como desejávamos.

(iii) *Para todo* $v \in V$, *vale* $-1 \cdot v = -v$.

Com efeito, $v + (-1)v = 1 \cdot v + (-1)v = (1 + (-1))v = 0 \cdot v = 0$. Logo, em vista de (i), temos $-1 \cdot v = -v$.

Segue-se que $(-\alpha)v = \alpha(-v) = -(\alpha v)$, para todo escalar α, pois $(-\alpha)v = (-1 \cdot \alpha)v = -1 \cdot (\alpha v) = -(\alpha v)$ e $\alpha(-v) = \alpha(-1 \cdot v) = (\alpha(-1))v = (-\alpha)v$.

(iv) *Se* $\alpha \cdot v = 0$, *então* $\alpha = 0$ *ou* $v = 0$.

Com efeito, se fosse $\alpha \neq 0$ e $v \neq 0$, então $\frac{1}{\alpha}(\alpha \cdot v) = v \neq 0$, logo não poderia ser $\alpha \cdot v = 0$, por causa de (ii).

Exemplos de Espaços Vetoriais

1) Para cada inteiro $n > 0$, indicaremos com R^n o conjunto de todas as n-uplas ordenadas $(\alpha^1 \ldots \alpha^n)$ de números reais.

Dados $u = (\alpha^1, \ldots, \alpha^n)$ e $v = (\beta^1 \ldots \beta^n)$ em R^n, definiremos a soma $u + v$ e o produto λu, de u por um escalar λ, como

$$u + v = (\alpha^1, \beta^1, \ldots, \alpha^n + \beta^n),$$
$$\lambda u = (\lambda\alpha^1, \ldots, \lambda\alpha^n).$$

Considerando $0 = (0, \ldots, 0) \in R^n$, temos $u + 0 = u$ para todo $u \in R^n$. Pondo $-u = (-\alpha^1, \ldots, -\alpha^n)$, temos $u+(-u) = 0$. Os demais axiomas de um espaço vetorial são facilmente verificados. R^n é portanto um espaço vetorial. Quando $n = 1$, temos $R^1 = R =$ reta real. Para $n = 2$, temos $R^2 =$ plano numérico. R^3 é o "espaço numérico tri-dimensional" da geometria analítica clássica.

2) Seja $C(X)$ o conjunto das funções reais contínuas $f\colon X \to R$, definidas em um subconjunto X qualquer da reta. Definimos a soma de duas funções $f, g \in C(X)$ e o produto λf de $f \in C(X)$ por um escalar λ da maneira óbvia:

$$(f + g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda \cdot f(x).$$

Deste modo, $C(X)$ torna-se um espaço vetorial, cujo zero é a função identicamente nula.

Para fixar as idéias, podemos imaginar, no exemplo

acima, $X = [0,1]$ = intervalo fechado unitário. Escreveremos então simplesmente C em vez de $C([0,1])$.

3) Seja P o conjunto dos polinômios reais $p = a_0 + a_1 t + \cdots + a_m t^m$. Com a soma e o produto por um escalara definidos da maneira habitual, P torna-se um espaço vetorial. Do mesmo modo, se fixarmos um inteiro n e considerarmos apenas o conjunto P_n dos polinômios de grau $\leq n$, veremos que P_n é ainda um espaço vetorial relativamente às mesmas operações.

4) Consideremos o plano euclidiano Π, da Geometria Clássica. Como se sabe, cada par de pontos $A, B \in \Pi$ determina um segmento de reta, cujas extremidades são estes pontos (quando $A = B$ este "segmento" reduz-se ao próprio ponto A.) Diremos que tal segmento é um segmento *orientado* quando escolhermos um dos seus extremos para chamá-lo de *origem* e chamarmos o outro extremo de *fim*. Convencionaremos escrever $\overline{AB}$ para indicar o segmento orientado de origem A e fim B. Se quisemos tomar B como origem, escreveremos $\overline{BA}$.

Diremos que os segmentos orientados $\overline{AB}$ e $\overline{CD}$ são *equipolentes* se eles forem "paralelos, de mesmo comprimento e de mesmo sentido", isto é, se $\overline{AB}$ e $\overline{CD}$ são lados opostos de um paralelogramo, no qual AC e BD formam o outro par de lados opostos.

Escreveremos $\overline{AB} \equiv \overline{CD}$ para indicar que o segmento orientado $\overline{AB}$ é equipolente a $\overline{CD}$. Verifica-se que a relação $\overline{AB} \equiv \overline{CD}$ é reflexiva, simétrica e transitiva. Indicaremos com $\overrightarrow{AB}$ o conjunto de todos os segmentos orientados equipolentes a $\overline{AB}$. Assim $\overrightarrow{AB} = \overrightarrow{CD}$ é o mesmo que

$\overline{AB} \equiv \overline{CD}$.

Chamaremos o conjunto $\overrightarrow{AB}$ o *vetor livre* determinado por A e B. Indicaremos com L_2 o conjunto de todos os vetores livres do plano Π. Dados dois vetores livres $u, v \in L_2$, e escolhido um ponto arbitrário $A \in \Pi$, pode-se sempre escrever $u = \overrightarrow{AB}$, $v = \overrightarrow{AC}$, com $B, C \in \Pi$ univocamente determinados.

Definimos então $u+v = \overrightarrow{AD}$, onde D é o quarto vértice do paralelogramo que tem A, B, C como vértices restantes. Definimos também o produto λu de um escalar λ por um vetor $u = \overrightarrow{AB}$ como o vetor $u = \overrightarrow{AC}$, onde C está sobre a reta AB, o comprimento $\lambda|AC|$ é igual a $\lambda|AB|$ e A está entre B e C ou não, conforme $\lambda < 0$ ou não. Estas operações fazem de L_2 um espaço vetorial: *o espaço dos vetores livres do plano.*

1.2 Bases de um espaço vetorial

Um sistema de vetores $\{x_1, \ldots, x_k\}$ em V é dito *linearmente dependente* se existe uma k-upla $(\alpha^1, \ldots, \alpha^k)$ de escalares nem todos nulos, tais que $\alpha^1 x_1 + \cdots + \alpha^k x_k = 0$. Se $\alpha^1 x_1 + \cdots + \alpha^k x_k = 0$ implica $\alpha^i = 0$, $i = 1, \ldots, k$, o sistema $\{x_1, \ldots, x_k\}$ diz-se *linearmente independente.*

Proposição 1. *Um sistema $\{x_1, \ldots, x_k\}$, com $k \geq 2$ é linearmente dependente se, e somente se, existe pelo menos um vetor x_h , $h \geq 2$ que é combinação linear dos vetores precedentes.*

Demonstração: Como a suficiência é óbvia, demonstraremos somente a necessidade da condição. Seja $(\alpha^1, \ldots, \alpha^k)$ uma k-upla de escalares não todos nulos tais que $\sum_{i=1}^{k} \alpha^i x_i = 0$ e α^h o último α^i não-nulo, então $\alpha^h x_h = \sum_{i<h} -\alpha^i x_i$ e portanto $x_h = \sum_{i<h} -\frac{\alpha^i}{\alpha^h} x_i$. Agora basta tomar $\beta^i = -\frac{\alpha^i}{\alpha^h}$ para chegarmos ao resultado desejado.

Um sistema finito $\{e_1, \ldots, e_n\}$ de vetores de um espaço vetorial V diz-se uma *base* de V se

1) é linearmente independente e

2) gera V, isto é, todo vetor $v \in V$ se exprime como uma combinação linear dos elementos do sistema.

Quando existe uma base $\{e_1, \ldots, e_n\} \subset V$, dizemos que V é um espaço vetorial de *dimensão finita.* No que se segue, consideraremos apenas espaços vetoriais reais, de dimensão finita, exceto quando for feita uma menção explícita em contrário.

Proposição 2. *Um sistema de vetores $\{e_1, \ldots, e_n\}$ de um espaço vetorial V constitui uma base se, e somente se, qualquer vetor v do espaço as exprime de maneira única como combinação linear dos elementos do sistema.*

Demonstração: De fato, se $v = \sum_{i=1}^{n} \alpha^i e_i$ e $v = \sum_{i=1}^{n} \beta^i e_i$, por subtração obtemos $0 = \sum_{i=1}^{n} (\alpha^i - \beta^i) e_i$ e, como o sistema $\{e_1, \ldots, e_n\}$ é linearmente independente, $\alpha^i - \beta^i = 0$,

$i = 1, \ldots, n$, portanto $\alpha^i = \beta^i$. Isso demonstra metade da proposição. Reciprocamente, se o sistema $\{e_1, \ldots, e_n\}$ de vetores de V é tal que todo vetor $v \in V$ se exprime, de maneira única, como combinação linear dos seus elementos, então o sistema $\{e_1, \ldots, e_n\}$ é linearmente independente e portanto uma base. Isso resulta do fato de que o vetor 0 se exprime de uma maneira única como combinação linear dos elementos do sistema $\{e_1, \ldots, e_n\}$. De fato, escrever $\sum_{i=1}^{n} \alpha^i e_i = 0$, é exprimir o vetor 0 como combinação linear dos e_i. Por outro lado, $0e_1 + \cdots + 0e_n = 0$. Pela unicidade da maneira de exprimir o vetor 0 como combinação linear dos e_i, obtemos $\alpha^i = 0$, $i = 1, \ldots, n$.

Nossa intenção é definir a *dimensão* de um espaço vetorial, de dimensão finita V, como o número de elementos de uma ase qualquer de V. Para isso devemos estar certos de que este número é o mesmo, qualquer que seja a base tomada. É o que nos assegura o seguinte

Teorema 1. *Duas bases quaisquer do mesmo espaço vetorial de dimensão finita têm o mesmo número de elementos.*

Demonstração: Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base do espaço vetorial V. Mostraremos que todo conjunto $\{f_1, \ldots, f_m\}$ com um número $m > n$ de vetores de V é linearmente dependente e portanto não pode ser uma base de V. Daí resultará que duas bases quaisquer de V possuem o mesmo número de elementos. Procuraremos, portanto, achar escalares $\beta^1, \ldots, \beta^m$, não todos nulos, tais que $\sum_{j=1}^{m} \beta^j f_j = 0$. Como $\mathcal{E}$ é uma base, temos $f_j = \sum_{i=1}^{n} \alpha_j^i e_i$, para cada

$j = 1, 2, \ldots, m$. Devemos então obter os escalares β^j de modo que $\sum_{j=1}^{m} \beta^j \left(\sum_{i=1}^{n} \alpha_j^i e_i \right) = 0$, ou seja,

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{m} \beta^j \alpha_j^i \right) e_i = 0.$$

Como os vetores e_i são linearmente independentes, deveremos ter $\sum_{j=1}^{m} \beta^j \alpha_j^i = 0$, $i = 1, \ldots, n$. Isto é:

$$\begin{aligned} \alpha_1^1 \beta^1 + \cdots + \alpha_m^1 \beta^m &= 0 \\ \alpha_1^2 \beta^1 + \cdots + \alpha_m^2 \beta^m &= 0 \\ \vdots \qquad\qquad & \qquad \vdots \\ \alpha_1^n \beta^1 + \cdots + \alpha_m^n \beta^m &= 0 \end{aligned}$$

Usando a hipótese $m > n$, esse sistema de n equações nas m incógnitas β^j possui pelo menos uma solução não trivial β_0^j, $j = 1, \ldots, m$. Portanto, $\sum_{i=1}^{m} \beta_0^f f_j = 0$ sem que todos os coeficientes β_0^j sejam nulos. □

Agora, podemos definir *dimensão* de um espaço vetorial V, de dimensão finita, como o *número de elementos de uma base de* V.

Uma demonstração do Teorema 1, onde não se usam resultados de equações lineares, pode ser encontrada nos livros de Birkhoff-Mac Lane e Halmos citados no início deste capítulo. Colocamos aqui esta demonstração a fim de chamar a atenção do leitor pra a estreita relação existente entre os espaços vetoriais de dimensão finita e os sistemas de equações lineares com um número finito de incógnitas.

Exemplo 5. O leitor poderá verificar com facilidade que o sistema $\{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$ constitui uma base do espaço vetorial do Exemplo 1. Também o conjunto $\{1,t,t^2,\ldots,t^n\}$, formado de polinômios, constitue uma base do espaço vetorial P_n do Exemplo 3. Essas bases são denominadas *bases canônicas* ou *naturais*, pois, como podemos observar, elas se impõem naturalmente a partir da própria definição dos espaços R^n e P_n. Aproveitamos a oportunidade para observar que o espaço vetorial C do Exemplo 2 não é de dimensão finita.

1.3 Isomorfimos

Consideremos dois espaços vetoriais V e W. Uma aplicação $T\colon V \to W$ é chamada *transformação linear* quando

$T(u+v) = T(u) + T(v)$ para todo par $u, v \in V$;

$T(\lambda u) = \lambda T(u)$ para todo escalar λ e todo $u \in V$.

Proposição 3. *Seja $T\colon V \to W$ uma transformação linear. Então:*

a) $T(0) = 0$

b) *T é biunívoca se, e só se, $T(x) = 0$ implica $x = 0$.*

Demonstração: a) Temos que: $T(0) = T(0,0) = 0$, $T(0) = 0$.

b) Em primeiro lugar, T biunívoca e $x \neq 0$ implicam que $T(x) \neq T(0) = 0$. Para a recíproca, se $T(x) = T(y)$, segue-se que $T(x-y) = 0$, donde $x - y = 0$, e $x = y$.

Exemplos: 6) Tomando $V = W = R^n$ a aplicação $T(x) = ax$, onde a é um escalar, é uma transformação linear (chamada *homotetia* de razão a).

7) A aplicação $\pi_i \colon R^n \to R$, que ao vetor $x \in R^n$ associa sua i-ésima coordenada, é também uma transformação linear (chamada *i-ésima projeção*).

Uma transformação linear $I \colon V \to W$ satisfazendo às condições:

1) $I(x) = 0$ implica $x = 0$;

2) para todo $w \in W$, existe um $v \in V$ tal que $I(v) = w$; é chamada de *isomorfismo*. Neste caso, V e W são ditos *isomorfos*.

Um isomorfismo $I \colon V \to W$ é portanto uma transformação linear biunívoca de V *sobre* W.

Exemplos: 8) A transformação linear do Exemplo 6 é um isomorfismo, desde que tomemos o escalar a não-nulo, enquanto a aplicação $\pi_i \colon R^n \to R$ com $n > 1$, do Exemplo 7 não é um isomorfismo, pois não é biunívoca, como facilmente se verifica.

9) Escolhamos uma base no espaço L_2, do Exemplo 4. Então, vê-se facilmente que L_2 é isomorfo a R^2, pela aplicação que a cada $u \in L_2$ associa suas coordenadas na base prefixada em L_2.

Um isomorfismo I entre dois espaços vetoriais preserva todas as propriedades desses espaços que são definidas a partir dos axiomas. Por exemplo, se $I \colon V \to W$ é um isomorfismo e $\{e_1, \ldots, e_n\}$ é uma base de V, então o sistema

$$\{f_1 = I(e_1), \ldots, f_n = I(e_n)\}$$

é uma base de W. Com efeito, o sistema $\{f_1, \ldots, f_n\}$ é linearmente independente pois, se $\alpha_1 f_1 + \cdots + \alpha_n f_n = 0$, ou

seja $\alpha_1 I(e_1) + \cdots + \alpha_n I(e_n) = 0$, pela linearidade de I, obtemos $I(\alpha_1 e_1 + \cdots + \alpha_n e_n) = 0$. Como I é um isomorfismo, $I(\alpha_1 e_1 + \cdots + \alpha_n e_n) = 0$ implica $\alpha_1 e_1 + \cdots + \alpha_n e_n = 0$; logo $\alpha_i = 0$, pois $\{e_1, \ldots, e_n\}$ é uma base. Dado $w \in W$, como I é um isomorfismo, existe um

$$v = \sum_{i=1}^{n} \beta^i e_i$$

pertencente a V tal que $I(v) = W$, isto é,

$$w = I(v) = I\left(\sum_{i=1}^{n} \beta^i e_i\right) = \sum_{i=1}^{n} \beta^i I(e_i) = \sum_{i=1}^{n} \beta^i f_i\,;$$

w se escreve, portanto, como combinação linear dos elementos do sistema $\{f_1, \ldots, f_n\}$ e nossa afirmação está totalmente demonstrada.

Proposição 4. *Sejam V e W espaços vetoriais de mesma dimensão n e $A\colon V \to W$ uma transformação linear. As seguintes afirmações são equivalentes:*

i) *A é um isomorfismo;*

ii) *A é biunívoca;*

iii) *A é sobre W.*

Demonstração: É evidente que i) implica ii). Mostremos que ii) implica iii). Suponhamos, por absurdo, que exista um vetor $w \in W$ que não seja da forma $w = A(v)$ com $v \in V$. Então, se $\mathcal{E} = \{e_1, \ldots, e_n\}$ é uma base de V, w não é combinação linear dos vetores $f_1 = A(e_1), \ldots, f_n = A(e_n)$,

pois $w = \Sigma\, \beta^i f_i$ implicaria $w = A(v)$, com $v = \Sigma\, \beta^i e_i$. Ora, a biunivocidade de A acarreta imediatamente que os f_i são linearmente independentes e portanto, pela Proposição 1, $\{f_1, \ldots, f_n, w\}$ é um sistema de $n+1$ vetores linearmente independentes no espaço W, que tem dimensão n. Isto contraria a demonstração do Teorema 1. Logo A é sobre W. De maneira semelhante, mostra-se que iii) implica ii). Como i) é a reunião de ii) e iii), segue-se que i), ii) e iii) são equivalentes.

Convém observar que este teorema só é válido se V e W têm mesma dimensão. As projeções constituem um exemplo de aplicações sobre que não são biunívocas. A imersão de uma reta no plano mostra uma aplicação linear biunívoca que não é sobre.

Mostraremos agora que dois espaços vetoriais de mesma dimensão são isomorfos. Como o composto de dois isomorfismos ainda é um isomorfismo, é suficiente provar o seguinte

Teorema 2. *Todo espaço vetorial V de dimensão n é isomorfo ao R^n.*

Demonstração: Seja $\{e_1, \ldots, e_n\}$ uma base de V. Definamos uma aplicação I que ao vetor $v = \sum_{i=1}^{n} \alpha^i e_i$ de V associa a n-upla $(\alpha^1, \ldots, \alpha^n)$ de R^n. Esta aplicação está bem definida, pois v se exprime de maneira única como combinação linear dos e_i, e estabelece o isomorfismo desejado, como facilmente se verifica.

Como dois espaços vetoriais isomorfos não podem ser distinguidos por nenhuma propriedade definida a partir dos axiomas, o teorema acima parece indicar que basta estudar

os espaços vetoriais R^n. Tal porém não é o caso. Dado um espaço vetorial abstrato V, de dimensão n, os vários isomorfismos entre V e R^n dependem da escolha de bases em V. Se a escolha de uma base $\mathcal{E}$ define, como no teorema acima, um isomorfismo I de V em R^n, que leva um certo vetor $v \in V$ em $I(v) = (\alpha^1, \ldots, \alpha^n)$, então a escolha de outra base $\mathcal{F}$ define um novo isomorfismo J tal que $J(v) = (\beta^1, \ldots, \beta^n)$ onde os β^i podem ser distintos dos α^i. Assim, não é permitido identificar V com R^n, pois dado um vetor $v \in V$ não se pode saber exatamente a n-upla que a ele corresponde, já que nenhuma base de V se destaca das demais para a definição do isomorfiswmo em questão.

Diremos que um isomorfismo I de V em W é *canônico* se a definição de I não envolve escolhas arbitrárias. Por exemplo, o isomorfismo entre V e R^n definido no teorema anterior não é canônico, pois depende da escolha de uma base. Apresentaremos a seguir um exemplo de isomorfismo canônico.

Exemplo 10. O conjunto $\mathcal{C}$ dos números complexos constitui um espaço vetorial com as operações de soma e produto por um número real. A aplicação $I: \mathcal{C} \to R^2$ que ao complexo $a + bi$ associa o par $(a, b) \in R^2$ é um isomorfismo canônico. Assim, podemos dizer que um número complexo pode ser pensado como um par de número reais.

Estudaremos adiante um outro exemplo importante de espaços isomorfos canonicamente a saber: um espaço vetorial de dimensão finita V com o seu "bi-dual" $= V^{**}$.

1.4 Mudanças de coordenadas

Vejamos, agor, qual a relação que existe entre as coordenadas de um vetor relativamente a uma base e as coordenadas do mesmo vetor relativamente a outra base.

Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ duas bases de um espaço vetorial V. Exprimindo os f_j com combinação linear dos e_i, obtemos

$$f_j = \sum_i \lambda_j^i e_i\,, \quad j = 1, \ldots, n.$$

O quadro formado pelos n^2 números λ_j^i, $i, j = 1, \ldots, n$, de acordo com o arranjo abaixo indicado, é o que se chama uma *matriz.*

$$\begin{pmatrix} \lambda_1^1 & \lambda_2^1 \ldots \lambda_n^1 \\ \lambda_1^2 & \lambda_2^2 \ldots \lambda_n^2 \\ \vdots & \vdots \\ \lambda_1^n & \lambda_2^n \ldots \lambda_n^n \end{pmatrix}$$

Note-se que λ_j^i está na i-ésima *linha* e na j-ésima *coluna.*

Nestas notas, adotaremos sistematicamente a seguinte convenção sobre as maneiras de escrever uma matriz: quando a matriz $\lambda = (\lambda_j^i)$ tiver seus elementos com um índice inferior e outro superior, o índice superior indicará sempre a *linha* e o inferior a *coluna* em que se encontra o elemento λ_j^i. Quando, porém, tivermos que escrever uma matriz $\lambda = (\lambda_{ij})$ com dois índices inferiores, ou $\lambda = (\lambda^{ij})$ com índices superiores, o primeiro índice dirá a linha e o segundo índice dirá a coluna em que se encontra o elemento λ_{ij} (ou λ^{ij}).

Mais exatamente, λ é chamada a *matriz de passagem* da base $\mathcal{E}$ para a base $\mathcal{F}$.

Tomemos um vetor $v \in V$ e exprimâmo-lo sucessivamente como combinação linear dos e_i e dos f_j:

$$v = \sum_i \alpha^i \, e_i \,, \qquad v = \sum_j \beta^j \, f_j \,.$$

Na segunda dessas igualdades, substituamos f_j pelo seu valor $f_j = \sum_i \lambda^i_j \, e_i$, obtendo

$$v = \sum_j \beta^j \left(\sum_i \lambda^i_j \, e_i \right) = \sum_j \left(\sum_j \lambda^i_j \, \beta^j \right) e_i \,.$$

Como v se exprime de maneira única como combinação linear dos e_i, comparando $v = \sum_i \alpha^i \, e_i$ com $v = \sum_i \left(\sum_j \lambda^i_j \, \beta^j \right) e_i$, podemos escrever:

$$\alpha^i = \sum_j \lambda^i_j \, \beta^j, \quad i = 1, \ldots, n.$$

Esta é a relação entre as coordenadas α^i, de v na base $\mathcal{E}$, e as coordenadas β^j, do mesmo vetor v na base $\mathcal{F}$. Note-se que a matriz λ, de passagem de $\mathcal{E}$ para $\mathcal{F}$ dá a passagem das coordenadas de v na base $\mathcal{F}$ para as coordenadas relativamente a $\mathcal{E}$. Houve uma inversão de sentido, por isso, os vetores de um espaço V são, às vezes, chamadas *vetores contravariantes.*

Em algumas situações, a noção de *base*, ou *referencial*, precede a de vetor. Quando isso se dá, os vetores são

definidos a partir dos referenciais, de tal maneira que as suas coordenadas nos diversos referenciais satisfaçam a "lei de transformação de coordenadas" expressa pela fórmula acima. Daremos, no Capítulo 4, um exemplo de uma situação concreta em que este fenômeno ocorre.

1.5 Espaço Dual

Uma função $f\colon V \to R$, com valores escalares, definida num espaço vetorial V, satisfazendo as condições

$$\begin{aligned} f(u+v) &= f(u) + f(v) \quad && u, v \in V \\ f(\alpha u) &= \alpha f(u) && u \in V \text{ e } \alpha \text{ escalar,} \end{aligned}$$

é denominada *funcional linear* ou *forma linear* sobre o espaço vetorial V. Chamaremos de V^* o conjunto dos funcionais lineares sobre o espaço vetorial V. É fácil verificar que se $f, g \in V^*$ e α é um escalar, então $f + g \in V^*$, $\alpha f \in V^*$, onde

$$\begin{aligned} (f+g)(u) &= f(u) + g(u), \\ (\alpha f)(u) &= \alpha f 9u). \end{aligned}$$

Com estas operações, V^* constitui um espaço vetorial denominado *espaço dual* de V. Da mesma maneira que os elementos de um espaço vetorial V são chamados *vetores contravariantes*, os elementos de V^* são chamados *vetores covariantes.*

Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base do espaço vetorial V. Um funcional linear $f \in V^*$ fica determinado quando conhecemos os n números:

$$f(e_1) = \alpha_1, \ldots, f(e_n) = \alpha_n\,.$$

De fato, se $v = \sum \xi^i e_i$ então

$$f(v) = f\left(\sum \xi^i e_i\right) = \sum \xi^i f(e_i) = \sum \xi^i \alpha_i .$$

Por outro lado, dada uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V, uma n-upla arbitrária de números $(\alpha_1, \ldots, \alpha_n)$ determina um funcional linear f a saber: o funcional definido por

$$f(v) = \sum \xi^i \alpha_i$$

onde os ξ^i são as coordenadas do vetor v na base $\mathcal{E}$.

A cada base $\mathcal{E} = \{e_1, \ldots, e_n\}$ do espaço vetorial V, corresponde uma base $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ do espaço vetorial V^*, chamada *base dual* de $\mathcal{E}$, assim definida: se $v = \sum \xi^i e_i$ então

$$e^i(v) = \xi^i, \qquad i = 1, \ldots, n,$$

isto é, e^i é o funcional linear que ao vetor v associa a sua i-ésima coordenada na base $\mathcal{E}$. Devemos demonstrar que $\mathcal{E}^*$ é, de fato, uma base. O sistema $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ é linearmente independe, pois $\sum \lambda_i e^i = 0$ significa

$$\left(\sum \lambda_i e^i\right)(v) = \sum \lambda_i e^i(v) = 0$$

qualquer que seja o vetor $v \in V$. Fazendo sucessivamente $v = e_j$, $j = 1, \ldots, n$, obtemos $\sum_i \lambda_i e^i(e_j) = 0$, $j = 1, \ldots, n$, e como

$$e^i(e_j) = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$$

a soma do primeiro membro se reduz a λ_j , portanto $\lambda_j = 0$, $j = 1, \ldots, n$. Por outro lado, todo funcional linear f é

combinação linear dos elementos de $\mathcal{E}^*$, pois de $f(v) = f\left(\sum \xi^i e_i\right) = \sum \xi^i f(e_i)$, como $\xi^i = e^i(v)$, obtemos $f(v) = \sum \alpha_i e^i(v) = \left(\sum \alpha_i e^i\right)(v)$, (fazendo $f(e_i) = \alpha_i$). Isto é, $f = \sum \alpha_i e^i$, o que demonstra totalmente nossa afirmação.

Exemplo 11. Sejam A um aberto do R^n e f uma função real, diferenciável, definida em A. A maneira mais conveniente de definir a *diferencial* de f em um ponto $p \in A$ é como um funcional linear sobre o espaço vetorial R^n. É o que faremos a seguir.

A derivada $\frac{\partial f}{\partial v}(p)$ da função f, no ponto p, relativamente a um vetor $v \in R^n$ é definida por

$$\lim_{t \to 0} \frac{f(p+tv) - f(p)}{t}.$$

Se $v = (\alpha^1, \ldots, \alpha^n)$, demonstra-se facilmente, usando a regra de derivação das funções compostas, que

$$\frac{\partial f}{\partial v}(p) = \sum_j \frac{\partial f}{\partial x^i}(p)\, \alpha^i.$$

Usando-se esta última expressão, é fácil verificar que $\frac{\partial f}{\partial v}(p)$ goza das seguintes propriedades:

1) $$\frac{\partial f}{\partial (v+w)}(p) = \frac{\partial f}{\partial v}(p) + \frac{\partial f}{\partial w}(p), \quad v, w \in R^n,$$

2) $$\frac{\partial f}{\partial (\lambda v)}(p) = \lambda \frac{\partial f}{\partial v}(p), \quad v \in R^n \text{ e } \lambda \text{ escalar.}$$

Assim, dados p e f, a aplicação $v \to \frac{\partial f}{\partial v}(p)$, de R^n em R é um funcional linear. Esta é, por definição, a diferencial

de f no ponto p, a qual se indica com df_p. Desta maneira, df_p aplicada no vetor $v = (\alpha^1, \ldots, \alpha^n)$ assume o valor

$$\frac{\partial f}{\partial v}(p) \sum_i \frac{\partial f}{\partial x^i}(p)\, \alpha^i, \quad \text{isto é,}$$

$$df_p(v) = \sum_i \frac{\partial f}{\partial x^i}\, \alpha^i.$$

Como sabemos, o espaço R^n possui uma base *canônica*, formada pelos vetores $e_1{=}(1, 0, \ldots, 0) \ldots, e_n{=}(0, \ldots, 0, 1)$. As *funções coordenadas* $x^i\colon A \to R$, que associam a cada ponto $q \in A$ sua i-ésima coordenada $x^i(q)$, fornecem as diferenciais $dx^1, \ldots, dx^n$, as quais, no ponto $p \in A$, são, como já vimos, funcionais lineares $dx^j \in (R^n)^*$. Afirmamos que $dx^1, \ldots, dx^n$, é a base dual da base canônica $\{e_1, \ldots, e_n\}$. Com efeito, se $v = (\alpha^1, \ldots, \alpha^n)$, vem que

$$dx^j(v) = \sum_i \frac{\partial x^j}{\partial x^i}\, \alpha^i, \quad j = 1, \ldots, n.$$

Como

$$\frac{\partial x^j}{\partial x^i} = \begin{cases} 1 & \text{se} \quad i = j \\ \\ 0 & \text{se} \quad i \neq j \end{cases},$$

obtemos $dx^j(v) = \alpha^j$, portanto $\{dx^1, \ldots, dx^n\}$ constitui a base dual da base $\{e_1, \ldots, e_n\}$ como queríamos demonstrar. A expressão conhecida

$$df_p = \sum_i \frac{\partial f}{\partial x^i}(p)\, dx^i,$$

que exprime a diferencial de uma função f como combinação linear dos dx^i, pode ser obtida substituindo-se em $df_p(v) = \sum \frac{\partial f}{\partial x^i}(p)\,\alpha^i$, α^i pelo seu valor $dx^i(v)$, $i = 1, \ldots, n$. De fato, assim procedendo, obtemos

$$df_p(v) = \sum \frac{\partial f}{\partial x^i}(p)\, dx^i(v)$$

para todo vetor $v \in R^n$, o que significa

$$df_p = \sum_i \frac{\partial f}{\partial x^i}(p)\, dx^i.$$

Como a base $\mathcal{E}$ e sua dual $\mathcal{E}^*$ têm o mesmo número de elementos, a dimensão do espaço dual de um espaço vetorial V é igual à dimensão de V, portanto V e V^* são isomorfos. Se $\mathcal{E} = \{e_1, \ldots, e_n\}$ e $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ é a base dual de $\mathcal{E}$, a aplicação $T_{\mathcal{E}} \colon V \to V^*$ que ao vetor $v = \sum \alpha^i\, e_i$ associa o funcional $T_{\mathcal{E}}(v) = \sum \alpha^i\, e^i$ é um isomorfismo que depende da base $\{e_1, \ldots, e_n\}$, pois se tomássemos uma outra base $\mathcal{F} = \{f_1, \ldots, f_n\}$ a aplicação $T_{\mathcal{F}} \colon V \to V^*$ definida de maneira análoga a T seria também um isomorfismo mas não necessariamente o mesmo. Não nos é sempre possível construir um isomorfismo canônico entre V e V^* e portanto não devemos identificá-los. Entretanto, é admissível identificar V com $(V^*)^*$, como mostraremos agora.

Escreveremos V^{**}, em vez de $(V^*)^*$, para indicar o espaço dual de V^*.

Teorema 3. *Existe um isomorfismo canônico I entre V e V^{**}.*

Demonstração: Definimos uma transformação linear I assim: se $v \in V$, $I(v)$ é o funcional que a cada $\varphi \in V^*$

associa o escalar $\varphi(v)$, isto é, $I(v)(\varphi) = \varphi(v)$. A aplicação I é biunívoca, isto é, $I(v) = 0$ implica $v = 0$. De fato, $I(v) = 0$ significa que $I(v)(\varphi) = \varphi(v) = 0$ qualquer que seja $\varphi \in V^*$. Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V, e suponhamos que $v = \sum \alpha^i e_i$. Então

$$\varphi(v) = \sum \alpha^i \varphi(e_i) = 0$$

qualquer que seja $\varphi \in V^*$. Considerando a base dual $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ e tomando sucessivamente $\varphi = e^1$, $\varphi = e^2, \ldots, \varphi = e^n$, obteremos

$$\alpha^1 = 0,\ \alpha^2 = 0, \ldots, \alpha^n = 0,$$

donde $v = 0$. Como V e V^{**} têm a mesma dimensão segue-se, da Proposição 4, que I é um isomorfismo. O leitor poderá observar facilmente que a definição de I não envolve nenhuma escolha arbitrária, portanto I é um isomorfismo canônico.

Observação: Somente para espaços de dimensão finita existe o isomorfismo $V \approx V^{**}$. Isto se exprime dizendo que tais espaços são *reflexivos*. Quando a dimensão de V é infinita, dada uma base $\mathcal{E} = \{e_\alpha\}$ de V, o conjunto $\mathcal{E}^* = \{e^\alpha\}$, definido de modo análogo ao do texto, é formado por funcionais e^α linearmente independentes, mas *nunca* gera V^*, logo não é uma base de V^*. Assim, a aplicação linear $I\colon V \to V^{**}$ é sempre biunívoca mas não é sobre V^{**}, a menos que dim V seja finita. Somente os espaços de dimensão finita gozam da propriedade de reflexividade. O leitor poderá encontrar, em Análise Funcional, a afirmação de que certos espaços de dimensão infinita são reflexivos,

mas é bom ter em mente que ali tais espaços possuem topologia, e os funcionais lineares considerados são apenas os *contínuos*. No caso puramente algébrico (único que estudamos aqui), o número cardinal de uma base de V^{**}, *quando infinito*, é sempre maior do que o número cardinal de uma base de V, isto é, $\dim V^{**} > \dim V$.

1.6 Subespaços

Um subconjunto W de um espaço vetorial V chama-se *subespaço vetorial* de V quando goza das seguintes propriedades:

1) se $u, v \in W$, então $u + v \in W$;
2) se $u \in W$ e λ é um escalar, então $\lambda u \in W$.

Proposição 5. *A interseção $W = \cap V_i$ de uma família qualquer de subespaços $V_i \subset V$ é ainda um subespaço de V.*

Demonstração: Sejam $u, v \in W$; então, $u, v \in V_i$, para todo i, e por conseguinte $u + v \in V_i$, para todo i. Assim, $u + v \in W$. A demonstração de que $\lambda u \in W$, λ escalar e $u \in W$, é análoga.

Dado um conjunto qualquer X, de vetores de um espaço vetorial V, consideremos a interseção $W = \cap V_i$ de todos os subespaços de V que contêm esse conjunto. O subespaço $W \subset V$ é denominado o *subespaço gerado* pelo conjunto X.

Proposição 6. *Seja X um conjunto qualquer de vetores de um espaço vetorial V. O conjunto de todos os vetores obtidos combinando linearmente os valores de X é o subespaço de V gerado por X.*

Exemplo 12. Se $\varphi\colon V \to R$ é um funcional linear, então $\varphi^{-1}(0)$ é um subespaço de V. De fato, se $u, v \in \varphi^{-1}(0)$, então $\varphi(u) = \varphi(v) = 0$, portanto $\varphi(u+v) = \varphi(u)+\varphi(v) = 0$ e $\varphi(\alpha u) = \alpha\varphi(u) = 0$, onde α é um escalar, ou seja, $u+v$ e $\alpha u \in \varphi^{-1}(0)$, o que demonstra nossa afirmação. No caso em que $\varphi = 0$, isto é, φ é o funcional identicamente nulo, é claro que $\varphi^{-1}(0) = V$. Se $\varphi \neq 0$, então o subespaço $\varphi^{-1}(0)$ chama-se um *hiperplano* do espaço vetorial V. Demonstraremos adiante que quando $\dim V = n$ e $\varphi \neq 0$ então $\dim \varphi^{-1}(0) = n - 1$, o que justifica a denominação de hiperplano. Aqui, nos limitaremos a exprimir em coordenadas o subespaço $\varphi^{-1}(0)$. Tomemos $\{e_1, \ldots, e_n\}$ uma base de V; aplicando φ a $v = \sum \xi^i e_i$ e fazendo $\varphi(e_i) = \alpha_i$ obtemos $\varphi(v) = \sum \alpha_i \xi^i$ onde pelo menos um dos α_i não é nulo, pois $\varphi \neq 0$. Assim, podemos escrever

$$\varphi^{-1}(0) = \left\{ v = \sum \xi^i e_i \,;\, \sum \alpha_i \xi^i = 0 \right\},$$

ou seja, $\varphi^{-1}(0)$ é o conjunto dos vetores $v = (\xi^1, \ldots, \xi^n)$ que satisfazem à equação

$$\alpha_1 \xi^1 + \cdots + \alpha_n \xi^n = 0.$$

Mais geralmente, se $T\colon V \to W$ é uma transformação linear, então $T^{-1}(0)$ é um subespaço de V (núcleo de T) e $T(V)$ é um subespaço de W (imagem de T).

A noção de subespaço vetorial nos fornece uma nova razão para não nos restringirmos a considerar os espaço R^n como únicos espaços vetoriais. Com efeito, um subespaço k-dimensional do R^n não é, necessariamente, um espaço R^k. (Por exemplo, uma reta passando pela origem em R não é o conjunto dos números reais e sim um conjunto de *pares* de números *reais*.)

1.7 Espaços Euclidianos

Um *produto interno* num espaço vetorial V é uma correspondência que a cada par u, v de vetores de V associa um escalar $u \cdot v$, satisfazendo as seguintes condições:

1) $u \cdot v = v \cdot u$;
2) $(u + v) \cdot w = u \cdot w + v \cdot w$;
3) $(\lambda u) \cdot v = \lambda(u \cdot v)$;
4) $u \cdot u \geq 0$; $\quad u \cdot u > 0$ se $u \neq 0$

(Onde $u, v, w \in V$ e λ é um escalar.)

Um *espaço vetorial euclidiano* é um par $(V, \cdot)$, isto é, um espaço vetorial munido de um produto interno. Dado um vetor u num espaço vetorial euclidano V, denominaremos *norma* de u, e indicaremos com $|u|$, o escalar $\sqrt{u \cdot u}$ (aqui só é considerado o valor positivo da raiz).

Exemplos: 13) Os exemplos mais conhecidos de espaços vetoriais euclidianos são os R^n com o produto interno $u \cdot v = \sum \alpha^i \beta^i$, onde $u = (\alpha^1, \ldots, \alpha^n)$ e $v = (\beta^1, \ldots, \beta^n)$.

14) Dado um espaço vetorial V, podemos torná-lo euclidiano de uma infinidade de maneiras distintas, uma das quais é a seguinte: fixamos primeiro uma base $\{e_1, \ldots, e_n\}$ de V e, se $u = \sum \alpha^i e_i$ e $v = \sum \beta^i e_i$, definimos $u \cdot v = \sum \alpha^i \beta^i$. Isto mostra que, num certo sentido, a classe dos espaços vetoriais euclidianos não constitui uma classe mais restrita que a classe dos espaços gerais.

Num espaço vetorial euclidiano V, dois vetores u, v dizem-se *ortogonais* quando $u \cdot v = 0$. Se $u_1, \ldots, u_k \in V$ são vetores dois a dois ortogonais (isto é, $u_i \cdot u_j = 0$ para $i \neq j$)

e não-nulos, então o conjunto $\{u_1, \ldots, u_k\}$ é linearmente independente. Com efeito, se fosse

$$\sum \alpha^i \, u_i = 0$$

então, tomando um u_j fixo qualquer, teríamos

$$0 = u_j \cdot \left(\sum_i \alpha^i u_i \right) = \sum_i \alpha^i (u_j \cdot u_i) = \alpha^j, |u_j|^2.$$

Como $u_j \neq 0$, segue-se que $\alpha^j = 0$. Assim todos os coeficientes da combinação linear $\sum \alpha^i u_i$ são nulos e os u_i são linearmente independentes.

Proposição 7. *Em um espaço vetorial euclidiano V, de dimensão n, todo conjunto $\{u_1, \ldots, u_n\}$ de n vetores não nulos, dois a dois ortogonais em V, é uma base de V.*

Uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de um espaço vetorial euclidiano chama-se *ortonormal* se

$$e^i \cdot e_j = \begin{cases} 0 & \text{quando} \quad i \neq j \\ 1 & \text{quando} \quad i = j \end{cases}$$

isto é, uma base é ortonormal quando é formada por vetores unitários e dois a dois ortogonais.

Em relação a uma base ortonormal $\{e_1, \ldots, e_n\}$ de um espaço vetorial V, o produto escalar de dois valores $u = \sum \alpha^i \, e_i$ e $v = \sum \beta^j \, e_j$ toma uma forma particularmente simples:

$$u \cdot v = \left(\sum \alpha^i e_i \right) \cdot \left(\sum \beta^j e_j \right) = \sum_{i,j} \alpha^i \beta^j (e_i \cdot e_j) = \sum \alpha^i \beta^i$$

pois

$$e^i \cdot e_j = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

Veremos a seguir que toda base pode ser ortonormalizada, portanto em todo espaço euclidiano V existe uma base em que o produto interno é expresso da maneira acima: $u \cdot v = \sum \alpha^i \beta^i$. No entanto, às vezes somos forçados a trabalhar com bases não-ortonormais. Neste caso, o produto interno de u e v é dado por

$$u \cdot v = \sum_{i,j} g_{ij}\, \alpha^i\, \beta^j \quad \text{onde} \quad g_{ij} = e_i \cdot e_j\,.$$

O processo de ortonormalização de uma base $\{f_1, ..., f_n\}$ a que nos referimos acima, consiste essencialmente no seguinte: como $f_1 \neq 0$ escreveremos $e_1 = \frac{f_1}{|f_1|}$; supondo que tivéssemos definido $\{e_1, e_2, \ldots, e_{k-1}\}$ definimos

$$e_k = \frac{f_k - \sum_{i<k} (f_k \cdot e_i) e_i}{|f_k - \sum_{i<k} (f_k \cdot e_i) e_i|}\,.$$

Por indução, verifica-se facilmente que $\{e_1, \ldots, e_n\}$ é uma base ortonormal.

Teorema 4. *Se V é um espaço vetorial euclidiano, existe um isomorfismo canônico J entre V e o seu dual V^*.*

Demonstração: Definamos uma aplicação $J\colon V \to V^*$ que, ao vetor $v \in V$, associa o funcional linear $J(v) \in V^*$ definido por $J(v)(u) = u \cdot v$. Por ser o produto interno linear em u, $J(v)$ é de fato um funcional linear; por ser o

produto interno linear em v, J é uma transformação linear. Como V e V^* são espaços de mesma dimensão, teremos demonstrado o isomorfismo quando demonstrarmos que a aplicação J é biunívoca. Se $u \neq 0$ então $J(u)(u) = u \cdot u \neq 0$, portanto $J(u)$ não é identicamente nulo, pois existe um vetor $u \in V$ tal que $J(u)(u) \neq 0$. Isso demonstra a biunivocidade de J. Evidentemente, a definição de J não depende de uma base ou de qualquer outra escolha arbitrária em V, por isso J é um isomorfismo canônico.

Em virtude deste teorema, não há grande interesse em se considerar o dual de um espaço vetorial euclidiano. Quando V é um espaço vetorial euclidiano, pode-se sempre substituir um funcional linear $f \in V^*$ pelo vetor $u = J^{-1}(f) \in V$. O resultado $f(v)$ da operação de f sobre um vetor $v \in V$ ficará substituido pelo produto interno $u \cdot v$; por isso é que, no cálculo vetorial clássico, onde o único espaço vetorial considerado é o espaço euclidiano R^n, não intervém explicitamente a noção de funcional linear. Por isso também é que, em espaços com produto interno, pode-se deixar de fazer distinção entre *vetores covariantes* e *contravariantes.*

Consideremos um espaço vetorial euclidiano V e uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V. Determinaremos em seguida as coordenadas α_i do funcional $J(v)$ relativamente à base dual $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ em função das coordenadas α^i de $v \in V$ relativamente à base $\mathcal{E}$.

Já sabemos que a i-ésima coordenada de um funcional linear relativamente à base dual $\mathcal{E}^*$ é obtida aplicando o funcional no i-ésimo vetor da base $\mathcal{E}$; assim $\alpha_i = J(v)e_i$, $i = 1, \ldots, n$. Mas, de $J(v)(e_i) = e_i \cdot v = e_i \cdot \sum \alpha^j (e_i \cdot e_j)$,

segue-se que

$$\alpha_i = \sum_j g_{ij}\,\alpha^j \quad \text{onde} \quad g_{ij} = e_i \cdot e_j\,.$$

A passagem das coordenadas α^j para as coordenadas α_i, de acordo com a expressão acima, é conhecida como *operação de baixar índice.* Este sistema pode ser resolvido de maneira a nos permitir passar das α_i para as α^i. É a operação de *subir índice.* A fim de obter a expressão explícita das coordenadas α^i em termos das α_i, consideraremos a matriz (g^{ij}), inversa da matriz (g_{ij}). (Vide §9, adiante.)

Obtemos então:

$$\alpha^i = \sum_j g^{ij}\,\alpha_j\,.$$

Quando a base $\mathcal{E}$ é ortonormal, obtemos $\alpha_i = \alpha^i$.

Ou seja, neste caso, o funcional $J(v)$, expresso na base $\mathcal{E}^*$, tem mesmas coordenadas que o vetor v, expresso na base $\mathcal{E}$.

Como exemplo de aplicação do isomorfismo J, temos o *gradiente* de uma função real, diferenciável, $f\colon A \to R$, definida em um aberto A do R^n. Dado um ponto $p \in A$, sabemos que a diferencial de f no ponto p é o funcional linear $df_p \in (R^n)^*$ tal que $df_p(v) = \frac{\partial f}{\partial v}(p)$ para todo vetor $v \in R^n$. O *gradiente* de f (no ponto p) será agora definido como o vetor $\nabla f(p) \in R^n$ que corresponde a df_p pelo isomorfismo canônico $J\colon R^n \to (R^n)^*$ (induzido pelo produto interno natural do R^n). Assim:

$$\nabla f(p) = J^{-1}(df_p).$$

Portanto, o gradiente $\nabla f(p)$ fica caracterizado pela propriedade de ser

$$v \cdot \nabla f(p) = df_p(v), \text{ para todo vetor } v \in R^n.$$

Como a base canônica $\mathcal{E} = \{e_1, \ldots, e_n\}$ do espaço R^n é ortonormal, as componentes do vetor $\nabla f(p)$ relativamente a esta base são as mesmas do funcional $df_p = J(\nabla f(p))$ relativamente à base dual $\mathcal{E}^* = \{dx^1, \ldots, dx^n\}$. Assim, as componentes do gradiente $\nabla f(p)$ com respeito à base ortonormal $\mathcal{E}$ são os números

$$\left\{ \frac{\partial f}{\partial x^1}(p), \ldots, \frac{\partial f}{\partial x^n}(p) \right\}.$$

Convém salientar que, em muitos problemas de Análise e de Geometria, é conveniente tomar bases não-ortonormais no espaço R^n. Quando se procede assim, as coordenadas do vetor gradiente não têm uma expressão tão simples e concisa, de modo que a definição de ∇f que demos acima se torna mais cômoda, por ser intrínseca, isto é, independente de sistemas de coordenadas.

Tomemos um espaço vetorial euclidiano V. Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V e $v = \sum \alpha^i e_i$ um vetor de V. Os escalares $\alpha^1, \ldots, \alpha^n$ são denominados *componentes contravariantes* do vetor v relativamente à base $\mathcal{E}$, ao passo que os escalares $\alpha_i = e_i \cdot v$ são chamados *componentes covariantes* de v. Se $\mathcal{E}$ é uma base ortonormal, as componentes covariantes de um vetor coincidem com suas componentes contravariantes. De fato:

$$\alpha_j = e_j \cdot v = e_j \cdot \left(\sum_i \alpha^i e_i \right) = \sum_i \alpha^i (e_j \cdot e_i) = \alpha^j, \; j = 1, \ldots, n.$$

As componentes covariantes de v podem ser obtidas aplicando o funcional $J(v)$ aos elementos de $\mathcal{E}$: $J(v)(e_i) = e_i \cdot v = \alpha_i$. Mas, aplicando o funcional $J(v)$ à base $\mathcal{E}$, obtemos, como já vimos, as componentes (contravariantes) de $J(v)$ na base $\mathcal{E}^*$. Desta maneira, as componentes contravariantes de $J(v)$ na base $\mathcal{E}^*$ coincidem com as covariantes de v na base $\mathcal{E}$. Note-se que este resultado, ao contrário do precedente, não exige que a base tomada seja ortonormal.

O leitor deve ter percebido que empregamos sistematicamente as seguintes convenções de índice, como uma concessão ao método clássico de expor o cálculo tensorial:

Em toda sequência $v_1, v_2, v_3, \ldots$ de vetores ("contravariantes") os índices são colocados em baixo da letra, enquanto que os índices das coordenadas (contravarianes) $\xi^1, \ldots, \xi^n$ de um vetor $v = \sum \xi^i e_i$ são sempre superiores. Por outro lado, uma sequência $f^1, f^2, \ldots$ de funcionais lineares ("vetores covariantes") tem índices superiores, mas as coordenadas $\alpha_1, \ldots, \alpha_n$ de um funcional $f = \sum \alpha_i f^i$ são dotadas de índices inferiores. Isto faz com que, num somatório, um índice segundo o qual se soma ("índice mudo") apareça sempre superiormente e inferiormente. (Mais preciso seria dizer *quase sempre*: nos espaços euclidianos, a distinção entre índice superior e índice inferior desaparece.)

Entretanto, não adotaremos a chamada "convenção de Einstein", que consiste em indicar as somas do tipo $\sum \alpha^i \beta_i$ por $\alpha^i \beta_i$. Tal convenção não é auto-consistente e, numa apresentação intrínseca do cálculo tensorial, ela é inteiramente dispensável.

1.8 Soma direta e produto cartesiano

Sabemos que, dados dois conjuntos quaisquer X e Y, o seu produto cartesiano $X \times Y$ é o conjunto de todos os pares ordenados (x, y), onde x percorre X e y percorre Y. Assim:

$$X \times Y = \{(x, y);\ x \in X, y \in Y\}.$$

Analogamente, dados dois espaços vetoriais V e W definimos o seu produto cartesiano $V \times W$, considerando-os como conjunto que são. Acontece, porém, que não ficamos aqui, pois podemos introduzir em $V \times W$ uma estrutura de espaço vetorial definindo aí as operações

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

onde v, $v' \in V$, w, $w' \in W$ e λ é um escalar. Continuamos chamando este espaço vetorial de *produto cartesiano* de V por W.

Proposição 8. *Se* $\mathcal{E} = \{e_1, \ldots, e_n\}$ *e* $\mathcal{F} = \{f_1, \ldots, f_p\}$ *são bases de* V *e* W *respectivamente, então* $\mathcal{E}+\mathcal{F}=\{(e_1, 0), \ldots, (e_n, 0), (0, f_1), \ldots, (0, f_p)\}$ *é uma base de* $V \times W$.

Demonstração: De fato,

$$\alpha_1(e_1, 0) + \cdots + \alpha_n(e_n, 0) + \beta_1(0, f_1) + \cdots + \beta_p(0, f_p) = 0$$

significa

$$\alpha_1 e_1 + \cdots + \alpha_n e_n + \beta_1 0 + \cdots + \beta_p 0 = 0$$

e

$$\alpha_1 0 + \cdots + \alpha_n 0 + \beta_1 f_1 + \cdots + \beta_p f_p = 0.$$

Como os e_i e os f_j são linearmente independentes temos $\alpha_i = 0$ e $\beta_j = 0$. Portanto $\mathcal{E} + \mathcal{F}$ é um sistema linearmente independente. Por outro lado, se $z \in V \times W$, então $z = (v, w)$, onde $v \in V$ e $w \in W$; assim,

$$z = (v, w) = \Big(\sum \alpha^i e_i, \sum \beta^j f_j\Big) = \\ = \sum \alpha^i (e_i, 0) + \sum \beta^j (0, f_j),$$

portanto todo $z \in V \times W$ se exprime como combinação linear dos elementos de $\mathcal{E} + \mathcal{F}$. Do exposto, concluimos que $\dim(V \times W) = \dim V + \dim W$, isto é, a dimensão do produto cartesiano é igual à soma das dimensões dos fatores.

Consideremos, agora, dois *subespaços* V de W de um espaço vetorial Z. Dizemos que Z é a *soma direta* de V e W e escrevemos $Z = V \oplus W$ quando

1) todo vetor $z \in Z$ se exprime como soma de um vetor $v \in V$ com um vetor $w \in W$, isto é $z = v + w$;

2) esta maneira de exprimir z é única. isto é, se $z = v + w = v' + w'$, com $v, v' \in V$ e $w, w' \in W$ então $v = v'$ e $w = w'$.

Esta segunda condição, em presença de 1), equivale a dizer

2') $V \cap W = \{0\}$.

De fato, (2) implica (2') pois se $x \in V \cap W$, x pensado como elemento de Z se exprime de duas maneiras $x = x + 0$,

$x \in V$, $0 \in W$ e $x = 0 + x$, $0 \in V$ e $x \in W$; supondo que (2) é satisfeita, segue-se que $x = 0$. Para mostrar que (2') implica (2), suponhamos que (2) não seja satisfeita, isto é, existe um vetor $z \in Z$ que se exprime de duas maneiras distintas

$$z = v + w \quad \text{e} \quad z = v' + w'$$

onde v, $v' \in V$, w, $w' \in W$, mas, digamos, tem-se $v \neq v'$. De $v + w = v' + w'$ obtemos $v - v' = w' - w$ o que mostra que $v - v'$, além de pertencer a V, pertence também a W, portanto $v - v' \in V \cap W$. Por outro lado, como $v \neq v'$, $v - v'$ é não-nulo. Isto demonstra a implicação desejada.

Se $Z = V \oplus W$, então existe um isomorfismo canônico $T\colon Z \to V \times W$, assim definido:

$$T(z) = (v, w)$$

onde $z \in Z$ é o vetor que se exprime, de modo único, como $z = v + w$, $v \in V$, $w \in W$. Deixamos ao leitor o encargo de mostrar que T é, de fato, um isomorfismo. Como a dimensão é invariante por isomorfismos, concluimos que, se $Z = V \oplus W$ então $\dim Z = \dim V + \dim W$ porque, como já vimos, $\dim(V \times W) = \dim V + \dim W$.

Um produto cartesiano $V \times W$ se decompõe como soma direta dos subespaços $V' = \{(v, 0) \in V \times W; v \in V\}$ e $W' = \{(0, w) \in V \times W; w \in W\}$. De fato,

1) se $x \in V \times W$, $x = (v, w)$ então $x = (v, 0) + (0, w)$;

2) se $x \in V' \cap W'$, por um lado $x = (v, 0)$ e por outro $x = (0, w)$, logo $x = (0, 0)$.

É fácil verificar que as aplicações $V' \to V$ e $W' \to W$

definidas por

$$(v, 0) \to v$$
$$(0, w) \to w$$

constituem isomorfismos. Concluimos que um produto cartesiano $V \times W$ se decompõe como soma direta de subespaços isomorfos canonicamente a V e a W.

Definimos, aqui, somente o produto cartesiano de dois espaços; no entanto, de maneira análoga, podemos definir o produto cartesiano de n espaços. O mesmo se dá com a soma direta, a qual foi definida para o caso de dois subespaços. Neste caso, devemos fazer uma nova formulação da condição (2'). Sendo $V_1, \ldots, V_n$ os subespaços em que um certo espaço Z se decompõe, é a seguinte a reformulação de (2'):

2") A intersecção do subespaço V_i $(i = 1, \ldots, n)$ com o subespaço gerado pelos outros subespaços V_j, $j \neq i$, se reduz ao vetor zero.

Já vimos que se $\varphi\colon V \to R$ é um funcional linear, $\varphi^{-1}(0)$ é um subespaço do espaço vetorial V. Agora, como aplicação do conceito de soma direta, demonstraremos a seguinte

Proposição 9. *Seja* $\dim V = n$ *e* $0 \neq \varphi \in V^*$. *Então,* $\dim \varphi^{-1}(0) = n - 1$.

Demonstração: Como φ, por hipótese, é um funcional não identicamente nulo, existe um vetor $u \in V$ tal que $\varphi(u) \neq 0$. Consideremos o subespaço $U = \{\lambda u; \lambda \in R\}$, cuja dimensão é evidentemente 1. Vamos mostrar que $V = U \oplus \varphi^{-1}(0)$. Com efeito,

1) todo $v \in V$ pode ser escrito como

$$v = \frac{\varphi(v)}{\varphi(u)} \cdot u + \left(v - \frac{\varphi(v)}{\varphi(u)} \cdot u \right).$$

É claro que $\frac{\varphi(v)}{\varphi(u)} \cdot u \in U$ e como $\varphi\left[v - \frac{\varphi(v)}{\varphi(u)} \cdot u\right] = 0$, segue-se que $\left[v - \frac{\varphi(v)}{\varphi(u)} \cdot u\right] \in \varphi^{-1}(0)$. Concluimos que todo $v \in V$ pode ser escrito como soma de um elemento de U com um elemento de $\varphi^{-1}(0)$.

2) Se $v \in U \cap \varphi^{-1}(0)$ então, $v = \lambda u$ por v pertencer a U e $\varphi(v) = 0$ por v pertencer a $\varphi^{-1}(0)$, portanto $0 = \varphi(v) = \varphi(\lambda u) = \lambda\varphi(u)$ e como $\varphi(u) \neq 0$ temos $\lambda = 0$, logo $v = \lambda u = 0 \cdot u = 0$. Isso demonstra que $V = U \oplus \varphi^{-1}(0)$. Portanto $\dim V = \dim U + \dim \varphi^{-1}(0)$. Como $\dim V = n$ e $\dim U = 1$, segue-se que $\dim \varphi^{-1}(0) = n - 1$, como queríamos demonstrar.

1.9 Relação entre transformações lineares e matrizes

Sejam $\mathcal{E}=\{e_1, \ldots, e_n\}$ e $\mathcal{F}=\{f_1, \ldots, f_p\}$ bases dos espaços vetoriais V e W respectivamente. Uma transformação linear $A\colon V \to W$ fica determinada quando conhecemos pn escalares, que formam o que chamamos a matriz $\alpha = [A; \mathcal{E}, \mathcal{F}]$ da transformação A relativamente às bases $\mathcal{E}$ e $\mathcal{F}$. Obtemos estes pn escalares assim: aplicamos a transformação A aos elementos da base $\mathcal{E}$ e obtemos elementos do espaço vetorial W que, portanto, são combinações line-

ares dos elementos de $\mathcal{F}$:

$$A(e_j) = \sum_i \alpha_j^i f_i\,, \qquad \begin{cases} i = 1, \ldots, p \\ \\ j = 1, \ldots, n \end{cases} .$$

As coordenadas α_j^i dos elementos $A(e_j)$ são os pn escalares formadores da matriz α:

$$\alpha = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^p & \ldots & \alpha_n^p \end{pmatrix}$$

Se $v = \sum \xi^j e_j$, então as coordenadas η^i de $A(v)$ relativametne à base $\mathcal{F}$ são os escalares $\sum\limits_k \alpha_j^i \xi^j$, isto é,

$$\eta^i = \sum_j \alpha_j^i \xi^j\,, \qquad \begin{cases} j = 1, \ldots, n \\ \\ i = 1, \ldots, p \end{cases} .$$

De fato,

$$A(v) = \sum \xi^j\, A(e_j) = \sum_j \xi^j \left(\sum_i \alpha_j^i f_i \right) = \sum_i \left(\sum_j \alpha_j^i \xi^j \right) f_i\,,$$

o que demonstra nossa afirmação. Como as coordenadas de $A(v)$ na base $\mathcal{F}$ dependem somente das coordenadas de v, ξ^j, e da matriz $\alpha = (\alpha_j^i)$ esta, de fato, determina a transformação A. Os elementos da matriz $\alpha = [A; \mathcal{E}, \mathcal{F}]$ são, por assim dizer, as "coordenadas" da transformação linear A relativamente às bases $\mathcal{E}$ e $\mathcal{F}$.

O conjunto $\mathcal{L}(V, W)$ de todas as transformações lineares de V em W constitui, de modo natural, um espaço vetorial de dimensão np, se n e p forem as dimensões de V e de W respectivamente. São as seguintes as operações que fazem de $\mathcal{L}(V, W)$ um espaço vetorial:

$$\begin{aligned} &(A + B)(v) = A(v) + B(v) && v \in V \\ &(\lambda A)(v) = \lambda \cdot A(v) && v \in V \text{ e } \lambda \text{ escalar.} \end{aligned}$$

Por outro lado, o conjunto $M(p \times n)$ das matrizes reais de p linhas e n colunas constitui também um espaço vetorial de dimensão np, onde a soma e o produto por um escalar são as operações conhecidas. Toda vez que fazemos uma escolha de bases $\mathcal{E}$ e $\mathcal{F}$ em V e W respectivamente, estabelecemos um isomorfismo

$$\theta \colon \mathcal{L}(V, W) \to M(p \times n)$$

que à transformação A associa a sua matriz relativamente às bases $\mathcal{E}$ e $\mathcal{F}$,

$$\theta \colon A \to \alpha = [A; \mathcal{E}, \mathcal{F}].$$

A aplicação θ é biunívoca porque, como já vimos, uma transformação fica perfeitamente determinada quando conhecemos sua matriz relativamente a bases $\mathcal{E}$ e $\mathcal{F}$ prefixadas, portanto cada matriz α é imagem de, no máximo, uma transformação A. A aplicação θ é sobre, pois se $\alpha = (\alpha_j^i)$ é uma matriz, a transformação linear B definida por

$$B(e_j) = \sum_i \alpha_j^i f_i\,, \qquad \begin{cases} i = 1, \ldots, p \\ j = 1, \ldots, n \end{cases}$$

é tal que $\theta(B) = \alpha$. A linearidade de θ decorre do fato de que, se

$$A(e_j) = \sum_i \alpha_j^i f_i \quad \text{e} \quad B(e_j) = \sum_i \beta_j^i f_i$$

então

$$(A+B)(e_j) = A(e_j) + B(e_j) = \sum_i \alpha_j^i f_i + \sum_i \alpha_j^i f_i =$$

$$= \sum_i (\alpha_j^i + \beta_j^i) f_i$$

e

$$(\lambda A)(e_j) = \lambda \cdot A(e_j) = \lambda \cdot \sum_i \alpha_j^i f_i = \sum_i (\lambda \alpha_j^i) f_i \,.$$

A base canônica do espaço vetorial $M(p \times n)$ é constituida pelas pn matrizes que na posição i, j (i=1, ..., p, j = 1, ..., n) têm o escalar 1 e nas demais posições zeros:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & & 0 \\ 0 & & & \\ \vdots & & & \vdots \\ & & 0 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Uma base de $\mathcal{L}(V, W)$ é constituida pelas pn transformações lineares que por θ são aplicadas nas matrizes acima, ou seja, são as transformações lineares $E_{ij} \colon V \to W$, assim definidas

$$\begin{cases} E_{ij}(e_j) = f_i & i = 1, \dots, p, , \quad j = 1, \dots, n \\ E_{ij}(e_k) = 0 & k \neq j \end{cases}$$

Observemos que a base $\{E_1, \ldots, E_{ij}, \ldots, E_{pn}\}$, relacionada pelo isomorfismo θ com a base canônica de $M(p \times n)$ acima exibida, não é canônica. O isomorfismo θ estabelecido entre $\mathcal{L}(V, W)$ e $M(p \times n)$ não é canônico.

Dadas as transformações lineares $A\colon U \to V$ e $B\colon V \to W$, definimos a transformação linear $BA\colon U \to W$, chamada *produto* de B por A, como a transformação que ao elemento $u \in U$ associa o elemento $B(A(u)) \in W$:

$$BA(u) = B(A(u)),\ u \in U.$$

Consideremos bases $\mathcal{E} = \{e_1, \ldots, e_n\}$, $\mathcal{F} = \{f_1, \ldots, f_m\}$ e $\mathcal{G} = \{g_1, \ldots, g_h\}$ em U, V e W respectivamente. Sejam $(\alpha_j^k) = [A.\mathcal{E}, \mathcal{F}]$ e $(\beta_k^i) = [B; \mathcal{F}, \mathcal{G}]$ as matrizes das transformações A e B nas bases $\mathcal{E}$, $\mathcal{F}$ e $\mathcal{F}$, $\mathcal{G}$ respectivamente. Então

$$A(e_j) = \sum_k \alpha_j^k f_k \quad \text{e} \quad B(f_k) = \sum_i \beta_k^i g_i\,.$$

Portanto,

$$BA(e_j) = B\left(\sum_k \alpha_j^k f_k\right) = \sum_k \alpha_j^k\, B(f_k) =$$

$$= \sum_k \alpha_j^k \left(\sum_i \beta_k^i g_i\right) = \sum_i \left[\sum_k \beta_k^i\, \alpha_j^k\right] g_i\,,$$

Concluimos que a matriz $(\gamma_j^i) = [BA; \mathcal{E}, \mathcal{G}]$, da transformação BA relativamente às bases $\mathcal{E}$ e $\mathcal{G}$, é dada por

$$\gamma_j^i = \sum_{k=1}^{m} \beta_k^i\, \alpha_j^k\,, \quad i = 1, \ldots, h, \quad j = 1, \ldots, n.$$

A matriz (γ_j^i) é chamada *produto* da matriz (β_k^i) pela matriz (α_j^k). Observamos que, para definirmos o produto BA das transformações lineares B e A, é preciso que B esteja definida no espaço vetorial onde A toma valores. Assim é que, somente quando as transformações A e B são de um espaço vetorial V em si mesmo, podemos definir ambos os produtos AB e BA, que, por sinal, não são necessariamente iguais.

Agora, consideremos somente transformações lineares $A\colon V \to V$ de um espaço vetorial de dimensão n em si mesmo. Neste caso, convencionaremos tomar somente matrizes de A da forma $\alpha = [A; \mathcal{E}, \mathcal{E}]$ as quais indicaremos simplesmente por $\alpha = [A; \mathcal{E}]$ e chamaremos de *matriz de A relativamente à base $\mathcal{E}$.*

Uma transformação linear $A\colon V \to V$, biunívoca e sobre V, é chamada *não-singular* ou *invertível.* Neste caso, existe uma transformação A^{-1} chamada *inversa* de A, tal que

$$A^{-1}A = AA^{-1} = \text{ identidade.}$$

É imediato que a transformação inversa de uma transformação linear é também linear. Lembramos, além disso, que de acordo com a Proposição 4, para que uma transformação linear $A\colon V \to V$ seja invertível, basta que seja biunívoca (ou então se faça sobre V).

Qualquer que seja a base $\mathcal{E}$ escolhida em V, a matriz da transformação identidade $I\colon V \to V$ é a matriz

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

com unidades na diagonal e zeros nas outras posições. Esta é chamada *matriz identidade*. Se $\alpha = [A; \mathcal{E}]$ é a matriz de uma transformação linear invertível A relativamente a uma base $\mathcal{E}$ e $\beta = [A^{-1}; \mathcal{E}]$ é a matriz de sua inversa relativamente à mesma base, então como $AA^{-1} = A^{-1}A = I$,

$$\alpha\beta = \beta\alpha = I_n\,.$$

A matriz β é chamada *matriz inversa* de α e é indicada por α^{-1}. A matriz de uma transformação invertível é também denominada *invertível*.

Sejam $\alpha = (\alpha_j^i) = [A; \mathcal{E}]$ e $\beta = (\beta_j^i) = [A; \mathcal{F}]$ matrizes de uma transformação $A\colon V \to V$ relativamente às bases $\mathcal{E} = \{e_1, \dots, e_n\}$ e $\mathcal{F} = \{f_1, \dots, f_n\}$ respectivamente, então

$$A(e_i) = \sum_i \alpha_i^j e_j \quad \text{e} \quad A(f_r) = \sum_k \beta_r^k f_k\,;$$

onde $i, j, k, r = 1, \dots, n$.

Chamemos de $\lambda = (\lambda_j^i)$ a matriz de passagem da base $\mathcal{F}$ para a base $\mathcal{E}$. Assim

$$e_j = \sum_k \lambda_j^k\, f_k\,.$$

Portanto,

$$A(e_i) = \sum_j \alpha_i^j e_j = \sum_j \alpha_i^j \left(\sum_k \lambda_j^k f_k \right) = \sum_k \left(\sum_j \lambda_j^k \alpha_i^j \right) f_k\,.$$

Por outro lado, aplicando λ em $e_i = \sum_r \lambda_i^r f_r$, obtemos

$$A(e_i) = A\left(\sum_r \lambda_i^r f_r\right) == \sum_r \lambda_i^r\, A(f_r) =$$

$$= \sum_r \lambda_i^r \left(\sum_k \beta_r^k\, f_k\right) = \sum_k \left[\sum_r \lambda_r^k\, \lambda_i^r\right] f_k\,,$$

Comparando as duas expressões obtidas para $A(e_i)$, vem:

$$\sum_j \lambda_j^k\, \alpha_i^j = \sum_r \beta_r^k\, \lambda_i^r\,,$$

isto é, $\lambda\alpha = \beta\lambda$, ou seja $\alpha = \lambda^{-1}\beta\lambda$, pois toda matriz de passagem de bases é invertível, porque a transformação linear que ela define leva uma base noutra base e portanto é invertível. Podemos então enunciar a seguinte:

Proposição 10. *Sejam $A\colon V \to V$ uma transformação linear e $\mathcal{E}$, $\mathcal{F}$ bases de V. Se $\alpha = [A;\mathcal{E}]$ e $\beta = [A;\mathcal{F}]$, então*

$$\alpha = \lambda^{-1}\,\beta\lambda,$$

onde λ é a matriz de passagem da base $\mathcal{F}$ para a base $\mathcal{E}$.

Sejam $\mathcal{E} = \{e_1, \dots, e_n\}$ e $\mathcal{F} = \{f_1, \dots, f_n\}$ duas bases de um espaço vetorial V e $\lambda = (\lambda_j^i)$ a matriz de passagem da base $\mathcal{F}$ para a base $\mathcal{E}$; temos assim

$$e_j = \sum_i \lambda_j^i\, f_i\,.$$

A seguinte proposição nos mostra como podemos passar da base dual $\mathcal{F}^* = \{f^1, \dots, f^n\}$ para a base dual $\mathcal{E}^* = \{e^1, \dots, e^n\}$.

Proposição 11. *Sejam $\mathcal{E}$ e $\mathcal{F}$ bases de um espaço vetorial V e $\mathcal{E}^*$ $\mathcal{F}^*$ suas bases duais em V^*. Se λ é a matriz de passagem de $\mathcal{F}$ para $\mathcal{E}$ então a matriz de passagem de $\mathcal{F}^*$ para $\mathcal{E}^*$ é igual a λ^{-1}.*

Demonstração: Sabemos que, sendo (λ^i_j) a matriz de passagem de $\mathcal{F}$ para $\mathcal{E}$, as coordenadas $(\alpha^1, \ldots, \alpha^n)$ de um vetor v relativamente à base $\mathcal{E}$ estão relacionadas com suas coordenadas $(\beta^1, \ldots, \beta^n)$ relativamente à base $\mathcal{F}$ por

$$\beta^i = \sum_j \lambda^i_j \, \alpha^j.$$

Mas $\beta^i = f^i(v)$ e $\alpha^j = e^j(v)$, logo $f^i(v) = \sum_j \lambda^i_j \, e^j(v)$, ou seja

$$f^i = \sum_j \lambda^i_j \, e^j, \quad i, j = 1, \ldots, n.$$

Consequentemente, λ é a matriz de passagem da base $\mathcal{E}^*$ para a base $\mathcal{F}^*$, portanto a matriz de passagem da base $\mathcal{F}^*$ para a base $\mathcal{E}^*$ é λ^{-1}, isto é, se

$$e^j = \sum_i \mu^j_i \, f^i$$

então $(\mu^i_j) = \lambda^{-1}$. Por fim, se $\varphi \in V^*$ se escreve como

$$\varphi = \sum \alpha_j \, e^j = \sum \beta_i \, f^i \quad \text{e} \quad e^j = \sum_i \mu^j_i \, f^i$$

então

$$\beta_i = \sum_j \mu^j_i \, \alpha_j \,, \quad i, j = 1, \ldots, n.$$

A demonstração deste fato é análoga à que foi feita para o caso dos vetores contravariantes.

Atenção: A definição que demos para a matriz de passagem de uma base para outra depende essencialmente da posição dos índices dos elementos dessas bases. Mais explicitamente, se tomamos as bases $X = \{x_1, \ldots, x_n\}$ e $Y = \{y_1, \ldots, y_n\}$ e matriz de passagem de X para Y é a matriz $\lambda = (\lambda^i_j)$ tal que $y_j = \sum_i \lambda^i_j \, x_i$. Lembremos que o índice superior numa matriz indicará sempre a *linha*. Ora, se escrevermos os MESMOS elementos de X e Y com índices superiores: $X = \{x^1, \ldots, x^n\}$ e $Y = \{y^1, \ldots, y^n\}$ então teremos $y^j = \sum_i \mu^j_i \, x^i$. Se mantivermos a convenção (útil) de que os índices superiores de uma matriz indicam as linhas e os inferiores indicam sempre as colunas, veremos agora que a nova matriz $(\mu^j_i) = \mu$ de passagem da *mesma* base X para a *mesma* base Y não é igual a $\lambda = (\lambda^i_j)$. Na realidade μ é a *transposta* de λ, isto é, as linhas de μ coincidem com as colunas de λ: $\mu^j_i = \lambda^i_j$.

Capítulo 2

Álgebra Multilinear

Trataremos, neste capítulo, dos principais conceitos e resultados da Álgebra Multilinear, os quais giram em torno da noção de produto tensorial de espaços vetoriais. Com a finalidade de simplificar as discussões, consideraremos quase sempre aplicações bilineares, em vez de trabalharmos com o caso p-linear geral. A passagem de 2 a $p > 2$ se faz trivialmente.

2.1 Aplicações bilineares

Sejam U, V, W espaços vetoriais. Uma aplicação

$$\phi\colon U \times V \to W,$$

do produto cartesiano de U por V em W chama-se *bilinear* quando é linear em cada um dos seus argumentos separadamente. Ou, de modo mais preciso, quando goza das seguintes propriedades:

1) $\phi(u+u', v) = \phi(u, v) + \phi(u', v)$;

2) $\phi(u, v+v') = \phi(u, v) + \phi(u, v')$;

3) $\phi(\lambda u, v) = \phi(u, \lambda v) = \lambda\phi(u, v)$

quaisquer que sejam $u, u' \in U$, $v, v' \in V$ e λ escalar.

Uma aplicação bilinear $\phi\colon U \times V \to R$, do par U, V no espaço vetorial R dos números reais é chamada um *funcional bilinear*, ou uma *forma bilinear.*

Indicaremos com $\mathcal{L}(U, V; W)$ o conjunto das aplicações bilineares $\phi\colon U \times V \to W$ do par U, V no espaço vetorial W. $\mathcal{L}(U, V; W)$ constitui um espaço vetorial, de modo natural, onde as operações de soma de duas aplicações e produto de uma aplicação por um número real são definidas como se faz habitualmente, entre funções.

No caso de $W = R$, escreveremos $\mathcal{B}(U, V)$, em vez de $\mathcal{L}(U, V; R)$, para indicar o espaço das formas bilineares sobre o par U, V. E, finalmente, quando $U = V$, escreveremos $\mathcal{B}(U)$, em vez de $\mathcal{B}(U, U)$ para indicar o espaço das formas bilineares sobre U.

Proposição 1. *Sejam U, V e W espaços vetoriais, $\mathcal{E} = \{e_1, \ldots, e_m\}$ uma base de U e $\mathcal{F} = \{f_1, \ldots, f_n\}$ uma base de V. Dada arbitrariamente uma mn-upla de vetores $w_{ij} \in W$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$, existe uma única aplicação bilinear $\phi\colon U \times V \to W$ tal que $\phi(e_i, f_j) = w_{ij}$ para todo $e_i \in \mathcal{E}$ e todo $f_j \in \mathcal{F}$.*

Demonstração: Mostraremos primeiro a unicidade. Se $u = \sum \alpha^i e_i$ e $v = \sum \beta^j f_j$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ são vetores arbitrários de U e V respectivamente, a bilinearidade

de ϕ nos dá:

$$\phi(u,v) = \sum_{i,j} \alpha^i \beta^j (e_i, f_j) = \sum_{i,j} \alpha^i \beta^j \, w_{ij} \, ,$$

o que mostra ser a aplicação ϕ univocamente determinada pelas condições $\phi(e_i, f_j) = w_{ij}$. Reciprocamente, dados arbitrariamente os vetores $w_{ij} \in W$ (e feitas as escolhas prévias das bases $\mathcal{E}$ e $\mathcal{F}$), a expressão

$$\phi(u,v) = \sum_{i,j} \alpha^i \beta^j \, w_{ij} \, \left(u = \sum \alpha^i e_i \, , v = \sum \beta^j f_j \right)$$

define, sem ambiguidade, uma aplicação $\phi \colon U \times V \to W$, a qual é evidentemente bilinear e satisfaz $\phi(e_i, f_j) = w_{ij}$.

Proposição 2. *Sejam U, V, $\mathcal{E}$ e $\mathcal{F}$ como na Proposição 1. Seja ainda $\mathcal{G} = \{g_1, \ldots, g_p\}$ uma base de W. Então:*

1) *Para cada i, j, k, com $1 \leq i \leq m$, $1 \leq j \leq n$ e $1 \leq k \leq p$, existe uma única aplicação bilinear $\mathcal{E}_k^{ij} : U \times V \to W$ tal que $\mathcal{E}_k^{ij}(e_i, f_j) = g_k$ e $\mathcal{E}_k^{ij}(e_r, f_s) = 0$ se $r \neq i$ ou $s \neq j$;*

2) *As aplicações $\mathcal{E}_k^{ij}$ constituem uma base do espaço vetorial $\mathcal{L}(U, V; W)$. Por conseguinte,* $\dim \mathcal{L}(U, V; W) = mnp$

3) *As coordenadas de uma aplicação bilinear $\phi \in \mathcal{L}(U, V; W)$ relativamente à base $\{\mathcal{E}_k^{ij}\}$ são os números ξ_{ij}^k tais que $\phi(e_i, f_j) = \sum_k \xi_{ij}^k \, g_k$.*

Demonstração: Para comprovar a primeira afirmação, dados i_0, j_0, k_0, tomamos, na Proposição 1, $w_{ij} = g_{k_0}$ se $i = i_0$ e $j = j_0$, e $w_{ij} = 0$ se $i \neq i_0$ ou $j \neq j_0$. As condições $\mathcal{E}_{k_0}^{i_0 j_0}(e_i, f_j) = w_{ij}$ definem então, univocamente,

uma aplicação bilinear $\mathcal{E}_{k_0}^{i_0 j_0}$ que satisfaz às exigências da parte 1). Para demonstrar 2), mostremos primeiro que as aplicações $\mathcal{E}_k^{ij}$ são linearmente independentes. Ora, se $\sum_{i,j,k} \lambda_{ij}^k \mathcal{E}_k^{ij} = 0$, então, em particular

$$\sum_k \left(\sum_{i,j} \lambda_{ij}^k \mathcal{E}_k^{ij}(e_r, f_s) \right) = \sum_{i,j,k} \lambda_{ij}^k \mathcal{E}_k^{ij}(e_r, f_s) = 0,$$

quaisquer que sejam r, s, com $1 \leq r \leq m$ e $1 \leq s \leq n$. Em virtude da definição dos $\mathcal{E}_k^{ij}$, isto significa que

$$\sum_k \lambda_{rs}^k g_k = 0,$$

quaisquer que sejam r, s da forma acima. Como os g_k são linearmente independentes, isto acarreta $\lambda_{rs}^k = 0$ quaisquer que sejam r, s e k, o que nos dá a independência das aplicações $\mathcal{E}_k^{ij}$. Em seguida, constataremos que estas aplicações geram $\mathcal{L}(U, V; W)$. Com efeito, dado ϕ em $\mathcal{L}(U, V; W)$, consideremos os números ξ_{ij}^k definidos como no enunciado da parte 3). Mostraremos que se tem

$$\phi = \sum_{i,j,k} \xi_{ij}^k \mathcal{E}_k^{ij}.$$

Para isto, basta mostrar que, para cada $e_r \in \mathcal{E}$ e cada $f_s \in \mathcal{F}$ tanto ϕ como a aplicação bilinear do segundo membro da igualdade acima alegada assumem o mesmo valor no par (e_r, f_s). Ora, por um lado, temos $\phi(e_r, f_s) = \sum_k \xi_{rs}^k g_k$ e, por outro lado, como $\mathcal{E}_k^{ij}(e_r, f_s) = 0$ para $i \neq r$ ou $j \neq s$, temos também

$$\sum_{i,j,k} \xi_{ij}^k \mathcal{E}_k^{ij}(e_r, f_s) = \sum_k \xi_{rs}^k \mathcal{E}_k^{rs}(e_r, f_s) = \sum_k \xi_{rs}^k g_k,$$

o que comprova a igualdade alegada, e demonstra simultaneamente 2) e 3).

Seja $\phi\colon U \times V \to W$ uma aplicação bilinear. Para cada $u_0 \in U$ fixo, a aplicação $T(u_0)\colon V \to \phi(u_0, v)$ é uma aplicação linear de V em W, ou seja, um elemento de $\mathcal{L}(V, W)$, o qual depende linearmente do parâmetro $u_0 \in U$. Estas considerações serão formalizadas na proposição a seguir.

Proposição 3. *Sejam U, V, W espaços vetoriais. Indiquemos com*

$$K\colon \mathcal{L}(U, V; W) \to \mathcal{L}(U, \mathcal{L}(V, W))$$

a aplicação que associa a cada elemento $\phi \in \mathcal{L}(U, V; W)$ o elemento $T = K(\phi) \in \mathcal{L}(U, \mathcal{L}(V, W))$ assim definido: para $u \in U$, $T(u) \in \mathcal{L}(V, W)$ é tal que $[T(u)](v) = \phi(u, v)$, $v \in V$. Então K é um isomorfismo canônico.

Demonstração: Verifica-se imediatamente que K é bem definida, isto é, que, para cada $\phi\colon U \times V \to W$ bilinear, $K(\phi)$ pertence, de fato, ao espaço $\mathcal{L}(U, \mathcal{L}(V, W))$. Também é claro que $K(\phi + \phi') = K(\phi) + K(\phi')$ e que $K(\lambda\phi) = \lambda K(\phi)$. Para mostrar que K é um isomorfismo, basta verificar sua biunivocidade, já que o domínio e o contradomínio de K têm obviamente a mesma dimensão. Seja, pois, $\phi \in \mathcal{L}(U, V; W)$ tal que $K(\phi) = 0$. Então $[K(\phi)](u) = 0$ qualquer que seja $u \in U$, e portanto $\{[K(\phi)](u)\}(v) = 0$, quaisquer que sejam $u \in U$ e $v \in V$. Mas, pela definição de K, tem-se $\{[K(\phi)](u)\}(v) = \phi(u, v)$. Logo, $\phi(u, v) = 0$ para todo $u \in U$ e todo $v \in V$, o que significa $\phi = 0$. Isto conclui a demonstração.

É claro que existe também um isomorfismo canônico $\mathcal{L}(U, V; W) \approx \mathcal{L}(V, \mathcal{L}(U, W))$.

2.2 Produtos tensoriais

Sejam U e V dois espaços vetoriais de dimensões m e n respectivamente. Chamamos *produto tensorial* de U por V a todo par (Z, ϕ) que satisfaça os seguintes axiomas:

1) Z é um espaço vetorial e $\phi\colon U \times V \to Z$ é uma aplicação bilinear de par U, V em Z;

2) $\dim Z = \dim U \dim V$;

3) $\phi(U \times V)$ gera Z, isto é, todo elemento de Z pode ser obtido como combinação linear (e portanto como soma) de elementos de $\phi(U \times V)$.

Os axiomas 2) e 3), em presença do axioma 1), são equivalentes a um único axioma 2'), que assim se enuncia:

2') Se $\mathcal{E} = \{e_1, \ldots, e_m\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ são bases de U e V respectivamente, então a mn-upla $\phi(e_i, f_j)$, $1 \leq i \leq m$, $1 \leq j \leq n$, forma uma base de W.

Com efeito, admitamos 1), 2) e 3). Dados $u = \sum \alpha^i e_i$ em U e $v = \sum \beta^j f_j$ em V, então, pela bilinearidade de ϕ, obtemos

$$\phi(u, v) = \sum_{i,j} \alpha^i \beta^j \, \phi(e_i, f_j),$$

ou seja, para todo $u \in U$ e todo $v \in V$, $\phi(u, v)$ se exprime como combinação linear dos elementos $\phi(e_i, f_j)$. Portanto, se $\phi(U \times V)$ gera Z, então a mn-upla $\phi(e_i, f_j)$ também gera Z. Mas por 2), $\dim Z = mn$. Logo esta mn-upla é uma base de Z, donde se conclui 2'). A recíproca é evidente.

Em seguida, mostraremos que, dados dois espaços vetoriais U e V, existe um par (Z, ϕ) satisfazendo aos axiomas acima, e que tal par é único a menos de um isomorfismo canônico; ou por outra, mostraremos a existência e unicidade do produto tensorial de dois espaços vetoriais U e V.

Existência: Daremos, em seguida, três construções do produto tensorial.

Primeira construção. Sejam $\mathcal{E} = \{e_1, \ldots, e_m\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ bases em U e V respectivamente. Tomemos como Z um espaço vetorial qualquer de dimensão mn. Escolhamos

$$\mathcal{H} = \{h_{11}, \ldots, h_{ij}, \ldots, h_{mn}\}$$

uma base de Z definamos $\phi\colon U \times V \to Z$, nos pares (e_i, f_j), $e_i \in \mathcal{E}$, $f_j \in \mathcal{F}$, por $\phi(e_i, f_j) = h_{ij}$, e estendamos ϕ por bilinearidade para os pares restantes, de acordo com a Proposição 1. O fato de que o par (Z, ϕ) satisfaz os axiomas acima é evidente, a partir da sua própria construção.

Segunda construção. Tomaremos como Z o espaço vetorial $\mathcal{B}(U, V)^*$, dual do espaço das formas bilineares sobre o par U, V. Definiremos $\phi\colon U \times V \to Z$ como a aplicação que ao par (u, v) associa o elemento $\psi \in \mathcal{B}(U, V)^*$ $= Z$ tal que

$$\psi(w) = w(u, v) \quad \text{para todo} \quad w \in \mathcal{B}(U, V).$$

Em outras palavras, $[\phi(u, v)](w) = w(u, v)$. A bilinearidade de ϕ segue-se imediatamente desta definição e da bilinearidade de cada $w \in \mathcal{B}(U, V)$. Para verificar o axioma

2'), consideremos $\mathcal{E} = \{e_1, \ldots, e_m\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ bases de U e V respectivamente. Como $\dim Z = \dim \mathcal{B}(U, V)^* = \dim \mathcal{B}(U.V) = \dim U.\ \dim V = mn$, basta mostrar que os elementos de mn-upla $\phi(e_i, f_j)$ são linearmente independentes, para concluir que eles constituem uma base de Z. Ora, se $\sum_{i,j} \lambda^{ij}\, \phi(e_i, f_j) = 0$, então $\big[\sum_{i,j} \lambda^{ij}\, \phi(e_i, f_j)\big](w) = 0$ qualquer que seja o elemento $w \in \mathcal{B}(U, V)$. Em outros termos, $\sum_{i,j} \lambda^{ij}\, w(e_i, f_j) = 0$ para todo $w \in \mathcal{B}(U, V)$. Para cada $e_r \in \mathcal{E}$ e cada $f_s \in \mathcal{F}$, definamos um elemento $w = w_{rs} \in \mathcal{B}(U, V)$, pondo $w_{rs}(e_r, f_s) = 1$ e $w_{rs}(e_i, f_j) = 0$ se $i \neq r$ ou $j \neq s$. Segue-se então daquela relação geral que, para todas as escolhas de r e s, tem-se $0 = \sum_{i,j} \lambda^{ij}\, w_{rs}(e_i, f_j) = \lambda^{rs}$, donde se conclui que os elementos $\phi(e_i, f_j)$ são linearmente independentes e o espaço $\mathcal{B}(U, V)^*$, juntamente com a aplicação ϕ aqui definida, constitui um produto tensorial de U por V.

Terceira construção. Consideremos o espaço vetorial $\mathcal{L}(U^*, V)$, das aplicações lineaes do dual de U em V, e definamos a aplicação $\phi\colon U \times V \to \mathcal{L}(U^*, V)$, que associa ao par $*u, v)$ a aplicação linear $T \in \mathcal{L}(U^*, V)$ tal que

$$T(w) = w(u) \cdot v, \text{ para } w \in U^*, u \in U \text{ e } v \in V.$$

Em outras palavras, $[\phi(u, v)](w) = w(u) \cdot v$. É claro que ϕ é uma aplicação bilinear do par U, V em $\mathcal{L}(U^*, V)$. Aqui também, $\dim \mathcal{L}(U^*, V) = \dim U^* \cdot \dim V = \dim U.\ \dim V = mn$. Assim, basta mostrar que, se $\mathcal{E}=\{e_1, \ldots, e_m\}$ é uma base de U e $\mathcal{F} = \{f_1, \ldots, f_m\}$ é uma base de V, então os elementos da mn-upla $\phi(e_1, f_j)$ são linearmente

independentes. Com efeito, se $\sum_{i,j} \lambda^{ij} \cdot \phi(e_i, f_j) = 0$ então $\big[\sum_{i,j} \lambda^{ij}\, \phi(e_i, f_j)\big](w) = 0$ qualquer que seja $w \in U^*$. Em particular, tomando $w = e^k$ = k-ésimo elemento da base dual $\mathcal{E}^* = \{e^1, \ldots, e^m\}$, obtemos $0 = \sum_{i,j} \lambda^{ij}\, e^k(e_i) \cdot f_j = \sum_j \lambda^{kj}\, f_j$, para todo $k = 1, \ldots, m$. Como os f_j são linearmente independentes, temos $\lambda^{kj} = 0$ para todo k e todo j, como queríamos mostrar. Concluimos que o espaço vetorial $\mathcal{L}(U^*, V)$, munido da aplicação bilinear ϕ acima definida, é um produto tensorial de U por V.

De agora por diante, indicaremos um produto tensorial de U por V com $U \otimes V$; assim, $\phi(u, v)$ será substituido por $u \otimes v$ (lê-se "*u tensor v*"). Observamos que, na definição de produto tensorial, não se tem $\phi(U \times V) = Z$. Isto significa que nem todo elemento em $U \otimes V$ é da forma $u \otimes v$, embora todo $z \in U \otimes V$ se exprima como $z = \sum u_k \otimes v_k$. Os elementos do espaço vetorial $U \otimes V$ são chamados *tensores* (de segunda ordem). Os tensores da forma $u \otimes v$ chamam-se *decomponíveis*. A maneira de exprimir um elemento $z \in U \otimes V$ como soma de tensores decomponíveis não é única. Basta ver que se tem, por exemplo, $(1/2)u \otimes 2v = u \otimes v$, como também $u \otimes v + u' \otimes v' = (u + u') \otimes v + u' \otimes (v' - v)$.

Teorema 1. *Sejam $U \otimes V$ um produto tensorial de U por V, e W um espaço vetorial qualquer. Se $g\colon U \times V \to W$ é uma aplicação bilinear, então existe uma única aplicação linear $\widetilde{g}\colon U \otimes V \to W$ tal que $\widetilde{g}(u \otimes v) = g(u, v)$ para todo $u \in U$ e todo $v \in V$.*

Demonstração: Definimos $\widetilde{g}\colon U \otimes V \to W$ pondo, para cada $z = \sum u_k \otimes v_k$ em $U \otimes V$, $\widetilde{g}(z) = \sum_k \widetilde{g}(u_k, v_k)$. Em particular, quando se trata de um tensor decomponível $u \otimes v$, obtemos a propriedade acima requerida: $\widetilde{g}(u \otimes v) = g(u, v)$. Devemos agora mostrar que a aplicação $\widetilde{g}$:

a) Está bem definida, isto é, não depende da particular maneira de exprimir o tensor z como soma de tensores decomponíveis;

b) é linear;

c) é única, nas condições requeridas.

Consideremos as bases $\mathcal{E} = \{e_1, \ldots, e_n\}$ em U, e $\mathcal{F} = \{f_1, \ldots, f_n\}$ em V. Se

$$u_k = \sum_i \alpha_k^i e_i \text{ e } v_k = \sum_j \beta_k^j f_j \text{, então}$$

$$z = \sum_k u_k \otimes v_k = \sum_{i,j,k} \alpha_k^i \beta_k^j \, e_i \otimes f_j \, .$$

Fazendo $\xi^{ij} = \sum_k \alpha_k^i \beta_k^j$, podemos escrever $z = \sum u_k \otimes v_k = \sum_{i,j} \xi^{ij} \, e_i \otimes f_j$. Assim, os números ξ^{ij} são as coordenadas do tensor z relativamente à base formada pelos tensores $e_i \otimes f_j$ e, como tal, não dependem da particular maneira de representar $z = \sum u_k \otimes v_k$ como soma de tensores decomponíveis. Ora, temos $\widetilde{g}(z) = \sum_k g(u_k, v_k) = \sum_{i,j} \xi^{ij} \, g(e_i, f_j)$, em virtude da bilinearidade de g. Logo, a aplicação $\widetilde{g}$ tem sua definição independente da maneira de escrever z como soma de tensores decomponíveis. Da expressão $g(z) = \sum_{i,j} \xi^{ij} \, g(e_i, f_j)$, onde os ξ^{ij} são as coordenadas de

z na base $\{e_i \otimes f_j\}$, vê-se imediatamente que g é uma aplicação linear. Por fim, se existisse outra aplicação linear $\widehat{g}\colon U \otimes V \to W$ tal que $\widehat{g}(u \otimes v) = g(u, v)$ para todo $u \in U$ e todo $v \in V$, teríamos, para um $z = \sum u_k \otimes v_k$ arbitrário em $U \otimes V$, $\widehat{g}(z) = \widehat{g}\left(\sum u_k \otimes v_k\right) = \sum \widehat{g}(u_k \otimes v_k) = \sum g(u_k \otimes v_k) = \sum \widetilde{g}(u_k \otimes v_k) = \widetilde{g}\left(\sum u_k \otimes v_k\right) = \widetilde{g}(z)$, donde $\widehat{g} = \widetilde{g}$.

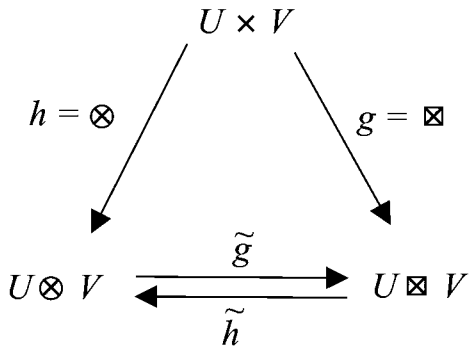
Corolário. *A correspondência $g \to \widetilde{g}$ constitui um isomorfismo canônico do espaço $\mathcal{L}(U, V; W)$ das aplicações bilineares do par U, V em W sobre o espaço $\mathcal{L}(U \otimes V, W)$ das aplicações lineares do produto tensorial $U \otimes V$ em W. Isto é:*

$$\mathcal{L}(U, V; W) \approx \mathcal{L}(U \otimes V, W).$$

Teorema 2. (Unicidade do produto tensorial.) *Sejam $U \otimes V$ e $U \boxtimes V$ dois produtos tensoriais de U por V. Existe um (único) isomorfismo canônico de $U \otimes V$ sobre $U \boxtimes V$ que leva $u \otimes v$ em $u \boxtimes v$, para todo $u \in U$ e $v \in V$.*

Demonstração: A aplicação bilinear $g\colon U \times V \to U \boxtimes V$ que, ao par (u, v), associa o elemento $g(u, v) = u \boxtimes v$, induz, de acordo com o Teorema 1, uma aplicação linear $\widetilde{g}\colon U \otimes V \to U \boxtimes V$ tal que $\widetilde{g}(u \otimes v) = u \boxtimes v$. Analogamente, a aplicação bilinear $h\colon U \times V \to U \otimes V$ tal que $h(u, v) = u \otimes v$ induz uma aplicação linear $\widetilde{h}\colon U \boxtimes V \to U \otimes V$ tal que $\widetilde{h}(u \boxtimes v) = u \otimes v$. A aplicação linear composta $\widetilde{h} \circ \widetilde{g}\colon U \otimes V \to U \otimes V$ é tal que $\widetilde{h} \circ \widetilde{g}(u \otimes v) = \widetilde{h}(u \boxtimes v) = u \otimes v$, isto é, $\widetilde{h} \circ \widetilde{g}$ coincide com a aplicação identidade no conjunto dos tensores decomponíveis de $U \otimes V$. Como tal conjunto gera $U \otimes V$, segue-se que $\widetilde{h} \circ \widetilde{g}$ é, ela própria, a aplicação identidade. Analogamente, $\widetilde{g} \circ \widetilde{h}$ é a aplicação identidade de

$U \boxtimes V$ em $U \boxtimes V$. Concluimos que $\widetilde{g}$ e $\widetilde{h}$ são isomorfismos, inversos um do outro. A unicidade de $\widetilde{g}$ resulta de ser sua ação especificada num conjunto de geradores de $U \otimes V$.



Observação: A demonstração do Teorema 2 mostra que a propriedade do produto tensorial expressa no Teorema 1 serve para caracterizá-lo. O enunciado do Teorema 1 é, muitas vezes, usado para definir produto tensorial.

2.3 Alguns isomorfismos canônicos

Comecemos com o mais simples. Considerando R como espaço vetorial (de dimensão 1) sobre si próprio, o produto tensorial $R \otimes V$ deverá ter dimensão igual à de V. Na realidade, existe um isomorfismo canônico $R \otimes V \approx V$, caracterizado por transformar $\alpha \otimes v$ em αv, para todo escalar α e $v \in V$. Basta notar que a aplicação bilinear $\phi\colon R \times V \to V$, dada por $\phi(\alpha, v) = \alpha v$ tem as propriedades requeridas para fazer do par (V, ϕ) um produto tensorial de R por V. Em seguida, aplique-se o Teorema 2 para obter o isomorfismo canônico desejado:

$$R \otimes V \approx V.$$

Já vimos que o espaço vetorial $\mathcal{L}(U^*, V)$ das aplicações lineares do dual de U em V (3ª construção), assim como o espaço vetorial $\mathcal{B}(U, V)^*$, dual do espaço das formas bilineares sobre o par U, V (2ª construção) constituem produtos tensoriais de U por V; portanto, em vista da unicidade acima demonstrada, valem os isomorfismos canônicos abaixo:

$$U \otimes V \approx \mathcal{B}(U, V)^* \approx \mathcal{L}(U^*, V).$$

Explicitamente, o isomorfismo canônico entre $U \otimes V$ e $\mathcal{B}(U, V)^*$, dado pelo Teorema 2, é a aplicação que, ao tensor decomponível $u \otimes v$, associa o funcional linear $\psi \in \mathcal{B}(U, V)^*$ tal que $\psi(w) = w(u, v)$, para $w \in \mathcal{B}(U, V)$, $u \in U$, $v \in V$. Com efeito, este funcional ψ é o "produto tensorial" de u por v de acordo com a Segunda Construção.

Analogamente, de acordo com a Terceira Construção, o produto tensorial de $u \in U$ por $v \in V$ é a transformação linear $A\colon U^* \to V$ tal que $A(f) = f(u) \cdot v$ para todo $f \in U^*$. Portanto o isomorfismo canônico entre $U \otimes V$ e $\mathcal{L}(U^*, V)$ leva o tensor decomponível $u \otimes v$ na aplicação linear $A\colon U^* \to V$ tal que $A(f) = f(u)v$, $f \in U^*$, $u \in U$, $v \in V$. Observemos que a imagem de U^* por uma tal aplicação é um subespaço de V de dimensão 1. Ou seja, os tensores decomponíveis em $U \otimes V$ correspondem às aplicações lineares "de posto 1" em $\mathcal{L}(U^*, V)$.

Proposição 4. *O espaço $\mathcal{L}(U, V)$ das aplicações lineares de U em V é canonicamente isomorfo a $U^* \otimes V$, isto é:*

$$\mathcal{L}(U, V) \approx U^* \otimes V.$$

Demonstração: Como $U \approx (U^*)^*$, segue-se que

$$\mathcal{L}(U, V) \approx \mathcal{L}((U^*)^*, V).$$

Por outro lado, substituindo U por U^* em $\mathcal{L}(U^*, V) \approx U \otimes V$, obtemos $\mathcal{L}((U^*)^*, V) \approx U^* \otimes V$. Combinando estes isomorfismos, obtemos o resultado desejado.

O isomorfismo acima leva o tensor decomponível $f \otimes v \in U^* \otimes V$ na aplicação linear $A \in \mathcal{L}(U, V)$ tal que $A(u) = f(u)v$, $f \in U^*$, $u \in U$, $v \in V$. Novamente, os tensores decomponíveis em $U^* \otimes V$ correspondem às aplicações lineares $A\colon U \to V$ que têm posto 1 (isto é, a imagem de U por A tem dimensão 1).

Um caso particular do isomorfismo anterior é

$$\mathcal{L}(U) \approx U^* \otimes U,$$

onde $\mathcal{L}(U)$ é o espaço vetorial das transformações lineares de U em si próprio.

Assim, as transformações lineares $A\colon U \to U$ podem ser identificadas aos *tensores mistos* de 2ª ordem sobre U, isto é, aos elementos do espaço $U^* \otimes U$. Os elementos de $U \otimes U$ são chamados *tensores contravariantes* de 2ª ordem, e os de $U^* \otimes U^*$ são os *tensores covariantes* de 2ª ordem.

Proposição 5. *Consideremos as bases $\mathcal{E} = \{e_1, \ldots, e_m\}$ em U, sua dual $\mathcal{E}^* = \{e^1, \ldots, e^m\}$ em U^*, e $\mathcal{E}^* \otimes \mathcal{E} = \{e^j \otimes e_1 \,; i, j = 1, \ldots, m\}$ em $U^* \otimes U$. Sejam $A\colon U \to U$ uma aplicação linear e $(\alpha^i_j) = [A, \mathcal{E}]$ sua matriz na base $\mathcal{E}$. Então as coordenadas, na base $\mathcal{E}^* \otimes \mathcal{E}$, do tensor $t \in U^* \otimes U$, correspondente a A no isomorfismo canônico $\mathcal{L}(U) \approx U^* \otimes U$, coincidem com os elementos da matriz (α^i_j).*

Demonstração: Lembramos, inicialmente, que a transformação linear $A \in \mathcal{L}(U)$ correspondente ao tensor

$$t = \sum_k f^k \otimes u_k$$

é definida por $A(u) = \sum_k f^k(u)u_k\,, \quad u \in U$. Se $u_k = \sum_i \beta_k^i e_i$ e $f^k = \sum_j \gamma_j^l e^j$, então, como já sabemos, $f^k(e_j) = \sum_i \gamma_i^k e^i(e_j) = \gamma_j^k$. Portanto,

$$A(e_j) = \sum_k \gamma_j^k \left(\sum_i \beta_k^i e_i \right) = \sum_i \left(\sum_k \gamma_j^k \beta_k^i \right) e_i\,.$$

Por outro lado, como $(\alpha_j^i) = [A; \mathcal{E}]$, segue-se que $A(e_j) = \sum_i \alpha_j^i e_i$. Comparando as expressões acima, obtemos $\alpha_j^i = \sum_k \gamma_j^k \beta_k^i$. Mas é fácil ver que o tensor $t = \sum f^k \otimes u_k$ se exprime, na base $\mathcal{E}^* \otimes \mathcal{E}$, como $t = \sum_{i,j} \left(\sum_k \gamma_j^k \beta_k^i \right) e_i \otimes f^j$, donde $t = \sum_{i,j} \alpha_j^i e_i \otimes f^j$, o que conclui a demonstração.

Definido no espaço vetorial $U^* \otimes U$, existe um funcional linear natural ϕ, caracterizado pela condição seguinte:

$$\phi(f \otimes u) = f(u), \quad f \in U^*, \quad u \in U.$$

Este funcional é a aplicação linear induzida em $U^* \otimes U$ pela aplicação bilinear $g\colon U^* \times U \to R$ tal que $g(f, u) = f(u)$. Assim, temos

$$\phi \left(\sum_k f^k \otimes u_k \right) = \sum_k f^k(u_k).$$

(Cfr. Teorema 1). O funcional linear ϕ composto com o isomorfismo canônico $\mathcal{L}(U) \approx U^* \otimes U$ dá um funcional linear

$$\tau\colon \mathcal{L}(U) \to R,$$

o qual é denominado *traço*. Se $A \in \mathcal{L}(U)$ é a transformação linear correspondente ao tensor $\sum f^k \otimes u_k$, então o traço de A é

$$\tau(A) = \sum f^k(u_k),$$

pois $\tau(A) = \phi\left(\sum f^k \otimes u_k\right)$.

Proposição 6. *Se (α^i_j) é a matriz de A numa base qualquer $\mathcal{E}$, então o traço de A é a soma dos elementos da diagonal dessa matriz, isto é,*

$$\tau(A) = \sum \alpha^i_i \,.$$

Demonstração: Usando as notações há pouco introduzidas, sejam

$$u_k = \sum_j \beta^j_k e_j \quad \text{e} \quad f^k = \sum_i \gamma^k_i e^i.$$

Então:

$$\begin{aligned}
\tau(A) &= \sum_k \left[\sum_i \gamma^k_i e^i \left(\sum_j \beta^j_k e_j \right) \right] = \\
&= \sum_k \left[\sum_{i,j} \gamma^k_i \beta^j_k e^i(e_j) \right] = \\
&= \sum_k \sum_i \gamma^k_i \beta^i_k = \sum_i \left(\sum_k \gamma^k_i \beta^i_k \right) = \sum_i \xi^i_i \,,
\end{aligned}$$

onde os ξ^i_j são as coordenadas do tensor $\sum f^k \otimes u_k$ relativamente à base $\mathcal{E}^* \otimes \mathcal{E}$, as quais, como já vimos, coincidem

com os elementos da matriz (α_j^i). Portanto, $\tau(A) = \sum \alpha_i^i$, como queríamos demonstrar.

Em consequência da proposição acima, podemos afirmar que, se (α_j^i) e (β_j^i) são matrizes da mesma aplicação linear $A\colon U \to U$ relativamente a duas bases distintas de U, então

$$\sum \alpha_i^i = \sum \beta_i^i \,.$$

Proposição 7. *O dual do produto tensorial de U por V é canonicamente isomorfo ao produto tensorial do dual de U pelo dual de V:*

$$(U \otimes V)^* \approx U^* \otimes V^*.$$

Demonstração: Temos sucessivamente:

$$(U \otimes V)^* \approx \mathcal{B}(U,V) \approx \mathcal{L}(U, \mathcal{L}(V,R)) = \mathcal{L}(U,V^*) \approx U^* \otimes V^*,$$

onde o primeiro isomorfismo é dado pelo Corolário do Teorema 1, o segundo pela Proposição 3, e o último pela Proposição 4.

Corolário. *$U^* \otimes V^* \approx \mathcal{B}(U,V)$. Em particular, os tensores covariantes de 2^a ordem sobre U (elementos de $U^* \otimes U^*$) identificam-se canonicamente às formas bilineares sobre U (elementos de $\mathcal{B}(U)$).*

O isomorfismo dado pela Proposição 7 associa ao tensor $f \otimes g \in U^* \otimes V^*$ o funcional linear $\psi \in (U \otimes V)^*$ tal que $\psi(u \otimes v) = f(u)g(v)$.

O isomorfismo $U^* \otimes U^* \approx \mathcal{B}(U)$, entre tensores covariantes de segunda ordem e formas bilineares, associa ao tensor

$f \otimes g$ a forma bilinear $f \cdot g$, tal que $(f \cdot g)(u, v) = f(u) \cdot g(v)$, a qual é chamada de *produto* das formas lineares f e g. Outra maneira (direta) de obter o mesmo isomorfismo é usar a unicidade do produto tensorial, tendo observado antes que a aplicação bilinear $\phi\colon U^* \times U^* \to \mathcal{B}(U)$ tal que $\phi(f, g) = f \cdot g$ goza das propriedades que fazem do par $(\mathcal{B}(U), \phi)$ um produto tensorial de U^* por U^*. O mesmo argumento pode ser usado para demonstrar diretamente o corolário acima.

Proposição 8. *O produto tensorial goza das seguintes propriedades formais:*

1) $U \otimes V \approx V \otimes U$;
2) $(U \otimes V) \otimes W \approx U \otimes (V \otimes W)$;
3) $(U \oplus V) \otimes W \approx (U \otimes W) \oplus (V \otimes W)$.

Demonstração: Os isomorfismos acima são as únicas aplicações lineares $\phi_1 : U \otimes V \to V \otimes U$, $\phi_2 : (U \otimes V) \otimes W \to U \otimes (V \otimes W)$ e $\phi_3 : (U \oplus V) \otimes W \to (U \otimes W) \oplus (V \otimes W)$ tais que

$$\phi_1(u \otimes v) = v \otimes u, \ \phi_2[(u \otimes v) \otimes w] = u \otimes (v \otimes w)$$

e $\phi_3 : [(u + v) \otimes w] = (u \otimes w) + (v \otimes w)$.

O leitor deverá demonstrar que as aplicações acima constituem, de fato, isomorfismos (usar a unicidade do produto tensorial). As propriedades 1), 2), e 3) podem ser denominadas "comutatividade", "associatividade" e "distributividade em relação à soma direta".

Atenção: Mesmo num produto tensorial $V \otimes V$, *não* se tem, em geral, $u \otimes v = v \otimes u$. A "comutatividade" existe para produto tensoriais *de espaços*, mas não para um produto $u \otimes v$ de vetores.

2.4 Produto tensorial de aplicações lineares

Sejam $A\colon U \to W$ e $B\colon V \to Z$ aplicações lineares. O *produto tensorial* de A por B é a aplicação linear $A \otimes B\colon U \otimes V \to W \otimes Z$ caracterizada pela igualdade

$$A \otimes B(u \otimes v) = A(U) \otimes B(v), \quad u \in U,\, v \in V.$$

$A \otimes B$ é obtida, por meio do Teorema 1, como $A \otimes B = \widetilde{g}$, sendo $g\colon U \times V \to W \otimes Z$ a aplicação bilinear dada por $g(u, v) = A(u) \otimes B(v)$.

Sejam $A, B\colon U \to U$ aplicações lineares, e $\mathcal{E}=\{e_1, ..., e_m\}$ uma base de U. Sejam ainda $\alpha = (\alpha^i_j) = [A; \mathcal{E}]$ e $\beta = (\beta^i_j) = [B; \mathcal{E}]$ as matrizes de A e B na base $\mathcal{E}$. A matriz $[A \otimes B, \mathcal{E} \otimes \mathcal{E}]$, da aplicação linear $(A \otimes B\colon U \otimes U \to U \otimes U$ na base $\mathcal{E} \otimes \mathcal{E} = \{e_i \otimes e_j; i, j = 1, \ldots, m\}$ é denominada *produto de Kronecker*, ou *produto tensorial* das matrizes α e β, e é indicada com $\alpha \otimes \beta$.

Como $A(e_k) = \sum_i \alpha^i_k e_i$ e $B(e_h) = \sum_j \beta^j_h e_j$, temos:

$$\begin{aligned}(A \otimes B)(e_k \otimes e_h) &= A(e_k) \otimes B(e_h) = \\ &= \left(\sum_i \alpha^i_k e_i\right) \otimes \left(\sum_j \beta^j_h e_j\right) = \\ &= \sum_{i,j} \alpha^i_k\, \beta^j_h\, e_i \otimes e_j\,,\end{aligned}$$

e portanto:

$$\alpha \otimes \beta = (\alpha_k^i \beta_h^j) =$$

$$= \begin{pmatrix} \alpha_1^1\,\beta_1^1 & \dots & \alpha_1^1\,\beta_n^1 & \dots & \alpha_n^1\,\beta_1^1 & \dots & \alpha_n^1\,\beta_n^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_1^1\,\beta_1^n & \dots & \alpha_1^1\,\beta_n^n & \dots & \alpha_n^1\,\beta_1^n & \dots & \alpha_n^1\,\beta_n^n \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_1^n\,\beta_1^1 & \dots & \alpha_1^n\,\beta_n^1 & \dots & \alpha_n^n\,\beta_1^1 & \dots & \alpha_n^n\,\beta_n^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_1^n\,\beta_1^n & \dots & \alpha_1^n\,\beta_n^n & \dots & \alpha_n^n\,\beta_1^n & \dots & \alpha_n^n\,\beta_n^n \end{pmatrix}$$

Mais abreviadamente, podemos escrever:

$$\alpha \otimes \beta = \begin{pmatrix} \alpha_1^1\beta & \dots & \alpha_n^1\beta \\ \vdots & & \vdots \\ \alpha_1^n\beta & \dots & \alpha_n^n\beta \end{pmatrix}$$

Proposição 9. *Sejam $A\colon U \to V$, $B\colon V \to W$, $A'\colon U' \to V'$, $B'\colon V' \to W'$ aplicações lineares. Então*

$$BA \otimes B'A' = (B \otimes B')(A \otimes A')\colon U \otimes U' \to W \otimes W'.$$

Demonstração: Temos $[BA \otimes B'A'](u \otimes u') = BA(u) \otimes B'A'(u') = (B \otimes B')[A(u) \otimes A'(u')] = [(B \otimes B')(A \otimes A')](u \otimes u')$, quaisquer que sejam $u \in U$ e $u' \in U'$, o que demonstra a proposição.

Corolário. *Se A e A' são isomorfismos, então $A \otimes A'$ é um isomorfismo, cujo inverso é $A^{-1} \otimes (A')^{-1}$.*

Segue-se do corolário acima que o produto de Kronecker de duas matrizes invertíveis é uma matriz invertível, e o inverso de $\alpha \otimes \beta$ é igual a $\alpha^{-1} \otimes \beta^{-1}$.

Como aplicação, podemos considerar as operações de baixar e subir índices num produto tensorial de espaços vetoriais euclidianos.

Vimos no Capítulo 1, §7, que para todo espaço vetorial euclidiano V existe um isomorfismo canônico $J\colon V \to V^*$, definido por $[J(u)](v) = u \cdot v$. Dada uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, se $u = \sum \alpha^i e_i$ então a expressão de $J(u)$ relativamente à base dual $\mathcal{E}^*$ é dada por $J(u) = \sum \alpha_i e^i$, onde

$$\alpha_i = \sum_j g_{ij}\, \alpha^j,$$

sendo $g_{ij} = e_i \cdot e_j$. Esta fórmula ensina a passar das coordenadas contravariantes α^i do vetor V na base $\mathcal{E}$ para as suas coordenadas covariantes $\alpha_i = v \cdot e_i$. Ela também pode ser interpretada como significando que a matriz da aplicação linear $J\colon V \to V^*$ relativamente às bases $\mathcal{E}$, $\mathcal{E}^*$ é a matriz (g_{ij}), $g_{ij} = e_i \cdot e_j$. Assim, a matriz da aplicação inversa $J^{-1}\colon V^* \to V$, relativamente às bases $\mathcal{E}^*$, $\mathcal{E}$, é a matriz inversa $(g_{ij})^{-1}$, a qual indicaremos sempre com (g^{ij}):

$$(g_{ij})^{-1} = (g^{ij}).$$

Portanto, conhecidas as coordenadas covariantes α_i de um vetor u na base $\mathcal{E}$, suas coordenadas contravariantes são dadas por

$$\alpha^i = \sum_j g^{ij}\, \alpha_j\,.$$

Examinamos agora o produto tensorial de dois espaços euclidianos. Por simplicidade, restringir-nos-emos a um

produto da forma $V \otimes V$, onde V é euclidiano. (Assim, o número de bases a considerar será reduzido à metade.)

Em virtude da Proposição 9, o produto tensorial do isomorfismo $J\colon V \to V^*$ por si mesmo é ainda um isomorfismo:

$$J \otimes J\colon V \otimes V \approx V^* \otimes V^*.$$

Dada a base $\mathcal{E}$ em V, a matriz de $J \otimes J$ relativamente à base $\mathcal{E} \otimes \mathcal{E}^*$ é o produto de Kronecker $(g_{ij}) \otimes (g_{rs}) = (g_{ij}g_{rs})$. Por conseguinte, se ξ^{ij} são as componentes (contravariantes) de um tensor $t = \sum \xi^{ij}\, e_i \otimes e_j \in V \otimes V$, suas componentes covariantes (isto é, as componentes de $[J \otimes J](t)$)

$$\xi_{ij} = \sum_{r,s} g_{ir} g_{js}\, \xi^{rs}.$$

De modo análogo, a matriz da aplicação inversa $(J \otimes J)^{-1} = J^{-1} \otimes J^{-1}\colon V^* \otimes V^* \to V \otimes V$, nas bases dadas, é o produto de Kronecker $(g^{ij}) \otimes (g^{ts})$, de modo que um tensor de componentes covariantes ξ_{ij} terá suas componentes contravariantes dadas por:

$$\xi^{ij} = \sum_{r,s} g^{ir} g^{js}\, \xi_{rs}\,.$$

Sendo V um espaço vetorial euclidiano, existe em $V \otimes V$ um produto interno, caracterizado pela propriedade:

$$(u \otimes v)\dot{(}u' \otimes v') = (u \cdot v)(u' \cdot v').$$

Com efeito, o produto interno de V, sendo uma forma bilinear g em V, induz um funcional linear $\widetilde{g}$ sobre $V \otimes V$, caracterizado pela relação $\widetilde{g}(u \otimes v) = u \cdot v$ (cfr. Teorema 1).

Consideremos a forma bilinear $\phi = \widetilde{g} \cdot \widetilde{g}$: $(V \otimes V) \times (V \otimes V) \to R$, produto de $\widetilde{g}$ por si próprio (vide definição deste produto entre as Proposições 7 e 8). Temos:

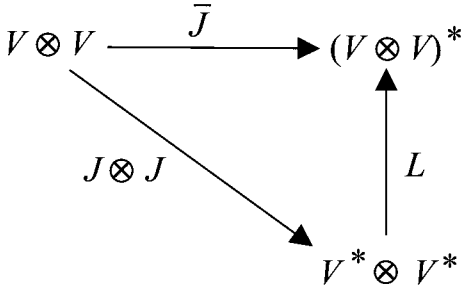
$$\phi(u \otimes v, u' \otimes v') = g(u, v)g(u', v') = (u \cdot v)(u' \cdot v').$$

É fácil verificar que a forma bilinear ϕ é um produto interno em $V \otimes V$. Como os tensores decomponíveis geram $V \otimes V$, este produto interno é o único que satisfaz a igualdade estipulada.

O produto interno induzido por V em $V \otimes V$ determina um isomorfismo canônico:

$$\overline{J}\colon V \otimes V \approx (V \otimes V)^*.$$

É fácil ver que o diagrama abaixo é comutativo (isto é, $L_o(J \otimes J) = \overline{J}$)



onde L indica o isomorfismo canônico da Proposição 7.

É claro que se podem considerar também os isomorfismos:

$$J \otimes \text{ id}\colon V \otimes V \to V^* \otimes V, \quad \text{id} \otimes J\colon V \otimes V \to V \otimes V^*,$$

onde id é a aplicação identidade. Se ξ^{ij} são as componentes contravariantes de um tensor $t \in V \otimes V$ na base $\mathcal{E} \otimes \mathcal{E}$, então suas componentes mistas são

$$\xi_i^j = \sum_r g_{ir}\, \xi^{rj}$$

(componentes de $J \otimes \mathrm{id}(t)$ na base $\mathcal{E}^* \otimes \mathcal{E}$) e

$$\xi^i_j = \sum_r g_{jr}\, \xi^{ir}$$

(componentes de $\mathrm{id} \otimes J(t)$ na base $\mathcal{E} \otimes \mathcal{E}^*$).

2.5 Mudança de coordenadas de um tensor

Será, talvez, útil dizer algumas palavras sobre a mudança de coordenadas num produto tensorial, assunto central das exposições clássicas do cálculo tensorial.

Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ bases do espaço vetorial V. Seja $\lambda = (\lambda^i_j)$ a matriz de passagem de $\mathcal{F}$ para $\mathcal{E}$. Temos portanto $e_j = \sum_i \lambda^i_j f_i$. Sabemos que, indicando com $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ e $\mathcal{F}^* = \{f^1, \ldots, f^n\}$ as bases duais de $\mathcal{E}$ e $\mathcal{F}$ respectivamente, temos $e^j = \sum_i \mu^j_i f^i$, onde $\mu = (\mu^j_i)$ é a matriz inversa de λ: $\mu = \lambda^{-1}$ (Capítulo 1, Proposição 11).

Considerando as bases $\mathcal{E} \otimes \mathcal{E}$, $\mathcal{F} \otimes \mathcal{F}$ em $V \otimes V$, $\mathcal{E}^* \otimes \mathcal{E}^*$, $\mathcal{F}^* \otimes \mathcal{F}^*$ em $V^* \otimes V^*$ e $\mathcal{E} \otimes \mathcal{E}^*$, $\mathcal{F} \otimes \mathcal{F}^*$ em $V \otimes V^*$, obtemos, com um cálculo simples:

$$e_i \otimes e_j = \sum_{r,s} \lambda^r_i\, \lambda^s_j\, f_r \otimes f_s\,; \quad e^i \otimes e^j = \sum_{r,s} \mu^i_r\, \mu^j_s\, f^r \otimes f^s;$$

$$e_i \otimes e^j = \sum_{r,s} \lambda^r_i\, \mu^j_s\, f_r \otimes f^s.$$

Ou seja, as matrizes de passagem de $\mathcal{F} \otimes \mathcal{F}$ para $\mathcal{E} \otimes \mathcal{E}$, de $\mathcal{F}^* \otimes \mathcal{F}^*$ para $\mathcal{E}^+ \otimes \mathcal{E}^*$ e de $\mathcal{F} \otimes \mathcal{F}^*$ para $\mathcal{E} \otimes \mathcal{E}^*$ são respectivamente $\lambda \otimes \lambda$, $\lambda^{-1} \otimes \lambda^{-1}$ e $\lambda \otimes \lambda^{-1}$. Consequentemente, se os tensores $t \in V \otimes V$, $t' \in V^* \otimes V^*$ e $t'' \in V \otimes V^*$ têm coordenadas ξ^{ij}, ξ_{ij} e ξ^i_j respectivamente, na base $\mathcal{E}$ (abuso de linguagem evidente), suas coordenadas na base $\mathcal{F}$ serão os números ζ^{ij}, ζ_{ij} e ζ^i_j, definidos por:

$$\zeta^{ij} = \sum_{r,s} \lambda^i_r \, \lambda^j_s \, \xi^{rs}; \quad \zeta_{ij} = \sum_{r,s} \mu^r_i \, \mu^s_j \, \xi_{rs}$$

$$\zeta^i_j = \sum_{r,s} \lambda^i_r \, \mu^s_j \, \xi^r_s \, .$$

As fórmulas de mudança de coordenadas acima obtidas constituem a própria definição de um tensor do ponto de vista clássico.

2.6 Produto tensorial de vários espaços vetoriais

Indicaremos a seguir, de modo breve, as modificações que devem ser feitas para tratar do produto tensorial de p espaços vetoriais, onde p é um inteiro positivo qualquer.

Dados os espaços vetoriais $V_1, ..., V_p$ e W, uma aplicação $\phi\colon V_1 \times \cdots \times V_p \to W$ chama-se *p-linear* quando é linear separadamente em cada variável.

O conjunto $\mathcal{L}(V_1, \ldots, V_p, W)$ das aplicações p-lineares de $V_1 \times \cdots \times V_p$ em W é um espaço vetorial de dimensão $n_1 \cdot n_2 \cdot \ldots \cdot n_p \cdot m$, onde $n_i = \dim V_i$ e $m = \dim W$. Valem os análogos das Proposições 1 e 2, mutatis mutandi. Qualquer

que seja a permutação $(i_1, \ldots, i_k, j_1, \ldots, j_{p-k})$ dos inteiros de 1 até p, tem-se um isomorfismo canônico:

$$\mathcal{L}(V_1, \ldots, V_p; W) \approx \mathcal{L}(V_{i_1}, \ldots, V_{i_k}; \mathcal{L}(V_{j_1}, \ldots, V_{j_{p-k}}; W)).$$

Quando se tem $V_1 = \cdots = V_p = V$, escreve-se $\mathcal{L}_p(V; W)$ em vez de $\mathcal{L}(V, \ldots, V; W)$.

Um *produto tensorial* dos espaços vetoriais $V_1, \ldots, V_p$ é um par (Z, ϕ) com as seguintes propriedades:

1) Z é um espaço vetorial e $\phi\colon V_1 \times \cdots \times V_p \to Z$ é uma aplicação p-linear;

2) $\dim Z = \dim V_1 \cdot \dim V_1 \cdot \cdots \cdot \dim V_p$;

3) $\phi(V_1 \times \cdots \times V_p)$ gera Z.

Os axiomas 2) e 3), em presença de 1), são equivalentes ao único axioma 2') seguinte:

2') Sejam $\mathcal{E}_1 = \{e_{11}, \ldots, e_{1n_1}\}, \ldots, \mathcal{E}_p = \{e_{p1}, \ldots, e_{pn_p}\}$ bases de $V_1, \ldots, V_p$ respectivamente. Então o conjunto $\mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_p = \{\phi(e_{1i_1}, e_{2i_2}, \ldots, e_{pi_p}); 1 \leq i_1 \leq n_1, \ldots, 1 \leq i_p \leq n_p\}$ constitui uma base de Z.

Dado um produto tensorial (Z, ϕ) dos espaços vetoriais $V_1, \ldots, V_p$, escreve-se $Z = V_1 \otimes \cdots \otimes V_p$ e $\phi(v_1, \ldots, v_p) = v_1 \otimes \cdots \otimes v_p$. O produto tensorial $V_1 \otimes \cdots \otimes V_p$ é único, a menos de um isomorfismo canônico. Isto decorre da seguinte propriedade:

Dada uma aplicação p-linear $g\colon V_1 \times \cdots \times V_p \to W$, existe uma única aplicação linear $\widetilde{g}\colon V_1 \otimes \cdots \otimes V_p \to W$ tal que $\widetilde{g}(v_1 \otimes \cdots \otimes v_p) = g(v_1, \ldots, v_p)$. A correspondência $g \to \widetilde{g}$ estabelece um isomorfismo canônico

$$\mathcal{L}(V_1, \ldots, V_p; W) \approx \mathcal{L}(V_1 \otimes \cdots \otimes V_p, W).$$

Como exemplos de produtos tensoriais de $V_1, \ldots, V_p$, podemos proceder de maneira análoga à da primeira construção do §2, ou então podemos tomar para Z o espaço

$$\mathcal{L}(V_1, \ldots, V_p; R)^*,$$

dual do espaço das formas p-lineares sobre $V_1 \times \cdots \times V_p$, sendo ϕ definida do modo natural: $[\phi(v_1, \ldots, v_p)](w) = w(v_1, \ldots, v_p)$ para todo $w \in \mathcal{L}(V_1, \ldots, V_p; R)$. Podemos ainda tomar $Z = \mathcal{L}(V_1^*, \ldots, V_{p-1}^*; V_p)$, sendo ϕ definida assim:

$$[\phi(v_1, \ldots, v_{p-1}.v_p)](f^1, \ldots, f^{p-1}) = \\ = f^1(v_1) f^2(v_2) \ldots f^{p-1}(v_{p-1}) v_p \,.$$

E, finalmente, admitindo que sabemos formar o produto tensorial de 2 espaços (o que é um fato), podemos dar uma definição indutiva de produto tensorial de p espaços por meio da fórmula:

$$V_1 \otimes \cdots \otimes V_p = (V_1 \otimes \cdots \otimes V_{p-1}) \otimes V_p \,.$$

(Bem entendido, $\phi(v_1, \ldots, v_p) = v_1 \otimes \cdots \otimes v_p$ é definido como $(v_1 \otimes \cdots \otimes v_{p-1}) \otimes v_p$.)

Todas estas construções do produto tensorial $V_1 \otimes \cdots \otimes V_p$ são canonicamente isomorfas, de modo que não há ambiguidade nem perda de generalidade em fixar-nos a uma qualquer delas, já que existe um modo natural de passar de um elemento dessa construção para um elemento bem definido de outra.

Não há dificuldade em estabelecer isomorfismos canônicos da forma $V_1^* \otimes \cdots \otimes V_p^* \approx (V_1 \otimes \cdots \otimes V_p)^* \approx \mathcal{L}(V_1, \ldots, V_p; R)$ ou $\mathcal{L}(V_1, \ldots, V_p; W) \approx V_1^* \otimes \cdots \otimes V_p^* \otimes W$.

Valem também, para p fatores, as propriedades comutativa e associativa do produto tensorial.

Usaremos a notação

$$V_s^r = \overbrace{V \otimes \cdots \otimes V}^{r} \otimes \overbrace{V^* \otimes \cdots \otimes V^*}^{s}$$

para indicar o produto tensorial de r cópias de V por s cópias do seu dual V^*. Assim, por exemplo, $V_0^2 = V \otimes V$, $V_1^1 = V \otimes V^*$. Escreveremos $V_0^0 = R$, $V_0^1 = V$ e $V_1^0 = V^*$. Com base na propriedade comutativa do produto tensorial, identificaremos com V_s^r todo produto tensorial de r espaços iguais a V por s espaços iguais a V^*, em qualquer ordem. Os elementos de V_s^r são chamados *tensores r vezes contravariantes e s vezes covariantes.* Toda base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V induz uma base $\mathcal{E}_s^r$ em V_s^r, formada por todos os tensores do tipo $e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$, onde os r primeiros fatores são tirados de $\mathcal{E}$ e os s últimos da base dual $\mathcal{E}^*$. As coordenadas de um tensor $t \in V_s^r$ relativamente à base $\mathcal{E}_s^r$ serão indicadas com $\xi_{j_1 \ldots j_s}^{i_1 \ldots i_r}$ de modo que

$$t = \sum \xi_{j_1 \ldots j_s}^{i_1 \ldots i_r} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$$

onde cada índice do somatório varia, independentemente dos outros, de 1 até n.

Dadas $\mathcal{E}$ e $\mathcal{F}$, bases em V, se λ é a matriz de passagem de $\mathcal{F}$ para $\mathcal{E}$, a matriz de passagem de $\mathcal{F}_s^r$ para $\mathcal{E}_s^r$ será $\lambda \otimes \cdots \otimes \lambda \otimes \lambda^{-1} \otimes \cdots \otimes \lambda^{-1}$ (r fatores iguais a λ e s iguais a λ^{-1}). Assim, as coordenadas do tensor $t \in V_s^r$, acima considerado, na base $\mathcal{F}_s^r$ são os números

$$\zeta_{j_1 \ldots j_s}^{i_1 \ldots i_r}$$

dados por:

$$\zeta_{j_1\ldots j_s}^{i_1\ldots i_r} = \sum \lambda_{k_1}^{i_1} \ldots \lambda_{k_r}^{i_r}\, \mu_{j_1}^{m_1} \ldots \mu_{j_1}^{m_s}\, \xi_{m_1\ldots m_s}^{k_1\ldots k_r}$$

onde o somatório estende-se a todos os índices $k_1, \ldots, k_r$, $m_1, \ldots, m_s$, os quais variam, independentemente, de 1 até n, sendo $\lambda = (\lambda_j^i)$ e $\lambda^{-1} = (\mu_j^i)$.

O espaço $V_{s+s'}^{r+r'}$, pode ser considerado, de modo natural, como produto tensorial de V_s^r por $V_{s'}^{r'}$. Logo existe uma aplicação bilinear canônica:

$$\phi\colon V_s^r \times V_{s'}^{r'} \to V_{s+s'}^{r+r'}$$

caracterizada pela igualdade

$$\begin{aligned}
&\phi(u_1 \otimes \cdots \otimes u_r \otimes f^1 \otimes \cdots \otimes f^s,\\
&\ v_1 \otimes \cdots \otimes v_{r'} \otimes g^1 \otimes \cdots \otimes g^{s'}) =\\
&= u_1 \otimes \cdots \otimes u_r \otimes v_1 \otimes \cdots \otimes v_{r'} \otimes\\
&\otimes f^1 \otimes \cdots \otimes f^s \otimes g^1 \otimes \cdots \otimes g^{s'}.
\end{aligned}$$

A aplicação ϕ chama-se a *multiplicação de tensores.* Escreveremos $tt' \in V_{s+s'}^{r+r'}$, em vez de $\phi(t, t')$, onde $t \in V_s^r$ e $t' \in V_{s'}^{r'}$. Note-se que tt' é a imagem de $t \otimes t' \in V_s^r \otimes V_{s'}^{r'}$ pelo isomorfismo canônico $V_s^r \otimes V_{s'}^{r'} \to V_{s+s'}^{r+r'}$. A multiplicação de tensores é associativa: se $t'' \in V_{s''}^{r''}$ é um terceiro tensor, então $(tt')t'' = t(t't'')$. Para verificar esta relação, basta considerar o caso em que t, t' e t'' são decomponíveis. Neste caso, a relação alegada segue-se imediatamente da expressão de ϕ, acima dada.

Uma operação importante entre tensores mistos é a *contração.* Ela generaliza a forma bilinear canônica $V \otimes V^* \to R$.

Consideremos o espaço V_s^r dos tensores r vezes contravariantes e s vezes convariantes, onde $r > 0$ e $s > 0$. Sejam i um inteiro entre 1 e r e j um inteiro entre 1 e s. Definiremos a contração c_j^i, do i-ésimo índice contravariante com o j-ésimo índice covariante, como a aplicação linear

$$c_j^i \colon V_s^r \to V_{s-1}^{r-1}$$

caracterizada pela igualdade:

$$c_j^i(v_1 \otimes \cdots \otimes v_r \otimes f^1 \otimes \cdots \otimes f^s) =$$
$$= f^j(v_i) v_1 \otimes \cdots \otimes \widehat{v_i} \otimes \cdots \otimes v_r \otimes f^1 \otimes \cdots \otimes \widehat{f^j} \otimes \cdots \otimes f^s,$$

onde o sinal $\widehat{\ }$ sobre um elemento numa fórmula significa que tal elemento deve ser omitido. Por exemplo:

$$c_s^1(v_1 \otimes \cdots \otimes v_r \otimes f^1 \otimes \cdots \otimes f^s) =$$
$$= f^s(v_1) v_2 \otimes \cdots \otimes v_r \otimes f^1 \otimes \cdots \otimes f^{s-1}.$$

A existência de contração c_j^i é assegurada pelo análogo do Teorema 1, com $c_j^i = \widetilde{g}$, onde

$$g \colon \overbrace{V \times \cdots \times V}^{r} \times \overbrace{V^* \times \cdots \times V^*}^{s} \longrightarrow V_{s-1}^{r-1}$$

é a aplicação $(r+s)$-linear dada por

$$g(v_1, \ldots, v_r, f^1, \ldots, f^s) =$$
$$= f^j(v_i) v_1 \otimes \cdots \times \widehat{v_i} \otimes \cdots \otimes v_r \otimes f^1 \otimes \cdots \otimes \widehat{f^j} \otimes \cdots \otimes f^s.$$

Em termos de coordenadas, se $t \in V_s^r$ é um tensor cujas coordenadas na base $\mathcal{E}$ são

$$\xi_{m_1 \ldots m_s}^{k_1 \ldots k_r},$$

então as coordenadas de $c_j^i(t) \in V_{s-1}^{r-1}$ na mesma base são os números

$$\zeta_{m_1 \ldots \widehat{m}_j \ldots m_s}^{k_1 \ldots \widehat{k}_i \ldots k_r} = \sum_{q=1}^{n} \xi_{m_1 \ldots m_{j-1} q m_{j+1} \ldots m_s}^{k_1 \ldots k_{i-1} q k_{i+1} \ldots k_r} .$$

Em particular, a contração $c_1^1 \colon V_1^1 \to R$ define o traço de uma aplicação linear $A \in \mathcal{L}(V, V) \approx V_1^1$.

2.7 A Álgebra tensorial $T(V)$

Neste parágrafo, ao contrário dos anteriores, a expressão "espaço vetorial" não deixará subentendido que a dimensão respectiva seja finita.

Seja $V_0, V_1, \ldots, V_i, \ldots$ uma sequência enumerável de espaços vetoriais V_i . Definiremos a *soma direta externa* ou, simplesmente, a *soma direta* desses espaços como sendo o espaço vetorial

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_i \oplus \ldots$$

cujos elementos são as sequências $v = (v_0, v_1, \ldots, v_i, \ldots)$ tais que $v_i \in V_i$ e *apenas um número finito* dos vetores v_i é diferente de zero. As operações em V são definidas componente a componente: dados $u = (u_0, u_1, \ldots, u_i, \ldots)$ e $v = (v_0, v_1, \ldots, v_i, \ldots)$ em V e α escalar, pomos

$$u + v = (u_0 + v_0, u_1 + v_1, \ldots, u_i + v_i, \ldots)$$
$$\alpha u = (\alpha u_0, \alpha u_1, \ldots, \alpha u_i, \ldots).$$

Para cada inteiro $i \geq 0$, existe uma aplicação linear biunívoca $\phi_i \colon V_i \to V$, que associa a um vetor $v_i \in V_i$ a

sequência $(0, \ldots, 0, v_i, 0, \ldots)$ cujos termos são todos nulos, com exceção talvez do i-ésimo, que é igual a v_i. Seja V_i' o subespaço de V formado pelas sequências onde somente o i-ésimo termo pode ser não nulo. Tem-se $V_i' = \phi_i(V_i)$.

Dado um elemento $v \in V$, seja $v=(v_0, v_1, \ldots, v_k, 0, \ldots)$, onde todos os termos seguintes a v_k são nulos. Temos

$$v=(v_0, 0, \ldots) + (0, v_1, 0, \ldots) + \cdots + (0, \ldots, 0, v_k, 0, \ldots)= \\ = v_0' + v_1' + \cdots + v_k'$$

onde cada parcela v_i' da soma acima é uma sequência com no máximo um termo diferente de zero, ou seja $v_i' \in V_i'$. É natural identificar $v_i' = \phi_i(v_i)$ com o próprio $v_i \in V_i$. Teremos então que cada elemento $v \in V$ se escreve, de modo único, como soma de um número finito de elementos dos V_i:

$$v \in V \Rightarrow v = v_0 + v_1 + \cdots + v_k\,, \quad v_i \in V_i\,,$$

sendo $v = 0$ se, e somente se, cada $v_i = 0$.

Em geral, o espaço vetorial $V = V_0 \oplus \cdots \oplus V_i \oplus \ldots$ tem dimensão infinita. (Isto só não ocorre quando apenas um número finito de espaços V_i tem dimensão não-nula.) Uma base de V é obtida tomando-se uma base em cada V_i e considerando todos estes vetores como elementos de V, do modo acima descrito. Podemos, então, dizer que toda reunião de bases dos V_i é uma base de V.

Uma *álgebra* é um par (A, m), onde A é um espaço vetorial e $m\colon A \times A \to A$ é uma aplicação bilinear, chamada a *multiplicação* da álgebra. É mais comum, numa álgebra, escrever $uv \in A$, em vez de $m(u, v) \in A$, para indicar o produto de dois elementos $u, v \in A$ e fazer referência à álgebra A, em vez de (A, m). A bilinearidade da

multiplicação significa que

$$u(v+w) = uv + uw,$$
$$(u+v)w = uw + vw,$$
$$(\lambda u)v = u(\lambda v) = \lambda(uv),$$

quaisquer que sejam $u, v, w \in A$ e λ escalar.

Uma álgebra A diz-se *associativa* quando se tem

$$u(vw) = (uv)w$$

quaisquer que sejam $u, v, w \in A$.

Diz-se que a álgebra A possui *unidade* quando existe um elemento $e \in A$ (a *unidade* da álgebra) tal que

$$eu = ue = u$$

para todo $u \in A$. É claro que uma álgebra A possui, no máximo, uma unidade.

Sejam A, A' álgebras. Uma aplicação linear $f\colon A \to A'$ chama-se um *homomorfismo* quando se tem $f(uv) = f(u)f(v)$ quaisquer que sejam $u, v \in A$. Quando existem unidades $e \in A$ e $e' \in A'$ e, além disso, tem-se $f(e) = e'$, diz-se então que f é um homomorfismo *unitário*.

Se a álgebra A possui uma unidade $\underline{e}$, existe um homomorfsimo unitário natural

$$E\colon R \to A$$

da álgebra R dos números reais em A, o qual leva um número real λ em $E(\lambda) = \lambda e \in A$. Sendo $e \neq 0$, E é biunívoco e fornece uma "imersão" canônica de R em A.

Assim, se identificarmos o escalar λ com o elemento $\lambda e \in A$ (o que equivale a identificar a unidade $e \in A$ com o escalar 1), podemos considerar $R \subset A$ para toda álgebra com unidade A.

Identificaremos sempre com o escalar 1 a unidade de uma álgebra A.

Seja V um espaço vetorial de dimensão finita. Para cada inteiro $r \geq 0$, temos o espaço $V_0^r = V \otimes \cdots \otimes V$ (r fatores) dos tensores r vezes contravariantes. Como sabemos, $V_0^0 = R$ e $V_0^1 = V$. Consideremos, em seguida, o espaço vetorial

$$T(V) = V_0^0 \oplus V_0^1 \oplus V_0^2 \oplus \cdots \oplus V_0^r \oplus \ldots$$

soma direta dos espaços V_0^r, $r = 0, 1, 2, \ldots$. Dois elementos genéricos de $T(V)$ são da forma

$$z = z_0 + z_1 + \cdots + z_k\,; \quad w = w_0 + w_1 + \cdots + w_m\,,$$

onde z_0, w_0 são escalares, z_1, w_1 são vetores em V, z_2, w_2 são tensores contravariantes de segunda ordem, etc.

Definiremos uma multiplicação em $T(V)$ pondo

$$zw = z_0w_0 + (z_0w_1 + z_1w_0) + (z_0w_2 + z_1w_1 + z_2w_0) + \cdots + z_kw_m\,,$$

onde cada produto $z_iw_j \in V_0^{i+j}$ é dado pela multiplicação dos tensores $z_i \in V_0^i$ e $w_j \in V_0^j$. (Na notação do parágrafo anterior, temos a aplicação bilinear $\phi\colon V_0^i \times V_0^j \to V_0^{i+j}$ e $z_iw_j = \phi(z_i, w_j)$.)

Como cada multiplicação parcial $(z_i, w_j) \to z_iw_j$ é bilinear, segue-se que a aplicação

$$m\colon T(V) \times T(V') \to T(V),$$

dada por $m(z,w) = zw$ (onde zw é definido pela fórmula acima) é bilinear e faz de $T(V)$ uma álgebra, chamada a *álgebra tensorial* de V (ou a *álgebra dos tensores contravariantes de* V).

Dados $t_k \in V_0^k$, $z_i \in V_0^i$ e $w_j \in V_0^j$, sabemos que $t_k(z_i w_j) = (t_k z_i)w_j \in V_0^{k+i+j}$. Segue-se que $t(zw) = (tz)w$ quaisquer que sejam t, z e w em $T(V)$. Portanto, a álgebra tensorial $T(V)$ é *associativa.*

Além disso, o escalar $1 \in V_0^0 = R$ é tal que $1z_i = z_i \in V_0^i$ para todo $z_i \in V_0^i$. Logo, $1z = z$ para todo $z \in T(V)$. A álgebra $T(V)$ possui, pois, uma *unidade.*

A dimensão da álgebra $T(V)$ é infinita, exceto no caso trivial em que $\dim V = 0$. Se $\mathcal{E} = \{e_1, \ldots, e_n\}$ é uma base de V, então os elementos

$$1,\ e_1, \ldots, e_n,\ e_1e_1, \ldots, e_{n-1}e_n.e_ne_n,\ e_1e_1e_1,\ e_1e_1e_2, \ldots$$

constituem uma base de $T(V)$.

Consideraremos também $V \subset T(V)$, identificando o vetor $v \in V$ com o elemento $0 + v + 0 + 0 + \cdots \in T(V)$.

Teorema 3. *Sejam V um espaço vetorial de dimensão finita, e A uma álgebra associativa com unidade (indicada com 1). Dada uma aplicação linear $f\colon V \to A$, existe um único homomorfismo unitário $F\colon T(V) \to A$ tal que $F(v) = f(v)$ para todo $v \in V$.*

Demonstração: Sejam as aplicações lineares $f_0\colon R \to A$, $v_1\colon V \to A$, $f_2\colon V_0^2 \to A, \ldots, f_r\colon V_0^r \to A, \ldots$ definidas

por:

$$\begin{aligned}
&f_0(\lambda) = \lambda \cdot 1\\
&f_1(v) = f(v),\\
&f_2(u \otimes v) = f(u)f(v),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&f_r(v_1 \otimes v_2 \otimes \cdots \otimes v_r) = f(v_1)f(v_2)\ldots f(v_r).
\end{aligned}$$

Cada uma das aplicações lineares f_r está bem definida em V_0^r, em virtude do análogo, para r fatores, do Teorema 1. De fato, tem-se $f_r = \widetilde{g}_r$, onde $g_r\colon V \times \cdots \times V \to A$ é a aplicação r-linear dada por $g_r(v_1, \ldots, v_r) = f(v_1)f(v_2)\ldots f(v_r)$. Definamos a aplicação linear $F\colon T(V) \to A$ pondo, para cada $z = z_0 + z_1 + \cdots + z_k \in T(V)$,

$$F(z) = f_0(z_0) + f_1(z_1) + \cdots + f_k(z_k).$$

Mostremos que F é um homomorfismo unitário. É claro que $F(1) = 1$. Verifiquemos agora que $F(zw) = F(z)F(w)$. Suponhamos primeiro que $z = u_1 \otimes \cdots \otimes u_r \in V_0^r$ e $w = v_1 \otimes \cdots \otimes v_s \in V_0^s$ sejam tensores decomponíveis. Então

$$\begin{aligned}
F(zw) &= F(u_1 \otimes \cdots \otimes u_r \otimes v_1 \otimes \cdots \otimes v_s) =\\
&= f_{r+s}(u_1 \otimes \ldots \otimes u_r \otimes v_1 \otimes \ldots \otimes v_s) =\\
&= f(u_1)\ldots f(u_r)f(v_1)\ldots f(v_s) =\\
&= f_r(u_1 \otimes \cdots \otimes u_r)f_s(v_1 \otimes \cdots \otimes v_s) = F(z)F(w).
\end{aligned}$$

Em seguida, consideremos dois elementos quaisquer $z, w \in T(V)$. Escrevendo cada um desses elementos como soma de tensores e decompondo, depois, cada um dos tensores obtidos como soma de tensores decomponíveis, teremos

$z = z_0 + \cdots + z_k$ e $w = w_0 + \cdots + w_m$ (onde, agora, as notações z_i e w_j não significam que $z_i \in V_0^i$ nem $w_j \in V_0^j$). Então teremos: $F(zw) = F\left(\sum z_i w_j\right) = \sum F(z_i w_j)$. Mas como cada produto $z_i w_j$ é um tensor decomponível, pelo que acabamos de ver, é $F(z_i w_j) = F(z_i)F(w_j)$ e portanto

$$F(zw) = \sum F(z_i)F(w_j) = \left[\sum F(z_i)\right]\left[\sum F(w_j)\right] = \\ = F\left(\sum z_i\right) F\left(\sum w_j\right) = F(z)F(w).$$

Assim, F é um homomorfismo. Finalmente, a álgebra $T(V)$ é gerada por V e 1, isto é, todo elemento de $T(V)$ é soma de um escalar com produtos de elementos de V. Segue-se daí que todo homomorfismo de $T(V)$ em A, que coincida com F em V e 1, coincidirá com F em toda a álgebra $T(V)$, o que conclui a demonstração do Teorema.

O Teorema 3 acima exprime que $T(V)$ é a *álgebra livre associativa* gerada por V e 1.

Observação: Seja A um espaço vetorial de dimensão finita. Uma estrutura de álgebra em A é dada por uma multiplicação, que é uma aplicação bilinear $m\colon A \times A \to A$, ou seja, um elemento de $\mathcal{L}(A, A; A)$. Como $\mathcal{L}(A, A; A) \approx \mathcal{L}(A \otimes A, A) \approx (A \otimes A)^* \otimes A \approx A \otimes A^* \otimes A^* = A_2^1$, segue-se que uma multiplicação em A é um tensor, uma vez contravariante e duas vezes covariante, $m \in A_2^1$. Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de A. O produto de dois elementos básicos e_i, e_j se escreve:

$$e_i e_j = \sum_k \gamma_{ij}^k \, e_k \, .$$

Os números γ_{ij}^k chamam-se as *constantes de estrutura* da álgebra A (relativamente à base $\mathcal{E}$). O conhecimento desses números determina inteiramente a multiplicação em A, pois se $u = \sum \alpha^i e_i$ e $v = \sum \beta^j e_j$, então

$$uv = \sum_{i,j} \alpha^i \beta^j \, e_i e_j = \sum_k \left[\sum_{i,j} \alpha^i \beta^j \, \gamma_{ij}^k \right] e_k \, .$$

As constantes de estrutura nada mais são do que as coordenadas do tensor multiplicação $m \in A_2^1$ relativamente à base definida por $\mathcal{E}$ em A_2^1.

Capítulo 3

Álgebra Exterior

Exporemos aqui os fatos básicos acerca das potências exteriores de um espaço vetorial. Isto equivale a estudar os tensores anti-simétricos, ou sejam, os p-vetores desse espaço. Tal estudo é rico em aplicações à Álgebra, à Geometria e à Análise. Mantendo, porém, nosso ponto de vista de introdução, daremos apenas algumas aplicações simples.

3.0 Permutações

Seja X um conjunto qualquer. Uma *permutação* de X é uma aplicação biunívoca $\sigma\colon X \to X$, de X sobre si mesmo.

O conjunto das permutações de X, munido da operação que associa a duas permutações σ, τ a sua composta $\sigma \circ \tau = \sigma\,\tau$, constitui um grupo, chamado o *grupo das permutações de* X. (Isto significa apenas que existe uma permutação identidadde $\varepsilon\colon X \to X$, que cada permutação $\sigma\colon X \to X$ possui uma inversa $\sigma^{-1}\colon X \to X$, e que as seguintes igualdades são sempre válidas: $\varepsilon\,\sigma = \sigma\,\varepsilon = \sigma$,

$\sigma\sigma^{-1} = \sigma^{-1}\,\sigma = \varepsilon$, $(\sigma\,\tau)\rho = \sigma(\tau\,\rho)$.)

Suponhamos que X possua pelo menos 2 elementos. Uma permutação $\tau\colon X \to X$ chama-se uma *transposição* quando existem $a \neq b$ em X tais que $\tau(a) = b$, $\tau(b) = a$ e $\tau(x) = x$ para todo $x \in X$ distinto de a e de b. Se τ é uma transposição, então $\tau\tau = \varepsilon$, mas a recíproca é falsa.

Quando X é um conjunto finito (único caso que consideraremos) com n elementos, então o grupo das permutações de X também é finito e tem $n!$ elementos.

Toda permutação $\sigma\colon X \to X$ de um conjunto finito X se exprime como produto de transposições $\sigma = \tau_1\tau_2 \ldots \tau_r$. Esta expressão não é absolutamente única, mas o número r de transposições cujo produto é σ tem sempre a mesma paridade: se σ é escrita uma vez como produto de um número par (respectivamente: ímpar) de transposições, qualquer outra decomposição de σ deverá conter um número par (resp. ímpar) de transposições. (Para demonstração desses fatos, veja N. Jacobson, Lectures in Abstract Algebra, vol. I, pag. 36).

Diremos que uma permutação $\sigma\colon X \to X$ (X finito) é *par* ou *ímpar* conforme σ se escreva como produto de um número par ou ímpar de transposições. Por exemplo, a permutação identidade é par e toda transposição é ímpar.

Usaremos o símbolo ε_σ para indicar o *sinal* de permutação σ: $\varepsilon_\sigma = +1$ se σ for par e $\varepsilon_\sigma = -1$ se σ for ímpar.

O produto $\sigma\rho$ é par se, e somente se, as permutações σ e ρ têm a mesma paridade (isto é, são ambas pares ou ambas ímpares). Isto equivale a escrever:

$$\varepsilon_{\sigma\rho} = \varepsilon_\sigma\varepsilon_\rho\,.$$

Multiplicando-se cada permutação par por uma permutação ímpar fixa, obtemos todas as permutações ímpares. Resulta então que, entre as $n!$ permutações de um conjunto X com n elementos, $n!/2$ delas são pares e $n!/2$ são ímpares.

Indiquemos com $I_n = \{1, \ldots, n\}$ o conjunto dos inteiros positivos de 1 até n. Uma *p-upla* de elementos de I_n é uma aplicação $i\colon I_p \to I_n$. Indicando com i_r o valor da aplicação i no elemento $r \in I_p$, é comum representar a p-upla i por $(i_1, \ldots, i_p)$. Diremos que $i = (i_1, \ldots, i_p)$ é uma *p-upla de elementos distintos* quando a aplicação $i\colon I_p \to I_n$ for biunívoca, isto é, quando $r \neq s$ implicar $i_r \neq i_s$. No caso contrário, diremos que a *p-upla i tem elementos repetidos.*

Deve-se distinguir a p-upla $(i_1, \ldots, i_p)$ do *conjunto* $\{i_1, \ldots, i_p\}$ por ela determinado. Quando a p-upla dada tem elementos repetidos, o conjunto $\{i_1, \ldots, i_p\}$, malgrado a notação, tem menos de p elementos. E, mesmo que a p-upla dada, $(j_1, \ldots, j_p)$, seja de elementos distintos, para cada permutação σ do conjunto $J = \{j_1, \ldots, j_p\}$, temos

$$J = \{\sigma(j_1), \ldots, \sigma(j_p)\}$$

mas as $p!$ p-uplas $(\sigma(j_1), \ldots, \sigma(j_p))$, obtidas variando σ, são duas a duas diferentes. Por exemplo, temos $\{1, 3\} = \{3, 1\}$, mas $(1, 3) \neq (3, 1)$.

Escreveremos $J = \{j_1 < j_2 < \cdots < j_p\}$ para indicar que a numeração dos elementos do conjunto J foi escolhida segundo a ordem crescente dos mesmos. Então, as p-uplas cujo conjunto de elementos é J são as da forma $(\sigma(j_1), \ldots, \sigma(j_p))$, onde σ varia entre as permutações de J. Cada uma dessas p-uplas fica inteiramente caracterizada

pela permutação σ que a ela corresponde.

O conjunto I_n possui $\binom{n}{p} = \frac{n(n-1)\ldots(n-p+1)}{p!}$ subconjuntos com p elementos, de modo que existem exatamente $\binom{n}{p}\cdot p! = n(n-1)\ldots(n-p+1)$ p-uplas de elementos distintos em I_m ("arranjos" de n elementos p a p).

3.1 Aplicações multilineares alternadas

Seja $p \geq 1$ um inteiro, e sejam V, W espaços vetoriais. Uma aplicação p-linear $\phi\colon V \times \cdots \times V \to W$ chama-se *alternada* quando muda de sinal ao inverter-se a ordem de 2 de seus argumentos, isto é, quando se tem

$$\phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) = -\phi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_p)$$

quaisquer que sejam $v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p$ em V.

Segue-se imediatamente da definição que uma aplicação p-linear alternada $\phi\colon V \times \cdots \times V \to W$ anula-se sempre que dois dos seus argumentos são iguais, isto é:

$$\phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) = 0 \quad \text{se} \quad v_i = v_j\,.$$

Reciprocamente, se uma aplicação p-linear ϕ satisfaz à relação acima, então ela é alternada. Com efeito, dada a p-upla $(\ldots, v_i, \ldots, v_j, \ldots)$ em V, a hipótese feita e a p-linearidade de ϕ acarretam:

$$\begin{aligned} 0 &= \phi(\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) = \\ &= \phi(\ldots, v_i, \ldots, v_i, \ldots) + \phi(\ldots, v_i, \ldots, v_j, \ldots) + \\ &\quad + \phi(\ldots, v_j, \ldots, v_i, \ldots) + \phi(\ldots, v_j, \ldots, v_j, \ldots). \end{aligned}$$

Ora, sendo ϕ como é, a primeira e a última parcelas da soma acima são nulas. Concluímos, então, que

$$\phi(\dots, v_i, \dots, v_j, \dots) + \phi(\dots, v_j, \dots, v_i, \dots) = 0,$$

e portanto ϕ é alternada.

Outra maneira de se exprimir que uma aplicação $\phi \in \mathcal{L}_p(V, W)$ é alternada é a seguinte:

$$\phi(v_{\sigma(1)}, \dots, v_{\sigma(p)}) = \varepsilon_\sigma \, \phi(v_1, \dots, v_p)$$

qualquer que seja a permutação σ do conjunto $\{1, \dots, p\}$.

De fato, a propriedade acima implica que ϕ é alternada, pois a troca de posição de dois argumentos v_i, v_j corresponde à transposição τ com $\tau(i) = j$ e $\tau(j) = i$, sendo $\varepsilon_\tau = -1$. Reciprocamente, se ϕ é alternada, dada a permutação σ, escrevemos σ como produto de r transposições, donde $\phi(v_{\sigma(1)}, \dots, v_{\sigma(p)})$ é obtido de $\phi(v_1, \dots, v_p)$ através de r trocas de pares de argumentos. Em cada troca, ϕ muda de sinal, donde

$$\phi\left(v_{\sigma(1)}, \dots, v_{\sigma(p)}\right) = (-1)^r \, \phi(v_1, \dots, v_p) = \varepsilon_\sigma \, \phi(v_1, \dots, v_p).$$

É imediato que a soma de duas aplicações p-lineares alternadas e o produto de uma aplicação p-linear alternada por um escalar são ainda alternadas, de modo que o conjunto, que indicaremos com $\mathfrak{A}_p(V, W)$, das aplicações p-lineares alternadas de V em W é um subespaço vetorial do espaço $\mathfrak{A}_p(V, W)$ de todas as aplicações p-lineares de V em W.

Escreveremos $\mathfrak{A}_p(V)$, em vez de $\mathfrak{A}_p(V, R)$, para indicar o espaço das *formas* p-lineares alternadas, isto é, das aplicações p-lineares alternadas $\phi\colon V \times \cdots \times V \to R$, com valores reais.

Quando $p = 1$, tem-se $\mathfrak{A}_1(V) = V^*$ e $\mathfrak{A}_1(V, W) = \mathcal{L}(V, W)$.

Proposição 1. *Se $v_1, \ldots, v_p \in V$ são linearmente dependentes, então $\phi(v_1, \ldots, v_p) = 0$ qualquer que seja $0 \in \mathfrak{A}_p(V, W)$.*

Demonstração: Sendo $v_1, \ldots, v_p$ linearmente dependentes, existe um índice j_0, $1 < j_0 \leq p$, tal que $v_{j_0} = \sum_{i<j_0} \alpha^i v_i$ (vide Proposiçao 1, Cap. 1). Então

$$\phi(v_1, \ldots, v_{j_0}, \ldots, v_p) = \phi\left(v_1, \ldots, \sum_{i<j_0} \alpha^i v_i, \ldots, v_p\right) =$$

$$= \sum_{i<j_0} \alpha^i \, \phi(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_p) = 0,$$

pois ϕ é alternada.

Corolário. *Se $p > \dim V$, então toda aplicação p-linear $\phi\colon V \times \cdots \times V \to W$ é identicamente nula.*

Com efeito, sendo $p > \dim V$, quaisquer p vetores em V são linearmente dependentes, e portanto $\phi(v_1, \ldots, v_p) = 0$ quaisquer que sejam $v_1, \ldots, v_p \in V$.

Por simplicidade, consideraremos inicialmente o espaço vetorial $\mathfrak{A}_2(V)$, das formas bilineares alternadas sobre um espaço V. Tem-se então o seguinte resultado:

Teorema 1'. *Sejam V um espaço vetorial, de dimensão $n > 1$, e $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V. Para cada par i, j de inteiros com $1 \leq i < j \leq n$, seja $\phi^{ij}\colon V \times V \to R$ a forma bilinear caracterizada pelas igualdades:*

$\phi^{ij}(e_i, e_j) = 1$, $\phi^{ij}(e_j, e_i) = -1$, *e* $\phi^{ij}(e_r, e_s) = 0$, *se os conjuntos* $\{r, s\}$ *e* $\{i, j\}$ *não coincidem.*
Então as formas bilineares ϕ^{ij}, $i < j$, *são alternadas e constituem uma base do espaço vetorial* $\mathfrak{A}_2(V)$. *Em particular,*

$$\dim \mathfrak{A}_2(V) - n(n-1)/2.$$

Demonstração: Dividiremos a demonstração em 3 partes.

Primeiro – As formas bilineares ϕ^{ij} são alternadas. Seja

$$u = \sum_i \alpha^i e_i \in V$$

um vetor qualquer. Devemos mostrar que, sejam quais forem $i < j$, temos $\phi^{ij}(u, u) = 0$.

Ora,

$$\phi^{ij}(u, u) = \sum_{k,m} \alpha^k \alpha^m \, \phi^{ij}(e_k, e_m).$$

Mas, exceto $\phi^{ij}(e_i, e_j) = 1$ e $\phi^{ij}(e_j, e_i) = -1$, todos os demais valores $\phi^{ij}(e_k, e_m)$ no somatório acima são nulos, donde $\phi^{ij}(u, u) = \alpha^i \alpha^j - \alpha^j \alpha^i = 0$.

Segundo – As formas ϕ^{ij}, $i < j$, geram $\mathfrak{A}_2(V)$. Seja $\pi \in \mathfrak{A}_2(V)$ uma forma bilinear alternada. Para cada par $i < j$, ponhamos $\xi_{ij} = \phi(e_i, e_j)$ e escrevamos $\psi = \sum_{i<j} \xi_{ij} \, \phi^{ij}$. Devemos mostrar que $\phi = \psi$. Sendo ϕ e ψ bilineares, basta verificar que $\phi(e_k, e_m) = \psi(e_k, e_m)$ sejam quais forem $e_k, e_m \in \mathcal{E}$. Ora, se $k < m$, então $\phi(e_k, e_m) = \xi_{km}$, por definição, enquanto $\psi(e_k, e_m) = \sum_{i<j} \xi_{ij} \, \phi^{ij}(e_k, e_m) =$

$\xi_{km}\,\phi^{km}(e_k, e_m) = \xi_{km}$ também. Sendo ϕ e ψ ambas alternadas, temos $\phi(e_k, e_m) = -\xi_{mk} = \psi(e_k, e_m)$ quando $m < k$ e $\phi(e_k, e_m) = 0 = \psi(e_k, e_m)$ se $k = m$, donde $\phi = \psi$.

Terceiro – As formas ϕ^{ij}, $i < j$, são linearmente independentes. Com efeito, se $\sum_{i<j} \lambda_{ij}\,\phi^{ij} = 0$, então, para cada par $k < m$, temos:

$$0 = \left(\sum_{i<j} \lambda_{ij}\,\phi^{ij}\right)(e_k, e_m) = \sum_{i<j} \lambda_{ij}\,\phi^{ij}(e_k, e_m) = \lambda_{km}\,,$$

o que mostra a independência das ϕ^{ij}, $i < j$.

Calcularemos agora a dimensão do espaço vetorial $\mathfrak{A}_p(V)$, onde p é um inteiro qualquer ≥ 1. O teorema geral abaixo engloba, naturalmente, o Teorema 1' como caso particular. A demonstração é análoga: apenas a notação é mais complicada. Somente por razões didáticas separamos o caso bilinear.

Teorema 1. *Seja V um espaço vetorial de dimensão $n > 1$. Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V. Dado $p \leq n$, para cada subconjunto $J = \{j_1 < \cdots < j_p\} \subset I_n$, seja $\phi^J : V \times \cdots \times V \to R$ a forma p-linear caracterizada por:*

$$\begin{cases} \phi^J(e_{j_1}, \ldots, e_{j_p}) = 1; \\ \phi^J(e_{\sigma(j_1)}, \ldots, e_{\sigma(j_p)}) = \varepsilon_\sigma\,, \text{ seja qual for a} \\ \qquad\qquad\qquad\qquad \text{permutação } \sigma : J \to J; \\ \phi^J(e_{i_1}, \ldots, e_{i_p}) = 0 \text{ se } \{i_1, \ldots, i_p\} \neq \{j_1, \ldots, j_p\}. \end{cases}$$

Então as formas p-lineares ϕ^J são alternadas e constituem uma base do espaço vetorial $\mathfrak{A}_p(V)$. Em particular,

$$\dim \mathfrak{A}_p(V) = \binom{n}{p}.$$

Demonstração: Segue as linhas do caso bilinear.

Primeiro – As formas ϕ^J são alternadas. Tomemos uma ϕ^J e uma p-upla de vetores $v_1, \ldots, v_p$, onde supomos que $v_i = v_j = u$. Em termos da base $\mathcal{E}$, temos:

$$v_1 = \sum_{k_1} \beta_1^{k_1} e_{k_1}, \ldots, v_p = \sum_{k_p} \beta_p^{k_p} e_{k_p}, v_i = v_j = u = \sum_r \alpha^r e_r .$$

Segue-se que

$$\phi^J(v_1, \ldots, u, \ldots, u, \ldots, v_p) =$$
$$= \sum_{k_1,\ldots,k_p} \beta_1^{k_1} \ldots \beta_p^{k_p} \left[\sum_{r,s} \alpha^r \alpha^s \phi^J(e_{k_1}, \ldots, e_r \ldots e_s \ldots e_{k_p}) \right]$$

onde o primeiro somatório se estende a todos os índices $k_1, \ldots, k_{i-1}, k_{i+1}, \ldots, k_{j-1}, k_{j+1}, \ldots, k_p$, compreendidos entre 1 e n. Provaremos que o segundo somatório é nulo, quaisquer que sejam $k_1, \ldots, k_p$. Abaixo, na primeira igualdade, desprezamos os termos $\phi^J(\ldots e_r \ldots e_r \ldots)$, pois são nulos. Na segunda, trocamos os nomes dos índices r e s no segundo somatório. Na terceira igualdade, usamos o fato

de que $\phi^J(\dots e_r \dots e_s \dots) = -\phi^J(\dots e_s \dots e_r \dots)$.

$$\sum_{r,s} \alpha^r \alpha^s \, \phi^J(\dots e_r \dots e_s \dots) =$$

$$= \sum_{r<s} \alpha^r \alpha^s \, \phi^J(\dots e_r \dots e_s \dots) +$$

$$+ \sum_{s<r} \alpha^r \alpha^s \, \phi^J(\dots e_r \dots e_s \dots) =$$

$$= \sum_{r<s} \alpha^r \alpha^s \, \phi^J(\dots e_r \dots e_s \dots) +$$

$$+ \sum_{r<s} \alpha^s \alpha^r \, \phi^J(\dots e_s \dots e_r \dots) =$$

$$= \sum_{r<s} \alpha^r \alpha^s \, \phi^J(\dots e_r \dots e_s \dots) -$$

$$- \sum_{r<s} \alpha^s \alpha^r \, \phi^J(\dots e_r \dots e_s \dots) =$$

$$= \sum_{r<s} (\alpha^r \alpha^s - \alpha^s \alpha^r) \phi^J(\dots e_r \dots e_s \dots) = 0.$$

Logo, $\phi^J(\dots u \dots u \dots) = 0$, mostrando que cada ϕ^J é alternada.

Segundo – As formas ϕ^J geram $\mathfrak{A}_p(V)$. Dada ϕ $\mathfrak{A}_p(V)$, para cada subconjunto $J = \{j_1 < \cdots < j_p\} \subset I_n$, pomos $\xi_J = \phi(e_{j_1}, \dots, e_{j_p})$. Definimos $\psi = \sum_J \xi_J \, \phi^J$, a soma estendida a todos os subconjuntos J, com p elementos, do conjunto $I_n = \{1, \dots, n\}$. Mostra-se, como no teorema anterior, que $\phi = \psi$.

Terceiro – As formas ϕ^J são linearmente independentes. De uma relação do tipo $\sum_J \lambda_J \, \phi^J = 0$ segue-se que, para

toda escolha de $K = \{k_1 < \cdots < k_p\}$, tem-se

$$0 = \left(\sum_J \lambda_J \, \phi^J \right) (e_{k_1}, \ldots, e_{k_p}) = \lambda_K \, .$$

Corolário 1. *Se* $\dim V = n$, *então* $\dim \mathfrak{A}_n(V) = 1$.

Corolário 2. (Recíproca da Proposição 1.)
Se $\phi(v_1, ..., v_p) = 0$, *seja qual for a forma p-linear alternada* ϕ, *então* $v_1, ..., v_p$ *são linearmente dependentes.*

Com efeito, se fossem $v_1, ..., v_p$ linearmente independentes, então existiria uma base $\mathcal{E}=\{e_1, ..., e_n\}$ em V com $e_1=v_1, ..., e_p=v_p$. Então, tomando $J=\{1, ..., p\}$, a forma ϕ^J constuida no Teorema 1 seria tal que $\phi^J(v_1, ..., v_p)=1$.

Corolário 3. *Seja* $\dim V = n$. *Se existe uma forma n-linear alternada* $\phi \neq 0$ *tal que* $\phi(v_1, \ldots, v_n) = 0$, *então* $v_1, \ldots, v_n \in V$ *são linearmente dependentes.*

Com efeito, como $\dim \mathfrak{A}_n(V) = 1, \{\phi\}$ é uma base de $\mathfrak{A}_n(V)$; toda forma n-linear alternada ψ sobre V é da forma $\psi = \alpha \cdot \phi$, α escalar. Logo $\psi(v_1, \ldots, v_n) = \alpha \cdot \phi(v_1, \ldots, v_n) = 0$. Segue-se do Corolário 2 que $v_1, \ldots, v_n$ são linearmente dependentes.

Corolário 4. *Sejam* $\dim V = n$, $\dim W = r$ *e* $0 < p \leq n$. *Então* $\dim \mathfrak{A}_p(V, W) = \binom{n}{p} \cdot r$.

Com efeito, tomando uma base $\mathcal{F} = \{f_1, \ldots, f_n\}$ em W, uma aplicação qualquer $\phi \in \mathfrak{A}_p(V, W)$ é tal que

$$\phi(v_1, \ldots, v_p) = \sum_i \phi^i(v_1, \ldots, v_p) f_i \, , \quad v_1, \ldots, v_p \in V.$$

É imediato que as r aplicações

$$(v_1, \ldots, v_p) \to \phi^i(v_1, \ldots, v_p) \in R,$$

determinadas por ϕ, são formas p-lineares alternadas. A correspondência $\phi \to (\phi^1, \ldots, \phi^r)$, assim estabelecida, define um isomorfismo de $\mathfrak{A}_p(V, W)$ sobre a soma direta de r cópias de $\mathfrak{A}_p(V)$. Segue-se que $\dim \mathfrak{A}_p(V, W) = \binom{n}{p} \cdot r$.

3.2 Determinantes

Como conseqüência do fato de que $\dim \mathfrak{A}_n(V) = 1$, quando $n = \dim V$, mostraremos uma maneira de definir intrinsecamente o determinante de uma aplicação linear $A\colon V \to V$, de um espaço vetorial V em si próprio.

A aplicação linear $A\colon V \to V$ induz uma aplicação

$$A^{\#}\colon \mathfrak{A}_n(V) \to \mathfrak{A}_n(V),$$

definida do seguinte modo: se $\phi \in \mathfrak{A}_n(V)$ é uma forma n-linear alternada, $A^{\#}(\phi)\colon V \times \cdots \times V \to R$ é a forma tal que

$$[A^{\#}(\phi)](v_1, \ldots, v_n) = \phi(Av_1, \ldots, Av_n).$$

É óbvio que $A^{\#}(\phi) \in \mathfrak{A}_n(V)$. Se $B\colon V \to V$ é outra aplicação linear, verifica-se sem dificuldade que

$$(AB)^{\#} = B^{\#}A^{\#}\colon \mathfrak{A}_n(V) \to \mathfrak{A}_n(V).$$

Ora, toda aplicação linear de um espaço vetorial de dimensão 1 em si mesmo é a multiplicação por um escalar fixo. Segue-se então que, dada $A\colon V \to V$ linear, existe um

único escalar δ tal que $A^{\#}(\phi) = \delta \cdot \phi$ para toda $\phi \in \mathfrak{A}_n(V)$. Escrevemos $\delta = \det A$ e chamamos este escalar de *determinante* da aplicação linear A.

Assim, o determinante de A fica caracterizado pela igualdade

$$\phi(Av_1, \ldots, Av_n) = \det A \cdot \phi(v_1, \ldots, v_n),$$

válida quaisquer que sejam os vetores $v_1, \ldots, v_n \in V$ e qualquer que seja a forma n-linear alternada ϕ sobre V. Com a notação acima introduzida, esta igualdade se escreve

$$A^{\#}(\phi) = \det A \cdot \phi, \text{ para toda } \phi \in \mathfrak{A}_n(V).$$

Além de fornecer uma definição de $\det A$ que não utiliza a escolha de uma base em V, este método permite ainda demonstrar, de modo simples, as propriedades fundamentais dos determinantes, como veremos abaixo.

Proposição 2. *Sejam $A, B\colon V \to V$ aplicações lineares. Então:*

a) $\det(AB) = \det A \cdot \det B$;

b) $\det A \neq 0$ *se, e somente se, A é invertível.*

Demonstração: (a) Tomemos $\phi \neq 0$ em $\mathfrak{A}_n(V)$. Vem:

$$\det(AB)\phi = (AB)^{\#}(\phi) = (B^{\#}A^{\#})(\phi) = B^{\#}[A^{\#}(\phi)] =$$
$$= B^{\#}(\det A \cdot \phi) = \det A \cdot B^{\#}(\phi) = \det A \cdot \det B \cdot \phi.$$

Segue-se que $\det(AB) = \det A \cdot \det B$.

(b) Se $I\colon V \to V$ é a aplicação identidade, então é claro que $\det I = 1$. Se $A\colon V \to V$ possui inversa A^{-1}, então, por (a), $AA^{-1} = I$ dá $\det A \cdot \det(A^{-1}) = 1$, donde

det $A \neq 0$ e $\det(A^{-1}) = (\det A)^{-1}$. Reciprocamente, seja $A\colon V \to V$ tal que $\det A \neq 0$. Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V e $\phi \neq 0$ uma aplicação n-linear alternada. Então $\phi(e_1, \ldots, e_n) \neq 0$. Daí concluimos que $\phi(Ae_1, \ldots, Ae_n) = \det A \cdot \phi(e_1, \ldots, e_n) \neq 0$. Logo, $Ae_1, \ldots, Ae_n$ são linearmente independentes (vide Proposição 1). Assim, A transforma toda base de V noutra base de V, donde é invertível.

Para mostrar que esta definição de determinante coincide com a definição clássica de $\det A$ em termos de uma matriz de A, tomemos uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V e seja $(\alpha_j^i) = [A, \varepsilon]$ a matriz da aplicação A na base $\mathcal{E}$. Escolhamos a forma n-linear $\phi \in \mathfrak{A}_n(V)$ tal que $\phi(e_1, \ldots, e_n) = 1$. Então $\phi(e_{i_1}, \ldots, e_{i_n}) = \varepsilon_\sigma$ quando $(i_1, \ldots, i_n) = (\sigma(1), \ldots, \sigma(n))$ é uma n-upla de números todos distintos, Com tal ϕ, temos $\det A = \phi(Ae_1, \ldots, Ae_n)$. Ora,

$$Ae_1 = \sum_{i_1} \alpha_1^{i_1}\, e_{i_1}, \ldots, Ae_n = \sum_{i_n} \alpha_n^{i_n}\, e_{i_n}\,.$$

Assim:

$$\begin{aligned}\det A &= \phi\left(\sum_{i_1} \alpha_1^{i_1}\, e_{i_1}, \ldots, \sum_{i_n} \alpha_n^{i_n}\, e_{i_n}\right) = \\ &= \sum_{i_1, \ldots, i_n} \alpha_1^{i_1}\, \alpha_2^{i_2} \ldots \alpha_n^{i_n}\, \phi(e_{i_1}, \ldots, e_{i_n}) = \\ &= \sum_{\sigma} \varepsilon_\sigma\, \alpha_1^{\sigma(1)}\, \alpha_2^{\sigma(2)} \ldots \alpha_n^{\sigma(n)},\end{aligned}$$

a última soma sendo estendida a todas as permutações σ do conjunto $\{1, \ldots, n\}$. Na última igualdade, suprimimos todas as parcelas correspondentes a n-uplas $(i_1, \ldots, i_n)$ onde

há elementos repetidos pois, para cada uma delas, tem-se $\phi(e_{i_1}, \ldots, e_{i_n}) = 0$.

A expressão

$$\det A = \sum_{\sigma} \varepsilon_\sigma \, \alpha_1^{\sigma(1)} \, \alpha_2^{\sigma(2)} \ldots \alpha_n^{\sigma(n)}$$

constitui a definição clássica de $\det A$, em termos da matriz (α_j^i) de A na base $\mathcal{E}$.

As aplicações lineares $A\colon V \to V$ não são os únicos objetos que possuem determinante. Podemos, ainda, considerar:

a) O determinante de uma matriz quadrada $\alpha = (\alpha_j^i)$;

b) O determinante de n vetores $v_1, \ldots, v_n \in V$ relativamente a uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V.

Dada a matriz quadrada $\alpha = (\alpha_j^i)$, de ordem n, seu determinante pode ser definido diretamente pela fórmula clássica acima, ou então como o determinante da aplicação linear $A\colon R^n \to R^n$, tal que $Ae_j = \sum_i \alpha_j^i \, e_i$, onde $\{e_1, \ldots, e_n\}$ é a base canônica do R^n. Isto quer dizer que α é a matriz de A relativamente à base canônica. Usam-se as notações $\det \alpha$, ou $\det(\alpha_j^i)$.

Dados os vetores $v_1, \ldots, v_n \in V$ e a base $\mathcal{E}=\{e_1, \ldots, e_n\}$ de V, o determinante desses vetores relativamente à base dada é indicado com $[v_1, \ldots, v_n, \mathcal{E}]$ ou, simplesmente, com $[v_1, \ldots, v_n]$, e é definido como o determinante da matriz (α_j^i) tal que $v_j = \sum_i \alpha_j^i \, e_i$, $j = 1, \ldots, n$. Tem-se, evidentemente, $[v_1, \ldots, v_n] = \det A$, onde $A\colon V \to V$ é a aplicação linear tal que $Ae_i = v_i$, $i = 1, \ldots, n$.

Se $\alpha = (\alpha_j^i)$ é uma matriz $n \times n$, indicando com $v_1 = (\alpha_1^1, \ldots, \alpha_1^n), \ldots, v_n = (\alpha_n^1, \ldots, \alpha_n^n)$ os vetores-coluna de α,

vemos que $\det \alpha = [v_1, \ldots, v_n]$. Tomemos a forma n-linear alternada $\phi_0 \in \mathfrak{A}_n(R^n)$ tal que $\phi_0(e_1, \ldots, e_n) = 1$, onde os e_i representam a base canônica do R^n. Então:

$$\det(\alpha_j^i) = [v_1, \ldots, v_n] = \phi_0(v_1, \ldots, v_n).$$

Assim, o determinante de uma matriz α é uma forma n-linear alternada nos vetores colunas dessa matriz, forma essa que é "normalizada" pela condição de assumir o valor 1 nas colunas da matriz identidade. Reciprocamente, toda função numérica $M(n \times n) \to R$, definida entre as matrizes $n \times n$, que é uma forma n-linear alternada nas colunas da matriz, é um múltiplo (constante) da função $\det\colon M(n \times n) \to R$. Se a função dada assume o valor 1 da matriz identidade, então ela é igual ao determinante. Esta é uma caracterização clássica (devida a Weierstrass) do determinante. Ela resulta do fato de ser $\dim \mathfrak{A}_n(R^n) = 1$.

Seja $\alpha = (\alpha_j^i)$ agora uma matriz $m \times k$, onde m pode ser diferente de k. Seja ainda $p \leq m$, $p \leq k$. Usaremos a notação α_J^I para indicar a matriz $p \times p$, obtida de $\alpha = (\alpha_j^i)$ do seguinte modo: $I = \{i_1 < \cdots < i_p\} \subset I_m$ e $J = \{j_1 < \cdots < j_p\} \subset I_k$ são conjuntos de p inteiros e $\alpha_J^I = (\alpha_{j_s}^{i_r})$ é formada pelos elementos de α cujos índices superiores pertencem a I e cujos índices inferiores pertencem a J. Em outras palavras: o conjunto I determina a escolha de p linhas em α e J determina a escolha de p colunas de α. A "submatriz" α_J^I é formada pelos elementos de α que estão simultaneamente numa das p linhas e numa das p colunas escolhidas. O determinante $\det(\alpha_J^I)$ é o *menor* de α relativo às p linhas I e às p colunas J.

Ao tomar uma submatriz $m \times m$ de uma matriz α do tipo $m \leq k$, basta escolher as m colunas $(m \times k)$ segundo

um subconjunto $J = \{j_1 < \cdots < j_m\} \subset I_k$. Usa-se então a notação α_J, em vez de $\alpha_J^{I_m}$. Do mesmo modo, escreve-se α^J para indicar uma submatriz $k \times k$ de uma matriz $m \times k$.

Examinemos agora o Teorema 1 no caso particular em que $V = R^n$, sendo $\mathcal{E} = \{e_1, \ldots, e_n\}$ a base canônica do espaço euclidiano R^n. Seja $p \leq n$.

O espaço $\mathfrak{A}_p(R^n)$ terá uma base canônica, constituida por formas p-lineares ϕ^J, uma delas para cada subconjunto $J = \{j_1 < \cdots < j_p\} \subset I_n$. Ora, uma p-upla de vetores em R^n corresponde a uma matriz $n \times p$, da qual os vetores dados são as colunas. Assim, um elemento $\phi \in \mathfrak{A}_p(R^n)$ corresponde a uma aplicação $\phi\colon M(n \times p) \to R$, tal que $\phi(\alpha)$ é uma função p-linear das colunas da matriz.

Mostraremos que, pensando em cada $\phi \in \mathfrak{A}_p(R^n)$ como uma função de matrizes $n \times p$, as ϕ^J são os menores, isto é, $\phi^J(\alpha) = \det(\alpha^J)$ para toda $\alpha \in M(n \times p)$ e $J = \{j_1 < \cdots < j_p\} \subset I_n$. Quando fizermos isto, poderemos enunciar:

Proposição 3. *Toda aplicação $\phi\colon M(n \times p) \to R$, que é uma função p-linear alternada das colunas de uma matriz, se escreve, de modo único, como combinação linear dos menores de ordem p dessa matriz.*

Demonstração: Basta lembrar que, para cada J, temos $\phi^J(e_{i_1}, \ldots, e_{i_p}) = 0$ quando $\{i_1, \ldots, i_p\} \neq J$. Isto significa que se $\alpha \in M(n \times p)$ é uma matriz na qual cada coluna possui apenas um elemento não nulo (igual a 1) então $\phi^J(\alpha) \neq 0$ precisamente quando as linhas que contêm os elementos não-nulos de α constituem o conjunto J. Sendo ϕ^J uma função p-linear das colunas de α, segue-se daí que $\phi^J(\alpha)$ depende apenas da submatriz α^J, formada com as p linhas de índices em J. Podemos então escrever $\phi^J(\alpha) = \phi(\alpha^J)$, onde

ϕ: $M(p\times p) \to R$ é uma função p-linear alternada das colunas das matrizes $p \times p$. Como $\phi^J(e_{j_1}, \ldots, e_{j_p}) = 1$, segue-se que $\phi(\alpha^J) = 1$ quando α^J é a matriz identidade $p \times p$. Logo $\phi(\alpha^J) = \det(\alpha^J)$. Concluimos que $\phi^J(\alpha) = \det(\alpha^J) = J$-ésimo menor de α, como queríamos demonstrar.

Corolário. *Seja* g: $V \times \cdots \times V \to W$ *uma aplicação* p-*linear alternada. Dada uma base* $\mathcal{E} = \{e_1, \ldots, e_n\}$ *de* V, *ponhamos* $w_J = g(e_{j_1}, \ldots, e_{j_p})$ *sempre que* $J = \{j_1 < \cdots < j_p\}$. *Sejam* $v_1 = \sum_i \alpha_1^i\, e_i, \ldots, v_p = \sum_i \alpha_p^i\, e_i$ *vetores em* V. *Considerando a matriz* $\alpha = (\alpha_j^i) \in M(n \times p)$, *cujas colunas são as coordenadas dos vetores* v_i, *temos*

$$g(v_1, \ldots, v_p) = \sum_J \det(\alpha^J) w_J\,.$$

Com efeito, tomando uma base $\mathcal{F} = \{f_1, \ldots, f_m\}$ em W, temos $g(v_1, \ldots, v_p) = \sum_i \phi^i(v_1, \ldots, v_p) f_i$, onde cada $\phi^i \in \mathfrak{A}_p(V)$. Logo, para cada $i = 1, \ldots, m$, temos $\phi^i = \sum_J \xi_J^i\, \phi^J$. Segue-se da definição dos ϕ^J e dos w_J que $w_J = \sum_I \xi_J^i\, f_i$. Assim, para $v_1, \ldots, v_p \in V$ quaisquer, temos $g(v_1, \ldots, v_p) = \sum_J \phi^J(v_1, \ldots, v_p) w_J$. Mas já vimos acima que $\phi^J(v_1, \ldots, v_p) = \det(\alpha^J)$. O corolário fica, então, demonstrado.

Na definição do determinante de uma matriz quadrada α, as colunas e as linhas de α não desempenham, formalmente, o mesmo papel. Com efeito, det α é o determinante da aplicação linear A: $R^n \to R^n$ tal que $Ae_i \in R^n$ é a i-ésima *coluna* da matriz α (onde e_i é o i-ésimo elemento da

base canônica do R^n). Daí recorre a predominância das colunas sobre as linhas nas proposições deste parágrafo. Por exemplo, $\det \alpha$ é uma função n-linear alternada das colunas de α mas, até agora, não sabemos como $\det \alpha$ depende das linhas de α.

Para estabelecer a simetria que falta, demonstraremos a seguir que o valor do determinante de uma matriz não se altera quando se trocam suas linhas por suas colunas. Lembramos que a *transposta* de uma matriz $\alpha = (\alpha_j^i)$ é a matriz $\alpha^t = (\beta_j^i)$ tal que $\beta_j^i = \alpha_i^j$. Assim, as linhas de α^t coincidem com as colunas de α.

Proposição 4. *Seja* $\alpha \in M(n\times n)$. *Então* $\det \alpha = \det(\alpha^t)$.

Demonstração: Como já vimos anteriormente, vale a expressão:

$$\det \alpha = \sum_\sigma \varepsilon_\sigma \, \alpha_1^{\sigma(1)} \, \alpha_2^{\sigma(2)} \ldots \alpha_n^{\sigma(n)},$$

onde a soma é estendida a todas as permutações do conjunto $I_n = \{1, \ldots, n\}$. Trocando a ordem dos fatores em cada parcela desta soma, de modo que os índices superiores fiquem em ordem crescente, temos:

$$\det \alpha = \sum_\sigma \varepsilon_\sigma \, \alpha_{\sigma^{-1}(1)}^1 \, \alpha_{\sigma^{-2}(2)}^2 \ldots \alpha_{\sigma^{-1}(n)}^n.$$

Quando σ percorrre o conjunto das permutações de I_n, σ^{-1} percorre-o também. Além disso, $\varepsilon_\sigma = \varepsilon_{\sigma^{-1}}$. Logo, pondo $\rho = \sigma^{-1}$, vem:

$$\det \alpha = \sum_\rho \varepsilon_\rho \, \alpha_{\rho(1)}^1 \, \alpha_{\rho(3)}^2 \ldots \alpha_{\rho(n)}^n$$

ou seja, $\det \alpha = \det(\alpha^t)$.

Resulta agora que, em todas as proposições referentes ao determinante de uma matriz, linhas e colunas se comportam igualmente.

3.3 Potências exteriores de um espaço vetorial

Seja V um espaço vetorial de dimensão n, e p um inteiro positivo $\leq n$. Uma *p-ésima potência exterior* de V é um par (Z, ϕ) tal que:

1) Z é um espaço vetorial e $\phi\colon V \times \cdots \times V \to Z$ é uma aplicação p-linear alternada:

2) $\dim Z = \binom{n}{p}$, onde $n = \dim V$;

3) $\phi(V \times \cdots \times V)$ gera Z.

Os aciomas 2) e 3), em presença de 1), são equivalentes ao único axioma seguinte:

2') Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V, os vetores $\phi(e_{j_1}, \ldots, e_{j_p})$, tais que $1 \leq j_1 < \cdots < j_p \leq n$, formam uma base de Z.

A verificação é imediata.

Devemos mostrar que existem as potências exteriores de V, e que duas potências exteriores p-ésimas de V são canonicamente isomorfas. Daremos, primeiramente, três construções que demonstram a existência.

Primeira construção: Seja Z um espaço vetorial qualquer, de dimensão igual a $\binom{n}{p}$, onde $n = \dim V$. Escolhamos uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, uma base $\mathcal{H}$ em Z,

e indiquemos os elementos $h_J \in \mathcal{H}$ com índices $J = \{j_1 < \cdots < j_p\}$ que variam entre os subconjuntos de p elementos do conjunto $I_n = \{1, \ldots, n\}$. Definamos a aplicação p-linear $\phi\colon V \times \cdots \times V \to Z$ pondo $\phi(e_{j_1}, \ldots, e_{j_p}) = h_J$ e $\phi(e_{\sigma(j_1)}, \ldots, e_{\sigma(j_p)}) = \varepsilon_\sigma\, h_J$ se $\{j_1 < \cdots < j_p\} = J$, e $\phi(e_{i_1}, \ldots, e_{i_p}) = 0$ se a p-upla $(i_1, \ldots, i_p)$ tem elementos repetidos. É claro que ϕ é alternada e o axioma 2') é evidentemente satisfeito.

Segunda construção: Tomemos $Z = \mathfrak{A}_p(V)^*$, dual do espaço das formas p-lineares alternadas sobre V. Definamos $\phi\colon V \times \cdots \times V \to Z$, pondo $[\phi(v_1, \ldots, v_p)](w) = w(v_1, \ldots, v_p)$, para todos $v_1, \ldots, v_p \in V$ e $w \in \mathfrak{A}_p(V)$. É imediato que valem os axiomas 1) e 2). Resta verificar que, dada uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, os elementos $z_J = \phi(e_{j_1}, \ldots, e_{j_p})$, com $J = \{j_1 < \cdots < j_p\}$, são linearmente independentes. Ora, se fosse $\sum_J \lambda^J z_J = 0$ viria, para cada forma p-linear ϕ^K, construída a partir da base $\mathcal{E}$, como no Teorema 1:

$$0 = \Big(\sum_J \lambda^J z_J\Big)(\phi^K) = \sum_J \lambda^J \phi^K(e_{j_1}, \ldots, e_{j_p}) = \lambda^K.$$

Logo os z_J são independentes, como queríamos mostrar.

Terceira construção: Tomaremos Z como o subespaço de $V_0^p = V \otimes \cdots \otimes V$ formado pelos tensores *anti-simétricos* p vezes contravariantes e definiremos $\phi\colon V \times \cdots \times V \to Z$ por meio da operação de *anti-simetrização*. Mais precisamente: para cada permutação σ do conjunto $\{1, \ldots, p\}$ consideraremos a aplicação linear $\sigma^*\colon V_0^p \to V_0^p$, caracterizada por:

$$\sigma^*(v_1 \otimes \cdots \otimes v_p) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)}\,.$$

A aplicação linear σ^* é induzida, de acordo com o Teorema 2 do Capítulo 2, pela aplicação p-linear

$$(v_1, \ldots, v_p) \to v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)} .$$

É imediato que se σ, ρ são permutações de I_p, então $(\sigma\rho)^* = \rho^*\sigma^*$ e, se ε é a permutação identidade, então ε^* é a aplicação identidade de V_0^p. Segue-se que cada σ^* é um isomorfismo de V_0^p sobre si próprio e que $(\sigma^*)^{-1} = (\sigma^{-1})^*$.

Diremos que um tensor $t \in V_0^p$ é anti-simétrico se $\sigma^*(t) = \varepsilon_\sigma t$ para toda permutação σ de $I_p = \{1, \ldots, p\}$. O tensor t é anti-simétrico se, e somente se, $\tau^*(t) = -t$ para toda transposição τ de I_p.

O conjunto Z dos tensores anti-simétricos é evidentemente um subespaço vetorial de V_0^p. Desejamos determinar a dimensão de Z e, mais precisamente, obter uma base de Z a partir de uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V. Primeiramente, mostraremos que se

$$t = \sum \xi^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p}$$

é um tensor anti-simétrico, então suas coordenadas $\xi^{i_1 \ldots i_p}$ são anti-simétricas: mudam de sinal quando se trocam 2 dos seus índices. Daí se concluirá que $\xi^{i_1 \ldots i_r} = 0$ quando houver dois índices iguais, e que $\{j_1 < \cdots < j_p\} = J$ dará

$$\xi^{\sigma(j_1) \ldots \sigma(j_p)} = \varepsilon_\sigma \, \xi^{j_1 \ldots j_p}$$

seja qual for a permutaçáo $\sigma \colon J \to J$.

Seja, então,

$$t = \sum \xi^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p} \text{ anti-simétrico.}$$

Tomemos uma transposição τ arbitrária dos inteiros $1, ..., p$. Seja $\tau(k) = m$, $\tau(m) = k$. Teremos $\tau^*(t) = -t$. Assim,

$$t = \sum \xi^{\ldots i_k \ldots i_m \ldots} \ldots \otimes e_{i_k} \otimes \cdots \otimes e_{i_m} \otimes \qquad \text{e}$$
$$-t = \sum \xi^{\ldots i_k \ldots i_m \ldots} \ldots \otimes e_{i_m} \otimes \cdots \otimes e_{i_k} \otimes \ .$$

Mudando os nomes de k e m na última expressão:

$$t = -\sum \xi^{\ldots i_m \ldots i_k \ldots} \ldots \otimes e_{i_k} \otimes \cdots \otimes e_{i_m} \ldots \ .$$

Segue-se que $\xi^{\ldots i_k \ldots i_m \ldots} = -\xi^{\ldots i_m \ldots i_k \ldots}$, como queríamos mostrar.

Assim, se t é um tensor anti-simétrico, na sua expressão em termos de uma base $\mathcal{E}$, podemos omitir as parcelas $\xi^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p}$ onde há índices repetidos e, nas parcelas restantes, temos $\xi^{\sigma(j_1) \ldots \sigma(j_p)} = \varepsilon_\sigma \xi^{j_1 \ldots j_p}$, onde $\{j_1 < \cdots < j_p\} = J$ e $\sigma \colon J \to J$ é uma permutação. Levando esses dois fatos em conta, podemos escrever, para todo tensor anti-simétrico t:

$$t = \sum_J \xi^{j_1 \ldots j_p} \left(\sum_\sigma \varepsilon_\sigma \, e_{\sigma(j_1)} \otimes \cdots \otimes e_{\sigma(j_p)} \right),$$

onde o primeiro somatório se estende a todos os subconjuntos $J = \{j_1 < \cdots < j_p\}$ e, para cada J, o segundo somatório se estende a todas as permutações $\sigma \colon J \to J$.

Se, para cada subconjunto $J = \{j_1 < \ldots, < j_p\} \subset I_n$, escrevermos

$$e_J = \sum_\sigma \varepsilon_\sigma (e_{\sigma(j_1)} \otimes \cdots \otimes e_{\sigma(j_p)}),$$

veremos que todo tensor anti-simétrico $t \in Z$ se escreve como combinação linear dos $\binom{n}{p}$ tensores e_J: $t = \sum \xi^J e_J$, onde as coordenadas ξ^J, são as coordenadas $\xi^{j_1 \cdots j_p}$ de t na base $\mathcal{E}$ que têm índices $j_1 < \cdots < j_p$ em ordem estritamente crescente.

Ora, a cada permutação $\sigma\colon J \to J$ corresponde, de modo natural, uma permutação $\sigma'\colon I_p \to I_p$ tal que $\sigma(j_m) = j_{\sigma'(m)}$. Identificando σ com σ', vemos que

$$e_J = \sum_\sigma \varepsilon_\sigma \, \sigma^*(e_{j_1} \otimes \cdots \otimes e_{j_p}).$$

Daí se conclui que os tensores e_J são anti-simétricos. Com efeito, seja ρ uma permutação qualquer de I_p. Então

$$\begin{aligned} \rho^*(e_J) &= \sum_\sigma \varepsilon_\sigma \rho^* \sigma^*(e_{j_1} \otimes \cdots \otimes e_{j_p}) = \\ &= \varepsilon_\rho \sum_\sigma \varepsilon_{\sigma\rho} (\sigma\rho)^*(e_{j_1} \otimes \cdots \otimes e_{j_p}). \end{aligned}$$

Sendo ρ fixa, ao fazer σ percorrer o conjunto das permutações de I_p, $\sigma\rho$ percorre igualmente o mesmo conjunto e assim podemos escrever:

$$\rho^*(e_J) = \varepsilon_\rho \sum_\mu \varepsilon_\mu \, \mu^*(e_{j_1} \otimes \cdots \otimes e_{j_p}) = \varepsilon_\rho \, e_J \,,$$

mostrando que e_J é, de fato, um tensor anti-simétrico.

Assim, vemos que os tensores e_J constituem um sistema de $\binom{n}{p}$ geradores do espaço Z dos tensores anti-simétricos. Mostraremos agora que os e_J são linearmente independentes, e portanto formam uma base de Z. A independência linear dos e_J resulta porém diretamente do fato de que os

produtos $e_{i_1} \otimes \cdots \otimes e_{i_p}$ formam uma base de V_0^p e, se $J \neq K$, nenhum elemento dessa base que entra na confecção de e_J comparece na expressão de e_K.

Para finalizar a terceira construção da p-ésima potência exterior de V, definimos uma aplicação p-linear alternada $\phi\colon V \times \cdots \times V \to Z$ pondo

$$\phi(v_1, \ldots, v_p) = \sum_\sigma \varepsilon_\sigma \, \sigma^*(v_1 \otimes \cdots \otimes v_p).$$

O par (Z, ϕ), assim definido, satisfaz os axiomas requeridos, como logo se vê. O tensor $\phi(v_1, \ldots, v_p)$ chama-se o "anti-simetrizado" de $v_1 \otimes \cdots \otimes v_p$. A aplicação linear $A\colon V_0^p \to V_0^p$ definida por $A = \sum_\sigma \varepsilon \, \sigma^*$ chama-se "operação de anti-simetrização" e $(1/p!)A$ é uma projeção de V_0^p sobre o subespaço Z dos tensores anti-simétricos p vezes contravariantes.

Prosseguimos, a fim de mostrar que a p-ésima potência exterior de um espaço vetorial V é única, a menos de um isomorfismo canônico. Antes mesmo de demonstrar a unicidade, já é conveniente adotar a notação definitiva. Se (Z, ϕ) é uma p-ésima potência exterior de V, escreveremos

$$\overset{p}{\wedge} V = V \wedge \cdots \wedge V$$

em lugar de Z e $v_1 \wedge \cdots \wedge v_p \in \overset{p}{\wedge} V$ em lugar de $\phi(v_1, \ldots, v_p)$.

Os elementos de $\overset{p}{\wedge} V$ chamam-se *p-vetores*. Os p-vetores da forma $v_1 \wedge \cdots \wedge v_p$ dizem-se *decomponíveis*. Cada p-vetor se escreve (de infinitas maneiras) como soma de p-vetores decomponíveis.

O p-vetor $v_1 \wedge \cdots \wedge v_p \in \overset{p}{\wedge} V$ chama-se *produto exterior* dos vetores $v_1, \ldots, v_p \in V$.

Sendo a aplicação $(v_1, \ldots, v_p) \to v_1 \wedge \cdots \wedge v_p$ p-linear alternada, segue-se do Corolário da Proposição 3 que, dada uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, e considerando-se a base de $\overset{p}{\wedge} V$ formada pelos produtos exteriores

$$e_J = e_{j_1} \wedge \cdots \wedge e_{j_p}\,, \quad J = \{j_1 < \cdots < j_p\},$$

temos, para $v_1 = \sum \alpha_1^i\, e_i, \ldots, v_p = \sum \alpha_p^i\, e_i$, a seguinte expressão do produto exterior $v_1 \wedge \cdots \wedge v_p$, em termos da base e_J:

$$v_1 \wedge \cdots \wedge v_p = \sum_J \det(\alpha^J) e_J\,,$$

onde $\alpha = (\alpha_j^i) \in M(n \times p)$ é a matriz cujas colunas são as coordenadas dos vetores v_i na base $\mathcal{E}$.

Teorema 2. *Seja $\overset{p}{\wedge} V$ uma potência exterior p-ésima de V. A toda aplicação p-linear alternada $g\colon V \times \cdots \times V \to W$ corresponde uma única aplicação linear $\widehat{g}\colon \overset{p}{\wedge} V \to W$ tal que $\widehat{g}(v_1 \wedge \cdots \wedge v_p) = g(v_1, \ldots, v_p)$ sejam quais forem $v_1, \ldots, v_p \in V$.*

Demonstração: Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V e consideremos a base correspondente $\{e_J\}$ em $\overset{p}{\wedge} V$. Seja $w_J = g(e_{j_1}, \ldots, e_{j_p}) \in W$, $J = \{j_1 < \cdots < j_p\}$. Definamos $\widehat{g}\colon \overset{p}{\wedge} V \to W$ pondo $\widehat{g}(e_J) = w_J$ e estendendo por linearidade. Dados $v_1 = \sum \alpha_1^i\, e_i, \ldots, v_p = \sum \alpha_p^i\, e_i$ em V, temos

$$v_1 \wedge \cdots \wedge v_p = \sum_J \det(\alpha^J) e_J\,, \quad \text{donde}$$

$$\widehat{g}(v_1 \wedge \cdots \wedge v_p) = \sum_p \det(\alpha^J) w_j = g(v_1, \ldots, v_p),$$

de acordo com o Corolário da Proposição 3. Verifica-se que $\widehat{g}$ cumpre a condição imposta. É claro que, nestas condições, $\widehat{g}$ é única, pois os p-vetores decomponíveis geram $\overset{p}{\wedge} V$. Em particular, a aplicação $\widehat{g}$, definida com auxílio de uma base, não depende da escolha desta.

Teorema 3. (Unicidade da p-ésima potência exterior.) *Sejam $\overset{p}{\wedge} V$ e $\overset{p}{\cap} V$ duas potências exteriores p-ésimas do mesmo espaço vetorial V. Existe um único isomorfismo $\widehat{g}\colon \overset{p}{\wedge} V \approx \overset{p}{\cap} V$ tal que $\widehat{g}(v_1 \wedge \cdots \wedge v_p) = v_1 \cap \cdots \cap v_p$.*

Demonstração: Faz-se usando o Teorema 2, nas mesmas linhas da demonstração do Teorema 2, Capítulo 2, seu análogo para produtos tensoriais.

Corolário. *A correspondência $g \to \widehat{g}$ estabelece um isomorfismo canônico $\mathfrak{A}_p(V, W) \approx \mathcal{L}(\overset{p}{\wedge} V, W)$.*

3.4 Algumas aplicações do produto exterior

Proposição 5. *Os vetores $v_1, \ldots, v_p \in V$ são linearmente independentes se, e somente se, seu produto exterior $v_1 \wedge \cdots \wedge v_p \in \overset{p}{\wedge} V$ é diferente de zero.*

Demonstração: Como a aplicação p-linear $(v_1, \ldots, v_p) \to v_1 \wedge \cdots \wedge v_p$ é alternada, $v_1 \wedge \cdots \wedge v_p \neq 0$ implica que $v_1, \ldots, v_p$ são linearmente independentes (Proposiçao 1). Reciprocamente, se tais vetores são independentes, existe uma base em V que os contém. Então $v_1 \wedge \cdots \wedge v_p$ é um elemento da base correspondente em $\overset{p}{\wedge} V$, donde é $\neq 0$.

Diremos que uma matriz $\alpha \in M(n \times k)$ tem *posto* p se ela possui p colunas linearmente independentes, mas $p+1$ quaisquer de suas colunas são linearmente dependentes.

Corolário. *Uma matriz $\alpha \in M(n \times k)$ tem posto p se, e somente se, possui um menor de ordem p não-nulo, mas todos os seus menores de ordem $p+1$ são iguais a zero.*

Com efeito, a condição necessária e suficiente para que r vetores-coluna de α, $v_1, \ldots, v_r$, sejam linearmente independentes é que $v_1 \wedge \cdots \wedge v_r \neq 0$. Ora, as coordenadas de $v_1 \wedge \cdots \wedge v_r$ relativamente à base canônica do R^n são os menores de ordem r extraídos das colunas de α correspondentes aos vetores v_i.

Segue-se do corolário acima que α tem posto p se, e somente se, possui p *linhas* linearmente independentes mas não possui $p+1$.

Proposição 6. *Sejam $u_1, \ldots, u_p \in V$ e $v_1, \ldots, v_p \in V$ duas p-uplas de vetores linearmente independentes. Essas p-uplas geram o mesmo subespaço de V se, e somente se, existe um escalar $\lambda \neq 0$, tal que $u_1 \wedge \cdots \wedge u_p = \lambda\, v_1 \wedge \cdots \wedge v_p$.*

Demonstração: Suponhamos que os u_i e os v_i geram o mesmo subespaço W de V. Como $\dim W = p$, temos $\dim \overset{p}{\wedge} W = 1$. Os p-vetores $u_1 \wedge \cdots \wedge u_p$ e $v_1 \wedge \cdots \wedge v_p$ formam duas bases de $\overset{p}{\wedge} W$, donde existe $\lambda \neq 0$ tal que $u_1 \wedge \cdots \wedge u_p = \lambda\, v_1 \wedge \cdots \wedge v_p$. Reciprocamente, se vale esta igualdade, dado $x \in V$, tem-se $x \wedge u_1 \wedge \cdots \wedge u_p = 0$ se, e somente se, $x \wedge v_1 \wedge \cdots \wedge v_p = 0$. Ou seja: x é combinação linear dos u_i se, e somente se, x é combinação linear dos v_i. Assim, os u_i e os v_i geram o mesmo subespaço de V.

Segue-se da Proposição 6 que a correspondência que associa a cada p-vetor decomponível não nulo $v_1 \wedge \cdots \wedge v_p \in \overset{p}{\wedge} V$ o subespaço de V gerado por $v_1, \ldots, v_p$ é bem definida. Além disso, a dois vetores decomponíveis corresponde o mesmo subespaço se, e somente se, um deles é múltiplo do outro. O subespaço W, gerado pelos vetores independentes $v_1, \ldots, v_p$, pode ser caracterizado como o conjunto dos vetores $x \in V$ tais que $x \wedge v_1 \wedge \cdots \wedge v_p = 0$. Pode-se ainda verificar que, se $u_1 \wedge \cdots \wedge u_p \to U$ e $w_1 \wedge \cdots \wedge w_k \to W$ segundo essa correspondência, então $U \cap W = \{0\}$ se, e somente se

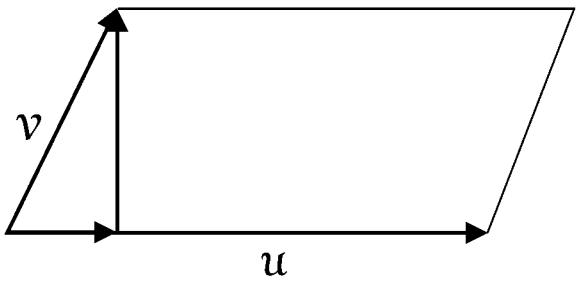
$$u_1 \wedge \cdots \wedge u_p \wedge w_1 \wedge \cdots \wedge w_k \neq 0$$

o que $U \subset W$ se, e somente se, existem $v_1, \ldots, v_{k-p}$ tais que $w_1 \wedge \cdots \wedge w_k = u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_{k-p}$.

Num espaço vetorial euclidiano V, podemos definir o *volume* (p-dimensional) do paralelepípedo gerado pelos vetores $v_1, \ldots, v_p \in V$. Poremos

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

igual à raiz quadrada do determinante da matriz $p \times p$ formada pelos produtos internos $v_1 \cdot v_j$. Esta fórmula é uma generalização direta dos casos conhecidos ($p = 2, 3$) da Geometria Analítica. Por exemplo, a área A do paralelogramo determinado por 2 vetores u, v é dada por



$A = |u| \cdot |h|$, onde $|h|^2 = |v|^2 - |w|^2$ e $w = (u \cdot v)u/|u|^2$ é a projeção ortogonal de v sobre u. Então $|u|^2\,|h|^2 = |u|^2\,|v|^2 - (u \cdot v)^2$ e portanto $A^2 = (u \cdot u)(v \cdot v) - (u \cdot v)^2$ isto é:

$$A^2 = \begin{vmatrix} u \cdot u & u \cdot v \\ u \cdot u & v \cdot v \end{vmatrix}.$$

A definição geral que demos requer dois cuidados. Em primeiro lugar, devemos mostrar que o determinante $\det(v_i \cdot v_j)$, conhecido como o "Gramiano" dos vetores v_i, é sempre ≥ 0. Em segundo lugar, como o volume p-dimensional de um paralelepípedo degenerado (isto é, contido num subespaço de dimensão $< p$) deve ser nulo, esse determinante só deve ser $\neq 0$ quando os vetores v_i forem linearmente independentes. Tais exigências sugerem que o Gramiano seja o quadrado do comprimento do p-vetor $v_1 \wedge \cdots \wedge v_p$.

Trataremos, pois, de introduzir na potência exterior $\overset{p}{\wedge} V$ um produto interno $(z, w) \to z \cdot w$ tal que

$$(u_1 \wedge \cdots \wedge u_p) \cdot (v_1 \wedge \cdots \wedge v_p) = \det(u_i \cdot v_j).$$

Para definir esse produto interno, começamos considerando a função de $2p$ variáveis $g\colon V \times \cdots \times V \to R$, dada por $g(u_1, \ldots, u_p, v_1, \ldots, v_p) = \det(u_i \cdot v_j)$. Como o determinante de uma matriz é uma função multilinear alternada de suas linhas e de suas colunas, segue-se que g é uma forma $2p$-linear, alternada relativamente aos u_i e aos v_j separadamente. Uma aplicação judiciosa do Teorema 2 mostra que existe uma forma bilinear

$$\widehat{g}\colon \overset{p}{\wedge} V \times \overset{p}{\wedge} V \to R$$

tal que $\widehat{g}(u_1 \wedge \cdots \wedge u_p,\, v_1 \wedge \cdots \wedge v_p) = \det(u_i \cdot v_j)$. Como o valor $\det(u_i \cdot v_j)$ não se altera quando se trocam as linhas pelas colunas, segue-se que $\widehat{g}$ é uma forma simétrica, isto é, $\widehat{g}(z, w) = \widehat{g}(w, z)$.

Escreveremos $z \cdot w = \widehat{g}(z, w)$, para $z, w \in \overset{p}{\wedge} V$.

Para mostrar que $(z, w) \to z \cdot w$ é um produto interno em $\overset{p}{\wedge} V$, resta somente verificar que, dado $z \in \overset{p}{\wedge} V$, temos $z \cdot z \geq 0$, e que $z \neq 0$ implica $z \cdot z > 0$.

Procederemos indiretamente, do seguinte modo: seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base ortonormal em V. $\mathcal{E}$ determina uma base em $\overset{p}{\wedge} V$, formada pelos p-vetores $e_J = e_{j_1} \wedge \cdots \wedge e_{j_p}$, $\{j_1 < \cdots < j_p\} = J$. É fácil verificar que cada produto $e_J \cdot e_K = \det(e_{j_r} \cdot e_{k_s})$, $K = \{k_1 < \cdots < k_p\}$, é zero se $J \neq K$ e igual a 1 se $J = K$. Segue-se então da bilinearidade de $z \cdot w$ que se $z = \sum_J \lambda^J e_J$ e $w = \sum_K \mu^K e_K$ são elementos de $\overset{p}{\wedge} V$, então $z \cdot w = \sum_J \lambda^J \mu^J$. Daí resulta imediatamente que $z \cdot z = \sum_J (\lambda^J)^2$, donde $z \cdot z \geq 0$ e $z \cdot z > 0$ quando $z \neq 0$.

Assim, $\overset{p}{\wedge} V$ fica munido de uma estrutura de espaço vetorial euclidiano, relativamente à qual o comprimento de um p-vetor $v_1 \wedge \cdots \wedge v_p$ coincide com o volume p-dimensional do paralelepípedo determinado em V pelos vetores $v_1, \ldots, v_p$.

Da maneira como o produto interno foi definido em $\overset{p}{\wedge} V$, segue-se que o Gramiano $\det(u_i \cdot u_j)$ é sempre ≥ 0 e é > 0 precisamente quando os vetores $u_1, \ldots, u_p$ são linearmente independentes. Para $p = 2$, obtemos a desigualdade de Cauchy-Schwarz: $(u \cdot v)^2 \leq (u \cdot u)(v \cdot v)$.

Escolhendo uma base ortonormal $\varepsilon = \{e_1, \ldots, e_n\}$ em V, pondo $u_1 = \sum_i \alpha_1^i \, e_i, \ldots, u_p = \sum_i \alpha_p^i \, e_i$, chamando de α a matriz $(\alpha_j^i) \in M(n \times p)$ dos coeficientes dos vetores u_j, temos $u_i \cdot u_j = \sum_k \alpha_i^k \, \alpha_j^k$, donde a matriz $(u_i \cdot u_j)$ é o produto $\alpha^t \cdot \alpha$ da transposta de α por α. Notando que $u_1 \wedge \cdots \wedge u_p = \sum_J \det(\alpha^J) e_J$ é a expressão do p-vetor $u_1 \wedge \cdots \wedge u_p$ relativamente à base ortonormal $\{e_J\}$ de $\overset{p}{\wedge} V$, segue-se que

$$\det(\alpha^t \cdot \alpha) = [\mathrm{vol}(u_1, \ldots, u_p)]^2 = |u_1 \wedge \cdots \wedge u_p|^2 = \\ = \sum_J [\det(\alpha^J)]^2.$$

Obtemos assim a *identidade de Lagrange*:

$$\det(\alpha^t \cdot \alpha) = \sum_J [\det(\alpha^J)]^2.$$

Acima, α é uma matriz $n \times p$, com $p \leq n$. Se fosse $n < p$, teríamos $\det(\alpha^t \cdot \alpha) = 0$. Com efeito, $\alpha^t \cdot \alpha$ seria então a matriz de um produto de aplicações lineares do tipo $R^p \to R^n \to R^p$ e, sendo $n < p$, a primeira dessas transformações não pode ser biunívoca, donde o produto também não é, e portanto o determinante desse produto é nulo.

Se V é euclidiano e $u_1 \wedge \cdots \wedge u_p = \alpha \cdot v_1 \wedge \cdots \wedge v_p \neq 0$, então

$$|\alpha| = \frac{|u_1 \wedge \cdots \wedge u_p|}{|v_1 \wedge \cdots \wedge v_p|} = \frac{\mathrm{vol}(u_1, \ldots, u_p)}{\mathrm{vol}(v_1, \ldots, v_p)} \, .$$

Logo $|\alpha|$ é a razão entre os volumes dos paralelepípedos determinados pelos u_i e pelos v_j. Note-se que, dadas essas duas p-uplas de vetores gerando o mesmo subespaço

$W \subset V$, a existência do número α não está subordinada a um produto interno em V. Mesmo num espaço V desprovido de produto interno, pode-se definir a razão entre os volumes de dois paralelepípedos "paralelos" (isto é, gerando o mesmo subespaço). Põe-se

$$\text{vol}(u_1, \ldots, u_p) : \text{vol}(v_1, \ldots, v_p) = |\alpha|, \quad \text{onde}$$
$$u_1 \wedge \cdots \wedge u_p = \alpha\, v_1 \wedge \cdots \wedge v_p\,.$$

Os p-vetores decomponíveis do espaço R^n podem ser definidos geometricamente, de maneira análoga aos vetores livres do plano. Um p-vetor decomponível $v_1 \wedge \cdots \wedge v_p$ é a classe de "equipolência" de uma p-upla $(v_1, \ldots, v_p)$ de vetores independentes, onde duas p-uplas $(u_1, \ldots, u_p)$ e $(v_1, \ldots, v_p)$ são *equipolentes* quando satisfazem as condições abaixo:

(1) Elas geram o mesmo subespaço $W \subset R^n$;

(2) Os paralelepípedos gerados pelas duas p-uplas dadas têm o mesmo volume, isto é, $\det(u_i \cdot u_j) = \det(v_i \cdot v_j)$;

(3) Elas estão "igualmente orientadas", isto é, a matriz de passagem de uma dessas p-uplas para a outra tem determinante > 0.

As condições (1) e (2) significam que $u_1 \wedge \cdots \wedge u_p = \pm\, v_1 \wedge \cdots \wedge v_p\,$. A condição (3) exclui o sinal menos na igualdade acima.

Num espaço vetorial V, de dimensão n, o importante conceito de *orientação* está intimamente relacionado com a álgebra exterior de V. Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ duas bases de V. Diremos que $\mathcal{E}$ e $\mathcal{F}$ estão *igualmente orientadas* quando a matriz de passagem de uma dessas bases para a outra tem determinante positivo.

A definição acima é motivada pela seguinte proposição, que não demonstraremos aqui: "A matriz de passagem de $\mathcal{E}$ para $\mathcal{F}$ tem determinante > 0 se, e somente se, existem n aplicações contínuas $g_i\colon [0,1] \to V$ tais que, para cada $t \in [0,1]$, $\{g_1(t), \ldots, g_n(t)\}$ é uma base de V, sendo $g_i(0) = e_i$ e $g_i(1) = f_i$, para $i = 1, \ldots, n$." Em outras palavras, $\mathcal{E}$ e $\mathcal{F}$ são igualmente orientadas se, e somente se, é possível deformar continuamente $\mathcal{E}$ em $\mathcal{F}$, na unidade de tempo, de modo que, durante toda a deformação, os n vetores em questão não deixem de formar uma base de V.

A relação "$\mathcal{E}$ e $\mathcal{F}$ estão igualmente orientadas" reparte o conjunto das bases de V em duas classes: duas bases quaisquer de uma classe estão igualmente orientadas, mas uma base de uma classe e uma base da outra não estão igualmente orientadas. Cada uma dessas duas classes chama-se uma *orientação* em V. Um espaço vetorial orientado é um par $(V, \mathcal{O})$, onde V é um espaço vetorial e $\mathcal{O}$ é uma orientação em V. Muitas vezes, porém, para evitar o pedantismo, referir-nos-emos ao espaço vetorial orientado V, deixando $\mathcal{O}$ subentendida.

Uma orientação num espaço vetorial V, de dimensão n, pode também ser definida como uma classe de equivalência no espaço vetorial $\overset{n}{\wedge} V$, de dimensão 1. Num espaço vetorial E de dimensão 1, os vetores $\neq 0$ se dividem em duas classes disjuntas. Dois vetores $e, f \in E$ ficam na mesma classe se, e somente se, $e = \lambda f$, com $\lambda > 0$. Se $e = \lambda f$ com $\lambda < 0$, então e, f pertencem a classes distintas. Quando $E = \overset{n}{\wedge} V$, todo vetor não-nulo em E é da forma $e = e_1 \wedge \cdots \wedge e_n$, sendo $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V. Se $\mathcal{F} = \{f_1, \ldots, f_n\}$ é outra base de V, então $e_1 \wedge \cdots \wedge e_n = \lambda f_1 \wedge \cdots \wedge f_n$, ou seja $e = \lambda f$, onde λ é o determinante da matriz de passagem

de $\mathcal{F}$ para $\mathcal{E}$.

Num espaço vetorial *orientado* V, de dimensão n, além do volume comum, podemos também definir o *volume orientado* de um paralelepípedo orientado n-dimensional. Se este paralelepípedo é determinado pelos vetores $v_1, \ldots, v_n$ (nesta ordem!), o vol$(v_1, \ldots, v_n)$ é definido do seguinte modo: escolhemos uma base ortonormal $\mathcal{E} = \{e_1, \ldots, e_n\}$ que pertença à orientação de V. Então $v_1 \wedge \cdots \wedge v_n = \lambda\, e_1 \wedge \cdots \wedge e_n$. O número λ, que pode ser positivo ou negativo, é o volume orientado do paralelepípedo. Em outras palavras, se $v_j = \sum_i \alpha_j^i\, e_i$, então $\lambda = \det(\alpha_j^i)$. É claro que $|\lambda|$ coincide com o volume anteriormente definido.

Como aplicação do produto exterior de vetores, deduziremos agora a *regra de Cramer*. Trata-se, como se sabe, da resolução de um sistema de equações

$$\begin{aligned}
&\alpha_1^1 x^1 + \alpha_2^1 x^2 + \cdots + \alpha_n^1 x^n = b^1 \\
&\alpha_1^2 x^1 + \alpha_2^2 x^2 + \cdots + \alpha_n^2 x^n = b^2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\alpha_1^n x^1 + \alpha_2^n x^2 + \cdots + \alpha_n^n x^n = b^n
\end{aligned}$$

onde o número de equações é igual ao número de incógnitas e a matriz $\alpha = (\alpha_j^i)$ tem, por hipótese, determinante $\neq 0$. Introduzindo os vetores-coluna $a_i = (\alpha_i^1, \alpha_i^2, \ldots, \alpha_i^n)$ e $b = (b^1, b^2, \ldots, b^n)$ em R^n, vemos que resolver o sistema acima significa encontrar n números $x^1, x^2, \ldots, x^n$ tais que $x^1 a_1 + x^2 a_2 + \cdots + x^n a_n = b$. Como $\det(\alpha) \neq 0$, os vetores $a_1, \ldots, a_n$ constituem uma base do espaço R^n. Daí resulta que os números x^i são as coordenadas de b relativamente a esta base, donde nosso problema tem solução única. Trata-se apenas de obter explicitamente as coordenadas x^i.

Multiplicando exteriormente ambos os membros da igualdade $\sum x^j a_j = b$, à esquerda por $a_1 \wedge \cdots \wedge a_{i-1}$ e à direita por $a_{i+1} \wedge \cdots \wedge a_n$, obtemos: (Para $i = 1$, não multiplicamos à esquerda; se $i = n$, não multiplicamos à direita.)

$$a_1 \wedge \cdots \wedge a_{i-1} \wedge \left(\sum x^j a_j\right) \wedge a_{i+1} \wedge \cdots \wedge a_n = \\ = a_1 \wedge \cdots \wedge a_{i-1} \wedge b \wedge a_{i+1} \wedge \cdots \wedge a_n ,$$

ou

$$x^i(a_1 \wedge \cdots \wedge a_n) = a_1 \wedge \cdots \wedge a_{i-1} \wedge b \wedge a_{i+1} \wedge \cdots \wedge a_n .$$

Seja $e = e_1 \wedge \cdots \wedge e_n$ a base de $\overset{n}{\wedge} R^n$ correspondente à base canônica de R^n. Então:

$$a_1 \wedge \cdots \wedge a_n = \det \alpha \cdot e$$

e

$$a_1 \wedge \cdots \wedge a_{i-1} \wedge b \wedge a_{i+1} \wedge \cdots \wedge a_n = \det[\alpha(b, i)] \cdot e,$$

onde indicamos com $\alpha(b, i)$ a matriz obtida de α substituindo-se a i-ésima coluna a_i por b. Segue-se que

$$x^i = \frac{\det[\alpha(b, i)]}{\det \alpha}, \quad i = 1, \ldots, n,$$

que é chamada regra de Cramer.

3.5 Formas exteriores

Sejam V um espaço vetorial e p um inteiro positivo. Sabemos que o espaço $V_p^0 = V^* \otimes \cdots \otimes V^*$ dos tensores p vezes

covariantes é canonicamente isomorfo ao espaço $\mathcal{L}_p(V)$ das formas p-lineares sobre V. Este isomorfismo leva o produto tensorial $f^1 \otimes \cdots \otimes f^p \in V_p^0$ de formas lineares $f^i \in V^*$ no produto comum $f^1 \cdot f^2 \ldots f^p \in \mathcal{L}_p(V)$, o qual, como á vimos, é definido por $(f^1 \cdot f^2 \ldots f^p)(v_1, \ldots, v_p) = f^1(v_1) \cdot \ldots \cdot f^p(v_p)$. Esquematicamente:

$$\mathcal{L}_p(V) \to V^* \otimes \cdots \otimes V^*$$
$$f^1 \cdot f^2 \ldots f^p \to f^1 \otimes f^2 \otimes \cdots \otimes f^p.$$

Isto nos diz que o estudo dos tensores puramente covariantes pode ser reduzido diretamente ao estudo das formas p-lineares, sem necessidade de introduzir-se o conceito geral de produto tensorial. Mais ainda, resulta daí que $V^* \otimes \cdots \otimes V^*$ é canonicamente isomorfo ao espaço $(V \otimes \cdots \otimes V)^*$, dual do espaço dos tensores p vezes contravariantes. Com efeito, temos $\mathcal{L}_p(V) \approx (V \otimes \cdots \otimes V)^*$ pois as formas p-lineares sobre V correspondem, pelo Teorema 1 do Capítulo 2, às formas lineares sobre $V \otimes \cdots \otimes V$.

Nessa mesma ordem de idéias, mostraremos agora que o espaço $\overset{p}{\wedge} V^*$ dos p-vetores covariantes sobre V é canonicamente isoformo ao espaço $\mathfrak{A}_p(V)$ das formas p-lineares alternadas. Mais precisamente, demonstraremos a

Proposição 7. *O isomorfismo canônico $\mathcal{L}_p(V) \approx V^* \otimes V^*$, acima referido, leva o subespaço $\mathfrak{A}_p(V)$ das formas p-lineares alternadas sobre o subespaço Z^* dos tensores anti-simétricos p vezes covariantes.*

Demonstração: Tomemos uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$; seja $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ a base dual em V^*. Pelo Teorema 1, obtemos uma base de $\mathfrak{A}_p(V)$, formada por formas

$$\varphi^J, J = \{j_1 < \cdots < j_p\} \subset \{1, \ldots, n\},$$

definidas no enunciado daquele teorema. Segue-se da definição que

$$\varphi^J = \sum_\sigma \varepsilon_\sigma \, e^{\sigma(j_1)} \, e^{\sigma(j_2)} \ldots e^{\sigma(j_p)}$$

soma estendida a todas as permutações $\sigma\colon J \to J$. O isomorfismo $\mathcal{L}_p(V) \approx V_p^0$ leva então φ^J em

$$\sum_\sigma \varepsilon_\sigma \, e^{\sigma(j_1)} \otimes e^{\sigma(j_2)} \otimes \cdots \otimes e^{\sigma(j_p)} = e^J.$$

Mas, como vimos no decorrer da Terceira Construção da p-ésima potência exterior, os vetores e^J constituem uma base de Z^*. Isto conclui a demonstração.

Sabemos que o espaço Z^* dos tensores anti-simétricos p vezes covariantes pode ser considerado como p-ésima potência exterior de V^*, sendo $f_1 \wedge \cdots \wedge f_p$

$$\sum_\sigma \varepsilon_\sigma \, f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(p)} \, .$$

Assim, temos o

Corolário 1. $\mathfrak{A}_p(V) \approx \overset{p}{\wedge} V^*$. *Por este isomorfismo canônico, ao produto exterior* $f^1 \wedge \cdots \wedge f^p$ *das formas lineares* $f^i \in V^*$, *corresponde a forma p-linear* $f \in \mathfrak{A}_p(V)$ *tal que*

$$f(v_i, \ldots, v_p) = \det[f^i(v_j)].$$

Com efeito,

$$f^1 \wedge \cdots \wedge f^p = \sum_\sigma \varepsilon_\sigma \, f^{\sigma(1)} \otimes \cdots \otimes f^{\sigma(p)}$$

é a imagem, pelo isomorfismo acima, de

$$f = \sum_{\sigma} \varepsilon_\sigma \, f^{\sigma(1)}, f^{\sigma(2)} \dots f^{\sigma(p)} .$$

donde

$$f(v_1, \dots, v_p) = \sum_{\sigma} \varepsilon_\sigma \big(f^{\sigma(1)}, f^{\sigma(2)} \dots f^{\sigma(p)}\big)(v_1, \dots, v_p) =$$
$$= \sum_{\sigma} \varepsilon_\sigma \, f^{\sigma(1)}(v_1) \dots f^{\sigma(p)}(v_p) = \det[f^i(v_j)].$$

Corolário 2. $\overset{p}{\wedge} V^* \approx (\overset{p}{\wedge} V)^*$ *pelo isomorfismo canônico que aplica o p-vetor* $f^1 \wedge \dots \wedge f^p$ *no funcional linear* $f\colon \overset{p}{\wedge} V \to R$ *tal que*

$$f(v_1 \wedge \dots \wedge v_p) = \det[f^i(v_j)].$$

Com efeito, temos o isomorfismo canônico

$$\mathfrak{A}_p(V) \approx \overset{p}{\wedge} V^*,$$

dado pelo Corolário do Teorema 2, que leva $f \in \mathfrak{A}_p(V)$ em $\widehat{f} \in (\wedge V)^*$, onde $\widehat{f}(v_1 \wedge \dots \wedge v_p) = f(v_1, \dots, v_p)$. Compondo este isomorfismo com o do Corolário acima, temos o resultado desejado.

Se quiséssemos fazer apenas o estudo dos p-vetores covariantes sobre um espaço V, poríamos $\overset{p}{\wedge} V = \mathfrak{A}_p(V)$ e, para $f^1, \dots, f^p \in V^*$, definiríamos $f^1 \wedge \dots \wedge f^p \in \mathfrak{A}_p(V)$ pondo

$$(f^1 \wedge \dots \wedge f^p)(v_1, \dots, v_p) = \det[f^i(v_j)].$$

Um p-vetor covariante $f \in \overset{p}{\wedge} V^*$ é também chamado uma *p-forma exterior* sobre V. O estudo das p-formas exteriores tem grande importância em Geometria Diferencial e Análise, onde foi introduzido por Élie Cartan, com o fito de tratar intrinsecamente a integração de formas diferenciais, e os chamados sistemas de Pfaff.

3.6 Potência exterior de uma aplicação linear

Seja $A\colon V \to W$ uma aplicação linear. Como a aplicação $(v_1, \ldots, v_p) \to Av_1 \wedge \cdots \wedge Av_p$ é p-linear alternada, existe uma única aplicação linear $\overset{p}{\wedge} V \to \overset{p}{\wedge} W$, que indicaremos com $\overset{p}{\wedge} A$ e chamaremos de *p-ésima potência exterior* de A, tal que:

$$\overset{p}{\wedge} A(v_1 \wedge \cdots \wedge v_p) = Av_1 \wedge \cdots \wedge Av_p\,.$$

Dadas as bases $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, $\mathcal{F} = \{f_1, \ldots, f_r\}$ em W, temos $Ae_j = \sum_i \alpha_j^i f_i$, onde $\alpha = (\alpha_j^i) \in M(r \times n)$ é a matriz de A relativamente a estas bases. Então, para cada subconjunto $J = \{e_1 < \cdots < j_p\} \subset I_n$, temos $e_J = e_{j_1} \wedge \cdots \wedge e_{j_p}$ e $(\overset{p}{\wedge} A)e_J = Ae_{j_1} \wedge \cdots \wedge Ae_{j_p} = \sum_K \det(\alpha_J^K) f_K$, onde $K = \{k_1 < \cdots < k_p\} \subset I_r$ e $f_K = f_{k_1} \wedge \cdots \wedge f_{k_p}$. Logo, a matriz de $\overset{p}{\wedge} A\colon \overset{p}{\wedge} V \to \overset{p}{\wedge} W$ relativamente às bases e_J e f_K, determinadas por $\mathcal{E}$ e $\mathcal{F}$ respectivamente, é a matriz

$$\big(\det(\alpha_J^K)\big) \in M\left[\binom{r}{p} \times \binom{n}{p}\right],$$

cujos elementos são os menores de ordem p da matriz de A.

Em particular, se $\dim V = n$ e $A\colon V \to V$, então $\overset{n}{\wedge} A\colon \overset{n}{\wedge} V \to \overset{n}{\wedge} V$ é tal que $(\overset{n}{\wedge} A)\, z = \det A \cdot z$, para todo $z \in \overset{n}{\wedge} V$. Esta igualdade pode ser usada para definir o determinante de A.

Segue-se que, quando $\dim V = n$, a razão entre os volumes dos paralelepípedos gerados por $\{u_1, \ldots, u_n\}$ e $\{Au_1, \ldots, Au_n\}$ é dada por

$$\mathrm{vol}(Au_1, \ldots, Au_n) : \mathrm{vol}(u_1, \ldots, u_n) = \det A.$$

Mais precisamente, esta é a razão entre os volumes orientados correspondentes. Ela não depende de uma orientação tomada em V. As bases $\{u_1, \ldots, u_n\}$ e $\{Au_1, \ldots, Au_n\}$ são igualmente orientadas ou não, conforme $\det A > 0$ ou $\det A < 0$. O valor absoluto $|\det A|$ fornece a escala segundo a qual os volumes em V são transformados pela aplicação A.

A forma da matriz de $\overset{p}{\wedge} A$ nos mostra ainda que o *posto* de uma aplicação linear $A\colon V \to W$ (dimensão do subespaço $A(V) \subset W$) é o maior inteiro p tal que $\overset{p}{\wedge} A \neq 0$. Com efeito, o posto de A, assim definido, é o posto de qualquer matriz de A.

Dada uma aplicação linear $A\colon V \to W$, sua *adjunta* é a aplicação linear $A^*\colon W^* \to V^*$ definida por $(A^* g)v = g(Av)$, $g \in W^*$, $v \in V$. Então o diagrama abaixo, onde L_1

e L_2 indicam isomorfismos canônicos, é comutativo.

$$\begin{array}{ccc} (\overset{p}{\wedge} W)^* & \xrightarrow{(\overset{p}{\wedge} A)^*} & (\overset{p}{\wedge} V)^* \\ {\scriptstyle L_1}\downarrow & & \downarrow{\scriptstyle L_2} \\ \overset{p}{\wedge} W^* & \xrightarrow{\overset{p}{\wedge} A^*} & \overset{p}{\wedge} V^* \end{array}$$

Em outras palavras, se fizermos as identificações $(\overset{p}{\wedge} V)^* = \overset{p}{\wedge} V^*$ e $(\overset{p}{\wedge} W)^* = \overset{p}{\wedge} W^*$, teremos $(\overset{p}{\wedge} A)^* = \overset{p}{\wedge} A^*$. A verificação deste fato é deixada a cargo do leitor.

Resulta diretamente da definição que, se $A\colon V \to W$ e $B\colon W \to Z$ são aplicações lineares, então

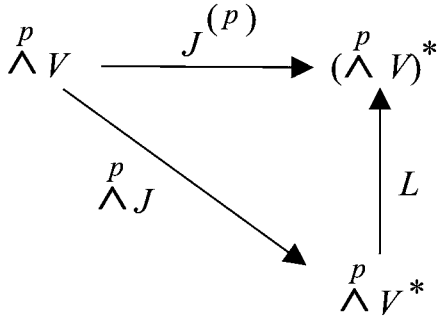
$$\overset{p}{\wedge}(BA) = \overset{p}{\wedge} B \circ \overset{p}{\wedge} A\colon \overset{p}{\wedge} V \to \overset{p}{\wedge} Z.$$

Em conseqüência, se $A\colon V \to W$ é invertível e $A^{-1}\colon W \to V$ é sua inversa, então $\overset{p}{\wedge} A\colon \overset{p}{\wedge} V \to \overset{p}{\wedge} W$ também é invertível, sendo $(\overset{p}{\wedge} A)^{-1} = \overset{p}{\wedge}(A^{-1})$.

Mais geralmente, se $A\colon V \to W$ é biunívoca, então $\overset{p}{\wedge} A$ também o é. Com efeito, existe nesse caso uma aplicação linear $B\colon B \to V$ tal que $BA =$ identidade: $V \to V$. (Basta tomar B como a composta de uma projeção $W \to A(V)$ com a inversa $A(V) \to V$, que existe pela biunivocidade de A.) Então teremos $\overset{p}{\wedge} B \cdot \overset{p}{\wedge} A = \overset{p}{\wedge}(BA) =$ identidade, donde $\overset{p}{\wedge} A$ é biunívoca. Tomando o caso particular em que $V \subset W$ e $A\colon V \to W$ é a aplicação de inclusão, vemos que $\overset{p}{\wedge} A\colon \overset{p}{\wedge} V \to \overset{p}{\wedge} W$ é biunívoca, donde $\overset{p}{\wedge} V$ pode ser considerado, de modo natural, como um subespaço de $\overset{p}{\wedge} W$, como sempre faremos.

Se $A\colon V \to W$ é sobre W, então $\overset{p}{\wedge} A\colon \overset{p}{\wedge} V \to \overset{p}{\wedge} W$ também é sobre $\overset{p}{\wedge} W$, como se vê por um raciocínio análogo.

Quando V é um espaço vetorial euclidiano, o isomorfismo canônico $J\colon V \to V^*$ induz um isomorfismo $\wedge J\colon \wedge V \to \wedge V^*$. Por outro lado, $\wedge V$ herda de V um produto interno, e portanto existe também um isomorfismo canônico $J^{(p)}\colon \overset{p}{\wedge} V \to (\overset{p}{\wedge} V)^*$. Seja $L\colon \overset{p}{\wedge} V^* \to (\overset{p}{\wedge} V)^*$ o isomorfismo canônico. O leitor pode verificar que $J^{(p)} = L \circ (\overset{p}{\wedge} J)$, ou seja, que o diagrama abaixo é comutativo:



3.7 Álgebra de Grassmann

Dadas as potências exteriores $\overset{p}{\wedge} V$, $\overset{q}{\wedge} V$ do espaço vetorial V, existe uma única aplicação bilinear

$$\overset{p}{\wedge} V \times \overset{q}{\wedge} V \to \overset{p+q}{\wedge} V$$

tal que $(u_1 \wedge \cdots \wedge u_p\, v_1 \wedge \cdots \wedge v_q) \to u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_q$. Ela é chamada a *multiplicação exterior de p-vetores por q-vetores*, e é induzida pela aplicação $(p+q)$-linear $(u_1, \ldots, u_p, v_1, \ldots, v_q) \to u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_q$, a qual é alternada em relação a $u_1, \ldots, u_p$ e a $v_1, \ldots, v_q$

separadamente. O produto exterior de um p-vetor por um q-vetor goza das seguintes propriedades:

(1) $z \wedge (w + t) = z \wedge w + z \wedge t$;
(2) $z \wedge (\alpha w) = \alpha(z \wedge w)$;
(3) $z \wedge w = (-1)^{pq}\, w \wedge z$, se $z \in \overset{p}{\wedge} V$ e $w \in \overset{q}{\wedge} V$;
(4) $z \wedge (w \wedge t) = (z \wedge w) \wedge t$.

As propriedades (1) e (2) exprimem simplesmente a bilinearidade da multiplicação. Em virtude de (3), basta postulá-las em relação a um dos fatores. Quando a (3), chamada *propriedade anti-comutativa*, é suficiente verificá-la quando $z = u_1 \wedge \cdots \wedge u_p$ e $w = v_1 \wedge \cdots \wedge v_q$ são decomponíveis. Neste caso, $z \wedge w = u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_q = (-1)^{pq}\, v_1 \wedge \cdots \wedge v_q \wedge u_1 \wedge \cdots \wedge u_p = (-1)^{pq}\, w \cdot z$. Do mesmo modo, a propriedade associativa (4) basta ser verificada para z, w, t decomponíveis, sendo evidente neste caso.

Como aplicação do produto exterior de um p-vetor por um q-vetor, obteremos agora o *desenvolvimento de Laplace* de um determinante.

Seja $\alpha = (\alpha^i_j)$ uma matriz quadrada de ordem $n = p + q$. Seja $H = \{h_1 < \cdots < h_p\}$ um conjunto de p inteiros compreendidos entre 1 e n. Indiquemos com $R^H \subset R^n$ o subespaço de dimensão p formado pelos vetores $x = (x^1, \ldots, x^n) \in R^n$ tais que $x^i = 0$ se $i \notin H$, e com $R^{H'}$ o subespaço de R^n de dimensão q, formado pelos vetores $y = (y^1, \ldots, y^n) \in R^n$ tais que $y^i = 0$ se $i \in H$. Cada vetor em R^n se escreve, de modo único, como soma de um vetor de R^H com um vetor de $R^{H'}$. Em particular, cada vetor-coluna $v_i = (\alpha^1_i, \alpha^2_i, \ldots, \alpha^n_i)$ da matriz α se escreve como $v_i = a_i + b_i$, onde $a_i \in R^H$ e $b_i \in R^{H'}$. Indicando com $\mathcal{E} = \{e_1, \ldots, e_n\}$ a base canônica de R^n e com $e = e_1 \wedge \cdots \wedge e_n$

a base correspondente de $\overset{n}{\wedge} R^n$, temos:

$$\det \alpha \cdot e = v_1 \wedge \cdots \wedge v_n = (a_1 + b_1) \wedge (a_2 + b_2) \wedge \cdots \wedge (a_n + b_n).$$

O produto à direita se distribui numa soma de parcelas do tipo $c_1 \wedge c_2 \wedge \cdots \wedge c_n$ onde cada $c_i = a_i$, ou $c_i = b_i$. Como $\dim R^H = p$ e $\dim R^{H'} = q$, cada produto desses que tiver mais de p fatores do tipo a ou mais de q fatores do tipo b, é igual a zero. Restarão, na soma, apenas as parcelas que possuem exatamente p fatores a_i e q fatores b_j . Agrupando os a_i no princípio, em ordem crescente de índices, e os b_j no fim, obtemos:

$$\det \alpha \cdot e = \sum_J \varepsilon(J, J') a_{j_1} \wedge a_{j_2} \wedge \cdots \wedge a_{j_p} \wedge b_{j'_1} \wedge \cdots \wedge b_{j'_q} ,$$

a soma estendendo-se a todos os subconjuntos $J = \{j_1 < \cdots < j_p\} \subset I_n$, onde $J' = \{j'_1 < \cdots < j'_q\}$ indica o complementar de J em I_n e $\varepsilon(J, J') = \pm 1$ é o sinal da permutação

$$\{1, 2, \ldots, n\} \to \{j_1, \ldots, j_p, j'_1, \ldots, j'_q\}$$

ou seja, -1 elevado ao número de pares (j, j'), onde $j \in J$, $j' \in J'$ e $j' < j$.

Escrevendo, como de costume,

$$e_H = e_{h_1} \wedge \cdots \wedge e_{h_p} \in \overset{p}{\wedge} R^H, \quad e_{H'} = e_{h'_1} \wedge \cdots \wedge e_{h'_q} \in \overset{p}{\wedge} R^{H'},$$

teremos

$$a_{j_1} \wedge \ldots a_{j_p} = \det(\alpha^H_J) e_H \, , b_{j'_1} \wedge \cdots \wedge b_{j'_q} = \det(\alpha^{H'}_{J'}) e_{H'} \, .$$

Logo:

$$\begin{aligned}\det\alpha\cdot e &= \sum_J \varepsilon(J,J')\det(\alpha_J^H)e_H\wedge\det(\alpha_{J'}^{H'})e_{H'} = \\ &= \sum_J \varepsilon(J,J')\det(\alpha_J^H)\det(\alpha_{J'}^{H'})e_H\wedge e_{H'} = \\ &= \varepsilon(H,H')\sum_J \varepsilon(J,J')\det(\alpha_J^H)\det(\alpha_{J'}^{H'})e.\end{aligned}$$

Assim, igualando os coeficientes:

$$\det\alpha = \varepsilon(H,H')\sum_J \varepsilon(J,J')\det(\alpha_J^H)\det(\alpha_{J'}^{H'}).$$

Esta é a fórmula que dá o *desenvolvimento de Laplace* para $\det\alpha$, relativo à escolha das linhas H: $\det\alpha$ é a soma dos produtos de cada menor $\det(\alpha_J^H)$ com H fixo, por seu "menor complementar" $\varepsilon(J,J')\det(\alpha_{J'}^{H'})$. Em particular, tomando $H=\{i\}$ obtemos o desenvolvimento de $\det\alpha$ segundo os elementos da i-ésima linha de α.

É claro que, trocando linhas por colunas, teremos um desenvolvimento de Laplace segundo um conjunto fixo K de colunas de α.

Como conseqüência da fórmula de Laplace, vemos que se α é da forma $\left(\begin{smallmatrix}\beta & \gamma\\ 0 & \delta\end{smallmatrix}\right)$, onde $\beta\in M(p\times p)$, $\delta\in M(q\times q)$, $\gamma\in M(p\times q)$ e 0 é a matriz nula $q\times p$, então $\det\alpha = \det\beta\cdot\det\delta$.

É natural considerar $\overset{0}{\wedge}V = R$. Além disso, da definição geral já segue que $\overset{1}{\wedge}V = V$. O produto exterior

$$\overset{p}{\wedge}V\times\overset{q}{\wedge}V\to\overset{p+q}{\wedge}V$$

reduz-se ao produto usual de um p-vetor (ou de um q-vetor) por um escalar quando $q = 0$ ou $p = 0$. Quando $p = q = 1$, o produto exterior $(u, v) \to u \wedge v$ reduz-se ao produto exterior usual de dois vetores.

Consideremos a soma direta

$$\wedge V = \overset{0}{\wedge} V \oplus \overset{1}{\wedge} V \oplus \cdots \oplus \overset{n}{\wedge} V = \sum_{p=0}^{n} (\overset{p}{\wedge} V), \quad n = \dim V.$$

(Como $\overset{p}{\wedge} V = \{0\}$ se $p > n$, poderíamos escrever $\wedge V = \sum_{p=0}^{\infty} \overset{p}{\wedge} V$.)

Cada elemento $z \in \wedge V$ se escreve, de modo único, como uma soma

$$z = z_0 = z_1 + \cdots + z_n \,,$$

onde cada parcela z_p é um p-vetor. É claro que $z_0 \in R$ e $z_1 \in V$. Dado outro elemento

$$w = w_0 + w_1 + \cdots + w_n \,,$$

definiremos o produto

$$z \wedge w = z_0 \wedge w_0 + (z_1 \wedge w_0 + z_0 \wedge w_1) + \cdots + z_n \wedge w_n \,.$$

Verifica-se facilmente que a aplicação

$$(\wedge V) \times (\wedge V) \to \wedge V$$

assim definida é bilinear e portanto introduz em $\wedge V$ uma estrutura de álgebra, com a qual V chama-se a *álgebra de Grassmann*, ou a *álgebra exterior* do espaço vetorial V.

A álgebra de Grassmann $\wedge V$ é associativa, possui uma unidade, e é uma *álgebra graduada anti-comutativa*, isto é,

cada elemento $z \in \wedge V$ se escreve, de modo único como soma $z = z_0 + \cdots + z_n$ de suas componentes "homogêneas" $z_p \in \overset{p}{\wedge} V$, sendo o produto $z_p \wedge z_q$ de dois elementos homogêneos um elemento homogêneo de grau $p+q$, (isto é, pertencente a $\overset{p+q}{\wedge} V$), e $z_p \wedge z_q = (-1)^{pq}\, z_q \wedge z_p$, quando z_p e z_q são homogêneos de graus p e q respectivamente.

Uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V determina uma base em $\wedge V$, formada pelos elementos $e_J = e_{j_1} \wedge \cdots \wedge e_{j_p}$, onde $J = \{j_1 < \cdots < j_p\}$ percorre *todos* os subconjuntos de $\{1, \ldots, n\}$. Põe-se $e_J = 1$ quando J é vazio. Assim, $\wedge V$ é uma álgebra de dimensão finita e, na realidade, $\dim \wedge V = 2^n$.

A álgebra de Grassmann $\wedge V$ é gerada por 1 e V, goza da propriedade de ser $v \wedge v = 0$ para todo $v \in V$, e a sua multiplicação não satisfaz nenhuma outra relação, salvo as que decorrem desta. Ou seja, $\wedge V$ é a *álgebra livre associativa e anti-comutativa* gerada por 1 e V. Esta afirmação significa, em termos matemáticos, que o teorema abaixo é verdadeiro.

Teorema 4. *Sejam V um espaço vetorial e A uma álgebra associativa com unidade 1. Se $f\colon V \to A$ é uma aplicação linear tal que $f(u)f(u) = 0$ seja qual for $u \in V$, então existe um único homomorfismo unitário $F\colon \wedge V \to A$ tal que $F(u) = f(u)$ para todo $u \in V$.*

Demonstração: Para cada inteiro $p > 0$, seja $f_p\colon V \times \cdots \times V \to A$ a aplicação p-linear definida por $f_p(u_1, \ldots, u_p) = f(u_1)f(u_2)\ldots f(u_p)$. Pela hipótese sobre f, cada f_p é alternada. Logo existe, para cada $p > 0$, uma aplicação linear $\widehat{f_p}\colon \overset{p}{\wedge} V \to A$ tal que $\widehat{f_p}(u_1 \wedge \cdots \wedge u_p) = f(u_1)f(u_2)\ldots$

$f(u_p)$. Ponhamos ainda $\widehat{f_0}\colon R \to A$, com $\widehat{f_0}(\lambda) = \lambda 1 \in A$. Definamos agora $F\colon \wedge V \to A$ pondo $F(z_0+z_1+\cdots+z_n)+\cdots+\widehat{f_n}(z_n)$. Verifica-se imediatamente que F é um homomorfismo unitário de $\wedge V$ em A. É o único homomorfismo unitário que coincide com f em V porque $\wedge V$ é gerado por V e R.

Observação: Na álgebra de Grassmann $\wedge V$, tem-se $z \wedge z = 0$ sempre que z é um p-vetor decomponível, ou quando z é soma de componentes homogêneas de grau ímpar apenas. Mas, se tomarmos $z = e_1 \wedge e_2 + e_3 \wedge e_4$, onde os e_i são elementos de uma base de V, veremos que $z \wedge z = 2e_1 \wedge e_2 \wedge e_3 \wedge e_4$, donde $z \wedge z \neq 0$.

3.8 Produtos interiores

É possível definir o "produto" de um p-vetor covariante por um q-vetor contravariante, dando como resultado um $(p-q)$-vetor covariante se $p \geq q$ e um $(q-p)$-vetor contravariante se $p \leq q$. Introduziremos pois as aplicações bilineares

$$\overset{p}{\wedge} V^* \times \overset{q}{\wedge} V \to \overset{p-q}{\wedge} V^*, \quad \text{se} \quad p \geq q,$$

$$\overset{p}{\wedge} V^* \times \overset{q}{\wedge} V \to \overset{q-p}{\wedge} V^*, \quad \text{se} \quad p \leq q.$$

A primeira delas será indicada com $(f, z) \to f \mathsf{L} z$ e será chamada o *produto interior à direita* do p-vetor covariante $f \in \overset{p}{\wedge} V^*$ pelo q-vetor contravariante $z \in \overset{q}{\wedge} V$, $(p \geq q)$, dando como resultado o $(p-q)$-vetor covariante $f \mathsf{L} z \in \overset{p-q}{\wedge} V^*$. Tudo se passa como se z estivesse retirando q das suas componentes covariantes de f.

A segunda dessas aplicações bilineares será indicada com $(f, z) \to f \lrcorner z$. Dados $f \in \overset{p}{\wedge} V^*$ e $z \in \overset{q}{\wedge} V$, onde $p \leq q$, o $(q-p)$-vetor $f \lrcorner z \in \overset{q-p}{\wedge} V$ será chamado o *produto interior à esquerda* de z por f.

Fazendo as identificações $\overset{p}{\wedge} V^* = (\overset{p}{\wedge} V)^*$ e $\overset{p-q}{\wedge} V^* = (\overset{p-q}{\wedge} V)^*$, $f \llcorner z$ será o funcional linear sobre $\overset{p-q}{\wedge} V$, definido pela condição:

$$(f \llcorner z)(w) = f(z \wedge w), \quad \text{seja qual for} \quad w \in \overset{p-q}{\wedge} V.$$

Como f é um funcional linear sobre $\overset{p}{\wedge} V$ e o produto $z \wedge w$ é bilinear, segue-se que o segundo membro é linear em f, z e w, donde $f \llcorner z$ é linear em f e z separadamente e pertence, de fato a $\overset{p-q}{\wedge} V^*$. Está, portanto, definido o produto interior à direita $\overset{p}{\wedge} V^* \times \overset{q}{\wedge} V \to \overset{p-q}{\wedge} V^*$. Note-se que, quando $p = q$, $f \llcorner z = f(z) \in R$ e, quando $q = 0$, $f \lrcorner \alpha = \alpha f$.

Se $p \leq q$, dados $f \in \overset{p}{\wedge} V^*$ e $z \in \overset{q}{\wedge} V$, o produto interior $f \lrcorner z \in \overset{q-p}{\wedge} V$ é definido pela condição

$$g(f \lrcorner z) = (g \wedge f)(x), \quad \text{seja qual for} \quad g \in \overset{q-p}{\wedge} V^*.$$

Como um vetor fica determinado univocamente quando se conhecem os valores que todos os funcionais lineares assumem nele (desde que tais valores dependam linearmente dos funcionais!), $f \lrcorner z$ está bem definido. O produto interior à esquerda

$$\overset{p}{\wedge} V^* \times \overset{q}{\wedge} V \to \overset{p-q}{\wedge} V,$$

assim introduzido, é bilinear, e no caso em que $p = q$, coincide com o produto interior à direita, e com a aplicação bilinear natural $\overset{p}{\wedge} V^* \times \overset{p}{\wedge} V \to R$.

Tem-se evidentemente $(f \llcorner z) \llcorner w = f \llcorner (z \wedge w)$ e $g \lrcorner (f \lrcorner z) = (g \wedge f) \lrcorner z$.

Sejam $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V e $\mathcal{E}^* = \{e^1, \ldots, e^n\}$ a base dual em V^*. Ficam determinadas bases em $\overset{p}{\wedge} V^*$ e em $\overset{p}{\wedge} V$, as quais são constituidas pelos produtos exteriores $e^J = e^{j_1} \wedge \cdots \wedge e^{j_p}$, $J = \{j_1 < \cdots < j_p\}$ e $e_K = e_{k_1} \wedge \cdots \wedge e_{k_q}$, $K = \{k_1 < \cdots < k_q\}$ respectivamente. Os produtos interiores satisfazem as seguintes relações:

$$e^J \lrcorner e_K = \begin{cases} 0, & \text{se} \quad J \not\subset K \\ \varepsilon(J', J) e_{J'}, & \text{se } J \subset K \text{ e } J' = K - J \end{cases}$$

$$e^J \llcorner e_K = \begin{cases} 0, & \text{se} \quad K \not\subset J \\ \varepsilon(K, K') e^{K'}, & \text{se } K \subset J \text{ e } K' = J - K \end{cases}$$

onde $\varepsilon(K, K')$ indica, como antes, o número de "inversões" na seqüência formada por K seguido de K'.

As igualdades acima são fáceis de obter diretamente, ou como conseqüência das seguintes fórmulas mais gerais:

1) $(f^1 \wedge \cdots \wedge f^p) \llcorner (v_1 \wedge \cdots \wedge v_q) =$
$$= \sum_J \varepsilon(J, J') f^J (v_1 \wedge \cdots \wedge v_q) f^{J'} \quad \text{se} \quad p \geq q;$$

2) $(f^1 \wedge \cdots \wedge f^p) \lrcorner (v_1 \wedge \cdots \wedge v_q) =$
$$= \sum_J \varepsilon(J', J) (f^1 \wedge \cdots \wedge f^p)(v_J) v_{J'} \quad \text{se} \quad p \leq q;$$

onde J percorre, na primeira fórmula, os subconjuntos de I_p com q elementos e $J' = I_p - J$, sendo $f^J = f^{j_1} \wedge \cdots \wedge f^{j_q}$, $J = \{j_1 < \cdots < j_q\}$. As demais notações são análogas.

Deduziremos a primeira fórmula. Qualquer que seja o $(p-q)$-vetor decomponível $u_1 \wedge \cdots \wedge u_{p-q}$, temos, por definição:

$$\begin{aligned}&[(f^1 \wedge \cdots \wedge f^p) \mathsf{L} (v_1 \wedge \ldots v_q)](u_1 \wedge \cdots \wedge u_{p-q}) = \\ &= (f^1 \wedge \cdots \wedge f^p)(v_1 \wedge \cdots \wedge v_q \wedge u_1 \wedge \cdots \wedge u_{p-q}) = \\ &= \det[f^i(w_j)],\end{aligned}$$

onde $w_j = v_j$ se $j \leq q$ e $w_j = u_{j-q}$ se $j > q$. Desenvolvendo este último determinante segundo a fórmula de Laplace relativa ao conjunto $K = \{1, \ldots, q\}$ das primeiras colunas da matriz $(f^i(w_j))$, obtemos:

$$\begin{aligned}&\det[f^i(w_j)] = \\ &= \sum_J \varepsilon(J, J') f^J(v_1 \wedge \cdots \wedge v_q) \cdot f^{J'}(u_1 \wedge \cdots \wedge u_{p-q}) = \\ &= \left[\sum_J \varepsilon(J, J') f^J(v_1 \wedge \cdots \wedge v_p) f^{J'}\right](u_1 \wedge \cdots \wedge u_{p-q}).\end{aligned}$$

Assim, o primeiro e o segundo membro da fórmula (1) são funcionais lineares sobre $\overset{p-q}{\wedge} V$ que assumem o mesmo valor em todos os $(p-q)$-vetores decomponíveis $u_1 \wedge \cdots \wedge u_{p-q}$. Logo esses funcionais coincidem.

A fórmula (2) se demonstra da mesma maneira.

Resulta destas fórmulas que, interpretando os elementos $f \in \overset{p}{\wedge} V^*$ e $z \in \overset{q}{\wedge} V$ como tensores anti-simétricos (covariantes e contravariantes, respectivamente), o produto interior $f \mathsf{L} z$ é o tensor anti-simétrico covariante que se obtém quando se contraem, de todas as maneiras possíveis, os q índices covariantes $j_1 < \cdots < j_q$ de f, e depois se somam

todos esses tensores contraídos, antecedendo cada um deles com o sinal adequado. O produto interior $f \lrcorner z$ tem interpretação semelhante.

Seja $\mathcal{E} = \{e_1, \ldots, e_n\}$ uma base de V. Como de costume, sejam $e = e_1 \wedge \cdots \wedge e_n$ a base correspondente em $\overset{n}{\wedge} V$, $e^* = e^1 \wedge \cdots \wedge e^n$ a base de $\overset{n}{\wedge} V^*$ correspondente à dual de $\mathcal{E}$. As aplicações

$$\phi_{e^*} \colon \overset{p}{\wedge} V \to \overset{n-p}{\wedge} V^*, \quad \phi_e \colon \overset{n-p}{\wedge} V^* \to \overset{p}{\wedge} V,$$

definidas por $\phi_{e^*}(z) = e^* \llcorner z$ e $\phi_e(f) = f \lrcorner e$, onde $z \in \overset{p}{\wedge} V$ e $f \in \overset{n-p}{\wedge} V^*$, são isomorfismos, na realidade inversos um do outro. Com efeito, dado um subconjunto $K = \{k_1 < \cdots < k_p\} \subset I_n$, temos $\phi_{e^*}(e_K) = e^* \llcorner e_K = \varepsilon(K, K')e^{K'}$ enquanto

$$\phi_e(e^{K'}) = e^{K'} \lrcorner e = \varepsilon(K, K')e_K \,.$$

Segue-se daí que a aplicação composta $\phi_e \circ \phi_{e^*} \colon \overset{p}{\wedge} V \to \overset{p}{\wedge} V$ coincide com a identidade na base $\{e_K\}$. Logo, $\phi_e \circ \phi_{e^*} =$ identidade. Do mesmo modo se vê que $\phi_{e^*} \circ \phi_e$ coincide com a aplicação identidade de $\overset{n-p}{\wedge} V^*$.

Observamos que se $\mathcal{F} = \{f_1, \ldots, f_n\}$ é outra base de V e $f = f_1 \wedge \cdots \wedge f_n$ é o n-vetor correspondente, temos $f = \alpha \cdot e$, onde α é o determinante da matriz de passagem de $\mathcal{F}$ para $\mathcal{E}$. Segue-se daí que $e^* = \alpha\, f^*$, donde os isomorfismos

$$\phi_{e^*} \colon \overset{p}{\wedge} V \to \overset{n-p}{\wedge} V^*$$

e

$$\phi_{f^*} \colon \overset{p}{\wedge} V \to \overset{n-p}{\wedge} V^*$$

estão relacionados por $\phi_{e^*} = \alpha\,\phi_{f^*}$. Semelhantemente, $\phi_e = (1/\alpha)\phi_f$.

Note-se que cada isomorfismo $\phi_e\colon \overset{p}{\wedge} V^* \to \overset{n-p}{\wedge} V$, levando a p-forma decomponível $e^J = e^{j_1} \wedge \cdots \wedge e^{j_p}$ no $(n-p)$-vetor decomponível $\varepsilon(J, J')e_{J'}$, levará também, por este motivo, toda p-forma decomponível num $(n-p)$-vetor decomponível.

Com efeito, isto é claro se a p-forma dada é zero. Se $f^1 \wedge \cdots \wedge f^p$ é uma p-forma decomponível $\neq 0$, então os funcionais $f^1, \ldots, f^p$ são linearmente independentes. Existe portanto uma base $\mathcal{F} = \{f_1, \ldots, f_n\}$ em V cuja dual $\mathcal{F}^* = \{f^1, \ldots, f^n\}$ tem como primeiros p elementos os funcionais dados. Segue-se que $\phi_f(f^1 \wedge \cdots \wedge f^p) = \pm\, f_{p+1} \wedge \cdots \wedge f_n$; logo

$$\phi_e(f^1 \wedge \cdots \wedge f^p) = \pm\,\lambda\, f_{p+1} \wedge \cdots \wedge f_n$$

é um $(n-p)$-vetor decomponível, onde λ é o determinante da matriz de passagem da base $\mathcal{E}$ para a base $\mathcal{F}$.

Em particular, para $p = 1$, como todo funcional linear é uma 1-forma decomponível, concluimos que todo $(n-1)$-vetor num espaço vetorial de dimensão n é decomponível.

Observação: Se o p-vetor $u_1 \wedge \cdots \wedge u_p \neq 0$ "determina" o subespaço $W \subset V$ (lembramos: isto significa que W é gerado por $u_1, \ldots, u_p$) então a $(n-p)$-forma

$$\phi_{e^*}(u_1 \wedge \cdots \wedge u_p) = f^1 \wedge \cdots \wedge f^{n-p}$$

determina em V^* o subespaço $W^0 = \{f \in V^*;\, f(w) = 0$ para todo $w \in W\}$. O subespaço $W^0 \subset V^*$ chama-se o *anulador* de W, ou o subespaço de V^* ortogonal a W. Isto é fácil de ver quando $e = e_1 \wedge \cdots \wedge e_n$ é tal que $e_1 = u_1, \ldots, e_p =$

u_p. Então temos $f^1 = e^{p+1}, \ldots, f^p = e^n$, sendo claro que $\{e^{p+1}, \ldots, e^n\}$ constitui uma base do subespaço formado pelos funcionais f que se anulam sobre todos os vetores $e_1, \ldots, e_p$, ou seja, sobre todos os vetores de W. Quando $\overline{e} = \overline{e}_1 \wedge \cdots \wedge \overline{e}_n$ é uma base qualquer, $\phi_{\overline{e}^*}$ difere de ϕ_{e^*} por um fator escalar, o que não altera o resultado.

Notemos que se $\{u_1, \ldots, u_p\}$ é uma base de W e

$$\phi_{e^*}(u_1 \wedge \cdots \wedge u_p) = f^1 \wedge \cdots \wedge f^{n-p},$$

sendo

$$W = \{x \in V; f^1(x) = \cdots = f^{n-p}(x) = 0\},$$

ou seja

$$f^1(x) = 0, \ldots, f^{n-p}(x) = 0$$

são as *equações* que definem W implicitamente. Assim, o isomorfismo ϕ_{e^*} estabelece a dualidade entre os subespaços de V e os sistemas de equações lineares que definem estes subespaços.

Seja agora V um espaço vetorial euclidiano orientado, de dimensão n. Se $\mathcal{E} = \{e_1, \ldots, e_n\}$ e $\mathcal{F} = \{f_1, \ldots, f_n\}$ são bases ortonormais pertencentes à orientação de V, então

$$e = e_1 \wedge \cdots \wedge e_n = f_1 \wedge \cdots \wedge f_n = f.$$

Logo, existe um isomorfismo canônico

$$\phi\colon \overset{p}{\wedge} V \to \overset{n-p}{\wedge} V$$

definido como o composto $\overset{p}{\wedge} V \to \overset{n-p}{\wedge} V^* \to \overset{n-p}{\wedge} V$, onde o primeiro é o isomorfismo ϕ_{e^*}, associado a uma base ortonormal positivamente orientada qualquer em V, e o segundo é a potência exterior $\overset{n-p}{\wedge} J^{-1}$, onde $J\colon V \to V^*$ é dado pelo produto interno de V.

Mais precisamente, deveríamos escrever $\phi_p\colon \overset{p}{\wedge} V \to \overset{n-p}{\wedge} V$, notando que $(\phi_p)^{-1} = (-1)^{p(n-p)}\,\phi_{n-p}$. [O sinal vem de $\varepsilon(J, J') = (-1)^{p(n-p)}\,\varepsilon(J', J)$.]

Dada uma base ortonormal positiva $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V, temos $\phi(e_J) = \varepsilon(J, J')e_{J'}$, seja qual for

$$J = \{j_1 < \cdots < j_p\} \subset I_n\,,$$

com $J' = I_n - J$. Segue-se daí, e da linearidade de ϕ, que se $u_1, \ldots, u_p \in V$ são linearmente independentes e geram o subespaço $W \subset V$, então $\phi(u_1 \wedge \cdots \wedge u_p) = v_1 \wedge \cdots \wedge v_{n-p}$ é caracterizado pelas seguintes propriedades:

1), $v_1, \ldots, v_{n-p}$ geram o complemento ortogonal $W^\perp$ de W em V;

2) $\mathrm{vol}(v_1, \ldots, v_{n-p}) = \mathrm{vol}(v_1, \ldots, v_p)$;

3) $\{u_1, \ldots, u_p, v_1, \ldots, v_{n-p}\}$, nesta ordem, constitui uma base positivamente orientada em V.

O $(n-p)$-vetor $v_1 \wedge \cdots \wedge v_{n-p} = \phi(u_1 \wedge \cdots \wedge u_p)$ é, às vezes, indicado com a notação $v_1 \wedge \cdots \wedge v_{n-p} = (u_1 \wedge \cdots \wedge u_p)^*$ e ϕ fica sendo chamado a *operação estrela*, de Hodge.

O $(n-p)$-vetor $(u_1 \wedge \cdots \wedge u_p)^*$ é uma generalização natural do *produto vetorial* para os p vetores $u_1, \ldots, u_p$. Quando $p = n-1$, $(u_1 \wedge \cdots \wedge u_{n-1})^*$ é de fato um vetor de V, que podemos chamar o produto vetorial de $u_1, \ldots, u_{n-1}$. Quando $p = 2$, o produto vetorial $(u \wedge v)^*$ de dois vetores $u, v \in V$ é, em geral, um $(n-2)$-vetor: $(u \wedge v)^* \in \overset{n-2}{\wedge} V$. Apenas nos espaços de 3 dimensões, o produto vetorial $(u \wedge v)^*$ de dois vetores é um vetor, coincidência esta que ocorre no espaço euclidiano ordinário R^3. Em virtude disto,

não há necessidade de considerarem-se bi-vetores no cálculo elementar, já que $\phi\colon \overset{2}{\wedge} R^3 \approx R^3$.

3.9 Observações sobre a álgebra simétrica

Voltamos a considerar, no espaço $V_0^r = V \otimes V \otimes \cdots \otimes V$ dos tensores r vezes contravariantes, a aplicação linear $\sigma^*\colon V_0^r \to V_0^r$, induzida por uma permutação σ dos inteiros $1, \ldots, r$. Como sabemos, σ^* é caracterizada pela igualdade

$$\sigma^*(v_1 \otimes \cdots \otimes v_r) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(r)}\,.$$

Se ρ é outra permutação, temos $(\rho\sigma)^* = \sigma^*\rho^*$. Como $(\mathrm{id})^* =$ identidade, segue-se que cada aplicação linear σ^* é invertível, e $(\sigma^*)^{-1} = (\sigma^{-1})^*$.

Um tensor r vezes contravariante $t \in V_0^r$ diz-se *simétrico* quando $\sigma^*(t) = t$ para toda permutação σ dos inteiros $1, \ldots, r$. Por exemplo, $u \otimes v + v \otimes u \in V_0^2$ é um tensor simétrico. Mais geralmente, se considerarmos a aplicação linear $S = \sum_\sigma \sigma^*$ de V_0^r em si mesmo, chamada a *operação de simetrização*, veremos que $S(t) = \sum_\sigma \sigma^*(t)$ é um tensor simétrico, seja qual for $t \in V_0^r$. Na realidade, t é simétrico se, e somente se, $S(t) = r!t$, e $(1/r!)S$ é uma projeção de V_0^r sobre o subespaço dos tensores simétricos.

Se $\xi^{i_1 \ldots i_n}$ são as coordenadas do tensor t relativamente a uma base $\mathcal{E}$ de V, então t é simétrico se, e somente se, $\xi^{i_{\sigma(1)} \ldots i_{\sigma(r)}} = \xi^{i_1 \ldots i_r}$ para toda permutação σ. Por exemplo, $t = \sum \xi^{ij}\, e_i \otimes e_j$ é simétrico se, e somente se $\xi^{ij} = \xi^{ji}$.

Uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de V determina em V_0^r a base formada pelos produtos tensoriais $e_{m_1} \otimes \cdots \otimes e_{m_r}$ onde $(m_1, m_2, \ldots, m_r)$ percorre todas as r-uplas de inteiros compreendidos entre 1 e n. Entre estas, tomaremos as r-uplas $M = (m_1 \leq m_2 \leq \cdots \leq m_r)$, cujos elementos estão em ordem *não decrescente* (podendo haver repetições). O número dessas r-uplas é o número de combinações com repetição de n elementos r a r, ou seja, $\binom{n-r+1}{r}$. Para cada r-upla $M = (m_1 \leq m_2 \leq \cdots \leq m_r)$, de inteiros compreendidos entre 1 e n, seja

$$e_{[M]} = S\big(e_{m_1} \otimes \cdots \otimes e_{m_r}\big) = \sum_{\sigma} e_{m_{\sigma(1)}} \otimes \cdots \otimes e_{m_{\sigma(r)}}$$

o tensor simetrizado de $e_{m_1} \otimes \cdots \otimes e_{m_r}$. Verifica-se sem dificuldade que os $\binom{n-r+1}{r}$ tensores $e_{[M]}$ assim obtidos constituem uma base para o subespaço dos tensores simétricos de V_0^r.

Podemos axiomatizar os conceitos acima, introduzindo a seguinte definição:

Chama-se *r-ésima potência simétrica* de um espaço vetorial V a todo par $(S^r(V), \phi)$ com as seguintes propriedades:

1) $S^r(V)$ é um espaço vetorial e $\phi\colon V \times \cdots \times V \to S^r(V)$ é uma aplicação r-linear *simétrica* (definição evidente);

2) $\dim S^r(V) = \binom{n-r+1}{r}$, onde $n = \dim V$;

3) A imagem $\phi(V \times \cdots \times V)$ gera $S^r(V)$.

As condições 2) e 3), em presença de 1), equivalem a

2') Se $\mathcal{E} = \{e_1, \ldots, e_n\}$ é uma base de V, então os elementos da forma $\phi\big(e_{m_1}, e_{m_2}, \ldots, e_{m_r}\big)$, onde $m_1 \leq m_2 \leq \cdots \leq m_r$, formam uma base de $S^r(V)$.

A imagem $\phi(v_1, \ldots, v_r)$ chama-se o *produto simétrico* dos vetores $v_1, \ldots, v_r$. Mais comumente, escreve-se $v_1 \cdot v_2 \cdot \ldots \cdot v_r$ em vez de $\phi(v_1, v_2, \ldots, v_r)$ para indicar o produto simétrico dos vetores $v_1, \ldots, v_r$.

Um exemplo de r-ésima potência simétrica de V é dado pelo espaço $S^r(V)$ dos tensores simétricos r vezes contravariantes sobre V, sendo $\phi(v_1, \ldots, v_r) = S(v_1 \otimes \cdots \otimes v_r) = \sum_\sigma v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(r)}$.

Outro exemplo é o dual do espaço das formas r-lineares simétricas sobre V, sendo $\phi(v_1, \ldots, v_r)(w) = w(v_1, \ldots, v_r)$, onde w é uma forma r-linear simétrica qualquer.

O significado intuitivo de $S^r(V)$ é o que transparece do seguinte exemplo. Se $\dim V = n$, tomamos para $S^r(V)$ o espaço dos polinômios homogêneos de grau r nas n indeterminadas $X_1, \ldots, X_n$. Para definirmos $\phi\colon V \times \cdots \times V \to S^r(V)$, escolhemos uma base $\mathcal{E} = \{e_1, \ldots, e_n\}$ em V. Como ϕ deve ser r-linear, basta definir $\phi\big(e_{i_1}, \ldots, e_{i_r}\big)$ onde $e_{i_1}, \ldots, e_{i_r} \in \mathcal{E}$. Poremos então:

$$\phi\big(e_{i_1}, e_{i_2}, \ldots, e_{i_r}\big) = X_{i_1} X_{i_2} \ldots X_{i_r}.$$

Como a multiplicação entre polinômios é comutativa, vemos que ϕ é simétrica e as demais propriedades são facilmente verificadas.

Quando $V = R^n$, existe uma base canônica, e $S^r(V)$ é de fato o conjunto dos polinômios homogêneos de grau r com n indeterminadas. No caso geral, $S^r(V)$ é um objeto intrínseco, que desempenha o papel dos polinômios homogêneos de grau r sem fazer referência explícita à base dos monômios $X_{i_1} X_{i_2} \ldots X_{i_r}$.

Demonstra-se que, dadas duas potências simétricas $(S^r(V), \phi)$, $(\overline{S}^r(V), \overline{\phi})$ do mesmo espaço V, existe um único isomorfismo $L\colon S^r(V) \to \overline{S}^r(V)$ tal que $\overline{\Phi} = L \circ \Phi$.

Dados r e s, existe uma única aplicação bilinear simétrica

$$S^r(V) \times S^s(V) \to S^{r+s}(V)$$

que leva o par $(u_1 \cdot u_2 \ldots u_r,\ v_1 \cdot v_2 \ldots v_s)$ no produto simétrico $u_1 \cdot u_2 \ldots u_r \cdot v_1 \cdot v_2 \ldots v_s$. Obtém-se então, na soma direta

$$S(V) = \sum_{r=0}^{\infty} S^r(V)$$

uma estrutura de álgebra, chamada a *álgebra simétrica* do espaço vetorial $V \cdot S(V)$ é uma álgebra associativa, comutativa, de dimensão infinita, com unidade. Mais precisamente, $S(V)$ é a *álgebra comutativa livre* gerada por V e 1.

A álgebra simétrica $S(V)$ desempenha o papel de uma álgebra de polinômios, onde os polinômios do primeiro grau são os vetores $v \in V$, e o número de indeterminadas desses polinômios é a dimensão de V. Trata-se porém de "polinômios intrínsecos", onde não foi feita a escolha explícita de uma base de monômios.

Podemos também considerar a álgebra simétrica *covariante* do espaço vetorial V, a qual nada mais é do que

$$S(V^*) = \sum S^r(V^*).$$

A potência simétrica $S^r(V^*)$ pode ser tomada como o espaço das formas r-lineares simétricas sobre V, ou seja, dos tensores simétricos r vezes covariantes. Assim, por exemplo, uma forma bilinear simétrica sobre V é um elemento

de $S^2(V^*)$. Um produto interno em V é um especial tensor simétrico de segunda ordem: deve ser "positivo definido".

Capítulo 4

Formas Diferenciais

4.1 Variedades diferenciáveis

Faremos uma exposição rápida dos conceitos e resultados, a respeito das variedades diferenciáveis, que serão utilizados no decorrer do capítulo.

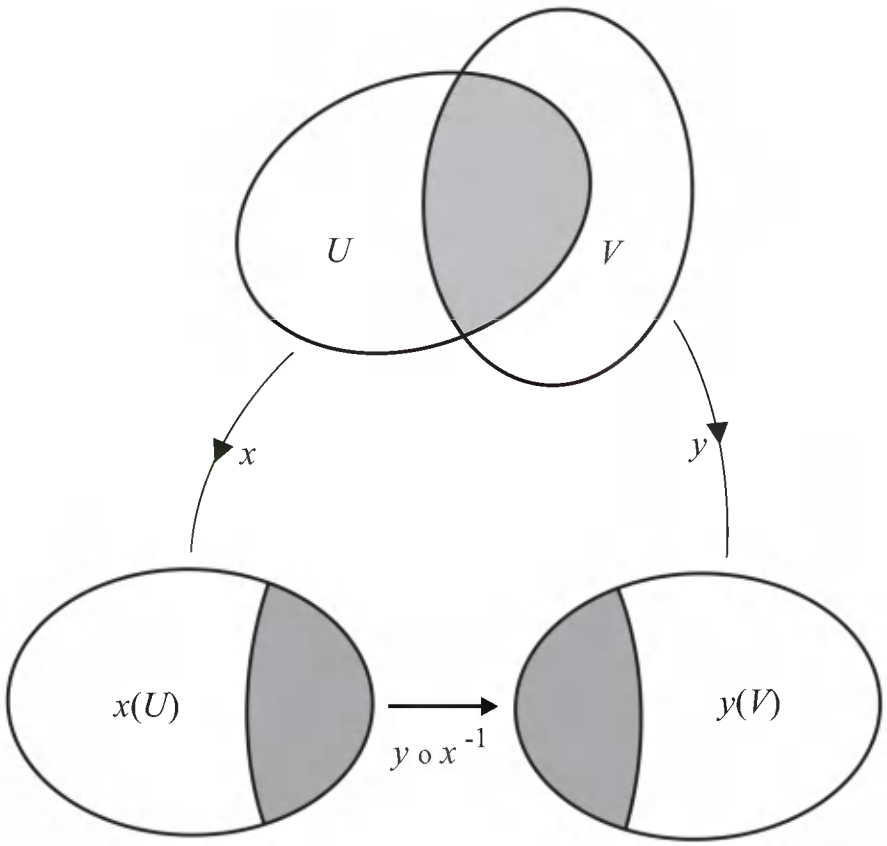
Um *atlas diferenciável*, de classe C^k e dimensão n, sobre um espaço topológico M, é uma coleção $\mathfrak{A}$ de homeomorfismos $x\colon U \to R^n$, do aberto U de M no aberto $x(U)$ do espaço euclidiano R^n, de modo que:

a. os domínios U cobrem M;

b. se $x, y \in \mathfrak{A}$ e $x\colon U \to R^n$, $y\colon V \to R^n$ com $U \cap V \neq \phi$, então a aplicação

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

do aberto $x(U \cap V) \subset R^n$ no aberto $y(U \cap V) \subset R^n$ é diferenciável, de classe C^k.



Os elementos de $\mathfrak{A}$ são chamados *sistemas de coordenadas locais*, os domínios U *vizinhanças coordenadas* e as aplicações diferenciáveis $y \circ x^{-1}$ *mudanças de coordenadas.* Dado um ponto $p \in M$ e um sistema $x \in \mathfrak{A}$, definido numa vizinhança de p, se $x(p) = (x^1, \ldots, x^n)$, os números $x^1, \ldots, x^n$ são as *coordenadas* do ponto p no sistema x.

Um atlas diferenciável $\mathfrak{A}$, sobre M, é dito *máximo* se, além das condições acima, estiver satisfeita ainda esta:

c. se $z\colon W \to R^n$ é um homeomorfismo de um aberto W de M no aberto $z(W)$ do R^n de modo que, para todo sistema de coordenadas $x \in \mathfrak{A}$, $x\colon U \to R^n$, com $U \cap W \neq \phi$, se tenha:

$$z \circ x^{-1}\colon x(U \cap W) \to z(U \cap W)$$

diferenciável de classe C^k, então também $z \in \mathfrak{A}$.

Uma *variedade diferenciável* $M = M^n$, de classe C^k e dimensão n, é um espaço topológico de Hausdorff, com base enumerável, munido de um atlas $\mathfrak{A}$ diferenciável de classe C^k e dimensão n. Por conveniência, suporemos que $\mathfrak{A}$ é máximo.

Dado um ponto $p \in M$, chamemos de $\mathfrak{X}_p$ o conjunto de todos os sistemas de coordenadas $x \in \mathfrak{A}$ definidos em alguma vizinhança de p:

$$\mathfrak{X}_p = \{x \in \mathfrak{A};\ x\colon U \to R^n \ \text{ e } \ p \in U\}.$$

Um *vetor tangente* à variedade M, no ponto $p \in M$, é uma função

$$v\colon \mathfrak{X}_p \to R^n$$

que a cada sistema de coordenadas $x \in \mathfrak{X}_p$ associa uma n-upla $v(x) = (\alpha^1, \ldots, \alpha^n) \in R^n$, de tal modo que, se $y \in \mathfrak{X}_p$ e $v(y) = (\beta^1, \ldots, \beta^n)$, então

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

onde, por $\left(\frac{\partial y^j}{\partial x^i}(p)\right)$ deve-se entender a matriz jacobiana da mudança de coordenadas $y \circ x^{-1}$, calculada no ponto $x(p)$. Os números $\alpha^1, \ldots, \alpha^n$ são as *coordenadas do vetor* v relativamente ao sistema x.

Sobre o conjunto M_p, de todos os vetores tangentes a M no ponto p, pode-se definir uma estrutura de espaço vetorial de modo natural, isto é, se $v, w \in M_p$ e α é um número real qualquer, definimos $v + w$ e $\alpha\, v$ como:

$$(v + w)(x) = v(x) + w(x)$$
$$(\alpha\, v)(x) = \alpha\, v(x)$$

para qualquer $x \in \mathfrak{X}_p$. O vetor nulo $0 \in M_p$ é a função

$$0\colon \mathfrak{X}_p \to R^n$$

tal que $0(x) = (0, \ldots, 0)$ para todo $x \in \mathfrak{X}_p$.

É fácil verificar que $v + w$ e αv são ainda vetores tangentes de M_p e, quanto à dimensão, vale a

Proposição 1. *M_p é um espaço vetorial de dimensão n.*

Demonstração: Basta construir uma base de M_p com n elementos. Na realidade, a cada sistema de coordenadas $x \in \mathfrak{X}_p$ (válido numa vizinhança de p) faremos corresponder uma base

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

do espaço M_p. Dado x, definiremos:

$$\left[\frac{\partial}{\partial x^i}(p)\right](y) = \left(\frac{\partial y^1}{\partial x^i}(p), \ldots, \frac{\partial y^n}{\partial x^i}(p)\right) \quad i = 1, \ldots, n,$$

qualquer que seja $y \in \mathfrak{X}_p$. Isto é, $\frac{\partial}{\partial x^i}(p)$ é o vetor de M_p cujas coordenadas no sistema y são os números

$$\frac{\partial y^1}{\partial x^i}(p), \ldots, \frac{\partial y^n}{\partial x^i}(p).$$

A regra de derivação das funções compostas mostra que, de fato, os $\frac{\partial}{\partial x^i}(p)$ assim definidos são vetores tangentes. Além disso, dado um vetor $v \in M_p$ qualquer, tem-se

$$v(x) = (\alpha^1, \ldots, \alpha^n).$$

Então não é difícil verificar que

$$v = \sum_{i=1}^{n} \alpha^i \frac{\partial}{\partial x^i}(p),$$

e portanto os vetores $\frac{\partial}{\partial x^i}(p)$ geram M_p. Finalmente, esses vetores são linearmente independentes pois se $w \equiv \sum \lambda^i \frac{\partial}{\partial x^i}(p) = 0$ então, em particular $w(x) = (0, 0, \ldots, 0)$. Mas

$$w(x) = \sum \lambda^i \left[\frac{\partial}{\partial x^i}(p)\right](x)$$

e $\left[\frac{\partial}{\partial x^i}(p)\right](x) = e_i = i$-ésimo vetor da base canônica do R^n. Assim, temos

$$0 = w(x) = \sum \lambda^i e_i \in R^n.$$

Segue-se que $\lambda^1 = \cdots = \lambda^n = 0$, o que conclui a demonstração.

Observação: Decorre da demonstração que, se $x\colon U \to R^n$ é um sistema de coordenadas em M^n, cujo domínio é o aberto $U \subset M^n$, então a cada ponto $q \in U$ corresponde a base

$$\left\{\frac{\partial}{\partial x^1}(q), \ldots, \frac{\partial}{\partial x^n}(q)\right\} \subset M_q\,.$$

Quando, num dado argumento, não houver perigo de confusão sobre o ponto q, escreveremos simplesmente

$$\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\}.$$

Exemplos:

1. Os espaços euclidianos R^n são variedades de classe C^k, para qualquer $k \geq 0$. Bastando para isso, considerar o atlas máximo diferenciável $\mathfrak{A}_k$, de classe C^k, que contém o sistema de coordenadas

$$x = \text{ identidade } : R^n \to R^n.$$

Verifica-se que o espaço R^n_p dos vetores tangentes a esta variedade, em um ponto $p \in R^n$, é canonicamente isomorfo ao R^n. Este isomorfismo canônico decorre da existência de uma base privilegiada de R^n_p, aquela associada ao sistema de coordenadas dado pela identidade sobre o R^n.

2. Todo subconjunto aberto A de uma variedade diferenciável M^n possui uma estrutura de variedade diferenciável, "herdada" da estrutura de M^n, de mesma classe e dimensão que M^n. Um atlas sobre A será dado pelos sistemas de coordenadas, de M^n, cujos domínios estejam contidos em A.

Verifica-se que, em cada ponto $p \in A$, o espaço M_p dos vetores tangentes a M em p, é canonicamente isomorfo ao espaço A_p dos vetores tangentes a A em p.

3. Definimos *variedade produto* de duas variedades $M = M^m$ e $N = N^n$ como sendo o produto $M \times N$ dos espaços topológicos dotado da seguinte estrutura de variedade diferenciável: se $\mathfrak{A}$ e $\mathcal{B}$ são os atlas sobre M e N, respectivamente, construímos, a partir destes, o atlas $\mathfrak{A} \times \mathcal{B}$, sobre $M \times N$, cujos elementos serão homeomorfismos

$$x \times y \colon U \times V \to R^{m+n},$$

com $x \colon U \to R^m$ percorrendo $\mathfrak{A}$, $y \colon V \to R^n$ percorrendo $\mathcal{B}$ e $(x \times y)(p, q) = (x(p), y(q))$.

O espaço $(M \times N)_{(p,q)}$ tangente à variedade produto, no ponto (p, q), é isomorfo, canonicamente, à soma direta dos espaços tangentes M_p e N_q. Com efeito, dados os sistemas de coordenadas x em M e y em N, em torno dos pontos p e q, respectivamente, sejam $\alpha^1, \ldots, \alpha^m$, $\beta^1, \ldots, \beta^n$ as coordenadas do vetor $v \in (M \times N)_{(p,q)}$ no sistema $x \times y$.

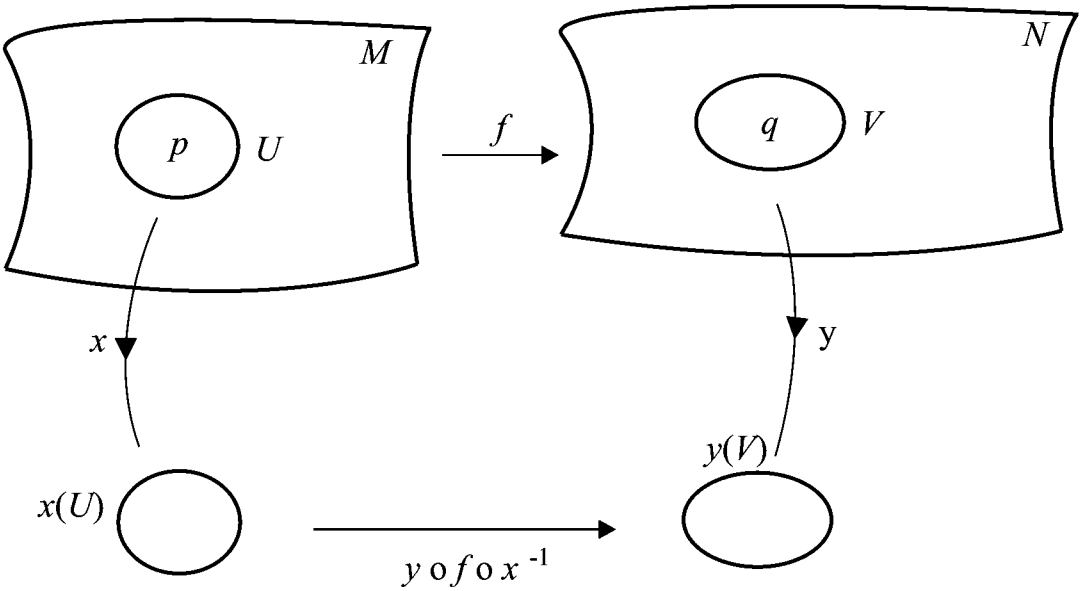
Vê-se, sem dificuldade, que as primeiras m coordenadas $\alpha^1, \ldots, \alpha^m$, de y, dependem somente do sistema x e as últimas coordenadas $\beta^1, \ldots, \beta^n$ dependem somente de y. Seja $v' \in M_p$ o vetor cujas coordenadas no sistema x são $\alpha^1, \ldots, \alpha^m$ e $v'' \in N_q$ o vetor cujas coordenadas no sistema y são $\beta^1, \ldots, \beta^n$. A correspondência $v \to v' \oplus v''$ estabelece o isomorfismo

$$(M \times N)_{(p,q)} \approx M_p \oplus N_q .$$

4. Um outro exemplo de variedade diferenciável é fornecido pelas superfícies regulares do espaço euclidiano, onde se consideram os inversos das parametrizações como sistemas de coordenadas.

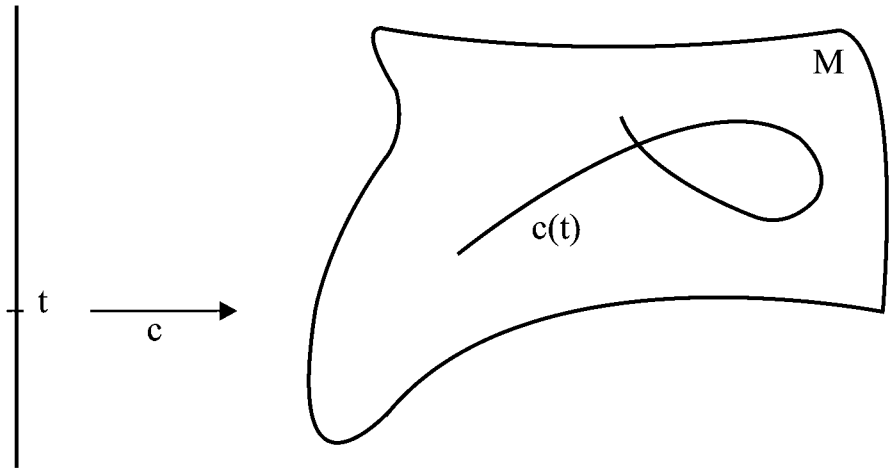
4.2 Aplicações diferenciáveis

Sejam $M = M^m$ e $N = N^n$ variedades diferenciáveis, de classe C^k, e $f\colon M \to N$ uma aplicação de M em N. Dizemos que f é uma *aplicação diferenciável*, de classe C^k, no ponto $p \in M$, se, dado um sistema de coordenadas y em N, definido numa vizinhança V de $q = f(p)$, $y\colon V \to R^n$, existe um sistema de coordenadas x, sobre M, em torno de p, $x\colon U \to R^m$, tal que $f(U) \subset V$ e a aplicação $y \circ f \circ x^{-1}$ do aberto $x(U) \subset R^m$ no aberto $y(V) \subset R^n$ seja diferenciável.



Se f for diferenciável em todos os pontos de M ela será dita diferenciável em M.

Uma *curva parametrizada* em $M = M^n$ é uma aplicação $C\colon I \subset R \to M$, de um intervalo aberto da reta real na variedade M. Se C for diferenciável, a curva é dita *diferenciável.*



As curvas parametrizadas diferenciáveis dão origem a vetores tangentes. Sejam $a \in I$ e $C(a) = p$; à curva C e ao

valor $a \in I$ está associado o vetor de M_p, que indicamos por $C'(a)$ que atua sobre $x \in \mathfrak{X}_p$ assim:

$$C'(a)(x) = \left(\frac{dx^1}{dt}(a), \ldots, \frac{dx^n}{dt}(a) \right),$$

onde $x^i(t) = x^i(C(t))$.

Diz-se que $v = C'(a)$ é o vetor tangente à curva $C(t)$ no ponto $p = C(a)$.

Neste novo sentido, cada vetor básico $\frac{\partial}{\partial x^i}(p)$, associado a um sistema $x \in \mathfrak{X}_p$, é o vetor tangente, no ponto p, à i-ésima "curva coordenada" de x:

$$C_i(t) = x^{-1}(x(p) + t \cdot e_i)$$

onde $e_1{=}(1, 0, \ldots, 0), \ldots, e_n{=}(0, 0, \ldots, 0, 1)$ e $t{\in}(-\varepsilon, +\varepsilon)$, com $\varepsilon > 0$ convenientemente escolhido de modo que $C_i(t)$ esteja bem definida.

Sejam p um ponto de M, v um vetor tangente a M em p e $f\colon M \to R$ uma função real, diferenciável, definida em M. Define-se *derivada direcional da função f, no ponto p, relativamente ao vetor v* como sendo o número $\frac{\partial f}{\partial v}(p) = \frac{d(f \circ C)}{dt}(0)$, onde C é uma curva parametrizada tal que $C(0) = p$ e $C'(0) = v$.

A fim de que este número fique bem determinado é necessário verificar dois fatos: o primeiro é que quaisquer que sejam o ponto $p \in M$ e o vetor $v \in M_p$, existe sempre uma tal curva C e o segundo é que a derivada direcional não depende da particular curva C considerada.

Se $x \in \mathfrak{X}_p$ é um sistema de coordenadas em torno do ponto p, $x\colon U \to R^n$, com $p \in U$, pode-se construir a curva $C\colon I \to M$ cujo vetor tangente, em p, é v. Se

$\alpha^1, \ldots, \alpha^n$ são as coordenadas de v no sistema x, isto é, $v(x) = (\alpha^1, \ldots, \alpha^n)$, basta considerar $\varepsilon > 0$ suficientemente pequeno de modo que o ponto

$$x(p) + t(\alpha^1, \ldots, \alpha^n) \in x(U),$$

para qualquer $t \in (-\varepsilon, \varepsilon)$ e definir a curva C como:

$$C(t) = x^{-1}(x(p) + t(\alpha^1, \ldots, \alpha^n)).$$

A segunda objeção estará sanada quando fizermos o cálculo efetivo da derivada $\frac{\partial f}{\partial v}(p)$. Com efeito, a regra de derivação das funções compostas dá-nos:

$$\frac{\partial f}{\partial v}(p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p) \cdot \alpha^i,$$

onde

$$\frac{\partial f}{\partial x^i}(p) = \frac{\partial (f \circ x^{-1})}{\partial x^i}(x(p))$$

e $x^1, \ldots, x^n$ são as componentes de x.

Esta expressão, que independe da particular curva C considerada, mostra-nos também que a derivada direcional é linear em relação ao vetor v.

A forma $df_p \colon M_p \to R$, que a cada vetor $v \in M_p$ faz corresponder a derivada direcional de f, em p, relativamente a v, é linear, como acabamos de observar, e é chamada *diferencial da função* f no ponto p, isto é, df_p é um elemento do dual de M_p, precisamente aquele que atua assim:

$$df_p(v) = \frac{\partial f}{\partial v}(p).$$

Para considerarmos a diferencial da função f, no ponto p, basta que f esteja definida numa vizinhança de p.

Em particular, se $x \in \mathfrak{X}_p$ é um sistema de coordenadas definido numa vizinhança U de p e

$$x(p) = (x^1(p), \ldots, x^n(p))$$

não é difícil verificar que as diferenciais $dx_p^1, \ldots, dx_p^n$ das n funções reais (componentes de x)

$$x^i \colon U \to R$$

constituem uma base de $(M_p)^*$, justamente a base dual da base de M_p,

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

associada ao sistema x. Aqui, também, escreveremos somente $dx^1, \ldots, dx^n$ quando não houver perigo de confusão a respeito do ponto onde as diferenciais estão sendo tomadas.

Além disto, em termos da base $dx_p^1, \ldots, dx_p^n$, a diferencial da função $f \colon M \to R$ escreve-se como:

$$df_p = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p)\, dx_p^i,$$

onde, por $\frac{\partial f}{\partial x^i}(p)$ entende-se $\frac{\partial (f \circ x^{-1})}{\partial x^i}(x(p))$.

Consideremos, agora, uma aplicação diferenciável $f \colon M \to N$ de uma variedade $M = M^m$ numa outra $N = N^n$. Em cada ponto p da variedade M, f induz uma aplicação f_* do espaço tangente M_p no espaço tangente N_q, onde $q = f(p)$, do seguinte modo: se $v \in M_p$ é

um vetor tangente à variedade M, no ponto p, seja C uma curva parametrizada diferenciável, em M, tal que $C(0) = p$ e $C'(0) = v$. O vetor $f_*(v)$ será o vetor tangente à curva de N, $f \circ C$, no ponto $q = f(p)$, isto é:

$$f_*(v) = (f \circ C)'(0).$$

Sendo $x\colon U \to R^m$ e $y\colon V \to R^n$ sistemas de coordenadas em torno de p e q, respectivamente, e

$$v = \sum_{j=1}^{m} \alpha^j \frac{\partial}{\partial x^j}(p)$$

então, pela regra de derivação as funções compostas,

$$f_*(v) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \frac{\partial y^i}{\partial x^j}(p)\, \alpha^j \right) \frac{\partial}{\partial y^i}\,,$$

onde $\frac{\partial y^i}{\partial x^j}(p)$ é uma notação simplificada para

$$\frac{\partial (y^i \circ f \circ x^{-1})}{\partial x^j}\,(x(p)).$$

Isto significa que a definição de f_* é independente da curva C e que f_* é uma aplicação linear em v cuja matriz $\left[f_* \,;\, \frac{\partial}{\partial x^j} \,;\, \frac{\partial}{\partial y^i} \right]$, relativamente às bases $\left\{ \frac{\partial}{\partial x^j} \right\}$ e $\left\{ \frac{\partial}{\partial y^i} \right\}$, é $\left(\frac{\partial y^i}{\partial x^j}(p) \right)$. A aplicação f é dita *aplicação linear induzida* pela f, no ponto p.

A adjunta (ou transposta) de f_* será a aplicação $f^*\colon N_q^* \to M_p^*$, do dual de N_q no dual de M_p, tal que, se $\omega \in N_q^*$ é uma forma linear sobre N_q e $v \in M_p$ é um vetor tangente a M no ponto p, então

$$f^*\,\omega(v) = \omega(f_*\, v).$$

Em termos dos sistemas de coordenadas x e y, acima considerados, se

$$\omega = \sum_{i=1}^{n} a_i \, dy^i \quad \text{e} \quad f^* \omega = \sum_{i=1}^{m} b_i \, dx^i,$$

então

$$b_i = f^*\omega \left(\frac{\partial}{\partial x^i} \right) = \omega \left(f_* \frac{\partial}{\partial x^i} \right) = \omega \left(\sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \right) =$$

$$= \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i} \, \omega \left(\frac{\partial}{\partial y^j} \right) = \sum_{j=1}^{n} a_j \frac{\partial y^j}{\partial x^i},$$

onde $\frac{\partial y^j}{\partial x^i}$ tem o sentido já mencionado e deve ser calculado em $x(p)$.

Sendo assim, $f^* \omega = \sum_{i=1}^{m} a_i \left(\sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j} \, dx^j \right)$.

Passando às potências exteriores dos espaços vetoriais M_p, N_q e de seus duais M_p^*, N_q^*, podemos considerar outras aplicações lineares induzidas pela f.

De acordo com a notação introduzida no capítulo anterior, $\overset{k}{\wedge} f^*$ seria uma aplicação entre k-formas. Neste capítulo, entretanto, vamos usar a mesma notação f^* para indicar qualquer uma destas aplicações:

$$f^* : \overset{k}{\wedge} N_q^* \to \overset{k}{\wedge} M_p^*$$

isto é, se ω é uma forma k-linear alternada sobre N_q, então $f^* \omega$ será uma forma k-linear alternada sobre M_p.

Em termos dos mesmos sistemas x e y, se

$$\omega = \sum_{j_1<\cdots<j_k} a_{j_1\ldots j_k}\, dy^{j_1} \wedge \cdots \wedge dy^{j_k}.$$

a expressão par $f^*\,\omega$ será

$$f^*\,\omega = \sum_{\substack{j_1<\cdots<j_k \\ r_1<\cdots<r_k}} a_{j_1\ldots j_k}\, \frac{\partial(y^{j_1}\ldots y^{j_k})}{\partial(x^{r_1}\ldots x^{r_k})}\, dx^{r_1} \wedge \cdots \wedge dx^{r_k},$$

onde

$$\frac{\partial(y^{j_1}\ldots y^{j_k})}{\partial(x^{r_1}\ldots x^{r_k})}$$

é o menor jacobiano das componentes aí indicadas da aplicação

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

calculado no ponto $x(p)$.

O ponto $p \in M$ é dito *ponto regular* de f se a transformação linear induzida

$$f_*\colon M_p \to N_q \quad (q = f(p))$$

é biunívoca. Se existir algum ponto regular, então $m \leq n$.

Por outro lado, um ponto $q \in N$ é um *valor regular* de f se, para todo ponto $p \in f^{-1}(q)$, a aplicação linear $f_*\colon M_p \to N_q$ é *sobre* N_q. Em particular, se $f^{-1}(q) = \varnothing$, q é valor regular. Analogamente, da existência de um valor regular q, com $f^{-1}(q) \neq \varnothing$, segue-se $m \geq n$.

Uma aplicação f é *regular* quando todo ponto $p \in M$ é regular. E f é uma *fibração* quando for uma aplicação sobre $N (f(M) = N)$ e todo ponto $q \in N$ for valor regular.

Um homeomorfismo regular é chamado *imersão.*

Exemplos:

1. A aplicação $f\colon R^m \to R^{m+k}$, que consiste em levar o ponto $(x^1, \ldots, x^m) \in R^m$ no ponto $(x^1, \ldots, x_m, 0, \ldots, 0) \in R^{m+k}$ é uma aplicação regular.

Além disto, toda aplicação regular f, localmente, é deste tipo. Isto é, dados p e $q = f(p)$, existem sistemas x e y, em torno de p e q, respectivamente, tal que $y \circ f \circ x^{-1}$ seja deste tipo (cfr. o Teorema do Posto).

2. A projeção $\pi\colon R^{m+k} \to R^m$, que consiste em desprezar as últimas k coordenadas do ponto, isto é,

$$\pi(x^1, \ldots, x^m, \ldots, x^{m+k}) = (x^1, \ldots, x^m),$$

é uma fibração. Como no caso anterior, observa-se que, localmente, toda fibração é deste tipo. (Vide loc. cit.)

3. A aplicação $f\colon R \to R^2$, definida como:

$$f(t) = (\cos t, \operatorname{sen} t)$$

é um exemplo de aplicação regular, e então localmente biunívoca, que não é imersão. A biunivocidade de f só se verifica quando se consideram intervalos de comprimento estritamente menor que 2π.

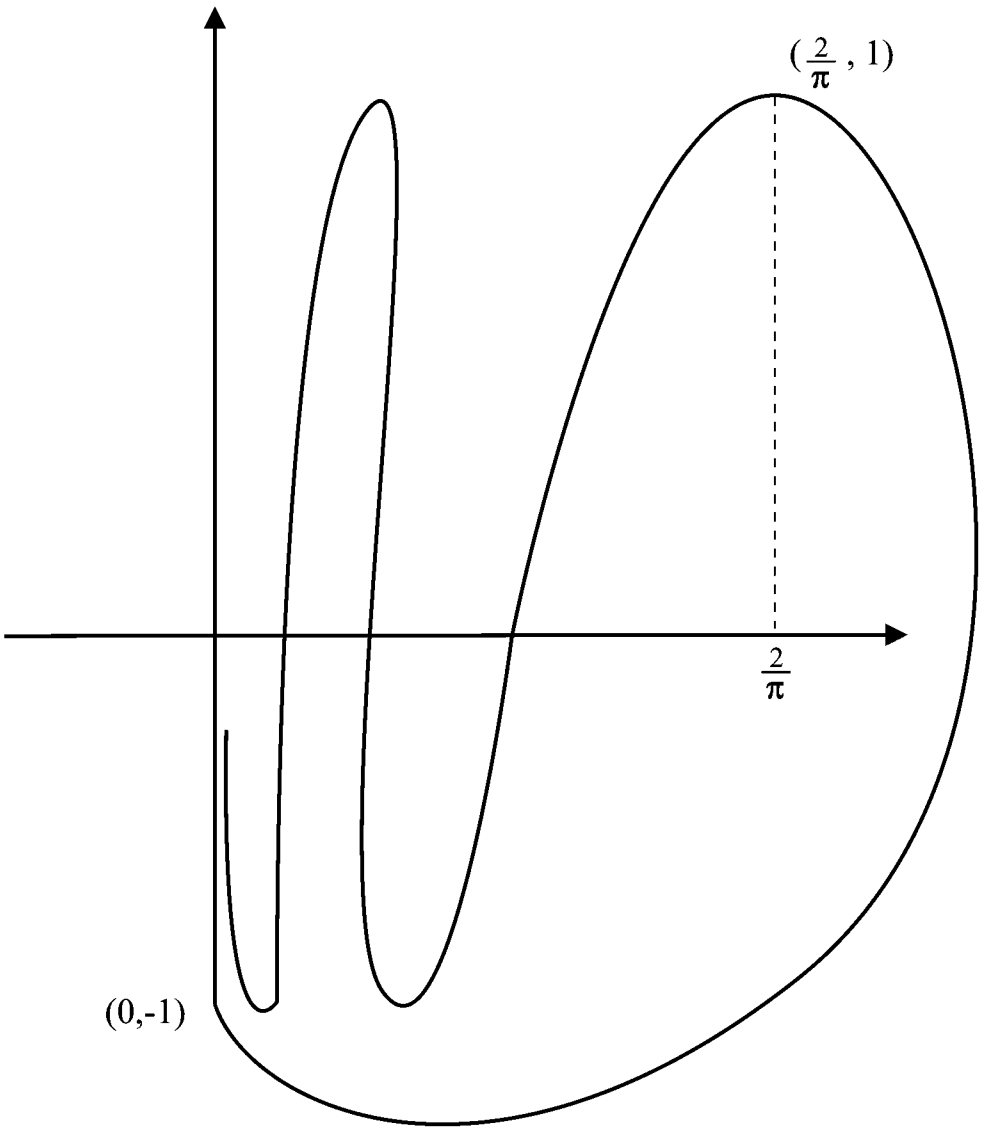
4. Existem, ainda, homeomorfismos – aplicação biunívocas, portanto – que não são regulares. Como exemplo, consideramos o homeomorfismo $f\colon R \to R$ que a cada $t \in R$ associa seu cubo t^3. A origem, $t = 0$, não é ponto regular de f.

5. Uma aplicação regular, mesmo sendo biunívoca, pode não ser uma imersão. Daremos, agora, um exemplo de aplicação regular e biunívoca que não é homeomorfismo.

Definimos $f\colon (0,\infty) \to R^2$ do seguinte modo:

se $\quad 0 < t \leq \dfrac{2}{\pi}, \quad f(t) = \left(t, \operatorname{sen} \dfrac{1}{t}\right)$

se $\quad 1 \leq t < \infty, \quad f(t) = (0, t-2)$



se $\dfrac{2}{\pi} < t < 1, \quad f(t)$ é uma curva simples, parametrizada, regular, ligando os pontos $(\frac{2}{\pi}, 1)$ e $(0, -1)$ sem tocar nos outros pontos $(t, f(t))$ já definidos e "concordando" com

os arcos já existentes de modo a tornar f uma aplicação regular em toda a semireta $(0, +\infty)$.

Os espaços $(0, +\infty)$ e $f(0, +\infty)$ não são homeomorfos, pois este último não é localmente conexo nos pontos da forma $(0, y)$, com $-1 \leq y \leq +1$.

Este fato não se dá quando a variedade M, onde f é definida, for compacta, pois então, se f é regular e biunívoca, será, conseqüentemente, um homeomorfismo.

4.3 Subvariedades

Um subconjunto $S^s \subset N^n$ de uma variedade diferenciável N^n é uma *subvariedade* de N^n, de dimensão s, se:

a. S^s possui uma estrutura de variedade diferenciável de dimensão s.

b. A aplicação de inclusão, $i\colon S^s \to N^n$, é uma imersão.

Em particular, a inclusão i é um homeomorfismo e, então, a topologia de S^s é a topologia induzida pela de N^n.

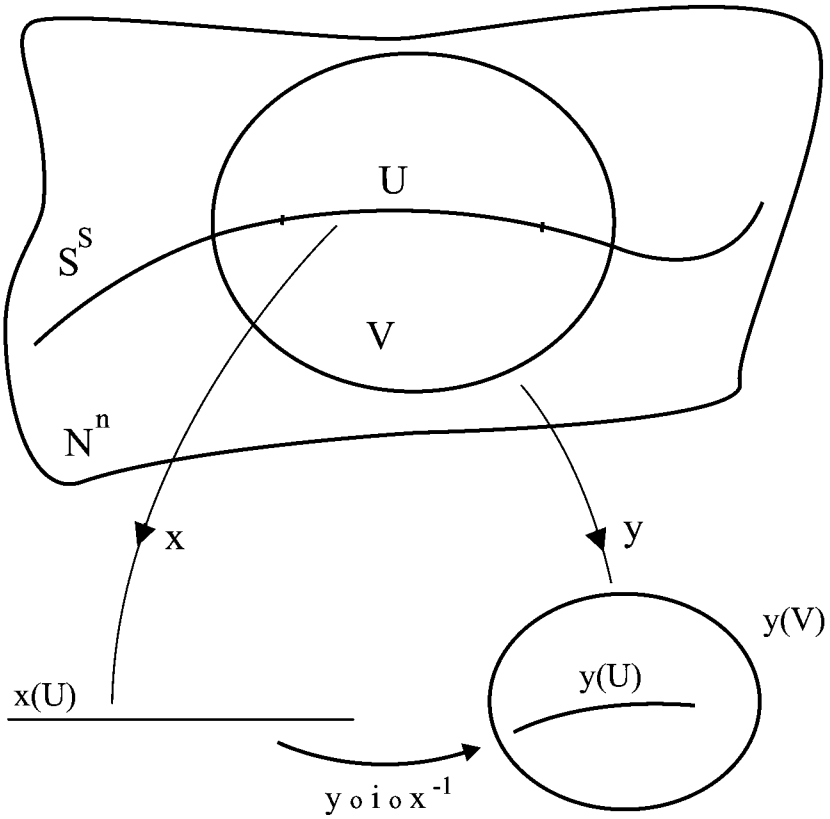
A inclusão é, por hipótese, também diferenciável. Isto significa que, se y é um sistema de coordenadas em M^n, com domínio V, existe, em correspondência, um sistema x de coordenadas, em S^s, com domínio $U \subset S^s \cap V$, de modo que a aplicação

$$y \circ i \circ x^{-1}\colon x(U \to y(v))$$

é diferenciável. Além disso, i é regular, isto é, a matriz

$$\left(\frac{\partial y^i}{\partial x^j}\right)$$

da aplicação $y \circ i \circ x^{-1}$ tem posto s em todos os pontos de $x(U)$.



Ora, estes dois fatos dizem-nos que a aplicação $y \circ i \circ x^{-1}$ é uma parametrização regular para $y(U)$.

Grosseiramente falando, portanto, podemos dizer que uma subvariedade é, localmente, uma superfície regular.

Observações:

1. A classe de diferenciabilidade da subvariedade S^s é a classe da aplicação de inclusão podendo, portanto, ser estritamente menor que a de N^n.

2. A literatura não é uniforme na definição de subvariedade. É necessário, em alguns casos, considerar outras

topologias, e não a induzida, na subvariedade S^s. Alguns autores consideram, então, um subconjunto $S^s \subset N^n$ como subvariedade quando a inclusão, $i\colon S^s \to N^n$, é uma aplicação regular e biunívoca – não necessariamente um homeomorfismo. Por coerência, tais autores definem imersão como uma aplicação $f\colon M \to N$, regular e biunívoca. No caso da variedade M ser compacta, estas definições coincidem com aquelas dadas anteriormente.

Exemplos:

1. Sejam M^m e N^n variedades diferenciáveis e $M \times N$ a variedade produto. Então os subespaços $p \times N^n$ e $M^m \times q$ são subvariedades do produto, para cada $p \in M^m$ e cada $q \in N^n$, respectivamente.

Verifica-se que um vetor v, tangente a $M \times N$ no ponto (p, q), pode ser decomposto, de um único modo, na soma de outros dois v_1, v_2, onde v_1 é tangente à subvariedade $M^m \times q$ e v_2 tangente a $p \times N^n$, ambos no ponto (p, q).

2. Sejam $f\colon M^m \to N^n$ uma aplicação diferenciável da variedade M^m na variedade N^n e $q \in N^n$ um valor regular de f. Então $f^{-1}(q)$ – se não for vazio – é uma subvariedade de M^m, de dimensão $m - n$.

A demonstração deste fato, que se reduz ao caso análogo, em espaços euclidianos, quando se consideram sistemas de coordenadas, é deixada a cargo do leitor. Este resultado é freqüentemente usado para demonstrar que certos subconjuntos do espaço euclidiano são variedades diferenciáveis. Consideremos a função $f\colon R^{n+1} \to R$ definida por

$$f(x^1, \ldots, x^{n+1}) = x^1 x^1 + \cdots + x^{n+1}\, x^{n+1}$$

ou seja, $f(x) = |x|^2$, $x \in R^{n+1}$. É claro que f é uma função real diferenciável. Temos ainda $f^{-1}(1) = \left\{x \in R^{n+1};\ |x| = 1\right\} = S^n$ = esfera unitária de dimensão n. Concluiremos que S^n é uma variedade de classe C^∞ (subvariedade de R^{n+1}) se mostrarmos que 1 é um valor regular de f. Ora, nos sistemas de coordenadas canônicos do R^{n+1} e R, a matriz de $f_*\colon R^{n+1} \to R$ é

$$\left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^{n+1}}\right) = (2x^1, \ldots, 2x^{n+1}).$$

Tal matriz só é nula no ponto $x = (0, \ldots, 0)$. Como $0 \notin f^{-1}(1)$, vemos que $f_* \neq 0$ em todos os pontos de $f^{-1}(1)$. Sendo $\dim R = 1$, isto basta para que f seja sobre R em todos os pontos de $f^{-1}(1)$, ou seja, 1 é valor regular de f.

Mais geralmente, se $f\colon M^n \to R$ é uma função real diferenciável definida em M^n e $a \in R$ é um valor regular de f, então a subvariedade $S^{n-1} = f^{-1}(a) \subset M^n$ chama-se uma *superfície de nível* da função f.

4.4 Campos de tensores sobre variedades

Um *campo de vetores* sobre uma variedade diferenciável M^n, de classe C^k, é uma função v que a cada ponto $p \in M^n$ faz corresponder um vetor $v_p \in M_p$ tangente à variedade nesse ponto:

$$v\colon p \in M^n \to v_p \in M_p\,.$$

Se $x \in \mathfrak{X}_p$ é um sistema de coordenadas definido numa

vizinhança U de p, então

$$v_p = \sum_{i=1}^{n} \alpha^i(p) \cdot \frac{\partial}{\partial x^i} \cdot$$

Quando se toma outro sistema de coordenadas, as n funções $\alpha^i(p)$ transformam-se de modo contravariante.

A definição clássica de campo de vetores focaliza este aspecto: ela define um campo de vetores como uma aplicação que a cada sistema de coordenadas $x\colon U \to R^n$ faz corresponder n funções, $\alpha^1, \ldots, \alpha^n\colon U \to R$ de modo que, se $y\colon V \to R^n$ é outro sistema em M^n, com $U \cap V \neq \phi$ e $v(y) = (\beta^1, \ldots, \beta^n)$, então, em cada ponto $p \in U \cap V$,

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

onde $\frac{\partial y^i}{\partial x^j}(p)$ tem significado já esclarecido. Como se vê, ela é essencialmente a definição que demos.

O campo de vetores v será contínuo ou diferenciável (de classe, no máximo, C^{k-1}) conforme o sejam as função α^i.

Se, ao invés de vetores, consideramos um campo de formas lineares, isto é, uma correspondência ω que a cada ponto $p \in M$ associa uma forma linear ω_p sobre o espaço vetorial tangente M_p, teremos o que chamamos de *forma diferncial* sobre M. Retomando o sistema de coordenadas $x \in \mathfrak{X}_p$,

$$\omega_p = \sum_{i=1}^{n} \alpha_i(p)\, dx^i$$

e os coeficientes $\alpha_i(p)$ mudam de modo covariante, isto é, se $y \in \mathfrak{X}_p$ e

$$\omega_p = \sum_{j=1}^{n} \beta_j(p)\, dy^j,$$

então

$$\beta_j(p) = \sum_{i=1}^{n} \alpha_i(p)\, \frac{\partial x^i}{\partial y^j}(p).$$

Observações análogas às que foram feitas para campo de vetores, relativas à definição clássica, têm lugar também aqui. Do mesmo modo, a forma diferencial será contínua ou diferenciável (de classe, no máximo, C^{k-1}) de acordo com as funções $\alpha_i(p)$.

Podemos considerar, de modo geral, um *campo t de tensores* do tipo (r, s), isto é, r vezes contravariante e s vezes covariante:

$$t\colon p \in M \to t_p \in \left(M_p\right)_s^r = \underbrace{M_p \otimes \ldots \otimes M_p}_{r} \otimes \underbrace{M_p^* \otimes \ldots \otimes M_p^*}_{s}.$$

Em termos do sistema de coordenadas $x \in \mathfrak{X}_p$,

$$t_p = \sum_{\substack{i_1 < \cdots < i_r \\ j_1 < \cdots < j_s}} \xi_{j_1 \ldots j_s}^{i_1 \ldots i_r}(p) \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}.$$

Novamente, o campo t diz-se contínuo ou diferenciável se os coeficientes $\xi_{j_1 \ldots j_s}^{i_1 \ldots i_r}(p)$ forem contínuos ou diferenciáveis.

Exemplos:

1. Sejam M^n uma variedade diferenciável e $f\colon M^n \to R$ uma função real diferenciável, definida em M^n. A correspondência que a cada ponto p associa a forma linear df_p é

um campo de formas diferenciais, df, ao qual se dá o nome de *diferencial da função* f.

2. Entre os campos de tensores, os mais importantes são os campos de tensores antissimétricos covariantes, isto é, os campos de formas diferenciais exteriores.

Uma *forma diferencial exterior*, de grau k, sobre a variedade diferenciável M, é, portanto, uma correspondência que a cada ponto p associa a forma ω_p, k-linear alternada,

$$\omega\colon p \in M \to \omega_p \in \overset{k}{\wedge}(M_p)^*.$$

A expressão "exterior" é, freqüentemente, omitida. Em relação ao sistema de coordenadas $x \in \mathfrak{X}_p$,

$$\omega_p = \sum_{j_1<\cdots<j_k} a_{j_1\ldots j_k}(p)\, dx^{j_1} \wedge \cdots \wedge dx^{j_k}$$

ou, usando a notação abreviada introduzida em capítulo anterior, onde escrevemos $J = \{j_1 < \cdots < j_k\}$, poremos também $dx^J = dx^{j_1} \wedge \cdots \wedge dx^{j_k}$ e teremos

$$\omega_p = \sum_J a_J(p)\, dx^J.$$

Se os coeficientes $a_J(p)$ forem diferenciáveis (contínuos), a forma ω será dita diferenciável (contínua).

Quando $k = 1$, ω será, simplesmente, uma forma diferencial. Definimos forma diferencial exterior de grau 0 como sendo uma função real qualquer, $f\colon M \to R$.

As definições de soma, produto por escalar e produto exterior estendem-se às formas diferenciais, sobre M, de

modo natural, isto é, em cada ponto $p \in M$,

$$\begin{aligned}(\omega + \overline{\omega})_p &= \omega_p + \overline{\omega}_p \\ (\alpha\omega)_p &= \alpha\omega_p \\ (\omega \wedge \overline{\omega})_p &= \omega_p \wedge \overline{\omega}_p\,;\end{aligned}$$

3. No plano R^2, as formas diferenciais exteriores de grau 0, 1 e 2 serão, respectivamente, dos seguintes tipos:

$$\begin{aligned}\omega^0 &= f(x,y) \\ \omega^1 &= a(x,y)dx + b(x,y)dy \\ \omega^2 &= c(x,y)dx \wedge dy.\end{aligned}$$

No espaço R^3, temos formas de grau 0, 1, 2 e 3, a saber:

$$\begin{aligned}\omega^0 &= f(x,y,z) \\ \omega^1 &= a(x,y,z)dx + b(x,y,z)dy + c(x,y,z)dz \\ \omega^2 &= m(x,y,z)dy \wedge dz + n(x,y,z)dz \wedge dx + p(x,y,z)dx \wedge dy \\ \omega^3 &= g(x,y,z)dx \wedge dy \wedge dz.\end{aligned}$$

Observação: Daqui por diante, deixaremos de usar o sinal $\wedge$, de produto exterior, sempre que não houver perigo de confusão.

4.5 Variedades riemannianas

Uma variedade diferenciável M diz-se uma *variedade riemanniana* quando o espaço vetorial tangente M_p, em cada ponto $p \in M$, possui um produto interno indicado com $u \cdot v$ que varia diferenciavelmente (ou continuamente) com

o ponto. Isto significa que, se x é um sistema de coordenadas, em M, definido em U, as funções $g_{ij}: U \to R$ dadas por

$$g_{ij}(p) = \frac{\partial}{\partial x^i}(p) \cdot \frac{\partial}{\partial x^j}(p), \quad \forall\, p \in U$$

são diferenciáveis (ou contínuas).

A matriz $(g_{ij}(p))$ é definida positiva – simétrica com todos os menores principais positivos – e, se $u, v \in M_p$,

$$u = \sum_{i=1}^{n} \alpha^i \frac{\partial}{\partial x^i} \quad \text{e} \quad v = \sum_{j=1}^{n} \beta^j \frac{\partial}{\partial x^j},$$

então

$$u \cdot v = \sum_{i,j=1}^{n} \alpha^i \beta^i \, g_{ij}(p).$$

O produto interno é um exemplo de campo de tensores duas vezes covariante sobre M.

Demonstraremos, mais adiante, que, dada uma variedade diferenciável de classe C^k, é sempre possível considerar sobre ela uma estrutura de variedade riemanniana, dada por uma métrica de classe C^{k-1}.

Numa variedade riemanniana M, pode-se definir comprimento de arco: seja $C: [a, b] \subset M \to M$ um arco de curva parametrizada diferenciável sobre M, define-se o comprimento $\ell(C)$, de C, por:

$$\ell(C) = \int_a^b |C'(t)|\, dt.$$

Dados dois pontos em M, define-se, também, a distância entre eles como sendo o extremo inferior dos comprimentos dos arcos que os unem. Verifica-se que esta métrica reproduz a topologia de M.

Em cada ponto p da variedade riemanniana M, existe um isomorfismo, $J_p\colon M_p \approx M_p^*$, do espaço tangente M_p no seu dual M_p^*. Esses isomorfismos J_p levam um campo de vetores v sobre M numa forma diferencial ω sobre M, isto é, em cada ponto $p \in M$,

$$\omega_p = J_p(v_p) \quad \text{ou} \quad v_p = J_p^{-1}(\omega_p).$$

Considerando um sistema de coordenadas $x\colon U \to R^n$, definido numa vizinhança de p, se

$$v_p = \sum_i \alpha^i(p)\,\frac{\partial}{\partial x^i}$$

e

$$\omega_p = \sum_i \alpha_i(p)\,dx^i$$

então

$$\alpha_i(p) = \sum_j g_{ij}(p)\,\alpha^j(p)$$

e

$$\alpha^i(p) = \sum_j g^{ij}(p)\,\alpha_j(p),$$

onde $(g^{ij}(p))$ é a matriz inversa de $(g_{ij}(p))$.

Já vimos que uma função diferenciáel $f\colon M \to R$, definida sobre uma variedade diferenciável M, determina uma forma diferencial df tal que:

$$df\colon p \in M \to df_p \in M_p^*.$$

Se M é riemanniana, à forma df corresponde um campo de vetores que chamaremos de *gradiente de* f e indicaremos com ∇f.

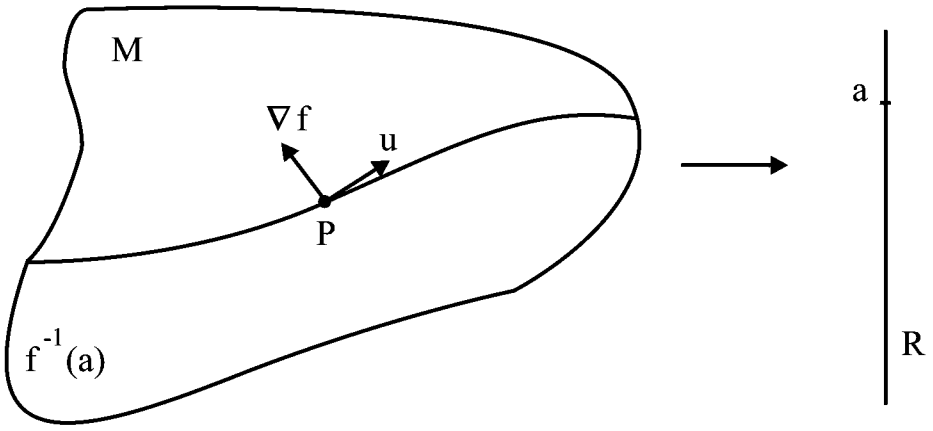
Lembrando a atuação do isomorfismo J_p, vemos que, se $u \in M_p$ e $\nabla f_p = \nabla f(p)$,

$$\nabla f_p \cdot u = df_p(u),$$

isto é,

$$\nabla f_p \cdot u = \frac{\partial f}{\partial u}(p).$$

Se $a \in R$ é um valor regular de f e $f^{-}(a) \neq \varnothing$, em cada ponto $p \in f^{-1}(a)$, o vetor gradiente de f, ∇f_p, é perpendicular à superfície de nível $f^{-1}(a)$, pois se u é tangente a $f^{-1}(a)$, no ponto p, então u é tangente a uma curva $C(t)$ contida em $f^{-1}(a)$, ao longo da qual f é constante.



No cálculo da derivada direcional, teremos então, $\frac{\partial f}{\partial u}(p) = \frac{d}{dt}(f \circ c) = 0$. Logo:

$$\nabla f_p \cdot u = \frac{\partial f}{\partial u}(p) = 0.$$

4.6 Diferencial exterior

As formas diferenciais de que vamos tratar, neste parágrafo, serão todas diferenciáveis de classe C^k $(k \geq 2)$, definidas numa variedade diferenciável M, de classe C^{k+1} e dimensão n.

A nossa intenção é definir o operador d, de *diferenciação exterior*, que leva uma forma diferencial ω, de grau r, numa outra $d\omega$, de grau $r+1$.

Seja $x\colon U \to R^n$ um sisema de coordenadas sobre M. Então, em cada ponto $p \in U$, ω_p será uma forma r-linear alternada que pode ser escrita como

$$\omega_p = \sum_K a_K(p)\, dx^K, \quad K = \{k_1 < \cdots < k_r\}.$$

Convencione-se que $dx^K = 1$ quando $K = \varnothing$, isto é, quando $r = 0$.

Introduziremos um operador que, de início, vamos indicar com d_x até que se demonstre a independência da definição relativamente ao sistema de coordenadas x, escolhido de partida.

Os coeficientes a_K são funções e para as funções já está definida a diferencial (V. Exemplo 1 do §4),

$$da_K(p) = \sum_{i=1}^{n} \frac{\partial a_K}{\partial x^i}(p)\, dx^i.$$

Define-se a forma diferencial d_x, no ponto $p \in U$, pela seguinte expressão:

$$d_x\omega_p = \sum_K da_K(p) \wedge dx^K.$$

Proposição 2. *O operador d_x goza das seguintes propriedades:*

a. $d_x(\omega + \overline{\omega}) = d_x\omega + dx\overline{\omega}$;

b. $d_x f = df$;

c. $d_x(\omega \wedge \overline{\omega}) = (d_x\omega) \wedge \overline{\omega} + (-1)^r \omega \wedge d_x\overline{\omega}$, *onde r é o grau de ω;*

d. $d_x(d_x\omega) = 0$.

Demonstração: As duas primeiras são conseqüências imediatas da definição, levando-se em conta a convenção adotada para o caso da forma de grau 0.

Demonstremos a terceira. Em virtude de a. e da distributividade do produto exterior, basta considerarmos formas do tipo

$$\omega = adx^K \quad \text{e} \quad \overline{\omega} = bdx^L.$$

Temos $\omega \wedge \overline{\omega} = adx^K \wedge bdx^L = abdx^K \wedge dx^L$. Logo

$$\begin{aligned} d_x(\omega \wedge \overline{\omega}) &= d(ab) \wedge dx^K \wedge dx^L = (bda + adb) \wedge dx^K \wedge dx^L = \\ &= bda \wedge dx^K \wedge dx^L + adb \wedge dx^K \wedge dx^L = \\ &= (da \wedge dx^K) \wedge bdx^L + (-1)^r adx^K \wedge (db \wedge dx^L) \end{aligned}$$

pois o fator b, sendo de grau 0, comuta com todas as formas, mas $db \wedge dx^K = (-1)^r dx^K \wedge db$ pois db tem grau 1 e de dx^K tem grau r. Segue-se portanto que

$$d_x(\omega \wedge \overline{\omega}) = d\omega \wedge \overline{\omega} + (-1)^r \omega \wedge d\overline{\omega}.$$

Esta propriedade pode ser generalizada para um número N de formas, $\omega_1, \ldots, \omega_N$, de graus $r_1, \ldots, r_N$, respectivamente:

$$\begin{aligned} d_x(\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_N) &= d_x\,\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_N + \\ &+ (-1)^{r_1}\,\omega_1 \wedge d_x\,\omega_2 \wedge \ldots \omega_N + \\ &+ (-1)^{r_1+r_2}\,\omega_1 \wedge \omega_2 \wedge d_x\,\omega_3 \wedge \cdots \wedge \omega_N + \ldots \\ &\cdots + (-1)^{r_1+\cdots+r_{N-1}}\,\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_{N-1} \wedge d_x\,\omega_N\,. \end{aligned}$$

Demonstremos agora a última propriedade para o caso particular da forma de grau 0, isto é, para uma função diferenciável $f\colon M \to R$. Ora,

$$d_x f_p = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p)\, dx^i,$$

donde:

$$\begin{aligned} d_x(d_x f)_p &= \sum_{i,j} \frac{\partial^2 f}{\partial x^i \partial x^j}(p)\, dx^j \wedge dx^i = \\ &= \sum_{i>j} \frac{\partial^2 f}{\partial x^i \partial x^j}(p)\, dx^j \wedge dx^i \\ &\quad + \sum_{i>j} \frac{\partial^2 f}{\partial x^i \partial x^j}(p)\, dx^j \wedge dx^i = \\ &= \sum_{i<j} \frac{\partial^2 f}{\partial x^i \partial x^j}(p)\, dx^j \wedge dx^i \\ &\quad - \sum_{i<j} \frac{\partial^2 f}{\partial x^j \partial x^i}(p)\, dx^j \wedge dx^i = 0. \end{aligned}$$

No caso de uma forma ω de grau r qualquer, podemos novamente supor que $\omega = a dx^K = a dx^{k_1} \wedge \cdots \wedge dx^{k_r}$. Sendo $a, x^{k_1}, \ldots, x^{k_r}$ funções, segue-se do caso anterior que $d_x(da) = 0$ e $d_x(dx^{k_1}) = \cdots = d_x(dx^{k_r}) = 0$. Temos então, pela propriedade c.:

$$\begin{aligned} d_x(d_x\,\omega) &= d_x(da \wedge dx^K) = \\ &= d_x(da) \wedge dx^K - da \wedge d_x(dx^K) = \\ &= d_x(da) \wedge dx^K - da \wedge \sum_{i=1}^{r} (-1)^i\, dx^{k_1} \wedge \cdots \wedge \\ &\quad \wedge \cdots \wedge d_x(dx^{k_i}) \wedge \cdots \wedge dx^{k_r} = 0. \end{aligned}$$

Proposição 3. *A definição do operador d_x é independente do sistema x, isto é, se $y\colon V \to R^n$ é outro sistema de coordenadas em M, com $U \cap V \neq \varnothing$, então, em todo ponto $p \in U \cap V$ e para toda forma ω sobre M, tem-se*

$$d_x\,\omega_p = d_y \omega_p\,.$$

Demonstração: Se $\omega_p = \sum_K a_K(p) dx^K = \sum_L b_L(p) dy^L$, então

$$d_y \omega = \sum_L db_L(p) \wedge dy^L,$$

mas

$$b_L(p) = \sum_K \frac{\partial x^K}{\partial y^L}(p)\, a_K(p),$$

onde $\frac{\partial x^K}{\partial y^L}(p)$ é uma notação simplificada, análoga às demais, para representar um menor jacobiano de ordem r da

mudança de coordenadas $x \circ y^{-1}$. Continuando,

$$db_L(p) = \sum_K d\left(\frac{\partial x^K}{\partial y^L}\right)_p a_K(p) + \sum_K \frac{\partial x^K}{\partial y^L}(p)\, da_K(p),$$

substituindo na expressão de $d_y\omega_p$, obtemos

$$d_y\omega_p = \sum_K \left[da_K(p) \wedge \left(\sum_L \frac{\partial x^K}{\partial y^L}(p)\, dy^L\right)\right] + \\ + \sum_K a_K(p) \left(\sum_L d\left(\frac{\partial x^K}{\partial y^L}\right)_p \wedge dy^L\right)$$

como

$$\sum_L \frac{\partial x^K}{\partial y^L}(p)\, dy^L = dx^K_p$$

e

$$\sum_L d\left(\frac{\partial x^K}{\partial y^L}\right)_p \wedge dy^L = d_y(dx^K)_p = 0,$$

(a última sendo conseqüência das propriedades c. e d. da Proposição 2) segue-se que

$$d_y\omega_p = \sum_K da_K(p) \wedge dx^K = d_x\omega_p\,.$$

Outro modo de demonstrar esta proposição seria verificando que as quatro propriedades de d_x, enunciadas na Proposição 2, caracterizam completamente tal operador.

Então, dada a forma ω, definimos a *diferencial exterior* de ω, $d\omega$, em cada ponto $p \in M$, como $d_x\omega_p$, onde $x \in \mathfrak{X}_p$ é qualquer sistema de coordenadas definido em torno de p.

Exemplos. Conservando as notações do Exemplo 3 do §4 para as formas definidas no plano R^2, teremos:

se $\omega^0 = f(x,y)$, $d\omega^o = \dfrac{\partial f}{\partial x}\,dx + \dfrac{\partial f}{\partial y}\,dy$,

se $\omega^1 = a\,dx + b\,dy$, $d\omega^1 = \left(\dfrac{\partial b}{\partial x} - \dfrac{\partial a}{\partial y}\right) dx\,dy$,

se $\omega^2 = c\,dx\,dy$, $d\omega^2 = 0$, o que é natural, pois não existem formas não nulas de grau maior que a dimensão do espaço.

Analogamente, no espaço R^3, teremos

$$\omega^0 = f(x,y,z), \qquad d\omega^0 = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz,$$

$$\omega^1 = a\,dx + b\,dy + c\,dz,$$

$$d\omega^1 = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dy\,dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right) dz\,dx + \\ + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx\,dy,$$

isto é, os coeficientes da diferencial exterior, $d\omega^1$, são as componentes do *rotacional* do vetor (a,b,c) de mesmas componentes que ω.

Prosseguindo, se $\omega^2 = m\,dy\,dz + n\,dz\,dx + p\,dx\,dy$, então $d\omega^2 = \left(\frac{\partial m}{\partial x} + \frac{\partial n}{\partial y} + \frac{\partial p}{\partial z}\right) dx\,dy\,dz$, cujo coeficiente é a *divergência* do vetor de componentes m, n e p.

Finalmente, se $\omega^3 = g(x,y,z)dx\,dy\,dz$, $d\omega^3 = 0$.

Se $f\colon M^m \to N^n$ é uma aplicação diferenciável da variedade diferencial M^m na variedade diferenciável N^n, vimos

no §2, que, em cada ponto $p \in M$, f induz uma aplicação linear f^* das formas sobre N^n nas formas sobre M^m.

Assim sendo, se ω é uma forma de grau r sobre N^n e $v_1, \ldots, v_r \in M_p$ são vetores tangentes a M^m, no ponto p, a forma diferencial f^*, sobre N^n, é também de grau r e atua do seguinte modo:

$$(f^*\omega)_p(v_1, \ldots, v_r) = \omega_{f(p)}(f_* v_1, \ldots, f_* v_r).$$

Em termos de coordenadas, se $x\colon U \to R^m$ e $y\colon V \to R^n$ são sistemas em torno de p e $f(p)$, respectivamente, e a forma ω é dada por

$$\omega = \sum_K a_K \, dy^K,$$

então

$$(f^*\omega)_p = \sum_{L,K} a_K(f(p)) \frac{\partial y^K}{\partial x^L} \, dx^L.$$

Por esta expressão, conclui-se a diferenciabilidade de $f^*\omega$ a partir da diferenciabilidade de ω e de f.

Em particular, se ω é de grau 0, isto é, ω é uma função $\varphi\colon N^n \to R$, teremos $f^*\varphi = \varphi \circ f$.

Veremos que, se ω e $\overline{\omega}$ são formas sobre N^n, então

$$f^*(\omega \wedge \overline{\omega}) = f^*\omega \wedge f^*\overline{\omega}.$$

Como toda forma é uma soma de formas decomponíveis, f^* é linear e o produto $\omega \wedge \overline{\omega}$ é distributivo, basta verificar esta propriedade quando ω e $\overline{\omega}$ são formas decomponíveis.

Este caso, por sua vez, é conseqüência da relação

$$f^*(a dy^1 \wedge \cdots \wedge dy^k) = f^*a \cdot f^* \, dy^1 \wedge \ldots f^* \, dy^k,$$

ou seja

$$f^*(a dy^1 \wedge \cdots \wedge dy^k) =$$
$$= (a \circ f) \sum_{\ell_1 < \cdots < \ell_k} \frac{\partial(y^1 \ldots y^k)}{\partial(x^{\ell_1} \ldots x^{\ell_k})}\, dx^{\ell_1} \ldots dx^{\ell_k} =$$
$$= (a \circ f) \left(\sum_{\ell_1} \frac{\partial y^1}{\partial x^{\ell_1}}\, dx^{\ell_1} \right) \wedge \cdots \wedge \left(\sum_{\ell_k} \frac{\partial y^k}{\partial x^{\ell_k}}\, dx^{\ell_k} \right),$$

que é fato conhecido, do capítulo anterior.

Isto posto, tem-se uma quinta propriedade do operador d de diferenciação exterior:

$$f^*(d\omega) = d(f^*\omega).$$

Seja, em primeiro lugar, $\omega = \varphi\colon N^n \to R$ uma função diferenciável (forma de grau 0) definida em N^n. Então, tomando os sistemas de coordenadas x e y acima considerados, teremos:

$$d\varphi_{f(p)} = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial y^i}\,(f(p)) dy^i$$

e

$$f^*(d\varphi)_p = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{\partial \varphi}{\partial y_i}\, f(p))\, \frac{\partial y^i}{\partial x^j}\,(p) dx^j =$$
$$= \sum_{j=1}^{m} \frac{\partial \varphi}{\partial x^j}(p)\, dx^j = d(\varphi \circ f)_p = d(f^*\varphi)_p\,,$$

para qualquer ponto $p \in M^m$.

Passando ao caso geral, em que ω é de grau r qualquer, teremos

$$\omega = \sum_K a_K \, dy^K \quad \text{e} \quad d\omega = \sum_K da_K \wedge dy^K,$$

portanto

$$f^*(d\omega) = \sum_{K,L} d(a_K \circ f) \wedge \frac{\partial y^K}{\partial x^L} \, dx^L,$$

aplicando as propriedades acima demonstradas de f^* e a decomposição de dy^K no produto de formas de grau 1.

Por outro lado,

$$f^*\omega = \sum_{K,L} (a_K \circ f) \, \frac{\partial y^K}{\partial x^L} \, dx^L$$

logo:

$$d(f^*\omega) = \sum_{K,L} d(a_K \circ f) \wedge \frac{\partial y^K}{\partial x^L} \, dx^L,$$

pois as demais parcelas são nulas.

4.7 Variedades orientáveis

Uma orientação de um espaço vetorial é, como já vimos, uma certa classe de equivalência de bases desse espaço.

Dizemos que uma variedade diferenciável M^n é *orientável* quando existe uma aplicação que a cada ponto $p \in M$ associa uma orientação $\mathcal{O}_p$ do espaço vetorial tangente em p, M_p, aplicação esta que varia "continuamente"

com o ponto p, no seguinte sentido: se U é um domínio conexo de um sistema de coordenadas x, então a base $\mathcal{E}_p = \left\{\frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\right\}$ de M_p, associada ao sistema x, ou é positiva para todo $p \in U$, ou seja,

$$(\mathcal{E}_p \in \mathcal{O}_p, \forall\, p \in U)$$

ou negativa para todo $p \in U$ $(\mathcal{E}_p \notin \mathcal{O}_p, \forall\, p \in U)$.

Uma tal aplicação $p \to \mathcal{O}_p$ chama-se uma *orientação* de M.

Quando $\mathcal{E}_p \in \mathcal{O}_p$ diremos que o sistema x é *positivo*. Caso contrário, dizemos que x é *negativo* (relativamente à orientação considerada em M).

Uma *variedade orientada* é uma variedade (orientável) na qual se fixou, de uma vez por todas, uma orientação. Mais precisamente, é um par $(M, \mathcal{O})$, onde M é uma variedade orientável e $\mathcal{O}$ é uma orientação de M.

Se a orientação $\mathcal{O}_p$ do espaço vetorial M_p varia "continuamente" com o ponto p, da maneira acima descrita, então, dado um sistema de coordenadas $x\colon U \to R^n$ em M, com U conexo, o conhecimento da orientação $\mathcal{O}_{p_0}$ para um ponto $p_0 \in U$ determina univocamente o conhecimento de $\mathcal{O}_p$ para todo $p \in U$. Em conseqüência, se M é conexa e orientável, o conhecimento da orientação $\mathcal{O}_p$ num ponto $p \in M$ determina $\mathcal{O}_q$ para todo $q \in M$. Com efeito, sendo M conexa, dados p e q, existe uma cadeia finita de vizinhanças conexas $U_1, \ldots, U_r$, domínios de sistemas de coordenadas, tais que $p \in U_1$, $q \in U_r$ e $U_i \cap U_{i+1} \neq \varnothing$. Como só existem duas orientações possíveis para o espaço M_p, *segue-se que uma variedade conexa orientável admite precisamente duas orientações.*

Há outras maneiras, todas equivalentes, de introduzir o conceito de orientação de uma variedade. Enunciaremos aqui algumas delas.

Um *atlas coerente*, sobre M, é um atlas $\mathfrak{A}$, cujas mudanças de coordenadas têm jacobiano positivo, isto é, os sistemas do atlas $\mathfrak{A}$ satisfazem à seguinte condição:

se $x\colon U \to R^n$ e $y\colon V \to R^n$ são sistemas de $\mathfrak{A}$ e $U \cap V \neq \varnothing$, a aplicação diferenciável

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

tem jacobiano positivo.

Um *atlas coerente máximo* em M é um atlas coerente não contido propriamente em outro da mesma natureza. Todo atlas coerente está contido num único atlas coerente máximo.

Proposição 4. *Uma variedade é orientável quando, e somente quando, admite um atlas coerente.*

Demonstração: A cargo do leitor.

Sendo assim, pode-se definir uma orientação numa variedade como sendo um atlas coerente máximo.

Dada a variedade diferenciável M, de dimensão n, vamos introduzir uma topologia e uma estrutura de variedade diferenciável, de mesma classe e dimensão que M, sobre o conjunto

$$\widetilde{M} = \{(p, \mathcal{O}_p); p \in M \text{ e } \mathcal{O}_p \text{ é uma orientação em } M_p\}.$$

Seja $\pi\colon \widetilde{M} \to M$ a aplicação definida como

$$\pi(p, \mathcal{O}_p) = p.$$

Sejam $\mathfrak{A}$ um atlas de M, $x\colon U \to R^n$ um elemento qualquer de $\mathfrak{A}$ e $\widetilde{U} \subset \widetilde{M}$ o subconjunto de $\widetilde{M}$ definido por

$$\widetilde{U} = \left\{(p, \mathcal{O}_p); p \in U \text{ e } \left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\} \in \mathcal{O}_p\right\}.$$

Um atlas $\widetilde{\mathfrak{A}}$ de $\widetilde{M}$ será constituído pelos sistemas de coordenadas $\widetilde{x}\colon \widetilde{U} \to R^n$, onde

$$\widetilde{x}(p, \mathcal{O}_p) = x(p).$$

Verifica-se que $\widetilde{\mathfrak{A}}$ define uma topologia (de Hausdorff e com base enumerável) sobre $\widetilde{M}$ e, relativamente a esta topologia, $\widetilde{\mathfrak{A}}$ é um atlas diferenciável. Com esta estrutura, a aplicação π é diferenciável. Mais ainda, π é um difeomorfismo local.

Se $\widetilde{x}\colon \widetilde{U} \to R^n$ e $\widetilde{y}\colon \widetilde{V} \to R^n$ são dois sistemas de $\widetilde{\mathfrak{A}}$ tais que $\widetilde{U} \cap \widetilde{V} \neq \varnothing$, seja $(p, \mathcal{O}_p) \in \widetilde{U} \cap \widetilde{V}$. Então, tanto $\left\{\frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\right\}$ como $\left\{\frac{\partial}{\partial y^1}(p), \ldots, \frac{\partial}{\partial y^n}(p)\right\}$ pertencem à mesma orientação $\mathcal{O}_p$ de M_p; isto significa que o determinante de passagem de uma base para outra é positivo. Ora, este determinante é, exatamente, o jacobiano da mudança de coordenadas $\widetilde{y} \circ \widetilde{x}^{-1}$ calculado em $\widetilde{x}(p) = x(p)$ e, sendo positivo para qualquer ponto $p \in U \cap V$, conclui-se que o atlas $\widetilde{\mathfrak{A}}$ é coerente, isto é, a variedade $\widetilde{M}$ é *sempre orientável.*

À variedade $\widetilde{M}$, acima definida, dá-se o nome de *recobrimento duplo* de M. Este é um exemplo de espaço de recobrimento.

Em geral, sendo M e N variedades diferenciáveis, N diz-se *recobrimento* de M se existe uma aplicação

$\pi\colon N \to M$, diferenciável, tal que cada ponto $p \in M$ possui uma vizinhança U, cuja imagem inversa $\pi^{-1}(U)$ seja reunião de abertos disjuntos de N, cada um dos quais é aplicado difeomorficamente, sobre U, por π.

Lema 1. *Sejam $\alpha\colon [a,b] \to M$ um arco contínuo, em M, $\alpha(a) = p_0$ e $p \in \pi^{-1}(p_0)$ um ponto de N que se projeta em p_0 por π. Existe um único arco $\widetilde{\alpha}\colon [a,b] \to N$, contínuo, em N, tal que $\widetilde{\alpha}(a) = p$ e cuja projeção seja α, isto é, $\pi \circ \widetilde{\alpha} = \alpha$.*

Esboço de demonstração: Considera-se uma divisão $a = t_0 < t_1 < \cdots < t_n = b$ do intervalo $[a,b]$ suficientemente fina, de modo que $\alpha([t_i, t_{i+1}]) \subset U_i$, onde cada conjunto U_i é um aberto de M, cuja imagem inversa por π é reunião de abertos disjuntos, homeomorfos a U_i por π. Partindo de $p_0 \in U_0$, consideramos o aberto $V_0 \subset \pi^{-1}(U_0)$ que contém o ponto p e é levado difeomorficamente, por π, sobre U_0. Se π_0 é a restrição de π a V_0, definimos $\widetilde{\alpha}$, em $[a,t_1]$, como $\widetilde{\alpha} = \pi_0^{-1} \circ \alpha$.

Procede-se da mesma forma relativamente aos pontos $\alpha(t_1)$ e $\widetilde{\alpha}(t_1)$, e assim por diante.

A unicidade é óbvia, dentro das condições exigidas.

Estamos, agora, em condições de demonstrar uma proposição que nos dará uma caracterização de variedade orientável.

Proposição 5. *Seja M uma variedade conexa. M é orientável se, e somente se, $\widetilde{M}$ é desconexa.*

Demonstração: Seja M uma variedade conexa orientável. Consideremos uma de suas orientações, isto é, a cada ponto

p fica associada uma orientação $\mathcal{O}_p$, que varia "continuamente" com p. Demonstraremos que $\widetilde{M}$ é desconexa, isto é, $\widetilde{M} = A \cup B$ com $A, B \neq \varnothing$, disjuntos e abertos.

Definimos

$$A = \{(p, \mathcal{O}_p); p \in M, \mathcal{O}_p \text{ é a orientação de } M_p \text{ coerente com a orientação de } M.\}$$

$$B = \{(p, \overline{\mathcal{O}}_p); p \in M, \overline{\mathcal{O}}_p \text{ é a orientação de } M_p, \text{ oposta a } \mathcal{O}_p.\}$$

É claro que A e B são disjuntos, não-vazios e que $\widetilde{M} = A \cup B$. Provemos que são abertos. Consideremos um deles, por exemplo A. Ora, se $(p, \mathcal{O}_p) \in A$, seja $x\colon U \to R^n$ um sistema de coordenadas em M, positivo, em torno do ponto p. De acordo com as notações introduzidas, o aberto

$$\widetilde{U} = \{(q, \mathcal{O}_q); q \in U, \mathcal{O}_q \text{ é a orientação a que } \left\{\frac{\partial}{\partial x^i}\right\} \text{ pertence}\}$$

está contido em A e contém $(p, \mathcal{O}_p)$.

Reciprocamente, suponhamos que M seja conexa e $\widetilde{M}$ desconexa. Fixemos um ponto $p \in M$. Seja

$$\pi^{-1}(p) = \{\widetilde{p}_1, \widetilde{p}_2\} \subset \widetilde{M}.$$

Dado um ponto arbitrário $\widetilde{q} = (q, \mathcal{O}_q)$ em $\widetilde{M}$, mostraremos que existe um arco contínuo em $\widetilde{M}$ ligando $\widetilde{q}$ a $\widetilde{p}_1$ ou a $\widetilde{p}_2$. Com efeito, sendo M conexa, existe um arco contínuo

α, em M, ligando q a p. Seja $\widetilde{\alpha}$ um levantamento de α (Lema 1) que comece em $\widetilde{q}$. Como $\alpha = \pi \circ \widetilde{\alpha}$ termina em p, $\widetilde{\alpha}$ termina em $\pi^{-1}(p)$, ou seja, em $\widetilde{p}_1$ ou em $\widetilde{p}_2$. Segue-se que $\widetilde{M}$ tem, no máximo, duas componentes conexas. Sendo desconexa, tem exatamente duas. Seja $\widetilde{M}_1$ a componente conexa de $\widetilde{p}_1$ e $\widetilde{M}_2$ a de $\widetilde{p}_2$. $\widetilde{M}_1$ e $\widetilde{M}_2$ são abertos conexos disjuntos e $\widetilde{M} = \widetilde{M}_1 \cup \widetilde{M}_2$.

Dado o ponto arbitrário $\widetilde{q} = (q, \mathcal{O}_q)$ em $\widetilde{M}$, seja $\overline{q} \in \widetilde{M}$ o outro ponto com a mesma projeção q. Se $\widetilde{q} \in \widetilde{M}_1$, então $\overline{q} \in \widetilde{M}_2$. Com efeito, seja $\widetilde{\alpha}\colon [0,1] \to \widetilde{M}$ um arco em $\widetilde{M}$ ligando $\widetilde{q}$ a $\widetilde{p}_1$. O arco $\widetilde{\alpha}$ é o levantamento de $\alpha = \pi\widetilde{\alpha}$ que começa em $\widetilde{q}$. Seja $\overline{\alpha}$ o levantamento de α que começa em $\overline{q}$. Então deve ser

$$\overline{\alpha}(1) = \widetilde{p}_2$$

pois se fosse $\overline{\alpha}(1) = \widetilde{p}_1$, os arcos $t \to \widetilde{\alpha}(1-t)$ e $t \to \overline{\alpha}(1-t)$ em $\widetilde{M}$ seriam dois levantamentos distintos do mesmo arco $t \to \alpha(1-t)$, ambos começando em p_1, o que contrariaria a unicidade no Lema 1. Assim, $\overline{\alpha}$ liga $\overline{q}$ a $\widetilde{p}_2$, donde $\overline{q} \in \widetilde{M}_2$. Segue-se que $\pi\colon \widetilde{M} \to M$ induz difeomorfismos $\pi\colon \widetilde{M}_1 \approx M$ e $\pi\colon \widetilde{M}_2 \approx M$.

Sabendo que as duas componentes conexas de $\widetilde{M}$ são ambas difeomorfas a M por meio de π, e que $\widetilde{M}_1$ e $\widetilde{M}_2$ são orientáveis (como subconjuntos abertos da variedade orientável $\widetilde{M}$) segue-se imediatamente que M é orientável.

Observe-se que, se indicamos com $\sigma\colon M \to \widetilde{M}$ o inverso do difeomorfismo $\widetilde{M}_1 \to M$, temos $\sigma(p) = (p, \mathcal{O}_p)$ e σ define uma orientação em M. A topologia de $\widetilde{M}$ permite falar da continuidade de σ e temos agora um sentido exato para a afirmação de que a escolha $p \to \mathcal{O}_p$ de uma orientação em cada ponto de M varia "continuamente" com o ponto p.

Mais adiante, caracterizaremos novamente o conceito de orientação, por meio de uma forma diferencial.

Exemplos. O espaço euclidiano R^n, sendo, em particular, um espaço vetorial, é orientável. Mais ainda, R^n possui uma orientação natural, definida pela base canônica $\{e_1, \ldots, e_n\}$, onde $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$.

A esfera S^n é orientável. Basta tomar uma orientação em R^n e orientar cada espaço tangente $(S^n)_p$ de modo que uma base positiva de $(S^n)_p$ seja tal que, quando se acrescente, como último elemento, a normal a $(S^n)_p$ que aponta para fora de S^n, a base de R^n obtida seja ainda positiva.

Mais geralmente, dados um aberto $\Omega \subset R^n$ e uma aplicação diferenciável $f : \Omega \to R^{n-r}$, se $a = (a^1, \ldots, a^{n-1}) \in R^{n-r}$ é um valor regular de f então a variedade $M^r = f^{-1}(a)$ é orientável. Para $x \in \Omega$, seja

$$f(x) = (f^1(x), \ldots, f^{n-r}(x)).$$

Então M^r é definida "implicitamente" pelas equações:

$$\begin{cases} f^1(x^1, \ldots, x^n) = a^1 \\ f^2(x^1, \ldots, x^n) = a^2, \quad x = (x^1, \ldots, x^n) \in \Omega \\ \ldots\ldots\ldots\ldots\ldots \\ f^{n-r}(x^1, \ldots, x^n) = a^{n-r} \end{cases}$$

Assim, toda subvariedade do R^n definida implicitamente por um sisema de equações cuja matriz jacobiana tem caracterísitca máxima (condição de regularidade do valor a) é uma variedade orientável.

A variedade não-orientável mais simples é a faixa de Möbius. Uma variedade de dimensão 2 que contenha uma

faixa de Möbius não é orientável. Por exemplo: o plano projetivo e a garrafa de Klein são superfícies (variedades de dimensão 2) compactas não-orientáveis. Os espaços projetivos de dimensão ímpar são orientáveis. Os de dimensão par não são.

4.8 Partição diferenciável da unidade

Uma família enumerável $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots : M \to R$ de funções reais diferenciáveis, definidas em M, constitui uma *partição diferenciável da unidade* quando são satisfeitas as condições seguintes:

a. $\varphi_i \geq 0, \quad \forall\, i = 1, 2, \ldots$

b. cada ponto $p \in M$ possui uma vizinhança na qual apenas um número finito de funções φ_i são diferentes de 0.

c. $\sum_{i=1}^{\infty} \varphi_i(p) = 1, \quad \forall\, p \in M.$

Diz-se que uma cobertura C de um espaço topológico X é *localmente finita* se qualquer ponto $p \in X$ possui uma vizinhança que interseta apenas um número finito de conjuntos da cobertura. Em particular, p pertence somente a um número finito de conjuntos de C.

Se definimos, ainda, *suporte de uma função* como sendo o fecho do conjunto dos pontos onde esta função é diferente de 0, as condições b. e c. garantem que a coleção dos suportes de φ_i é uma cobertura de M, localmente finita.

Uma partição diferenciável da unidade $\{\varphi_i\}$ é dita *subordinada a uma cobertura* $\mathcal{C}$ de M quando, para cada i, o

suporte de φ_i está contido em algum $A_i \in \mathcal{C}$, isto é, para cada i, existe $A_i \in \mathcal{C}$ tal que $\varphi_i(\overline{M - A_i}) = \{0\}$.

Iremos, neste parágrafo, construir, efetivamente, uma partição da unidade subordinada a uma cobertura aberta qualquer da variedade M.

Sejam X um espaço topológico qualquer e $\mathcal{C}$ uma cobertura de X. Diz-se que uma outra cobertura $\mathcal{C}'$ *refina* $\mathcal{C}$, ou é um *refinamento* de $\mathcal{C}$, se todo $A' \in \mathcal{C}'$ está contido em algum $A \in \mathcal{C}$. A relação "$\mathcal{C}'$ refina $\mathcal{C}$" é transitiva, mas não é uma relação de ordem por não ser antissimétrica, isto é, "$\mathcal{C}'$ refina $\mathcal{C}$" e "$\mathcal{C}$ refina $\mathcal{C}'$" não implicam $\mathcal{C} = \mathcal{C}'$.

Como exemplo, veremos que uma cobertura $\mathcal{C}$, qualquer, da variedade diferenciável M pode ser refinada por uma cobertura enumerável, $\{U_1, U_2, \dots\}$, onde cada U_i é o domínio de um sistema de coordenadas $x_i \colon U_i \to R^n$. Com efeito, para cada ponto $p \in M$, consideremos um sistema $x \colon U \to R^n$ definido em torno de p e um aberto $A \in \mathcal{C}$ que contenha p. Seja $U(p) = A \cap U$, então a cobertura $\{U(p); p \in M\}$ refina $\mathcal{C}$, onde, para, sistema de coordenadas, tomamos a restrição de x a $U(p)$. Além disto, como a variedade M tem base enumerável, podemos extrair desta última cobertura, pela propriedade de Lindelof, uma subcobertura enumerável $\{U_1, U_2, \dots\}$ que é a procurada.

Para nossos objetivos, é conveniente que os sistemas $x_i \colon U_i \to R^n$ sejam tais que $x_i(U_i) = B(3)$ (bola de raio 3 em R^n) e, mais ainda, pondo $V_i = x_i^{-1}(B(2))$ e $W_i = x_i^{-1}(B(1))$, os $\{W_i\}$ ainda sejam uma cobertura da variedade. Isto se obtém com uma ligeira modificação no argumento anterior. Basta considerar, em primeiro lugar, uma bola contida em $x(U(p))$, para cada $p \in M$, e compor x com uma homotetia que leve esta bola em $B(3)$. Continu-

amos chamando de x o sistema composto com esta homotetia e restrito à imagem inversa de $B(3)$. Pondo, agora, $W(p) = x^{-1}(B(1))$, extraímos a subcobertura enumerável da cobertura $\{W(p); p \in M\}$ e consideramos os U_i correspondentes.

Quando a variedade for compacta, pode-se substituir "enumerável" por "finita".

Um espaço topológico X diz-se *para-compacto* (conceito introduzido por J.A. Dieudonné) quando qualquer cobertura aberta de X pode ser refinada por uma cobertura aberta localmente finita.

Lema 2. *Toda cobertura aberta de uma variedade diferenciável M pode ser refinada por uma cobertura enumerável $\{U_i;\ i = 1, 2, \ldots\}$, localmente finita, tal que: existem sistemas de coordenadas $x_i\colon U_i \to R^n$, com $x_i(U_i) = B(3)$ e, pondo $V_i = x_i^{-1}(B(2))$ e $W_i = x_i^{-1}(B(1))$, os W_i, $i = 1, 2, \ldots$, ainda cobrem M.*

Demonstração: Com isso, estaremos demonstrando, inclusive, que uma variedade é um espaço para-compacto. Estudemos o caso por etapas. Observemos, antes, que toda variedade é um espaço localmente compacto.

1ª parte. Onde demonstramos que um espaço topológico X, localmente compacto e com base enumerável, pode ser considerado como reunião de uma seqüência crescente de compactos, isto é, existem compactos $L_1 \subset L_2 \subset \ldots$ cuja reunião

$$L_1 \cup L_2 \cup \ldots \quad \text{é igual a} \quad X.$$

Basta, em cada ponto $p \in X$, considerar uma vizinhança $J(p)$ cujo fecho seja compacto. Da cobertura

$\{J(p)\,;\ p \in X\}$ extrai-se uma subcobertura enumerável $J_1, J_2, \ldots$ e, para L_i^*, tomamos o fecho $\overline{J}_i$. A seguir, fazemos $L_1 = L_1^*$

$$L_2 = L_1^* \cup L_2^*,$$

e assim por diante.

2ª parte. O espaço X pode ser escrito como

$$X = K_1 \cup K_2 \cup \ldots,$$

onde os conjuntos K_i são compactos tais que $K_i \subset \text{int } K_{i+1}$ $(i = 1, 2, \ldots)$.

De fato, pomos $K_1 = L_1$. Consideramos, agora, para cada ponto $p \in L_2$, uma vizinhança de fecho compacto. Desta cobertura, extraímos uma subcobertura finita. Definimos K_2 como sendo o fecho da união desta subcobertura. Em resumo, K_2 é uma vizinhança compacta de L_2. Prosseguimos, tomando para K_3 uma vizinhança compacta de $K_2 \cup L_3$, em geral, K_i será uma vizinhança compacta de $K_{i-1} \cup L_i$.

Assim sendo $K_i \subset \text{int } K_{i+1}$ e $L_i \subset K_i$, então

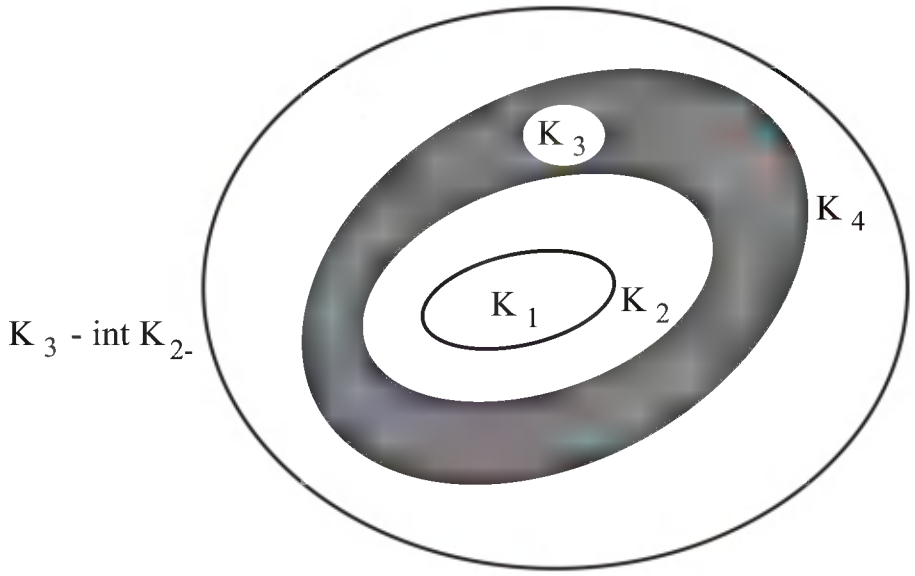
$$\{K_i\,,\ i = 1, 2, \ldots\}$$

é também uma cobertura de X.

3ª parte (Final). A conclusão anterior é válida para M, isto é,

$$M = K_1 \cup K_2 \cup \ldots,\ K_i \subset \text{int } K_{i+1} \text{ com } K_i \text{ compacto.}$$

Conservando as notações introduzidas acima, K_2 pode ser coberto por um número finito de abertos do tipo W_i de modo que cada U_i correspondente esteja contido no interior de K_3 e em algum aberto da cobertura dada inicialmente.



Do mesmo modo, a faixa compacta $K_3 -$ int K_2 pode ser coberta por um número finito de abertos W_i, com os U_i correspondentes contidos em $K_4 - K_1$ e cada um contido em algum aberto da cobertura.

Prosseguindo de maneira análoga, obtém-se a cobertura $\{W_i,\ i = 1, 2, \dots\}$ e, em correspondência, a cobertura

$$\{U_i\,;\ i = 1, 2, \dots\}.$$

A cobertura $\{U_i\,;\ i = 1, 2, \dots\}$ é localmente finita, pois, dado um ponto $p \in M$, existe uma vizinhança W_j que o contém e está contida em algum K_i, intersetando, conseqüentemente, apenas um número finito de conjuntos U_i.

Lema 3. *Sendo $B(r)$ a bola de raio r, em R^n, existe uma função diferenciável $\varphi\colon B(3) \to R$, de classe C^∞, tal que*

a. $0 \leq \varphi \leq 1$,
b. $\varphi(t) = 1$, *quando* $t \in B(1)$ *e*
$\varphi(t) = 0$, *quando* $t \in B(3) - B(2)$.

Demonstração: Para $n = 1$, o gráfico da função φ é o da figura que segue. Para $n = 2$, é a superfície de revolução gerada por essa curva.

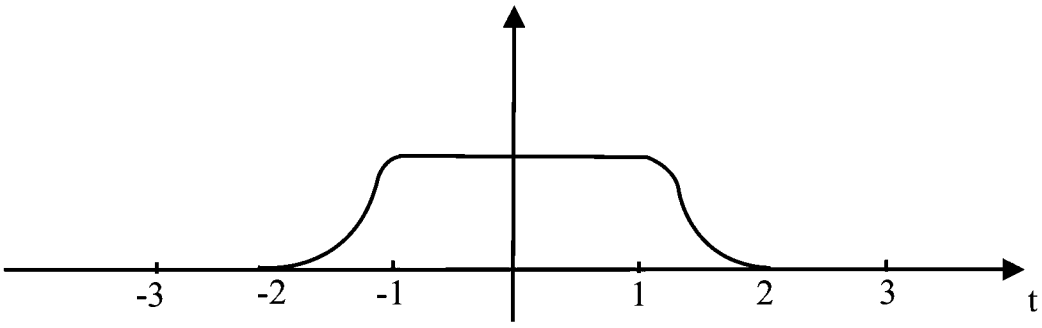


gráfico de φ

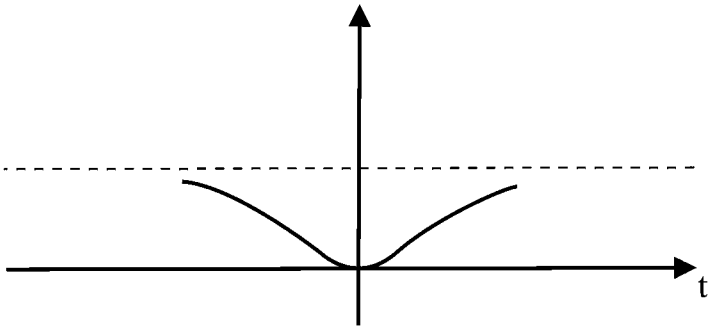


gráfico da função de Cauchy

A peculiaridade de φ é que nos pontos t, em que $|t| = 1$ ou $|t| = 2$, ela possui derivadas parciais todas nulas, mas não é constante em nenhuma vizinhança destes pontos. Isto nos lembra o exemplo clássico de Cauchy, que é a função definida, para $t \in R$, como e^{-1/t^2} $(t \neq 0)$ e como 0 para $t = 0$. Esta é de classe C^∞, tem todas as derivadas nulas na origem, mas não é constante em nenhuma vizinhança da origem.

Utilizaremos uma função do mesmo tipo que esta, mas que se anula em dois pontos. Se t é real, consideramos a

função

$$\xi(t) = \begin{cases} e^{-\frac{1}{(t+1)(t+2)}}, & \text{para} \quad -2 < t < -1 \\ 0 \quad \text{fora do intervalo} \quad (-2,-1) \end{cases}$$

Seja

$$a = \int_{-2}^{-1} \xi(t)\, dt = \int_{-\infty}^{+\infty} \xi(t)\, dt,$$

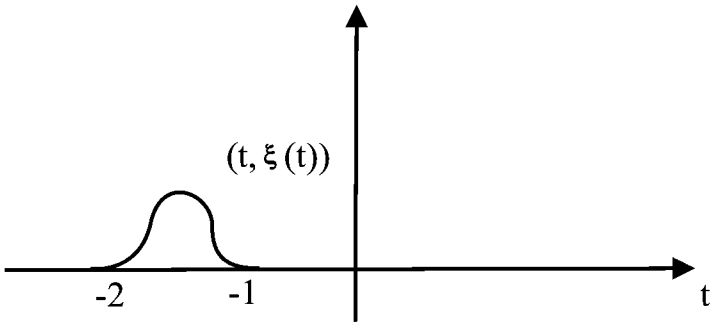


gráfico de ξ

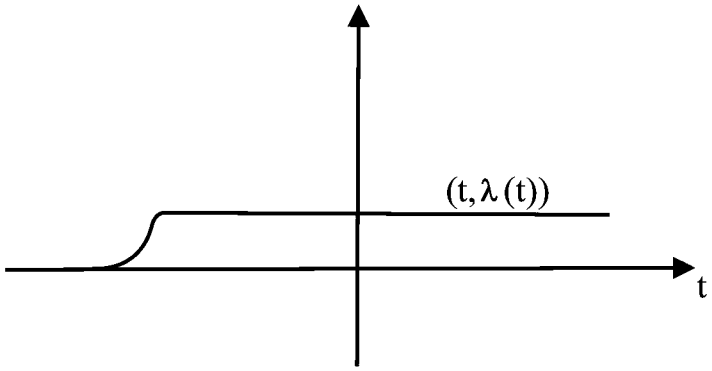


gráfico de λ

construímos a função $\lambda(t)$, diferenciável de classe C^∞, a partir de $\xi(t)$:

$$\lambda(t) = \frac{1}{a} \int_{-\infty}^{+\infty} \xi(s)\, dx.$$

Finalmente, se $t \in B(3) \subset R^n$ define-se

$$\varphi(t) = \lambda(-|t|) = \lambda\left(-\sqrt{(t_1)^2 + \cdots + (t_n)^2}\right)$$

que é a função procurada.

Estamos, agora, em condições de construir uma partição da unidade, o que será feito na

Proposição 6. *Seja $\mathcal{C}$ uma cobertura aberta da variedade diferenciável M. Existe uma partição diferenciável da unidade subordinada a $\mathcal{C}$.*

Demonstração: Na realidade, construiremos uma partição subordinada a um refinamento de $\mathcal{C}$ e, portanto, subordinada a $\mathcal{C}$. Sejam $\{U_i\,;\, i = 1, 2, \dots\}$ o refinamento de $\mathcal{C}$ construído como no Lema 2, e $\varphi\colon B(3) \to R$ a função construída no Lema 3.

Introduzimos as funções $\psi_i\colon M \to R$ definidas por

$$\psi_i = \begin{cases} \varphi \circ x_i\,, & \text{em} \quad U_i \\ 0, & \text{em} \quad M - U_i\,, \end{cases}$$

onde $x_i\colon U_i \to R^n$ são os sistemas de coordenadas definidos em U_i. Estas funções são diferenciáveis, pois já em $U_i - V_i$ elas são nulas.

Em seguida, definimos, em cada ponto $p \in M$,

$$\varphi_i(p) = \frac{\psi_i(p)}{\sum\limits_{j=1}^{\infty} \psi_j(p)}\,.$$

A soma do denominador, apesar da aparência, é soma de um número finito de parcelas, desde que a cobertura

U_i é localmente finita e as funções ψ_i anulam-se fora de U_i. As funções φ_i, $i = 1, 2, \ldots$, constituem a partição diferenciável da unidade subordinada a $\mathcal{C}$.

Aplicações:

1. Demonstremos que toda variedade diferenciável, de classe C^k, possui uma métrica riemanniana de classe C^{k-1}.

Sejam M uma variedade diferenciável de classe C^k, $\{U_i\,;\ i = 1, 2, \ldots\}$ uma cobertura de sistemas de coordenadas $x_i \colon U_i \to R^n$ e $\{\varphi_i\,,\ i = 1, 2, \ldots\}$ uma partição diferenciável da unidade subordinada a esta cobertura.

Se $p \in U_i$ é um ponto de U_i e $u, v \in M_p$ são vetores tangentes a M em p, sejam $u = \sum_{j=1}^{n} \alpha^j \frac{\partial}{\partial (x_i)^j}$ e

$$v = \sum_{j=1}^{n} \beta^j \frac{\partial}{\partial (x_i)^j}\,.$$

Pomos, como produto escalar, em U_i, de u por v

$$g_i(u, v) = \sum_{j=1}^{n} \alpha^j\, \beta^j.$$

A partir destes, definimos o produto escalar $u \cdot v$, na variedade M, por

$$u \cdot v = \sum_{i=1}^{\infty} \varphi_i(p)\, g_i(u, v).$$

Verifica-se que este é realmente um produto escalar e que varia diferenciavelmente com o ponto.

2. A observação a seguir fornece-nos mais uma definição de variedade orientável. Usaremos uma partição diferenciável da unidade para mostrar que *uma variedade M, de dimensão n, é orientável quando, e somente quando, existe em M uma forma diferencial contínua de grau n, diferente de* 0 *em todos os pontos.*

Suponhamos que exista, sobre M, a forma diferencial ω contínua e diferente de 0 em todos os pontos. Seja $p \in M$ um ponto de M, consideremos, em M_p, a seguinte orientação: uma base $\{v_1, \ldots, v_n\}$ de M_p será positiva se, e somente se, $\omega_p(v_1, \ldots, v_n) > 0$.

Ora, se U é um domínio conexo do sistema de coordenadas x e $p \in U$, o sinal de x será o de $\omega_p\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right)$ e este é o mesmo para qualquer $p \in U$ desde que ω é contínua e não se anula.

Reciprocamente, sejam M orientável e mais, orientada, $\{U_i\,;\, i = 1, 2, \ldots\}$ uma cobertura de M por domínios de sistemas de coordenadas positivos $x_i \colon U_i \to R^n$ e $\{\varphi_i\,;\, i = 1, 2, \ldots\}$ uma partição da unidade subordinada a $\{U_i\}$.

Em U_i, tem-se a forma diferencial de grau n

$$dx_i^1 \wedge dx_i^2 \wedge \cdots \wedge dx_i^n\,.$$

Por meio da função φ_i, pode-se estendê-la a toda a variedade M, pondo

$$\omega_i = \varphi_i\, dx_i^1 \wedge dx_i^2 \wedge \cdots \wedge dx_i^n\,.$$

A forma ω_i é identicamente nula fora de U_i, isto significa que a soma

$$\sum_{i=1}^{\infty} \omega_i$$

é finita em cada ponto $p \in M$. Define-se, então,

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

Falta-nos verificar que $\omega \neq 0$ em qualquer ponto de M. De fato, sejam $v_1, \ldots, v_n \in M_p$ vetores linearmente independentes que, nesta ordem, formam uma base positiva de M_p, então

$$\omega_p(v_1, \ldots, v_n) = \sum_{p \in U_i} \omega_i(v_1, \ldots, v_n) > 0$$

porque cada parcela $\omega_i(v_1, \ldots, v_n) \geq 0$ havendo alguma estritamente positiva.

4.9 Integral de uma forma diferencial

Seja ω uma forma diferencial contínua, de grau n, sobre uma variedade orientada M^n. Definiremos por etapas a integral de ω sobre M, que indicaremos com $\int_M \omega$.

1o. caso: Em que o suporte K de ω é compacto e está contido no domínio U de um sistema de coordenadas positivo $x \colon U \to R^n$. Em cada ponto $p \in U$, tem-se:

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

sendo a função $a \colon U \to R$ contínua. Definiremos então

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

onde $a(x^1, \ldots, x^n)$ significa $(a \circ x^{-1})(x^1, \ldots, x^n)$ e a integral do 2o. membro é tomada no sentido usual de Riemann no espaço euclidiano. A função $a(x^1, \ldots, x^n)$ sendo contínua em $x(U)$ e nula fora do compacto $x(K)$, esta integral sempre existe.

Devemos mostrar que a definição acima não depende da escolha do sistema de coordenadas positivo x. Seja pois $y\colon V \to R^n$ outro sistema de coordenadas positivo, cujo domínio V também contém o suporte K de ω. Para $p \in V$, tem-se

$$\omega_p = b(p) dy^1 \wedge \cdots \wedge dy^n, \text{ com}$$
$$b(p) = a(p) \frac{\partial(x^1, \ldots, x^n)}{\partial(y^1, \ldots, y^n)} \text{ se } p \in U \cap V.$$

Aí $b\colon V \to R$ é também contínua e o jacobiano da direita é calculado no ponto $y(p)$. A forma ω se anula fora do aberto $W = U \cap V$. Sendo o jacobiano acima > 0, podemos aplicar a fórmula de mudança de variáveis nas integrais múltiplas. Temos então:

$$\begin{aligned}\int_{y(V)} b\, dy^1 \ldots dy^n &= \int_{y(W)} b\, dy^1 \ldots dy^n = \\ &= \int_{y(W)} a \frac{\partial(x^1 \ldots x^n)}{\partial(y^1 \ldots y^n)} dy^1 \ldots dy^n = \\ &= \int_{x(W)} a\, dx^1 \ldots dx^n = \int_{x(V)} a\, dx^1 \ldots dx^n\end{aligned}$$

como queríamos demonstrar.

2o. caso: Em que M^n é compacta e ω é uma forma diferencial contínua qualquer, de grau n, sobre M.

Seja $\varphi_1, \ldots, \varphi_r$ uma partição da unidade, subordinada a uma cobertura $U_1, \ldots, U_r$ (como no Lema 2) formada por domínios de sistemas de coordenadas $x_i \colon U_i \to R^n$, que podemos supor positivos.

Cada forma $\omega_i = \varphi_i \omega$ tem suporte compacto, contido em U_i, donde satisfaz as condições do 1o. caso. Definimos então

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

É necessário verificar que esta definição não depende da partição da unidade considerada. Seja $\psi_1, \ldots, \psi_s$ outra partição. Tem-se

$$\sum_{i=1}^{r} \int_M \varphi_i \omega = \sum_{i=1}^{r} \int_M \left(\sum_{j=1}^{s} \psi_j \right) \varphi_i \omega = \sum_{i,j} \int_M \varphi_i \psi_j \omega =$$
$$= \sum_{j=1}^{s} \int_M \left(\sum_{i=1}^{r} \varphi_i \right) \psi_j \omega = \sum_{j=1}^{s} \int_M \psi_j \omega$$

onde a primeira e a última igualdades são válidas porque $\sum \varphi_i = \sum \psi_j = 1$ e as demais porque as integrais consideradas, são na realidade, integrais no R^n, onde essas transformações são válidas.

3o. caso: Em que ω é uma forma contínua de grau n numa variedade M^n que pode não ser compacta.

Tomando uma partição da unidade $\varphi_1, \ldots, \varphi_r, \ldots$ subordinada a uma cobertura enumerável (como no Lema 2) de domínios de sistemas de coordenadas $x_i \colon U_i \to R^n$,

todos os x_i positivos, podemos considerar a série

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

cujos termos são todos bem definidos, de acordo com o 1o. caso. Mas esta série pode ser convergente ou divergente, dando origem à subdivisão das formas contínuas de grau n sobre M^n em formas integráveis e formas não integráveis, as quais discutiremos sucintamente agora.

Em primeiro lugar digamos que uma forma ω (de grau n, sobre uma variedade orientada M^n) é *positiva* quando, em cada ponto $p \in M$, $\omega_p(v_1, \ldots, v_n) \geq 0$ se $\{v_1, \ldots, v_n\}$ é uma base positiva de M_p. Isto equivale a dizer que, para todo sistema de coordenadas positivo $x\colon U \to R^n$ em M, tem-se

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

com $a(p) \geq 0$ para qualquer $p \in U$. Dada a forma ω, chama-se *parte positiva* de ω à forma ω_+ que coincide com ω nos pontos em que ω é positiva e é igual a zero nos pontos em que ω não é ≥ 0. Se x é novamente um sistema de coordenadas positivo, tem-se $(\omega_+)_p = a_+(p)\, dx^1 \wedge \cdots \wedge dx^n$, $p \in U$, onde

$$a_+(p) = \max\{a(p), 0\}.$$

Analogamente, chama-se *parte negativa* de ω à forma ω_- que é igual a $-w$ nos pontos em que $w \leq 0$ e $w_- = 0$ nos pontos em que $\omega \geq 0$. Se ω é contínua, ω_+ e ω_- são formas positivas contínuas sobre M e tem-se

$$\omega = \omega_+ - \omega_- .$$

Diremos que uma forma contínua *positiva* ω é *integrável* quando existe uma partição da unidade $\{\varphi_i\}$, do tipo acima considerado, tal que a série (de termos ≥ 0)

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

seja convergente. Segue-se que esta série é absolutamente convergente e que são válidas as manipulações formais feitas no 2o. caso para demonstrar que a soma da série independe da partição. Mais ainda, o mesmo raciocínio mostra que a convergência desta série para uma partição implica convergência para qualquer partição. Definimos então a integral da forma positiva ω pondo

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

No caso de uma forma não-positiva ω, temos $\omega = \omega_+ - \omega_-$. Diremos que ω é *integrável* quando ω_+ e ω_- ambas o forem. Diremos então

$$\int_M \omega = \int_M \omega_+ - \int_M \omega_- .$$

Existe uma classe importante de formas contínuas integráveis numa variedade M^n. São as formas de *suporte compacto*. Para uma tal ω, existe um conjunto compacto $K \subset M$ tal que $\omega_p = 0$ para todo $p \in M - K$. Então ω_+ e ω_- têm ambas suporte compacto, de modo que podemos supor ω positiva. Quando se toma a partição da unidade para definir $\int \omega_+$, pode-se sempre tomá-la de forma que

apenas um número finito de vizinhanças coordenadas V_i intersetam o compacto K, e assim apenas um número finito das funções φ_i são $\neq 0$ em K. Logo, a série

$$\sum \int_M \varphi_i \omega$$

que define a integral de ω é, na realidade, uma soma finita e, em particular, ω é integrável.

Consideremos agora uma forma contínua ω, de grau r, definida numa variedade M^n. Seja $S^r \subset M^n$ uma subvariedade de dimensão r. A inclusão $i\colon S \to M$ induz um homomorfismo i^* das formas sobre M nas formas sobre S. Se $i^*\omega$ for integrável, diremos que ω é integrável sobre S e poremos

$$\int_S \omega = \int_S i^*\omega.$$

Exemplos:

1. Seja $M = R$. Uma forma contínua ω, de grau 1, sobre R é do tipo $f(t)dt$ e se identifica portanto à função real contínua $f\colon R \to R$. A integral $\int_R \omega$ reduz-se à integral $\int_{-\infty}^{+\infty} f(t)\,dt$ onde a convergência é tomada no sentido *absoluto*, isto é, consideramos convergentes apenas as integrais onde $\int_{-\infty}^{+\infty} |f(t)|\,dt < \infty$. Com efeito, temos $f = f_+ - f_-$ e $|f| = f_+ + f_-$ e a integrabilidae de ω (e portanto de f) exige que $\int f_+ < \infty$ e $\int f_- < +\infty$, donde $\int |f| < \infty$. Idênticas observações podem ser feitas quando ω é uma forma de grau n sobre R^n.

2. Sejam $M = R^3$ e S uma superfície orientada em R^3. Esta orientação pode ser dada, em cada ponto, do plano tangente ou pela normal. Suponhamos que a forma contínua

$$\omega = a\,dydz + b\,dzdx + c\,dxdy$$

esteja definida num aberto $\Omega \supset S$ e que se anule fora duma região, onde S é dada pela parametrização positiva $(x(u,v), y(u,v), z(u,v))$, com $(u,v) \in D$. De acordo com a definição vista

$$\int_S \omega = \int_D \left[a\,\frac{\partial(y,z)}{\partial(u,v)} + b\,\frac{\partial(z,x)}{\partial(u,v)} + c\,\frac{\partial(x,y)}{\partial(u,v)} \right] dudv,$$

que é uma integral de superfície da Análise Clássica.

Analogamente, teríamos, como caso particular, a definição de integral de linha.

3. Volume de uma variedade riemanniana. Seja M uma variedade riemanniana orientada, de dimensão n. Uma forma importante sobre M é o *elemento de volume* σ, de grau n, definida em cada ponto $p \in M$ como

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

onde $v_1, \ldots, v_n \in M_p$ e o sinal escolhido é o sinal da base $\{v_1, \ldots, v_n\}$. No capítulo anterior, vimos que este número é o volume orientado do paralelepípedo gerado por $v_1, \ldots, v_n$. Vimos ainda que, se $\{e_1, \ldots, e_n\}$ é uma base ortonormal positiva de M_p e

$$v_i = \sum_{j=1}^{n} \alpha_{ij}\, e_j\,,$$

então $\sigma_p(v_1, \ldots, v_n) = \det(v_i \cdot e_j) = \det(\alpha_{ij})$. Esta é uma forma positiva e, se $x\colon U \to R^n$ é um sistema de coordenadas positivo,

$$\sigma_p = \sqrt{g(p)}\, dx^1 \ldots dx^n,$$

onde $g = \det(g_{ij})$ e $g_{ij} = \frac{\partial}{\partial x^i} \cdot \frac{\partial}{\partial x^j}$. Isto é óbvio porque

$$\sigma_p\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right) = \sqrt{g(p)}$$

e, como $g(p) > 0$ para qualquer $p \in M$, $\sqrt{g(p)}$ é diferenciável de mesma classe que a métrica riemanniana.

Se a variedade M é compacta, toda forma contínua é integrável, em particular σ é integrável. Define-se o *volume* de M como $\int_M \sigma$. No caso em que M é uma curva ou superfície do espaço euclidiano, este volume coincide, respectivamente, com o comprimento ou área de M.

Se a variedade M não for compacta, nem sempre seu volume será finito. O fato do volume ser finito, ou não, depende da métrica. Por exemplo, pode-se introduzir uma métrica no plano fazendo uso do difeomorfismo entre este e o disco. Com esta métrica, a área do plano será finita.

Em geral, em qualquer variedade riemanniana, é possível introduzir uma métrica relativamente à qual o volume seja finito.

3. Definição de integral de uma função. Seja $f\colon M \to R$ uma função real, contínua, definida na variedade riemanniana diferenciável orientada M, cujo elemento de volume é σ. A forma diferencial $f\sigma$, sobre M, pode ser integrável ou não. No caso de $f\sigma$ ser integrável, pode-se definir a

integral de f sobre M como

$$\int_M f = \int_M f\sigma.$$

Observação: Para somas finitas, é evidente que, sendo $w_1, \ldots, w_r$ integráveis sobre M, então $w = w_1 + \cdots + w_r$ também é integrável e

$$\int_M w = \sum_{i=1}^{r} \int_M w_i \,.$$

Mas para séries, sendo w integrável e $w = w_1 + \cdots + w_r + \ldots$, onde cada w_i é integrável, *ainda mesmo que a seqüência w_i seja localmente finita*, pode acontecer que

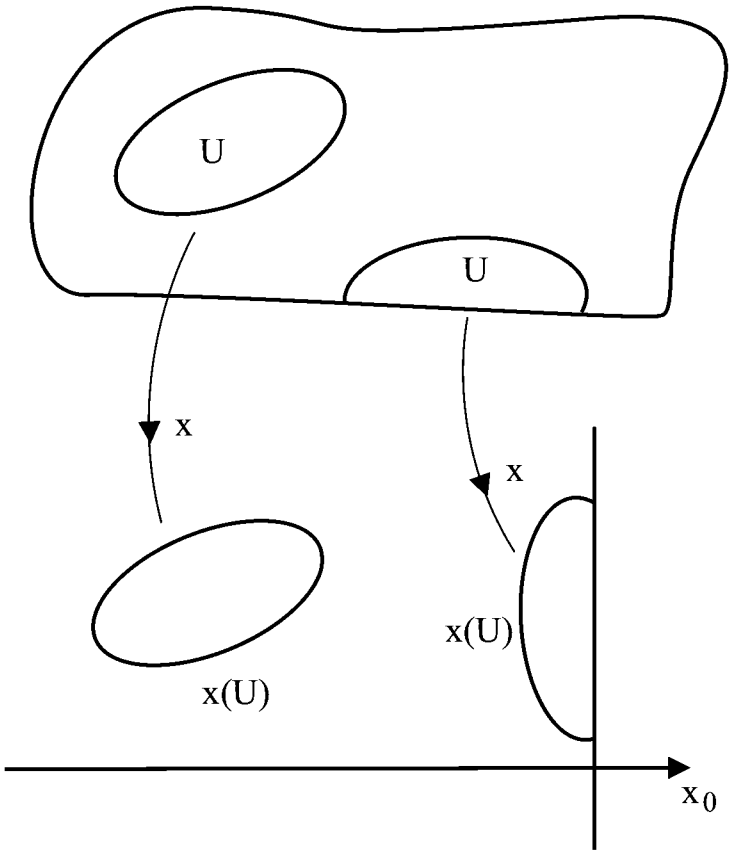
$$\int_M w \neq \sum \int_M w_i \,.$$

Esta discrepância ocorre mesmo que $\sum \int_M w_i$ seja uma série convergente. Como exemplo basta tomar $M = R =$ reta real, $w_1 = f_1$, $w_i = f_i - f_{i-1}$, $i > 1$, onde a função f_i é nula fora do intervalo $[i, 2i]$ e é constante, igual a $1/i$, neste intervalo. Então, pondo $w \equiv 0$, temos $w = \sum w_i = \lim_{i\to\infty} f_i$. (E a seqüência (w_i) é localmente finita, num sentido evidente). Temos $\int_R w = 0$ mas, para cada i,

$$\int_R f_i = 1, \quad \text{donde} \quad \sum_{i=1}^{\infty} \int w_i = \lim_{i\to\infty} \int f_i = 1.$$

4.10 Teorema de Stokes

O conceito de variedade diferenciável visto no início deste capítulo não abrange, por exemplo, o disco fechado. É nossa intenção demonstrar um teorema que contenha como caso particular as formas do teorema de Stokes para o plano e o espaço. Torna-se necessário, portanto, modificar as definições do §1 de modo a incluir em nossa teoria as situações em que o teorema é válido.



Um *sistema de coordenadas* de um espaço topológico M^{n+1} é um homeomorfismo

$$x\colon U \to R_0^{n+1}$$

do aberto U de M^{n+1} no aberto $x(U)$ do semi-espaço euclidiano

$$R_0^{n+1} = \left\{(x^0, \ldots, x^n) \in R^{n+1};\ x^0 \leq 0\right\}.$$

Esta definição contém a anterior como particular.

Daqui por diante, as definições de atlas, atlas máximo, vizinhança coordenada, orientação, campos de tensores, integral, etc... são análogos às que já foram vistas.

A única modificação a fazer diz respeito ao conceito de derivada parcial. Até aqui, só considerávamos derivadas parciais de funções definidas num subconjunto *aberto* do espaço euclidiano. Agora admitiremos também derivadas parciais de funções $f\colon \Omega \to R$, onde Ω é um subconjunto aberto do semi-espaço R_0^{n+1}. As derivadas parciais $\frac{\partial f}{\partial x^i}(p)$, para $i > 0$ são as mesmas que dantes. A derivada $\frac{\partial f}{\partial x^0}(p)$, quando $p = (0, a^1, \ldots, a^n)$ é definida como se esperava:

$$\frac{\partial f}{\partial x^0} = \lim_{\substack{t \to 0 \\ t<0}} \frac{f(t, a^1, \ldots, a^n) - f(0, a^1, \ldots, a^n)}{t}$$

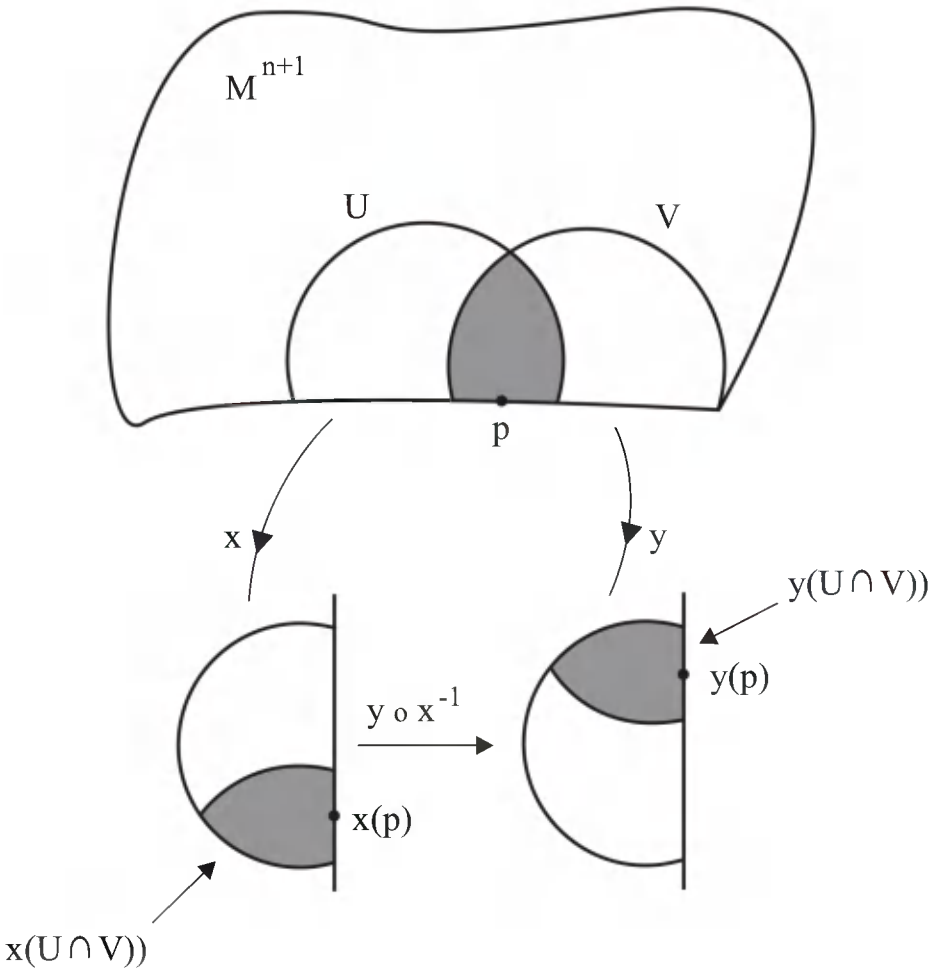
ou seja, é a derivada à esquerda $\frac{\partial f^-}{\partial x^0}(p)$. Isto é natural pois f está definida apenas para pontos $(t, a^1, \ldots, a^n)$ com $t \leq 0$.

Um ponto $p \in M^{n+1}$ é dito *ponto do bordo* se, e somente se, existe um sistema de coordenadas $x\colon U \to R_0^{n+1}$ com $p \in U$ e $x(p) = \big(0, x^1(p), \ldots, x^n(p)\big)$. O *bordo* da variedade M^{n+1} (conjunto dos pontos do bordo) é indicado com ∂M. Quando ∂M é vazio, M^{n+1} é uma variedade de dimensão $n+1$, de acordo com a definição do §1. Quando $\partial M \neq \varnothing$, M^{n+1} é dita *variedade com bordo.*

Observa-se que, se existe um sistema de coordenadas x, com $x^0(p) = 0$, então isto se dará para qualquer outro sistema definido em torno de p, isto é, se $x\colon U \to R_0^{n+1}$ e $y\colon V \to R_0^{n+1}$ são sistemas de coordenadas definidos em torno de $p(p \in U \cap V)$ e $x^0(p) = 0$, então também $y^0(p) = 0$. Sabemos que, se x e y são sistemas de coordenadas, a aplicação

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

é um difeomorfismo.



Suponhamos, por absurdo, que $y^0(p) < 0$. Indiquemos com Ω o interior de $y(U \cap V)$ em R^{n+1}. Então

$$\Omega = \{(y^0, \ldots, y^n) \in y(U \cap V);\ y^0 < 0\}$$

e $y(p) \in \Omega$. Consideremos a aplicação $\varphi\colon \Omega \to R^{n+1}$ dada por $\varphi(y(q)) = (x \circ y^{-1})(y(q)) = x(q)$, $q \in \Omega$. Ou seja, φ é a restrição de $x \circ y^{-1}$ ao aberto Ω. Pelo teorema das funções inversas (pois o jacobiano de φ é $\neq 0$ em todos os pontos de Ω), $\varphi(\Omega)$ é aberto em R^{n+1}. Mas é claro que $x(p) \in \varphi(\Omega) \subset x(U \cap V)$. Segue-se que $x(p)$ pertence ao interior de $x(U \cap V)$ em R^{n+1}, o que contraria a hipótese $x^0(p) = 0$.

Lembramos aqui que as derivadas parciais relativas à primeira coordenada, num ponto do bordo, devem ser calculadas somente à esquerda.

Verificaremos ainda que *o bordo ∂M da variedade M^{n+1} é uma variedade (sem bordo) de dimensão n.*

Para isso, começamos por observar que, se uma vizinhança coordenada U contém um ponto p do bordo, deve conter uma bola do bordo a que p pertença. Isto porque sua imagem, sendo aberta em R^{n+1}_0, é a intersecção de um aberto do espaço euclidiano R^{n+1} com o semi-espaço R^{n+1}_0.

Isto posto, construimos um atlas sobre ∂M a partir de um atlas sobre M: se $x\colon U \to R^{n+1}_0$ é um sistema de coordenadas sobre M, com $U \cap \partial M \neq \varnothing$, então $\overline{x}\colon U \cap \partial M \to R^n$, que consiste em tomar a restrição de x a $U \cap \partial M$ e desprezar a primeira coordenada (que é sempre 0), será um sistema de coordenadas em ∂M. Explicitamente, se $p \in U \cap \partial M$ e

$$x(p) = \big(0, x^1(p), \ldots, x^n(p)\big),$$

então

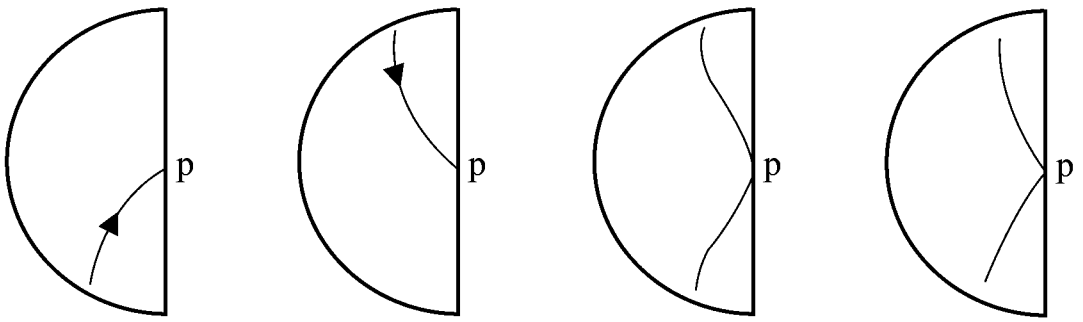
$$\overline{x}(p) = \big(x^1(p), \ldots, x^n(p)\big).$$

Exemplos. A bola fechada $B^n = \{t \in R^n;\ |t| \leq 1\}$ é uma variedade com bordo e $\partial B^n = S^{n-1}$.

Se M é uma variedade sem bordo e I é o intervalo fechado $[0, 1]$, a variedade produto $M \times I$ é uma variedade com bordo e $\partial(M \times I) = (M \times \{0\}) \cup (M \times \{1\})$.

Se M e N são ambas variedades com bordo, o produto $M \times N$ não possui uma estrutura natural de variedade diferenciável. Por exemplo, o quadrado $I \times I$ possui 4 pontos singulares (angulosos) no bordo.

Se $p \in \partial M$ é um ponto do bordo de M^{n+1}, as curvas por p são de diferentes tipos: há aquelas que começam (ou terminam) em p, aquelas que, em p, tangenciam o bordo ou aquelas que passam por p sem tangenciar (devem ter vetor tangente nulo, caso sejam diferenciáveis). Compostas com um sistema de coordenadas, elas têm os seguintes aspectos no espaço euclidiano:



Observe-se que o vetor tangente a cada uma das curvas acima no ponto p, aponta para fora de M^{n+1}, para dentro de M^{n+1}, tangencia M^{n+1} ou é nulo, respectivamente. Isto ilustra o fato de que o espaço vetorial tangente M_p, mesmo num ponto $p \in \partial M$, é um espaço vetorial de dimensão

$n+1$ (e não um semi-espaço ou um espaço de dimensão n). Na realidade, a definição do espaço M_p é absolutamente idêntica à que foi dada anteriormente no caso sem bordo.

É claro que para $p \in \partial M$, há dois espaços vetoriais que se podem considerar no ponto p: o espaço M_p, tangente a M e o espaço $(\partial M)_p$, tangente à subvariedade ∂M. O primeiro tem dimensão $n+1$ e o segundo tem dimensão n.

Proposição 7. *Se a variedade com bordo M for orientável, sem bordo também o será.*

Demonstração: Seja $\mathfrak{A}$ um atlas coerente sobre M. Partindo deste, introduzimos um atlas coerente $\mathfrak{A}'$ sobre ∂M do mesmo modo como fizemos acima para construir a estrutura de variedade diferenciável sobre ∂M.

Sejam x e y sistemas de $\mathfrak{A}$ e seus correspondentes em $\mathfrak{A}'$, x' e y'. Para verificar que $\mathfrak{A}'$ é um atlas coerente é necessário mostrar que o jacobiano $J(y' \circ (x')^{-1})$ da mudança de coordenadas $y' \circ (x')^{-1}$ é sempre positivo.

Ora, $J(y \circ x^{-1}) > 0$ porque $\mathfrak{A}$ é um atlas coerente, mas, num ponto $p \in \partial M$ da intersecção dos domínios de x e y, temos $x^0(p) = y^0(p) = 0$, donde

$$\frac{\partial y^0}{\partial x^1}(p) = \frac{\partial y^0}{\partial x^2}(p) = \cdots = \frac{\partial y^0}{\partial x^n}(p) = 0.$$

Logo:

$$J(y \circ x^{-1})_p = \frac{\partial y^0}{\partial x^0} \begin{vmatrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y^n}{\partial x^1} & \frac{\partial y^n}{\partial x^2} & \cdots & \frac{\partial y^n}{\partial x^n} \end{vmatrix} = \frac{\partial y^0}{\partial x^0} J\left(y' \circ (x')^{-1}\right)_p > 0.$$

De $x^0 \leq 0$ e $y^0 \leq 0$, segue $\frac{\partial y^0}{\partial x^0} \geq 0$, mas a igualdade não pode verificar-se pois $J(y \circ x^{-1}) \neq 0$. Conclui-se, finalmente, que $J(y' \circ (x')^{-1}) > 0$.

A orientação de ∂M definida por $\mathfrak{A}'$ diz-se *orientação induzida* pela orientação de M (determinada por $\mathfrak{A}$).

Exemplo: Seja M^n uma variedade diferenciável orientável sem bordo. Mostraremos agora como se pode obter no produto $M \times I$ um atlas coerente $\mathcal{B}$, a partir de um atlas coerente $\mathfrak{A}$ sobre M. Concluiremos, em particular que M orientável $\Rightarrow M \times I$ orientável.

Cada sistema de coordenadas positivo $x \in \mathfrak{A}$ em M, $x\colon U \to R^n$, dará origem a dois sistemas de coordenadas $\overline{x}$ e $\widetilde{x}$ em $M \times I$, e o atlas $\mathcal{B}$ será o conjunto de todos os sistemas $\overline{x}$ e $\widetilde{x}$, quando x percorre $\mathfrak{A}$. Primeiramente convencionemos indicar com $x^*\colon U \to R^n$ o sistema de coordenadas *negativo* em M, obtido de x mediante a mudança de sinal da 1a. coordenada. Isto é, $x^*(p) = (-x^1(p), \ldots, x^n(p))$. Em seguida, definamos $\overline{x}$ e $\widetilde{x}$. O domínio de $\overline{x}$ será o aberto $U \times (0,1]$ de $M \times I$ e o domínio de $\widetilde{x}$ será o aberto $U \times [0,1)$. Teremos pois $\overline{x}\colon U \times (0,1] \to R_0^{n+1}$ e $\widetilde{x}\colon U \times [0,1) \to R_0^{n+1}$ dados por

$$\begin{aligned} \overline{x}(p,t) &= (t-1, x(p)), \quad 0 < t \leq 1 \\ \widetilde{x}(p,t) &= (-t, x^*(p)), \quad 0 \leq t < 1. \end{aligned}$$

Na intersecção $U \times (0,1) = [U \times [0,1)] \cap [U \times (0,1]]$ dos domínios de $\widetilde{x}$ e $\overline{x}$, a mudança de coordenadas $\widetilde{x} \circ \overline{x}^{-1}$ tem jacobiano positivo, pois as mudanças de coordenadas $t - 1 \to -t$ e $x(p) \to x^*(p)$ têm ambas o jacobiano negativo. De modo análogo verifica-se que, em geral, se x, y são dois sistemas de coordenadas positivos arbitrários em

$\mathfrak{A}$, cujos domínios U e V têm intersecção não-vazia, então as mudanças de coordenadas $\overline{x} \circ \widetilde{y}^{-1}$, $\widetilde{x} \circ \overline{y}^{-1}$, etc. têm todas jacobiano positivo. É óbvio que os domínios de $\overline{x}$ e $\widetilde{x}$ cobrem $M \times I$, quando x percorre $\mathfrak{A}$. Segue-se que o conjunto $\mathcal{B}$ dos sistemas $\overline{x}$ e $\widetilde{x}$ é um atlas coerente sobre $M \times I$.

Observemos ainda um fato importante neste exemplo. Se identificarmos, como é natural, a subvariedade $M \times \{0\} \subset M \times I$ com a variedade M, através do difeomorfismo $i_0 \colon p \to (p, 0)$, e identificarmos também $M \times \{1\}$ com M pela correspondência $i_1 \colon p \to (p, 1)$, veremos que a orientação induzida por $M \times I$ no seu bordo $M \times \{0\} \cup M \times \{1\}$ (de acordo com o processo acima descrito de desprezar a primeira coordenada de um sistema positivo em $M \times I$) reproduz em $M \times \{1\}$ a orientação dada a M pelo atlas $\mathfrak{A}$, porém introduz em $M \times \{0\}$ a orientação oposta.

Por conseguinte, se indicarmos com M uma variedade conexa *orientada* e com $-M$ a mesma variedade com a orientação oposta, e indicarmos ainda com $M \times I$ a variedade munida da orientação dada por M, e com $\partial(M \times I)$ o bordo com a orientação induzida, poderemos escrever

$$\partial(M \times I) = (M \times \{1\}) - (M \times \{0\}).$$

Observação: Quando definimos variedade com bordo, exigimos que os sistemas de coordenadas locais tomem valores em subconjuntos abertos do semi-espaço

$$R_0^{n+1} = \left\{(x^0, x^1, \ldots, x^n) \in R^{n+1};\ x^0 \leq 0\right\}.$$

Isto é conveniente e simplifica muitos argumentos, como por exemplo na Proposição 7. Há porém uma pequena

desvantagem, para a qual chamamos a atenção do leitor. Considerando o intervalo $[0,1]$ como uma variedade com bordo, devendo os sistemas de coordenadas em torno dos pontos 0 e 1 tomar valores em $(-\infty, 0] = R_0^1$, vemos que um sistema de coordenadas em torno do 0 tem sempre derivada negativa, enquanto em torno do 1 a derivada é positiva. Tais sistemas não podem ser compatíveis. Como nós insistimos em considerar R_0^n como o único semi-espaço aceitável, o intervalo compacto $[0,1]$ não é, portanto, uma variedade orientável! Este é o inconveniente. Em dimensão maior que 1 não aparece nenhum outro absurdo desta natureza. Vale a pena, portanto, manter a definição adotada.

Teorema de Stokes. *Seja M^{n+1} uma variedade orientada, cujo bordo ∂M está dotado da orientação induzida. Se ω é uma forma de grau n e classe C^2, com suporte compacto em M, então*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Observações:

1. A hipótese de classe C^2 para ω é necessária a fim de que $d\omega$ tenha significado intrínseco. Ela foi usada quando demonstramos que a diferencial exterior estava definida invariantemente.

2. O teorema acima inclui como casos particulares as fórmulas da Análise Vetorial clássica conhecidas como Teoremas de Gauss, Green, Stokes e Ostrogradsky. Por exemplo, se $n+1 = 2$ e M^2 é uma superfície do espaço R^3, cujo bordo é a curva $\Gamma = \partial M^2$, e $\omega = a\,dx + b\,dy + c\,dz$ é uma

forma de grau 1 sobre M^2, nossa fórmula de Stokes fornece:

$$\int_M \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dydz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right) dzdx + \\ + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dxdy = \int_\Gamma a\,dx + b\,dy + c\,dz,$$

que é o teorema clássico de Stokes. (Vide R. Courant, "Differential and Integral Calculus, vol. II, pag. 386).

3. Quando M^{n+1} não é compacta, o Teorema de Stokes não é válido para uma ω de grau n e classe C^2 qualquer, mesmo que ω seja integrável sobre ∂M e $d\omega$ seja integrável sobre M. Por exemplo, consideremos o disco unitário do plano:

$$D^2 = \{(x, y) \in R^2;\ x^2 + y^2 \leq 1\}$$

e ponhamos $M^2 = D^2 - 0 =$ disco menos a origem. Temos $\partial D^2 = S^1 =$ círculo unitário $= \{(x, y) \in R^2;\ x^2 + y^2 = 1\}$. Tomemos agora a forma diferencial de grau 1 e classe C^∞ em $D^2 - 0$, definida por

$$\omega = \frac{-y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy.$$

Temos $d\omega = 0$, donde $d\omega$ é integrável sobre M^2 e $\int_{M^2} d\omega = 0$. Além disso, ω é integrável sobre $S^1 = \partial M^2$, pois S^1 é compacto. Parametrizando S^1 com $t \to$ (cost,sent, $0 \leq t \leq 2\pi$, temos

$$\int_{S^1} \omega = \int_0^{2\pi} [-\text{sent} \cdot d(\text{cost}) + \text{cost} \cdot d(\text{sent})] = \\ = \int_0^{2\pi} (\text{sen}^2\text{t} + \cos^2\text{t})dt = 2\pi.$$

Portanto, $\int_M d\omega \neq \int_{\partial M} \omega$.
Assim, hipóteses adicionais sobre o comportamento de ω são necessárias a fim de manter a validez da fórmula de Stokes. A hipótese de ω ter suporte compacto não é, com certeza, a mais geral possível, mas é suficiente para a maioria das aplicações. Note-se que sup.ω compacto $\Rightarrow$ sup.dω compacto.

Demonstração do teorema: Suponhamos, inicialmente que o suporte de ω esteja contido dentro do domínio U de um sistema de coordenadas positivo $x\colon U \to R_0^{n+1}$. Neste sistema, pode-se escrever

$$\omega = \sum_{j=0}^{n} a_j \, dz^0 \ldots d\widehat{x}^j \ldots dx^n,$$

(onde o sinal $\widehat{\ }$ sobre um objeto significa que este objeto deve ser omitido).
Então, para a diferencial exterior $d\omega$, tem-se

$$d\omega = \sum_{j=0}^{n} da_j \wedge dx^0 \ldots d\widehat{x}^j \ldots dx^n.$$

Mas $da_j = \sum_{k=0}^{n} \dfrac{\partial a_j}{\partial x^k} \, dx^k$, donde

$$d\omega = \sum_{j=0}^{n} \frac{\partial a_j}{\partial x^j} \, dx^j \, dx^0 \ldots d\widehat{x}^j \ldots dx^n.$$

Vamos distinguir dois casos, conforme U contenha ou não, pontos do bordo.

1º caso: $U \cap \partial M = \varnothing$. Neste caso, é óbvio que

$$\int_{\partial M} \omega = \int_{\partial M} i^*\omega = 0,$$

já que $\omega = 0$ fora de U.

Demonstramos que também o primeiro membro é nulo, isto é, que $\int_M d\omega = 0$. Sendo $x(U)$ limitado, pode-se considerá-lo contido num cubo fechado K, de arestas paralelas aos eixos, ao qual se estendem as funções a_j fazendo-as iguais a zero em $K - x(U)$. Isto não altera a continuidade de a_j.

Obtém-se então

$$\int_M d\omega = \sum_{j=0}^{n} \int_K \frac{\partial a_j}{\partial x^j}\, dx^j dx^0 \ldots d\widehat{x^j} \ldots dx^n.$$

Procedendo-se à integração, relativamente a x^j, aparecem as diferenças entre valores de a_j nas faces do cubo K e estes são nulos, portanto

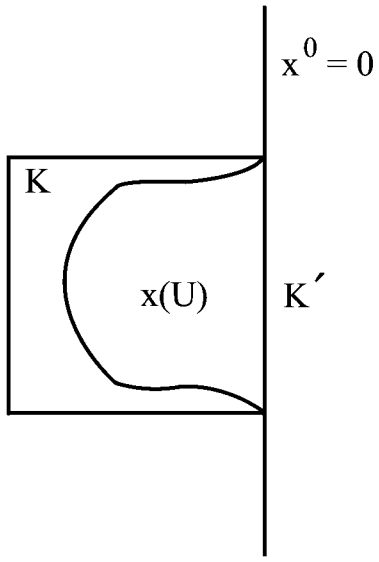
$$\int_M d\omega = 0.$$

2º caso: $U \cap \partial M \neq \varnothing$. Mais uma vez, a aplicação $i\colon \partial M \to M$ é a aplicação de inclusão do bordo em M, então $i^*\omega = a_0\, dx^1 \wedge \cdots \wedge dx^n$ porque restringir-se ao bordo significa fazer $x^0 = 0$ e, conseqüentemente, $dx^0 = 0$.

Então $\int_{\partial M} \omega = \int_{\partial M} a_0\, dx^1 \ldots dx^n$. Observe-se que estamos levando em consideração o fato de que a orientação de ∂M é a induzida pela de M.

Consideremos, agora, um cubo fechado K, de arestas paralelas aos eixos, que contenha $x(U)$ e que tenha uma de suas faces, K', no hiperplano $x^0 = 0$. Estendemos as funções a_j , analogamente ao caso anterior, fazendo-as nulas onde não estão definidas, então

$$\int_{\partial M} \omega = \int_{K'} a_0 \, dx^1 \ldots dx^n.$$



Por outro lado,

$$\int_M d\omega = \sum_{j=0}^{n} \int_K \frac{\partial a_j}{\partial x^j} \, dx^j \, dx^0 \ldots d\widehat{x^j} \ldots dx^n.$$

Efetuando a integração, relativamente a x^j, aparecem novamente as diferenças entre valores de a_j nas faces do cubo K e estes são nulos, com exceção de um – aquele correspondente à face K' $(j = 0)$, isto é

$$\int_M d\omega = \int_{K'} a_0 \, dx^1 \ldots dx^n.$$

Em seguida, mostremos que a situação geral, onde ω é uma forma com suporte compacto K arbitrário, reduz-se à anterior. Com efeito, seja $\varphi_1, \ldots, \varphi_i, \ldots$ uma partição da unidade subordinada a uma cobertura enumerável $U_1, \ldots, U_i, \ldots$ como no Lema 2. Existe um inteiro r tal que $U_1, \ldots, U_r$ cobrem o compacto K = suporte de ω. Como $\omega = 0$ fora de K, temos $\varphi_i \omega = 0$ para $i > r$. Logo:

$$\omega = \sum_{i=1}^{r} \varphi_i \omega = \sum_{i=1}^{r} \omega_i \,,$$

onde $\omega_i = \varphi_i \omega$. Assim, $\displaystyle\int_{\partial M} \omega = \sum_{i=1}^{r} \int_{\partial M} \omega_i$. Por outro lado $d\omega = \displaystyle\sum_{i=1}^{r} d\omega_i$, donde $\displaystyle\int_M d\omega = \sum_{i=1}^{r} \int_{\partial M} d\omega_i$. Além disso o suporte de cada ω_i está contido na vizinhança coordenada U_i. Então, de acordo com a primeira parte do teorema,

$$\int_{\partial M} \omega_i = \int_M d\omega_i \,, \quad i = 1, \ldots, r.$$

Logo $\displaystyle\int_{\partial M} \omega = \int_M d\omega$, como queríamos demonstrar.

Observações:

1. Freqüentemente, a variedade com bordo M^{n+1} está imersa numa variedade maior N^r e ω é uma forma sobre N. A fórmula de Stokes $\displaystyle\int_{\partial M} \omega = \int_M d\omega$ ainda é válida neste caso. Basta lembrar a definição de integral de uma forma

sobre uma subvariedade. Sejam $i\colon M \to N$ e $j\colon \partial M \to M$ as inclusões. Por definição é

$$\int_{\partial M} \omega = \int_{\partial M} j^* i^* \omega \quad \text{e} \quad \int_M d\omega = \int_M d\omega = \int_M i^* d\omega.$$

Mas $i^* d\omega = d(i^*\omega)$ e vale a fórmula de Stokes em M. Logo

$$\int_{\partial M} j^* i^* \omega = \int_M d(i^*\omega) = \int_M i^* d\omega, \text{ ou seja } \int_{\partial M} \omega = \int_M d\omega.$$

2. Mesmo no caso elementar de uma integral curvilínea no plano, encontram-se expressões do tipo

$$\int_\Gamma = \int a\,dx + b\,dy,$$

onde Γ não é necessariamente uma curva simples, isto é, uma subvariedade de dimensão 1 do plano. Muitas vezes, Γ possui auto-interseções: é simplesmente uma "curva parametrizada", definida por uma aplicação diferenciável $f\colon I \to R^2$, onde I é um intervalo da reta. Este caso não está exatamente incluido na definição geral que demos para $\int_M \omega$, onde exigimos que M seja uma variedade. Nas integrais curvilíneas, se Γ é definida por $t \to (x(t), y(t))$, põe-se

$$\int_\Gamma \omega = \int_\alpha^\beta \left[a(x(t), y(t)) \frac{dx}{dt} + b(x(t), y(t)) \frac{dy}{dt} \right] dt,$$

onde α e β são os extremos do intervalo I. Notemos que isto equivale a definir $\int_\Gamma \omega = \int_\alpha^\beta f^*\omega$, ou seja $\int_{f(I)} \omega = \int_I f^*\omega$.

Isto é o que faremos em geral. Sejam M, N variedades orientáveis (com ou sem bordo) e $f\colon M \to N$ uma aplicação diferenciável. Dada uma forma diferencial ω sobre N, definiremos $\int_{f(M)} \omega$ como sendo $\int_M f^*\omega$. Se escrevermos $\Gamma = f(M)$, Γ será uma espécie de "subvariedade" com auto-intersecções, pontos angulosos, etc. E f será uma "parametrização" de Γ. Ainda nesta situação geral, o Teorema de Stokes é válido. Se M possui um bordo ∂M, podemos escrever $\partial\Gamma = f(\partial M)$ e vale a fórmula $\int_{\partial\Gamma} \omega = \int_\Gamma d\omega$, significando que

$$\int_{\partial M} f^*\omega = \int_M f^*(d\omega) = \int_M d(f^*\omega).$$

(Supomos, naturalmente, que o grau de ω é igual à dimensão de ∂M).

3. Para uma forma arbitrária, a demonstração do Teorema de Stokes falha no seguinte ponto: seja $\omega = \sum \varphi_i\omega = \sum \omega_i$. Temos $\int_{\partial M} \omega = \sum_i \int_{\partial M} \omega_i$ e, sendo a soma $\sum \omega_i$ finita em cada ponto, vale também $d\omega = \sum d\omega_i$. Mas não é verdade que $\int_M d\omega = \sum \int_M d\omega_i$. (Cfr. observação em seguida à definição de integral.)

O Teorema de Stokes possui uma interpretação interessante em termos da dualidade existente entre as formas diferenciais (de classe ≥ 2 e suporte compacto) sobre uma variedade N^n e as subvariedades de N.

Sejam $\mathcal{F}$ o conjunto das formas diferenciais de classe C^2 e suporte compacto sobre N e $\mathcal{J}$ o conjunto das subvariedades orientadas de N. Temos $\mathcal{F} = \mathcal{F}^0 \cup \cdots \cup \mathcal{F}^n$ e

$\mathcal{J} = \mathcal{J}^0 \cup \cdots \cup \mathcal{J}^n$, onde $\mathcal{F}^r$ é o conjunto das formas de grau r e $\mathcal{J}^r$ é o conjunto das subvariedades de dimensão r. A dualidade acima referida consiste no seguinte: para cada r, $0 \leq r \leq n$, existe uma aplicação

$$\mathcal{J}^r \times \mathcal{F}^r \to R$$

que associa a cada subvariedade orientada M^r e cada forma $\omega \in \mathcal{F}^r$ o número real

$$\langle M^r, \omega \rangle = \int_M \omega.$$

Esta aplicação é bilinear no sentido de que

$$\langle M, \omega + \theta \rangle = \langle M, \omega \rangle + \langle M, \theta \rangle$$

e

$$\langle M \cup P, \omega \rangle = \langle M, \omega \rangle + \langle P, \omega \rangle$$

se M e P são disjuntas. O Teorema de Stokes exprime que o operador de bordo $\partial\colon \mathcal{J}^r \to \mathcal{J}^{r-1}$ e o operador de diferencial exterior $d\colon \mathcal{F}^r \to \mathcal{F}^{r+1}$ são *adjuntos* um do outro, relativamente a esta dualidade. (Estamos considerando em cada $\mathcal{J}^r$, a subvariedade vazia $\varnothing \in \mathcal{J}^r$, a fim de podermos falar em ∂M^r quando M^r é uma subvariedade sem bordo). Com efeito, na notação acima, o Teorema de Stokes se escreve

$$\langle \partial M^{r+1}, \omega \rangle = \langle M^{r+1}, d\omega \rangle, \ \ \omega \text{ de grau } r.$$

Por analogia, diremos que a forma diferencial ω é *fechada* quando $d\omega = 0$. Por exemplo, uma forma $\omega = a\,dx + b\,dy$ no plano é fechada se, e somente se, $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$.

Diremos que duas aplicações diferenciáveis $f, g\colon M^r \to N^n$ são *homotópicas* quando existir uma aplicação diferenciável $H\colon M \times I \to N$ (chamada *homotopia* entre f e g) tal que

$$H(p,0) = f(p) \quad \text{e} \quad H(p,1) = g(p), \text{ para todo } p \in M.$$

Dadas uma aplicação $f\colon M \to N$ e uma forma de suporte compacto sobre N, pode acontecer que $f^*\omega$ não tenha suporte compacto sobre M. Isto será verdade, porém, se f for uma aplicação *própria*, isto é, se $f^{-1}(K)$ for compacto para todo compacto $K \subset N$. Para evitar essas hipóteses adicionais, e simplificar os enunciados, é que suporemos, na proposição seguinte, que M é compacta.

Se M é orientada e compacta e $H\colon M \times I \to N$ é uma homotopia entre f e g, então tem-se evidentemente:

$$\int_{M\times 0} H^*\omega = \int_M f^*\omega \quad \text{e} \quad \int_{M\times 1} H^*\omega = \int_M g^*\omega.$$

Proposição 8. *Sejam $f, g\colon M^r \to N$ aplicações homotópicas, M orientada e compacta. Seja ω uma forma diferencial fechada de grau r, sobre N. Então*

$$\int_M f^*\omega = \int_M g^*\omega.$$

Demonstração: Como $d\omega = 0$, lembrando que

$$\partial(M \times I) = (M \times 1) - (M \times 0),$$

temos:

$$0 = \int_{M\times I} H^*(d\omega) = \int_{M\times I} d(H^*\omega) = \int_{\partial(M\times I)} H^*\omega =$$
$$= \int_{M\times 1} H^*\omega - \int_{M\times 0} H^*\omega = \int_M g^*\omega - \int_M f^*\omega$$

como queríamos demonstrar. Acima, $M \times I$ está munida da orientação fornecida por M, conforme foi explicado antes.

Considerando $f\colon M^r \to N$ e $g\colon M^r \to N$ como parametrizações dos conjuntos $\Gamma = f(M^r)$ e $\Lambda^r = g(M^r)$, de acordo com a notação anteriormente introduzida, podemos enunciar o teorema acima dizendo:

Para Γ^r e Λ^r homotópicos em N e ω fechada, tem-se

$$\int_{\Gamma^r} \omega = \int_{\Lambda^r} \omega.$$

Em particular, se Γ e Λ são curvas fechadas numa região N do plano, onde está definida a forma $\omega = a\,dx + b\,dy$, reobtemos o fato de que as integrais

$$\int_{\Gamma} a\,dx + b\,dy \quad \text{e} \quad \int_{\Lambda} a\,dx + b\,dy,$$

com $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$, coincidem quando as curvas Γ e Λ são homotópicas em N. Quando N é uma região *simplesmente conexa* (isto é: uma curva fechada em N é sempre homotópica a um ponto) tem-se então $\int_{\Gamma} a\,dx + b\,dy = 0$ sempre que Γ for uma curva fechada e $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$.

O Teorema de Stokes pode também ser considerado sob o ponto de vista da integração sobre uma "cadeia diferenciável", mais dentro do espírito da Topologia Algébrica. Para um tratamento sob este aspecto, o leitor poderá consultar o livro "Homologia Básica", do autor.

Capítulo 5

Sistemas Diferenciais

5.1 O colchete de Lie de 2 campos vetoriais

Sejam M uma variedade diferenciável e $v \in M_p$ um vetor tangente a um ponto $p \in M$. Para cada função real diferenciável f, definida numa vizinhança de p, podemos considerar a derivada direcional $\frac{\partial f}{\partial v}(p) = df(p) \cdot v$ a qual, como se recorda, é o limite

$$\lim_{t \to 0} \frac{f(\lambda(t)) - f(p)}{t}$$

onde $\lambda \colon (-\varepsilon, \varepsilon) \to M$ é qualquer caminho diferenciável tal que $\lambda(0) = p$ e o vetor velocidade $\lambda'(0)$ de λ no ponto p é v. Além disso, se $x \colon U \to R^n$ é um sistema de coordenadas válido na vizinhança de p e $v = \sum_{i=1}^{n} \alpha^i \frac{\partial}{\partial x^i}(p)$ é a expressão de v relativamente à base $\left\{\frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\right\} \subset M_p$ de-

terminada por x, tem-se

$$\frac{\partial f}{\partial v}(p) = \sum_{i=1}^{n} \alpha^i \frac{\partial f}{\partial x^i}(p)$$

onde $\frac{\partial f}{\partial x^i}(p) = \frac{\partial (f \circ x^{-1})}{\partial x^i}(x(p))$, por definição.

Seja agora X um campo vetorial sobre a variedade M, isto é, uma correspondência que associa a cada ponto $p \in M$ um vetor $X(p) \in M_p$, tangente a M no ponto p. Dado um sistema de coordenadas locais $x\colon U \to R^n$ em M, para cada $p \in U$, teremos

$$X(p) = \sum_{i=1}^{n} \alpha^i(p) \frac{\partial}{\partial x^i}(p).$$

O campo X determina assim, para cada sistema de coordenadas locais $x\colon U \to R^n$, n funções reais $\alpha^i\colon U \to R$, que são as coordenadas de X relativamente às bases definidas por x nos vários pontos de U. Diremos que X é um campo de classe C^r $(0 \leq r \leq \infty)$ se, para todo sistema x, as funções $\alpha^i\colon U \to R$ são de classe C^r. Este conceito tem sentido sempre que m é uma variedade de classe C^{r+1}. Com efeito, se noutro sistema de coordenadas $y\colon V \to R^n$ tivermos $X(p) = \sum_{i=1}^{n} \beta^i(p) \frac{\partial}{\partial y^i}(p)$, então, para cada $p \in U \cap V$,

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

onde $\frac{\partial y^i}{\partial x^j}\colon x(U \cap V) \to R$ são funções de classe C^r.

Dados um campo de vetores X de classe C^{r-1} e uma função de classe C^r, $f\colon A \to R$, definida num aberto $A \subset M$, podemos obter uma nova função, de classe C^{r-1},

$$\frac{\partial f}{\partial X}\colon A \to R,$$

definida em A, chamada a *derivada* de f relativamente ao campo X. Para cada ponto $p \in A$, poremos

$$\frac{\partial f}{\partial X}(p) = \frac{\partial f}{\partial X(p)}(p) = df(p) \cdot X(p).$$

(O campo X basta estar definido em A.)

Relativamente a um sistema de coordenadas $x\colon U \to R^n$, com $U \subset A$, temos

$$\frac{\partial f}{\partial X}(p) = \sum_{i=1}^{n} \alpha^i(p)\, \frac{\partial f}{\partial x^i}(p), \quad p \in U,$$

onde $X(p) = \sum \alpha^i(p)\, \frac{\partial}{\partial x^i}(p)$. Por exemplo: $\frac{\partial x^k}{\partial X} = \alpha^k$.

Se Y é outro campo vetorial de classe C^{r-1}, podemos ainda considerar a nova função real, de classe C^{r-2},

$$\frac{\partial^2 f}{\partial Y \partial X} = \frac{\partial}{\partial Y}\left(\frac{\partial f}{\partial X}\right) : A \to R,$$

que é simplesmente a derivada da função $\partial f/\partial X$ relativamente ao campo Y. Em termos do sistema de coordenadas x, teremos:

$$\frac{\partial^2 f}{\partial Y \partial X} = \sum_{i,j} \beta^i \left(\frac{\partial \alpha^j}{\partial x^i}\, \frac{\partial f}{\partial x^j} + \alpha^j\, \frac{\partial^2 f}{\partial x^i \partial x^j}\right), \qquad (1)$$

onde $Y(p) = \Sigma\, \beta^i(p)\, \frac{\partial}{\partial x_i}(p)$. Vê-se imediatamente que, em geral $\frac{\partial^2 f}{\partial X \partial Y} \neq \frac{\partial^2 f}{\partial Y \partial X}$.

Teorema 1. *Sejam X, Y campos vetoriais de classe C^r $(r \geq 1)$ sobre uma variedade M. Existe um único campo vetorial Z, de classe C^{r-1}, sobre M, tal que*

$$\frac{\partial f}{\partial Z} = \frac{\partial^2 f}{\partial X \partial Y} - \frac{\partial^2 f}{\partial Y \partial X}$$

para toda função $f \in C^2$, definida sobre um aberto de M.

Demonstração: Notemos inicialmente que se M é uma variedade de classe C^r e Z, Z' são campos vetoriais definidos sobre um aberto $A \subset M$, tais que $\partial f/\partial Z = \partial f/\partial Z'$ para toda função real f de classe C^r, cujo domínio está contido em A, então $Z = Z'$. Com efeito, as coordenadas de Z e Z' relativamente a um sistema de coordenadas locais x são $\partial x^i/\partial Z$ e $\partial x^i/\partial Z'$ respectivamente.

Resulta daí que se o campo Z existir, ele será único e, em virtude da igualdade (1) acima, em cada domínio U de um sistema de coordenadas locais $x\colon U \to R^n$, onde

$$X = \sum_{i=1}^{n} \alpha^i \frac{\partial}{\partial x^i} \quad \text{e} \quad Y = \sum_{i=1}^{n} \beta^i \frac{\partial}{\partial y^i},$$

deveremos ter

$$\frac{\partial f}{\partial Z} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha^i \frac{\partial \beta^j}{\partial x^i} - \beta^i \frac{\partial \alpha^j}{\partial x^i} \right) \frac{\partial f}{\partial x^j}. \tag{2}$$

Então, pondo $Z = \sum \gamma^i \frac{\partial}{\partial x^i}$, vem $\gamma^i = \frac{\partial x^i}{\partial Z}$ e a fórmula (2) dá:

$$\gamma^i = \sum_{j=1}^{n} \left(\alpha^j \frac{\partial \beta^i}{\partial x^j} - \beta^j \frac{\partial \alpha^i}{\partial x^j} \right). \tag{3}$$

Consideremos então, para cada sistema de coordenadas locais $x\colon U \to R^n$ em M, o campo vetorial

$$Z_x = \sum_{i=1}^{n} \gamma^i \frac{\partial}{\partial x^i},$$

definido apenas sobre U, onde os γ^i são obtidos através da fórmula (3). É imediato que, sobre U, tem-se

$$\frac{\partial f}{\partial Z_x} = \frac{\partial^2 f}{\partial X \partial Y} - \frac{\partial^2 f}{\partial Y \partial X}.$$

Por conseguinte, se $y\colon V \to R^n$ é outro sistema de coordenadas locais com $U \cap V \neq \varnothing$, tem-se em $U \cap V$ $\frac{\partial f}{\partial Z_x} = \frac{\partial f}{\partial Z_y}$ e portanto $Z_x = Z_y$ em $U \cap V$. Conclui-se daí que os vários Z_x definem um único campo vetorial Z sobre M, o qual cumpre as condições requeridas.

Definição. Dados dois campos vetoriais X, Y sobre uma variedade M, ambos de classe C^r $(r \geq 1)$, chama-se *colchete de Lie* desses campos ao campo vetorial $Z = [X, Y]$, de classe C^{r-1}, caracterizado pela propriedade

$$\frac{\partial f}{\partial Z} = \frac{\partial^2 f}{\partial X \partial Y} - \frac{\partial^2 f}{\partial Y \partial X}$$

onde f é uma função qualquer de classe C^2 sobre um aberto de M.

A existência e unicidade do colchete $[X, Y]$ estando estabelecidas no teorema anterior, passamos a registrar aqui algumas propriedades formais desta operação.

Sejam $a, b \in R$ *constantes*, X, X_1, Y, Y_1 campos de classe C^r, $r \geq 1$, e f, g funções diferenciáveis. Então:

$1^{\underline{o}}$) $[aX + bX_1, Y] = a[X, Y] + b[X_1, Y]$
$[X, aY + bY_1] = a[X, Y] + b[X, Y_1]$

$2^{\underline{o}}$) $[X, Y] = -[Y, X]$. Em particular:
$[X, X] = 0$

$3^{\underline{o}}$) $[fX, gY] = f \cdot g \cdot [X, Y] + f \cdot \frac{\partial g}{\partial X} \cdot Y - g \cdot \frac{\partial f}{\partial Y} \cdot X$

$4^{\underline{o}}$) $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$.

Exemplos:

1) Seja $x\colon U \to R^n$ um sistema de coordenadas locais numa variedade M. O aberto $U \subset M$ é evidentemente uma variedade, sobre a qual estão definidos os n campos vetoriais $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$. Para $1 \leq i, j \leq n$ quaisquer, temos

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0.$$

2) Sejam X, Y os campos vetoriais sobre o plano R^2, definidos por $X(x, y) = (x, y)$ e $Y(x, y) = (0, 1)$ para todo $(x, y) \in R^2$. Tem-se $[Y, X] = Y$.

3) Sejam X e Y *campos vetoriais lineares* no R^n, isto é, existem transformações lineares $A, B\colon R^n \to R^n$ tais que $X(x) = Ax$ e $Y(x) = Bx$ para todo $x \in R^n$. Então $[X, Y](x) = (AB - BA)x$.

4) Sobre o domínio U de um sistema de coordenadas $x\colon U \to R^n$, seja $X = \sum_{i=1}^{n} a^i \frac{\partial}{\partial x^i}$ um campo vetorial definido pelas funções reais $a^i\colon U \to R$. Para cada $k = 1, 2, \ldots, n$, tem-se

$$\left[\frac{\partial}{\partial x^k}, X\right] = \sum_{i=1}^{n} \frac{\partial a^i}{\partial x^k} \cdot \frac{\partial}{\partial x^i}.$$

Com efeito, a i-ésima coordenada de $Z = \left[\frac{\partial}{\partial x^k}, X\right]$ relativamente ao sistema x é

$$\frac{\partial x^i}{\partial Z} = \frac{\partial^2 x^i}{\partial x^k \partial X} - \frac{\partial^2 x^i}{\partial X \partial x^k} = \\ = \frac{\partial}{\partial x^k}\left(\frac{\partial x^i}{\partial X}\right) - \frac{\partial}{\partial X}\left(\frac{\partial x^i}{\partial x^k}\right) = \frac{\partial a^i}{\partial x^k}.$$

5.2 Relações entre colchetes e fluxos

Seja X um campo vetorial sobre uma variedade diferenciável M. Uma *trajetória* (ou *órbita*, ou *curva integral*) de X é um caminho $\varphi\colon J \to M$, de classe C^1, definido num intervalo aberto J da reta, cujo vetor velocidade $\frac{d\varphi}{dt} = \varphi'(t)$ em todo ponto $p = \varphi(t)$ é igual ao vetor $X(p)$ dado pelo campo X. Se, relativamente a um sistema de coordenadas $x\colon U \to R^n$, válido no aberto $U \subset M$, o caminho φ é definido por

$$t \to (x^1(t), \ldots, x^n(t))$$

e o campo vetorial X é representado por n funções reais α^i, então a condição para que φ seja uma trajetória de X se exprime por meio das equações

$$\frac{dx^i}{dt} = \alpha^i(x^1(t), \ldots, x^n(t)), \quad i = 1, 2, \ldots, n.$$

Diremos que uma trajetória $\varphi\colon J \to M$ tem *origem* p quando $0 \in J$ e $\varphi(0) = p$. Segue-se do teorema clássico de existência para equações diferenciais ordinárias que, dado

um campo vetorial de classe C^1 sobre uma variedade diferenciável M, por cada ponto $p \in M$ passa uma trajetória de origem p. O teorema clássico de unicidade interpreta-se, em nosso contexto, do seguinte modo: entre as trajetórias de X com origem p existe uma que é *máxima*, isto é, qualquer trajetória de X com origem p é uma restrição desta a um subintervalo menor. De agora em diante, quando nos referirmos à trajetória do campo X com origem p, estaremos querendo dizer a trajetória máxima.

Outro teorema básico sobre equações diferenciais diz que "as soluções de um sistema com dados diferenciáveis dependem diferenciavelmente das condições iniciais". Isto significa que se $X \in C^r$ e se $t \to \varphi(t,p) = \varphi_t(p)$ for a trajetória (máxima) de X com origem no ponto $p \in M$, a qual se caracteriza pelas propriedades $\varphi(0,p) = p$, $\frac{d}{dt}\varphi(t,p) = X(\varphi(t,p))$, então $\varphi(t,p) \in M$ depende diferenciavelmente de t e de p. Mais precisamente, para cada $p \in M$ existem uma vizinhança $V \ni p$ e um número $\varepsilon > 0$ tais que a trajetória máxima com origem em qualquer ponto $q \in V$ está definida pelo menos para $-\varepsilon < t < \varepsilon$ e $\varphi\colon (-\varepsilon, +\varepsilon) \times V \to M$, dada por $(t,q) \to \varphi(t,q)$, é uma aplicação de classe C^r.

Uma conseqüência da unicidade das trajetórias é que, para todo $p \in M$, $\varphi_s[\varphi_t(p)] = \varphi_{s+t}(p)$. Com efeito, fixando s arbitrariamente, consideremos o caminho $\lambda(t) = \varphi(s+t,p)$. Temos $\lambda(0) = \varphi_s(p)$ e $\lambda'(t) = \varphi'(s+t,p) = X(\varphi(s+t,p)) = X(\lambda(t))$. Logo λ é parte da trajetória máxima de X com origem $\varphi_s(p)$. Ora, tal trajetória é $t \to \varphi_t[\varphi_s(p)]$. Logo

$$\varphi(s+t,p) = \lambda(t) = \varphi_t[\varphi_s(p)],$$

como afirmamos.

Diremos que $\varphi_t(p)$ é o *fluxo local* definido, ou gerado, por X. Pode-se demonstrar que, quando M é compacta, as trajetórias (máximas) de X são definidas para $-\infty < t < \infty$. Então $(t,p) \to \varphi(t,p)$ é um *fluxo global* $\varphi\colon R \times M \to M$.

Teorema 2. *Seja X um campo vetorial de classe C^r $(r \geq 1)$ sobre uma variedade M. Seja $p \in M$ um ponto tal que $X(p) \neq 0$. Existe um sistema de coordenadas $x\colon U \to R^n$ de classe C^r em M, com $p \in U$, tal que $X(q) = \frac{\partial}{\partial x^1}(q)$ para todo $q \in U$.*

Demonstração: Seja $\varphi_t(q)$ o fluxo local definido por X. Escolhamos uma parametrização $\xi\colon A \to M$, ($A \subset R^n$ aberto, $\xi(A) \subset M$ aberto, $\xi =$ inversa de um sistema de coordenadas locais) tal que $0 \in A$, $\xi(0) = p$ e $\frac{\partial \xi}{\partial x^1}(0) = X(p)$. Em seguida, definamos uma nova aplicação $\zeta \cdot A \to M$ pondo $\zeta(t, x^2, \ldots, x^n) = \varphi_t(\xi(0, x^2, \ldots, x^n))$. No ponto $0 \in A$, a matriz jacobiana de ζ é não-singular, pois $\frac{\partial \zeta}{\partial t}(0) = X(p)$, $\frac{\partial \zeta}{\partial x^i}(0) = \frac{\partial \xi}{\partial x^i}(0)$, $i \geq 2$. Restrita então a uma vizinhança menor B de 0 em R^n, a aplicação $\zeta\colon B \to M$ é uma parametrização. Seja $U = \zeta(B)$. Pondo $x = \zeta^{-1}\colon U \to R^n$, temos o teorema demonstrado.

Corolário. *Seja X um campo vetorial de classe C^r $(r \geq 1)$ sobre uma variedade M. Seja $X(p) \neq 0$. Existe um sistema de coordenadas $x\colon U \to R^n$, de classe C^r, com $p \in U$ e tal que o fluxo de X, transportado para $x(U)$ (isto é, o fluxo $x[\varphi_t(q)]$) tem a forma*

$$(t, (x^1, \ldots, x^n)) \to (x^1 + t, x^2, \ldots, x^n).$$

Basta observar que, transportado para $x(U)$, o campo X tem a forma $X(x^1, \ldots, x^n) = (1, 0, \ldots, 0)$.

Teorema 3. *Sejam X, Y campos vetoriais de classe C^r $(r \geq 1)$ sobre uma variedade M. Indiquemos com ξ_s e η_t respectivamente os fluxos locais gerados por estes campos. Se o colchete de Lie $[X, Y]$ é identicamente nulo em M, então $\xi_s \eta_t = \eta_t \xi_s$, isto é $\xi_s(\eta_t(p)) = \eta_t(\xi_s(p))$. Reciprocamente, se os dois fluxos ξ e η "comutam" neste sentido, então $[X, Y] = 0$ em todos os pontos de M.*

Demonstração: Suponhamos que $[X, Y] = 0$. Tomemos $p \in M$ arbitrariamente. Se $X(p) = Y(p) = 0$ então, pelo teorema de unicidade das trajetórias, temos necessariamente

$$\xi_s(p) = \eta_t(p) = p$$

para todo s e todo t. Segue-se que $\xi_s \eta_t(p) = \eta_t \xi_s(p)$, trivialmente. Admitamos agora que os dois vetores não se anulam simultaneamente no ponto p. Digamos que $X(p) \neq 0$. Então existe um sistema de coordenadas locais $x \colon U \to R^n$, com $p \in U$, tal que $\xi_s(x^1, \ldots, x^n) = (x^1 + s, x^2, \ldots, x^n)$, ou seja, $X = \frac{\partial}{\partial x^1}$ em U. Escrevamos $Y(x^1, \ldots, x^n) = \sum_{i=1}^{n} a^i(x^1, \ldots, x^n) \frac{\partial}{\partial x^i}$. Então

$$0 = [X, Y] = \left[\frac{\partial}{\partial x^1}, \sum a^i \frac{\partial}{\partial x^i}\right] = \sum \frac{\partial a^i}{\partial x^1} \frac{\partial}{\partial x^i}.$$

Logo

$$\frac{\partial a^1}{\partial x^1} = \frac{\partial a^2}{\partial x^1} = \cdots = \frac{\partial a^n}{\partial x^1} = 0$$

em U. Isto significa que $a^i(x^1 + s, x^2, \ldots, x^n) = a^i(x^1, x^2, \ldots, x^n)$ para todo s, $i = 1, 2, \ldots, n$. Em outras

palavras,

$$Y(x^1 + s, x^2, \ldots, x^n) = Y(x^1, \ldots, x^n).$$

Daí resulta que $\eta_t(x^1 + s, x^2, \ldots, x^n) = \eta_t(x^1, x^2, \ldots, x^n) + (x, 0, \ldots, 0)$, em virtude do teorema de unicidade, já que ambos os membros desta igualdade, considerados como funções apenas de t, são trajetórias de origem $(x^1+s, x^2, \ldots, x^n)$ relativas ao campo Y. Isto significa, porém, que

$$\eta_t \xi_s(x^1, \ldots, x^n) = \xi_s \eta_t(x^1, \ldots, x^n),$$

como queríamos provar.

Reciprocamente, seja $\xi_s \eta_t = \eta_t \xi_s$. Para cada ponto $p \in M$, temos:

$$\frac{\partial^2 f}{\partial X \partial Y}(p) = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial Y}(\xi_s(p))\right) = \frac{\partial}{\partial s}\frac{\partial}{\partial t} f(\eta_t \xi_s(p))$$

e, analogamente

$$\frac{\partial^2 f}{\partial Y \partial X}(p) = \frac{\partial}{\partial t}\frac{\partial}{\partial s} f(\xi_s \eta_t(p)),$$

as derivadas relativamente a s e t sendo calculadas nos pontos $s = 0$, $t = 0$. Segue-se que $\partial^2 f/\partial X \partial Y = \partial^2 f/\partial Y \partial X$ para toda $f \in C^2$, donde $[X, Y] = 0$.

Definição: Dois campos vetoriais X, Y de classe C^r $(r \geq 1)$ sobre uma variedade M dizem-se *comutativos* se $[X, Y] = 0$ (identicamente em M). Isto significa, como acabamos de ver, que os fluxos gerados por X e Y comutam.

Teorema 4. *Sejam $X_1, \ldots, X_k$ campos vetoriais de classe C^r $(r \geq 1)$ tais que $[X_i, Y_j] = 0$ $(i, j = 1, \ldots, k)$ sobre a*

variedade M. Seja $p \in M$ um ponto tal que os vetores $X_1(p), \ldots, X_k(p)$ são linearmente independentes. Então existe um sistema de coordenadas $x\colon U \to R^n$, de classe C^r, com $p \in U$, tal que $X_1 = \partial/\partial x^1, \ldots, X_k = \partial/\partial x^k$ em todos os pontos de U.

Demonstração: Por simplicidade de notação, consideremos apenas o caso de 3 campos X, Y, Z. Sejam ξ_s , η_t , ζ_u respectivamente os fluxos por eles gerados. Tomemos um sistema de coordenadas y em M, válido numa vizinhança de p, tal que

$$y(p) = 0 \quad \text{e} \quad \left\{ X(p), Y(p), Z(p), \frac{\partial}{\partial y^4}(p), \ldots, \frac{\partial}{\partial y^n}(p) \right\}$$

seja uma base de M_p. A aplicação

$$\varphi\colon (s, t, u, y^4, \ldots, y^n) \to \xi_s \eta_t \zeta_u [y^{-1}(0, 0, 0, y^4, \ldots, y^n)]$$

é definida numa vizinhança de $0 \in R^n$ e toma valores em M. Temos

$$\varphi(0) = p, \ \frac{\partial \varphi}{\partial t}(0) = X(p), \ \frac{\partial \varphi}{\partial s}(0) = Y(p), \ \frac{\partial \varphi}{\partial u}(0) = Z(p)$$

e $\frac{\partial \varphi}{\partial y^i}(0) = \frac{\partial}{\partial y^i}(p)$ se $i \geq 4$. Segue-se que $\varphi_*(0)\colon R^n \to M_p$ é um isomorfismo. Pelo Teorema da Função Inversa, concluimos que φ aplica uma vizinhança de $0 \in R^n$ difeomorficamente sobre uma vizinhança U de p em M. Seja $x = \varphi^{-1}\colon U \to R^n$. Afirmamos que x é o sistema de coordenadas procurado. Com efeito, para cada $q \in U$, com

$x(q) = (s, t, u, y^4, \ldots, y^n)$, temos:

$$\frac{\partial}{\partial x^1}(q) = \frac{\partial \varphi}{\partial s}(x(q)) = \frac{\partial}{\partial s}\, \xi_s[\eta_t \zeta_u y^{-1}(0,0,0,y^4,\ldots,y^n)] = \\ = X(q),$$

$$\frac{\partial}{\partial x^2}(q) = \frac{\partial \varphi}{\partial t}(x(q)) = \frac{\partial}{\partial t}\, \eta_t[\xi_s \zeta_u y^{-1}(0,0,0,y^4,\ldots,y^n)] = \\ = Y(q),$$

e, analogamente, $\frac{\partial}{\partial x^3}(q) = Z(q)$. Note-se que, para esta verificação, foi usada pela primeira vez a hipótese de que os fluxos ξ, η, ζ comutam.

Corolário. *Relativamente ao sistema x do teorema acima, o fluxo local ξ_s^i, gerado pelo campo X_i, é dado por* $\xi_s^i(x^1,\ldots, x^n) = (x^1, \ldots, x^i + s, \ldots, x^n)$.

5.3 Sistemas diferenciais

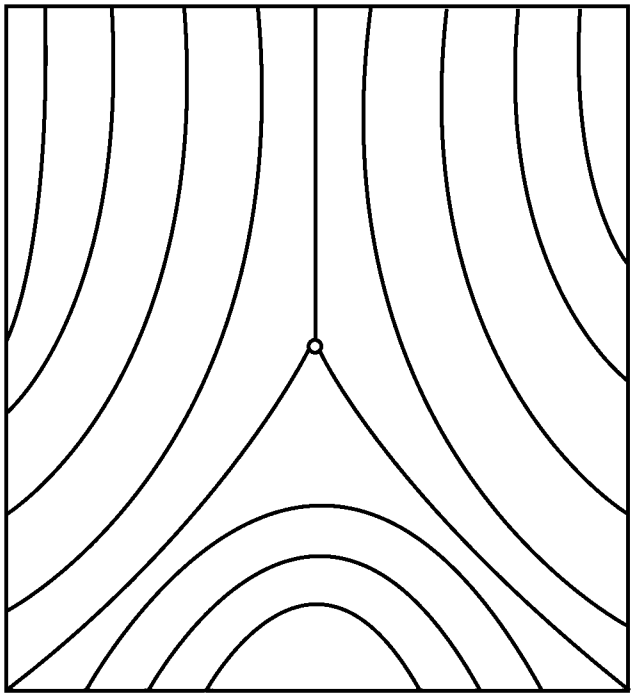
Um *sistema diferencial* de dimensão d numa variedade diferenciável M é uma aplicação L que associa a cada ponto $p \in M$ um subespaço $L(p)$ de dimensão d do espaço tangente M_p.

Um sistema diferencial L diz-se de *classe* C^r quando cada ponto de M possui uma vizinhança V na qual existem campos vetoriais $X_1, \ldots, X_d$, de classe C^r, tais que $\{X_1(q), \ldots, X_d(q)\}$ é uma base de $L(q)$ para todo $q \in V$.

Exemplos:

1) Um sistema diferencial de dimensão 1 chama-se também um *campo de direções*: a cada ponto $p \in M$ fica

associada uma reta $L(p) \subset M_p$, passando pela origem. Todo campo vetorial X sem singularidades (isto é, $X(p) \neq 0$ para todo $p \in M$) sobre M define um campo de direções L em M: $L(p) =$ reta que contém o vetor $X(p)$ em M_p. Note-se que se $X \in C^r$ então $L \in C^r$. Se a variedade M for simplesmente conexa, então todo campo de direções de classe C^r em M provém, dessa maneira, de um campo de vetores de classe C^r sem singularidades em M. (Vide a pag. 203 do livro "Grupo Fundamental e Espaços de Recobrimento", do autor.)



Se M não for simplesmente conexa, pode haver um campo de direções de classe C^∞ em M que não provém de nenhum campo contínuo de vetores não nulos em M. Isto se dá, por exemplo, para $M = R^2 - \{0\}$, com o campo de direções definido pelas tangentes às curvas indicadas na figura acima.

(Atenção: nem toda variedade diferenciável admite um campo contínuo de direções. Por exemplo, o toro admite, a esfera S^2 não).

2) Numa variedade riemanniana M^n, a todo sistema diferencial L, de dimensão d, corresponde um sisstema L^0, de dimensão $n-d$. Em cada ponto $p \in M$, $L^0(p)$ é o complemento ortogonal de $L(p)$ no espaço vetorial M_p. Tem-se $L \in C^r$ se, e somente se, $L^0 \in C^r$. Em particular, se existir em M um campo vetorial de classe C^r sem singularidades, existirá também um sistema diferencial de dimensão $n-1$ e classe C^r. Por exemplo, numa região U do plano um campo de direções é definido por uma equação da forma

$$a\,dx + b\,dy = 0, \tag{1}$$

onde $a, b\colon U \to R$ são funções reais. A direção $L(x,y)$ associada ao ponto $p = (x,y) \in U$ é a reta perpendicular ao vetor $(a(x,y), b(x,y))$. Exige-se, naturalmente, que $a^2 + b^2 \neq 0$ em todos os pontos de U. De maneira análoga, uma equação do tipo

$$a\,dx + b\,dy + c\,dz = 0, \tag{2}$$

onde a, b, c são funções reais num aberto $U \subset R^3$, define um campo de planos em U (sistema diferencial de dimensão 2). Em cada ponto $p = (x,y,z) \in U$, o plano $L(p)$ é perpendicular ao vetor $(a(p), b(p), c(p))$, o qual é suposto $\neq 0$ em todos os pontos de U. Por outro lado, um campo de direções no aberto $U \subset R^3$ é definido por um sistema de equações do tipo

$$\begin{aligned} a_1\,dx + b_1\,dy + c_1\,dz &= 0 \\ a_2\,dx + b_2\,dy + c_2\,dz &= 0 \end{aligned} \tag{3}$$

onde, em cada ponto $p \in U$, a matriz 2×3 dos coeficientes tem característica 2. A reta $L(p)$ é, para cada $p \in U$, a interseção dos dois planos fornecidos pelas equações, isto é, consiste de todos os vetores (X, Y, Z) tais que substituindo-se suas coordenadas no lugar de dx, dy e dz respectivamente, as equações acima são identicamente satisfeitas.

O problema que se deve resolver quando se tem um sistema diferencial é o de integrá-lo. Isto significa fazer passar por cada ponto da variedade uma subvariedade de dimensão d cujo espaço tangente em cada um dos seus pontos q é o subespaço $L(q)$ dado pelo sistema. Por exemplo, dado um campo de direções numa variedade M, integrá-lo é obter uma coleção de subvariedades de dimensão 1 (curvas) que cubram M, sendo a tangente a cada uma dessas curvas num ponto p a direção $L(p)$ dada pelo campo. Contraste-se com a integração de um campo de vetores, onde as curvas integrais são caminhos parametrizados: definindo não somente um conjunto de pontos como também um sentido de percurso e ainda um modo bem definido de percorrê-lo.

Assim, integrar o campo de direções definido pela equação (1) acima é obter, localmente, funções $x = x(t)$, $y = y(t)$ tais que, para $dx = x'(t)dt$ e $dy = y'(t)dt$, a igualdade $a(x(t), y(t))dx + b(x(t), y(t))dy = 0$ seja satisfeita identicamente. De modo análogo, integrar o campo de planos definido pela equação (2) é obter, localmente, funções $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, tais que, para $dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv$, $dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv$, $dz = \frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv$, a igualdade (2) seja válida quaisquer que sejam u, v em seu domínio.

Um campo de direções é sempre integrável, mas um campo de planos, por exemplo, nem sempre o é. Em outras

palavras: existem sistemas diferenciais L de dimensão 2 num aberto $U \subset R^3$ tais que não é possível fazer passar por cada ponto de U uma superfície tangente ao plano $L(p)$ em cada um dos seus pontos p.

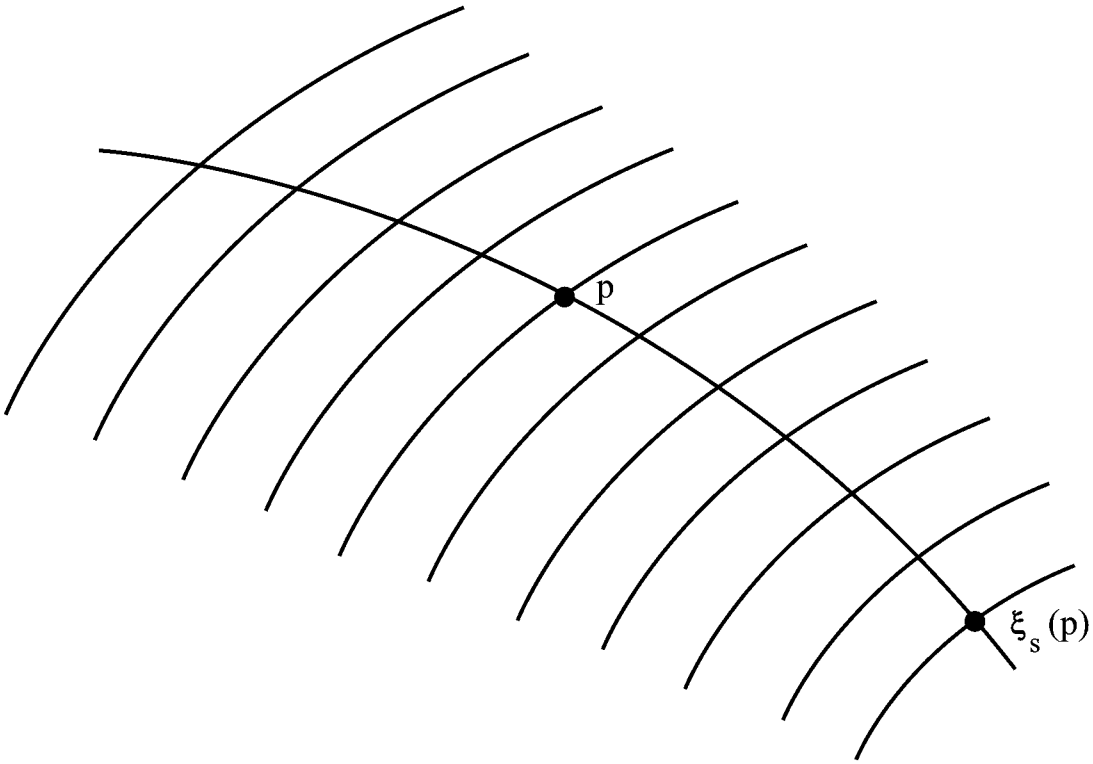
Faremos agora algumas considerações intuitivas com o fito de indicar razões geométricas devido às quais um campo de planos numa região $U \subset R^3$ pode deixar de ser integrável. Em primeiro lugar, daremos uma definição formal.

Definição: Seja L um sistema diferencial sobre uma variedade M. Diz-se que um campo vetorial X sobre M *pertence a* L, e escreva-se $X \in L$, quando $X(p) \in L(p)$ para cada $p \in M$.

Seja L um sistema diferencial integrável. Se um campo vetorial X pertence a L então, para cada ponto $p \in M$, a trajetória de X que tem origem p está contida na subvariedade integral de L que passa por P. Suponhamos, por exemplo, que $\dim L = 2$, isto é, que L seja um campo de planos. Seja Y outro campo vetorial pertencente a L, tal que $X(q)$ e $Y(q)$ formem uma base de $L(q)$ numa vizinhança de p. Indiquemos com ξ_s e η_t os fluxos locais determinados por X e Y respectivamente.

Por cada ponto da trajetória $s \to \xi_s(p)$ de X com origem p, façamos passar a trajetória $t \to \eta_t[\xi_s(p)]$, do campo Y, com origem $\xi_s(p)$. Todas estas trajetórias estão contidas na superfície integral S do sistema L que passa por p. Na realidade, para $\varepsilon > 0$ suficientemente pequeno, a aplicação $(s,t) \to \eta_t \xi_s(p)$, $-\varepsilon < s,t < \varepsilon$, é uma imersão de S em M. Consideremos agora um ponto $q_0 = \eta_{t_0}\, \xi_{s_0}(p) \in S$. A trajetória $s \to \xi_s(q_0)$, estando contida em S, deverá também

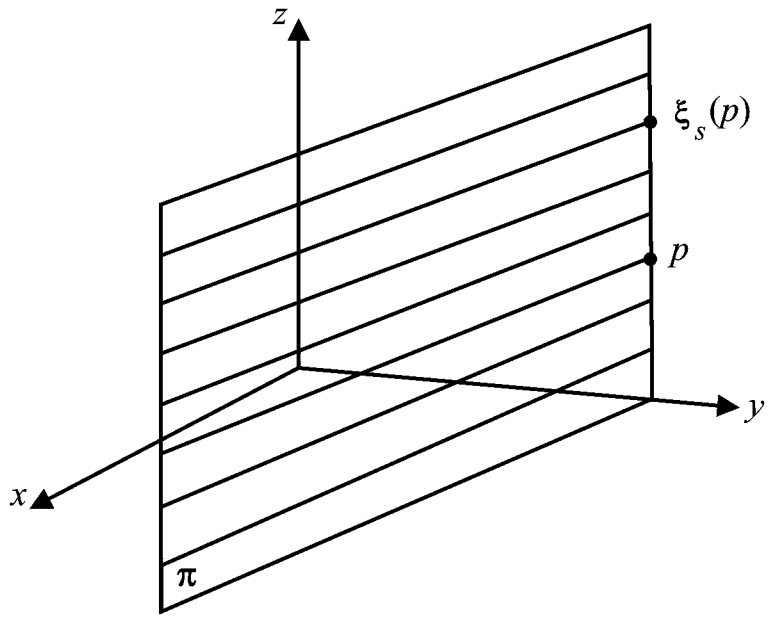
pertencer à superfície gerada pelas trajetórias $\eta_t \xi_s(p)$. Ora, dado um campo de planos arbitrário em R^3, nada obriga a que esta última condição se cumpra.



Consideremos, por exemplo, os campos vetoriais X, Y em R^3 definidos assim: $X(x, y, z) = (0, x, 1)$, $Y(x, y, z) = (1, 0, 0)$. Um sistema diferencial L de classe C^∞ e dimensão 2 (campo de planos) é definido pela condição de $L(p)$ ser o plano gerado por $X(p)$ e $Y(p)$ para cada $p = (x, y, z) \in R^3$. Afirmamos que o sistema diferencial L não é integrável.

Com efeito, consideremos um ponto fixo $p = (0, b, c)$ no plano yz. A trajetória de X que passa por este ponto é $\xi_s(p) = (0, b + s, c)$, reta vertical passando por p. A superfície gerada pelas trajetórias de Y que passam por esta reta é o plano π paralelo ao plano xz passando por p.

(Com efeito *todas* as trajetórias de Y são retas paralelas ao eixo dos x). Tal plano deve portanto ser a superfície integral de L que contém p. Entretanto, qualquer que seja o ponto q do plano π fora do eixo dos y, o vetor $X(q)$ não pertence a π, isto é $X(q) \notin L(q)$, o que contraria $X \in L$.



Por uma questão de conveniência técnica, daremos uma definição de integrabilidade que, à primeira vista, parece mais forte do que a apresentada informalmente acima.

Definição: Seja L um sistema diferencial de classe C^r e dimensão d numa variedade M^n. Diz-se que L é *integrável* quando, para cada ponto $p \in M$, existe um sistema de coordenadas $x\colon U \to R^n$, de classe C^r, tal que

$$\left\{ \frac{\partial}{\partial x^1}(q), \ldots, \frac{\partial}{\partial x^d}(q) \right\}$$

é uma base de $L(q)$ seja qual for $q \in U$.

Se $x = (x^1, \ldots, x^n)$ for um sistema de coordenadas locais adaptado, da maneira acima descrita, ao sistema diferencial L então, para cada $a = (a^{d+1}, \ldots, a^n) \in R^{n-d}$, o conjunto $S^a = \{q \in U;\ x^{d+1}(q) = a^{d+1}, \ldots, x^n(q) = a^n\}$ ou é vazio ou é uma subvariedade de M que tem $L(q)$ como espaço tangente em cada um dos seus pontos q.

Teorema de Frobenius. *A condição necessária e suficiente para que um sistema diferencial L, de classe C^r, seja integrável é que o colchete de Lie de dois campos de classe C^r pertencentes a L seja ainda um campo pertencente a L.*

Demonstração: Seja $\dim L = d$. Suponhamos primeiro que L seja integrável. Dados $X, Y \in L$, ponhamos $Z = [X, Y]$. Para mostrar que $Z \in L$, tomemos $p \in M$ e um sistema de coordenadas $x\colon U \to R^n$, adaptado a L de acordo com a definição de integrabilidade, com $p \in U$. Para $1 \leq i \leq n - d$, temos $(\partial x^{d+i}/\partial X)(q) = (\partial x^{d+i}/\partial Y)(q) = 0$, seja qual for $q \in U$. Isto implica imediatamente

$$\frac{\partial x^{d+i}}{\partial Z}(p) = \frac{\partial^2 x^{d+i}}{\partial X \partial Y}(p) - \frac{\partial^2 x^{d+i}}{\partial Y \partial X}(p) = 0$$

para $1 \leq i \leq n - d$, donde $Z(p) \in L(p)$. Isto mostra a necessidade da condição.

Para demonstrar que a condição é suficiente, seja L um sistema diferencial de classe C^r e dimensão d sobre M, tal que o colchete de dois campos quaisquer de classe C^r pertencentes a L seja ainda um campo pertencente a L.

Dado um ponto $p \in M$ arbitrário, seja $y\colon V \to R^n$ um sistema de coordenadas locais de classe C^r, com $p \in V$. Restringindo-se V, se necessário, podemos admitir a existência de campos vetoriais $X_1, \ldots, X_d$ de classe C^r,

definidos em V, tais que $\{X_1(q), \ldots, X_d(q)\}$ é, para todo $q \in V$, uma base de $L(q)$. Em relação à base definida por y temos:

$$X_i(q) = \sum_{j=1}^{n} a_i^j(q) \, \frac{\partial}{\partial y^j}(q), \quad i = 1, \ldots, d,$$

para todo $q \in V$.

Como os vetores $X_i(p)$ são linearmente independentes podemos, mediante um rearranjo de índices, supor que a matriz quadrada $(a_i^j(p))$, $1 \leq i, j \leq d$, é invertível. Por continuidade, $(a_i^j(q))$ será ainda invertível para todo ponto q numa vizinhança de p, a qual suporemos ser ainda V (restringindo-a mais, caso necessário). Seja $(b_j^i(q))$, $q \in V$, a matriz $d \times d$ inversa de $(a_i^j(q))$. As igualdades

$$Y_j = \sum_{i=1}^{d} b_j^i \, X_j \,, \quad j = 1, \ldots, d,$$

definem em V campos vetoriais $Y_1, \ldots, Y_d$, de classe C^r, tais que em cada ponto $q \in V$, $\{Y_1(q), \ldots, Y_d(q)\}$ é uma base de $L(q)$. Além disso, para cada $j = 1, \ldots, d$, temos

$$Y_j = \frac{\partial}{\partial y^j} + Z_j \,, \quad \text{com} \quad Z_j = \sum_{k \geq d+1} c_j^k \, \frac{\partial}{\partial y^k} \,.$$

Um cálculo simples mostra que o colchete de dois quaisquer dos campos $Z_1, \ldots, Z_d$ é uma combinação linear dos $\partial/\partial y^k$, $k \geq d+1$. Segue-se então das igualdades acima que os colchetes $[Y_i, Y_j]$ também são combinações lineares de

$\partial/\partial y^{d+1}, \ldots, \partial/\partial y^n$. Por outro lado, como os Y_j pertencem a L, tem-se, em virtude da hipótese feita, $[Y_i, Y_j] \in L$, ou seja:

$$[Y_i, Y_j] = \sum_{m \leq d} \lambda_{ij}^m \, Y_m, . \qquad (*)$$

Mais explicitamente:

$$[Y_i, Y_j] = \sum_{m \leq d} \left(\lambda_{ij}^m \frac{\partial}{\partial y^m} + \lambda_{ij}^m \, Z_m \right) =$$

$$= \sum_{m \leq d} \lambda_{ij}^m \frac{\partial}{\partial y^m} + \sum_{k \geq d+1} \mu_{ij}^k \frac{\partial}{\partial y^k} \, .$$

Conclui-se, portanto que $\lambda_{ij}^m = 0$ para $1 \leq i, j, m \leq d$. Voltando à igualdade (*), isto dá $[Y_i, Y_j] = 0$, $1 \leq i, j \leq d$. Em suma, na vizinhança de cada ponto de M existe uma base do sistema L formada por campos de vetores comutativos. O teorema segue-se então do Teorema 4, §2.

A condição "$X, Y \in L \Rightarrow [X, y] \in L$" chama-se *condição de integrabilidade* do sistema L. Agora vemos que o sistema de dimensão 2 definido em R^3 pelos campos vetoriais $X = (0, x, 1)$ e $Y = (1, 0, 0)$ não é integrável porque o colchete $[X, Y] = (0, 1, 0)$, em ponto algum do espaço é combinação linear dos vetores X e Y nesse ponto. (Bastava que não o fosse num único ponto.)

Verifica-se imediatamente que se um sistema L, de dimensão d, é integrável então na vizinhança de cada ponto $p \in M$ existe apenas uma subvariedade integral de L contendo p. (*Subvariedade integral* é uma subvariedade cujo espaço tangente em cada ponto coincide com L naquele ponto.)

Com efeito, basta notar o seguinte: se $x\colon U \to R^n$ é um sistema de coordenadas adaptado a L no sentido da definição de integrabilidade e $W \subset V$ é uma subvariedade integral conexa contida em V então, para cada $i = 1, \ldots, n-d$, pondo $\varphi^i = x^{d+i}|W$ = restrição da função real $x^{d+i}\colon V \to R$ a W, temos $(d\varphi^i)_q(v) = dx^{d+i}(v) = 0$ qualquer que seja $v \in W_q = L(q)$ pois dx^{d+i} anula-se em $L(q)$. Logo $d\varphi^i = 0$ em W. Como W é conexa, φ^i é constante em W, isto é $x^{d+1}, \ldots, x^n$ são constantes em W. Para simplificar, seja $x(p) = 0 \in R^n$. Então, como $p \in W$, $x^{d+1}(q) = \cdots = x^n(q) = 0$ para todo $q \in W$. Segue-se que W coincide, na vizinhança de p, com a subvariedade integral $\{x^{d+1} = \cdots = x^n = 0\}$.

Um raciocínio imediato mostra que, mais geralmente, se W e W' são subvariedades integrais de L, ambas conexas, ambas contendo o mesmo ponto p, então ou $W \subset W'$ ou $W' \subset W$. Isto conduz a um estudo global das subvariedades integrais de um sistema diferencial. Para maiores esclarecimentos ver o livro "Teoria Geométrica das Folheações", por Cesar Camacho e Alcides Lins Neto.