

Na termodinâmica trabalharemos frequentemente com diferenciais de funções de estado, como, por exemplo, a diferencial da energia interna dU ou a diferencial da temperatura dT . Além disso, usaremos também uma generalização do conceito de diferencial que é chamado *forma diferencial*. Para se aprender termodinâmica será extremamente útil entender estes conceitos matemáticos claramente. Portanto investiremos algum tempo para aprendermos conceitos básicos e simples de geometria diferencial. Usaremos frequentemente o gás ideal com sua equação de estado $PV = NRT$ como exemplo, mesmo não tendo ainda introduzido o conceito de temperatura.

2.1) Coordenadas

Primeiramente podemos notar que o espaço de estados de equilíbrio de um sistema termodinâmico não possui a geometria euclidiana do espaço comum nem nada parecido. Por exemplo, não existe uma noção de ortogonalidade para vetores neste espaço. Mas podemos, como no espaço comum, descrever os pontos com coordenadas $x_1, x_2, x_3, \dots, x_n$. Um sistema de n coordenadas num espaço é um conjunto de n funções $x_1, x_2, x_3, \dots, x_n$ que mapeiam uma região do dito espaço nos números reais¹ de tal forma que uma n -upla de valores a_1, a_2, \dots, a_n define um ponto P na dita região de forma única pelas equações $a_1 = x_1(P), a_2 = x_2(P), \dots, a_n = x_n(P)$.² No caso de um gás simples com um número de mols N fixo n seria 2 e poderíamos usar como coordenadas, por exemplo, $x_1 = P$ e $x_2 = T$. Frequentemente queremos mudar o sistema de coordenadas. No exemplo do gás poderíamos usar T e V ou V e P no lugar de P e T . Podemos então escrever as novas coordenadas \tilde{x}_k como funções das antigas coordenadas $\tilde{x}_k = \tilde{x}_k(x_1, x_2, \dots, x_n)$. No exemplo do gás ideal, teríamos com $\tilde{x}_1 = V$ e $\tilde{x}_2 = T$ a lei de transformação de coordenadas $\tilde{x}_1 = NRx_2/x_1$ e $\tilde{x}_2 = x_2$. Vamos limitar a classe de coordenadas permissíveis de tal forma que uma mudança de coordenadas sempre resulte em funções diferenciáveis; $\tilde{x}_k = \tilde{x}_k(x_1, x_2, \dots, x_n)$ diferenciáveis. Também temos que exigir que a transformação possa ser invertida. Isto significa que as coordenadas antigas podem ser escritas como funções das novas: $x_k = x_k(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$. Consequentemente podemos concluir que a matriz jacobiana tem inversa:

¹ Podemos também usar valores dimensionais, ou seja, os valores das coordenadas podem ter unidade.

² Podem existir n -uplas sem ponto correspondente no espaço, pois os mapeamentos x_k não precisam ser sobrejetivos.

$$\det \begin{pmatrix} \frac{\partial \tilde{x}_k}{\partial x_j} \end{pmatrix} \neq 0 \quad (2.1.1)$$

Ocasionalmente podemos ter um sistema de coordenadas que não é bem definido em todo o espaço de estados. Por exemplo, sabemos que a densidade da água possui um máximo numa temperatura de 4°C (a uma pressão de uma atmosfera). Podemos então concluir que existem estados com uma pressão de uma atmosfera e temperaturas ligeiramente acima e abaixo de 4°C que resultam no mesmo volume V . Desta forma existe uma região do espaço de estados da água onde o sistema de coordenadas (P, V) não funciona, já que dois estados diferentes resultam nos mesmos valores de P e V . Este tipo de falha de um sistema de coordenadas não é grave. Simplesmente temos que usar outro sistema de coordenadas para a região com problemas. Podemos imaginar que cobrimos o espaço com vários sistemas de coordenadas, como as cartas de um atlas cobrem todo o globo terrestre. O tipo de espaço que acabamos de descrever chama-se *variedade diferenciável*.

Freqüentemente usamos *funções de estado*. Uma função de estado U tem um valor que depende do estado E . Deveríamos então escrever $U(E)$. Mas esta notação não é empregada na prática; usando um sistema de coordenadas x_1, \dots, x_n podemos descrever o estado E pelos valores das coordenadas e escrever $U(x_1, \dots, x_n)$ no lugar de $U(E)$. O problema com esta notação aparece quando mudarmos de coordenadas. Descrevendo o mesmo ponto E por coordenadas $\tilde{x}_1, \dots, \tilde{x}_n$, a mesma função de estado $U(E)$ resulta agora numa outra função numérica $\tilde{U}(\tilde{x}_1, \dots, \tilde{x}_n)$. Por exemplo, a energia interna de um mol de um gás ideal monoatômico nas coordenadas T e V é $U(T, V) = \frac{3}{2}RT$, mas nas coordenadas P e V obtemos outra função $\tilde{U}(P, V) = \frac{3}{2}PV$. No entanto, em todos os textos de termodinâmica é costume escrever novamente o símbolo U e não \tilde{U} . Isto significa que estamos no fundo escrevendo $U(E)$. No nosso curso adotaremos este costume também. Então quando se vê uma função $U(T, V)$ não podemos interpretá-la como função dos valores T e V , mas temos que interpretá-la como U no ponto E que tem os valores de coordenadas T e V . Esta notação é prática, mas ela tem um inconveniente: nas derivadas parciais de uma função de estado, temos que anotar qual é o sistema de coordenadas usado, pois (como veremos logo em seguida) um $\partial / \partial T$ num sistema de coordenadas T e V tem outro significado que um $\partial / \partial T$ num sistema de coordenadas T e P .

Variando o valor de uma determinada coordenada x_m ($m = \text{fixo}$) e mantendo os valores de todas as outras coordenadas x_k ($k \neq m$) constantes, percorremos no espaço uma curva chamada de linha de coordenada m .

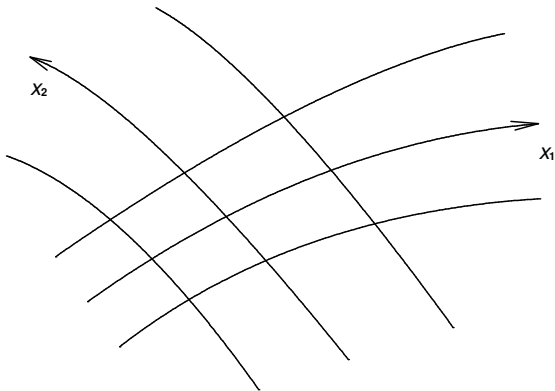


Fig. 2.1 Linhas de coordenadas

Veremos o que acontece com as linhas de coordenadas numa mudança de coordenadas: suponha que a figura 2.2 mostre as linhas das coordenadas x_1, x_2 num espaço bidimensional.

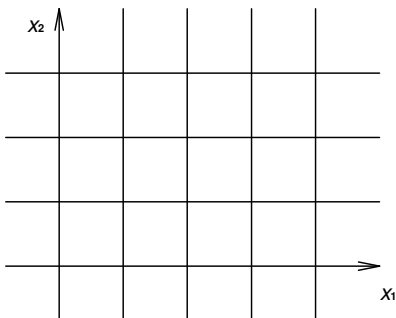


Fig. 2.2

Agora vamos introduzir as novas coordenadas

$$\tilde{x}_1 = x_1$$

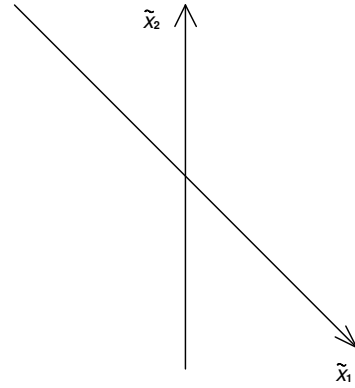
$$\tilde{x}_2 = \frac{1}{2}(x_1 + x_2)$$

Como são as linhas destas novas coordenadas? Poder-se-ia pensar que as linhas da coordenada 1 ficam inalteradas já que a coordenada 1 não muda. Mas isto não é o caso! As linhas da coordenada 1 no novo sistema de coordenadas são determinadas pelas equações $\tilde{x}_2 = \text{const.}$. Isto significa

$$const. = \frac{1}{2}(x_1 + x_2)$$

Fig. 2.3 Linhas das novas coordenadas

Estas equações determinam linhas inclinadas. Por outro lado as linhas da coordenada 2 são determinadas pelas equações $\tilde{x}^1 = const.$ que são idênticas as equações $x^1 = const.$ e, portanto as linhas da coordenada 2 não mudarão. O novo sistema de coordenadas tem então o seguinte aspecto: ->



Esta observação simples sobre o comportamento das linhas de coordenadas tem consequências importantes: a *derivada parcial* de uma função F com respeito à coordenada $x_1 = \tilde{x}_2$ depende de quem é a coordenada 2. A derivada parcial de uma função F com respeito à coordenada x_1 mede a taxa de variação de F quando andarmos na direção da linha de coordenada 1. Como vimos, a direção desta linha de coordenada 1 depende também das outras coordenadas $x_k, (k \neq 1)$. Consequentemente, a derivada parcial com respeito à x_1 vai depender da escolha das outras coordenadas também. Veremos um exemplo: a função

$$F(\tilde{x}_1, \tilde{x}_2) = \tilde{x}_1 \exp\{a\tilde{x}_2\} = x_1 \exp\left\{a \frac{x_1 + x_2}{2}\right\}$$

tem as derivadas parciais

$$\left(\frac{\partial F}{\partial \tilde{x}_1}\right)_{\tilde{x}_2} = \exp\{a\tilde{x}_2\}$$

$$\left(\frac{\partial F}{\partial x_1}\right)_{x_2} = \exp\{a\tilde{x}_2\} + \frac{a}{2} x_1 \exp\{a\tilde{x}_2\}$$

Então as derivadas parciais $\left(\frac{\partial F}{\partial x_1}\right)_{x_2}$ e $\left(\frac{\partial F}{\partial x_1}\right)_{\tilde{x}_2} = \left(\frac{\partial F}{\partial \tilde{x}_1}\right)_{\tilde{x}_2}$ diferem pelo termo adicional $\frac{a}{2} x_1 \exp\{a\tilde{x}_2\}$.

Se escrevermos, por exemplo, na eletrostática a componente x do campo elétrico como o negativo da derivada parcial do potencial elétrico $E_x = -\frac{\partial \phi}{\partial x}$, é mais ou menos óbvio que as outras coordenadas do sistema usado são y e z . Por isso ninguém escreve $\left(\frac{\partial \phi}{\partial x}\right)_{y,z}$.

na termodinâmica não é a priori claro quais as coordenadas empregadas. Portanto na termodinâmica é indispensável anotar sempre nas derivadas parciais quais as grandezas que se mantêm constantes. Por exemplo, $\frac{\partial F}{\partial T}$ não é bem definida; poder-se-ia tratar de $\left(\frac{\partial F}{\partial T}\right)_v$ ou de $\left(\frac{\partial F}{\partial T}\right)_p$ e estas grandezas são diferentes.

2.2) Vetores

Bernhard Lesche

No espaço comum podemos definir vetores como pares ordenados de pontos sendo que um par (A, B) representa o mesmo vetor $\overrightarrow{(A, B)}$ que um par (A', B') que difere do (A, B) apenas por um transporte paralelo.

$$\overrightarrow{(A, B)} = \overrightarrow{(A', B')}$$

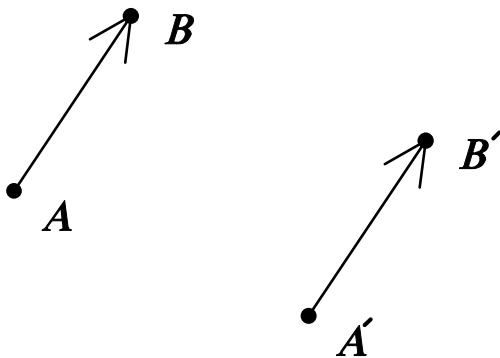


Fig. 2.4 Identificação de vetores.

Esta identificação de pares que diferem por um transporte paralelo é importante para poder somar vetores. Se queremos, por exemplo, somar os vetores \vec{a} e \vec{b} da figura 2.5a temos que transportar o vetor \vec{a} até a ponta do vetor \vec{b} para formar o vetor resultante $\vec{c} = \vec{a} + \vec{b}$.

Fig.2.5a

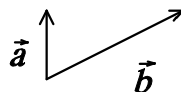
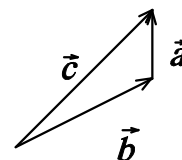


Fig. 2.5b



Mas infelizmente no espaço dos estados da termodinâmica a noção de transporte paralelo não é definida. Poderíamos tentar usar as coordenadas para definir o transporte paralelo: por exemplo, com a seguinte tentativa:

$$\left\{ \vec{(A,B)} = \vec{(A',B')} \right\} \Leftrightarrow \left\{ (x_k)_A - (x_k)_B = (x_k)_{A'} - (x_k)_{B'} \quad \text{para } k = 1, \dots, n \right\} \quad (2.2.1)$$

onde $(x_k)_A, (x_k)_B, (x_k)_{A'}, (x_k)_{B'}$ são as coordenadas dos pontos A, B, A', B' . Mas infelizmente esta definição mudaria seu sentido cada vez que mudássemos o sistema de coordenadas. A afirmação (2.2.1) ficaria invariante apenas sob mudanças de coordenadas do tipo inhomogêneo linear:

$$\tilde{x}_k = c_k + \sum_{j=1}^n M_{kj} x_j \quad (2.2.2)$$

com constantes c_k e M_{kj} sendo M_{kj} uma matriz que possui inversa. Para transformações não lineares (como por exemplo $V, P \rightarrow V, T$), a afirmação (2.2.1) muda de significado.

A invariância de (2.2.1) sob transformações do tipo inhomogêneo linear (2.2.2) sugere, no entanto, uma bela saída da dificuldade de definir vetores: se limitarmos os pontos A, B, A', B' a uma vizinhança infinitesimal de um ponto N , toda transformação diferenciável de coordenadas tem a forma (2.2.2). Veremos o que isto significa: sejam N_k e \tilde{N}_k as coordenadas do ponto N no sistema antigo (x_k) e novo (\tilde{x}_k) respectivamente. Para quaisquer números fixos $a_1, a_2, a_3, \dots, a_n$ e para $\varepsilon \in \mathbb{R}$ (variável) temos um ponto $A(\varepsilon)$ cujas coordenadas no sistema antigo (x_k) são

$$A_k = N_k + \varepsilon a_k \quad (2.2.3)$$

Este mesmo ponto A teria no sistema novo (\tilde{x}_k) as coordenadas

$$\tilde{A}_k = \tilde{N}_k + \sum_{j=1}^n \left(\left. \frac{\partial \tilde{x}_k}{\partial x_j} \right|_N \right)_{x_j} \cdot \varepsilon a_j + \Delta_k(\varepsilon) \quad (2.2.4)^3$$

onde o termo de erro $\Delta_k(\varepsilon)$ vai para 0 mais rapidamente que o termo linear em ε quando $\varepsilon \rightarrow 0$. Isto é,

$$\left| \frac{\Delta_k(\varepsilon)}{\varepsilon} \right| \rightarrow 0 \quad \text{para } \varepsilon \rightarrow 0 \quad (2.2.5)$$

Costumamos expressar o fato de que

³ O índice x_{-j} significa que os x_m com $m \neq j$ são mantidos constantes.

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \left\{ \tilde{A}_k - \left(\tilde{N}_k + \sum_{j=1}^n \left(\frac{\partial \tilde{x}_k}{\partial x_j} \Big|_N \right) \cdot \varepsilon a_j \right) \right\} = 0$$

de maneira mais simples dizendo: para ε *infinitesimal* vale a equação

$$\tilde{A}_k = \tilde{N}_k + \sum_{j=1}^n \left(\frac{\partial \tilde{x}_k}{\partial x_j} \Big|_N \right) \cdot \varepsilon a_j \quad (2.2.6)$$

Repare agora que com $x_k = N_k + \varepsilon a_k$ e $M_{kj} = \left(\frac{\partial \tilde{x}_k}{\partial x_j} \Big|_N \right)$ e com $c_k = \tilde{N}_k - \sum_{j=1}^n M_{kj} N_j$ a

equação (2.2.6) tem exatamente a forma da equação (2.2.2). O que expressamos aqui com fórmulas é no fundo uma observação bem intuitiva: numa vizinhança suficientemente pequena de um ponto fixo N , qualquer sistema de coordenadas é aproximadamente retilíneo em relação a qualquer outro sistema de coordenadas (compare a figura 2.6).

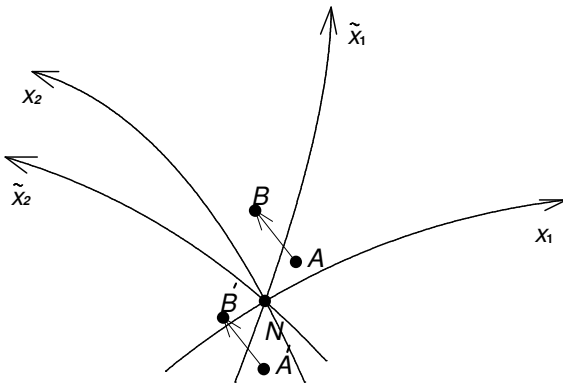


Fig. 2.6 Dois pares de pontos numa pequena vizinhança de um ponto N . O julgamento se estes pares definem o mesmo vetor, ou não, difere apenas por erros de ordem superior do tamanho da vizinhança, quando se julga com os dois sistemas de coordenadas.

Com pares ordenados de pontos numa vizinhança infinitesimal de um ponto fixo N podemos então definir vetores da forma usual usando o critério (2.2.1) para decidir se dois pares

definem o mesmo vetor. A única diferença entre a construção de vetores no nosso espaço comum e nos espaços da termodinâmica é que podemos associar ao espaço Euclidiano um único espaço vetorial, enquanto no espaço da termodinâmica temos um espaço vetorial diferente para cada ponto N . Este espaço é chamado espaço tangente no ponto N .⁴ Dentro do espaço tangente \mathcal{T}_N de um ponto N podemos somar vetores como em qualquer espaço vetorial, mas não podemos somar um vetor do espaço tangente \mathcal{T}_N do ponto N com um vetor no espaço tangente \mathcal{T}_M de um outro ponto M .

Como em qualquer espaço vetorial, os vetores nos espaços tangentes são caracterizáveis pela sua *direção*, sua *orientação* e seu *tamanho* (não usaremos a palavra *módulo* porque o

⁴Os matemáticos definem os espaços tangentes sem o uso de parâmetros infinitesimais. Usamos aqui uma definição com vetores infinitesimais que é intuitiva e é mais próxima dos procedimentos em todos os livros texto sobre termodinâmica. Os matemáticos definem os vetores como classes de equivalência de curvas diferenciáveis. Isto é mais prático para demonstrações matemáticas.

módulo de um vetor requer uma definição adicional⁵). Veremos um exemplo: um gás ideal encontra-se num estado com coordenadas P e T . A partir deste ponto submetemos o gás a um processo isocórico (volume=const.) que leva o sistema a um estado infinitesimalmente próximo. Este processo descreve um vetor no espaço tangente do ponto com pressão P e temperatura T . A direção deste vetor é caracterizada pela condição $V = \text{const.}$. Podemos descrever a orientação e o tamanho pela variação de algum outro parâmetro que se altera no processo. Por exemplo, podemos dizer que a temperatura muda de T para o novo valor $T + \varepsilon$. Vamos escrever este vetor da seguinte forma:

$$\overrightarrow{\text{tamanho e orientação} \setminus \text{direção}} = \overrightarrow{\delta T = \varepsilon \setminus V} \quad (2.2.7)$$

Dizendo em palavras, este vetor descreve um processo infinitesimal que mantém V constante e altera T pelo valor ε . Poderíamos escrever o mesmo vetor na forma $\overrightarrow{\delta P = \eta \setminus V}$ com $\eta = NR\varepsilon/V$.

Vamos escolher agora um sistema fixo de coordenadas. Por exemplo P e T . Em relação a este sistema de coordenadas, podemos definir os vetores $\overrightarrow{\delta P = \kappa \setminus T}$ e $\overrightarrow{\delta T = \varepsilon \setminus P}$. As direções destes vetores são determinadas pela escolha do sistema de coordenadas. As orientações e os tamanhos ainda dependem da escolha dos parâmetros infinitesimais ε e κ . Para fixar estes também podemos normalizar estes vetores da seguinte forma, definindo vetores

$$\vec{e}_1 = \frac{1}{\kappa} \overrightarrow{\delta P = \kappa \setminus T} \quad \text{e} \quad \vec{e}_2 = \frac{1}{\varepsilon} \overrightarrow{\delta T = \varepsilon \setminus P} \quad (2.2.8)$$

Note, no entanto, que dividindo por grandezas infinitesimais obtemos vetores tão grandes que eles não terão mais a interpretação de pares de pontos numa vizinhança infinitesimal do ponto P, T . Os vetores \vec{e}_1 e \vec{e}_2 são objetos um tanto formais, mas eles servem formalmente como vetores básicos. Podemos escrever os vetores que possuem uma interpretação como pares ordenados de pontos na vizinhança infinitesimal do ponto (P, T) como combinações lineares de \vec{e}_1 e \vec{e}_2 . $\vec{a} = \delta P \vec{e}_1 + \delta T \vec{e}_2$. Por exemplo, podemos escrever o vetor (2.2.7) na forma $\overrightarrow{\delta T = \varepsilon \setminus V} = \frac{NR}{V} \varepsilon \vec{e}_1 + \varepsilon \vec{e}_2$.

⁵ O módulo permite comparar tamanhos de vetores que apontam para diferentes direções. Na termodinâmica, podemos comparar somente tamanhos de vetores que apontam na mesma direção.

2.3) O espaço dual

Bernhard Lesche

Seja \mathfrak{V} um espaço vetorial. Dentro de \mathfrak{V} temos os vetores que costumamos representar por setas. Agora podemos definir dentro de \mathfrak{V} ainda outros objetos geométricos que chamaremos de vetores duais. Um vetor dual \bar{Q} é dado por um (hiper) plano em \mathfrak{V} que não contem o vetor zero. Um hiperplano num caso de um espaço bidimensional seria uma reta, num caso de um espaço tridimensional seria um plano comum e num espaço de n dimensões seria de $n-1$ dimensões. A figura 2.7 mostra um vetor dual para um espaço \mathfrak{V} bidimensional. A cruz representa o vetor zero.

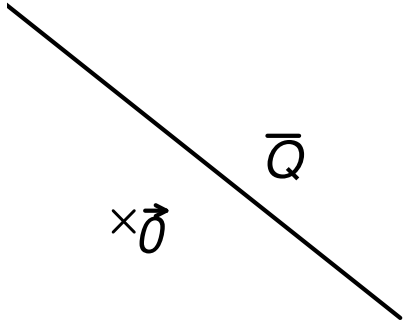


Fig. 2.7 Vetor dual num espaço bidimensional. Os vetores comuns são setas que partem da cruz (vetor zero). \bar{Q} é a coleção de vetores comuns cujas pontas estão em cima de uma reta (hiperplano em duas dimensões).

Podemos definir a soma de vetores duais da seguinte forma: sejam \bar{Q} e \bar{W} vetores duais⁶. Para somá-los, construímos o plano \underline{Q} que é paralelo ao plano \bar{Q} e que contém o vetor zero e construímos o plano \underline{W} que é paralelo ao plano \bar{W} e que contém o vetor zero. O vetor dual $\bar{Z} = \bar{Q} + \bar{W}$ é o plano que passa pela intersecção de \bar{Q} com \underline{W} e pela intersecção de \bar{W} com \underline{Q} .

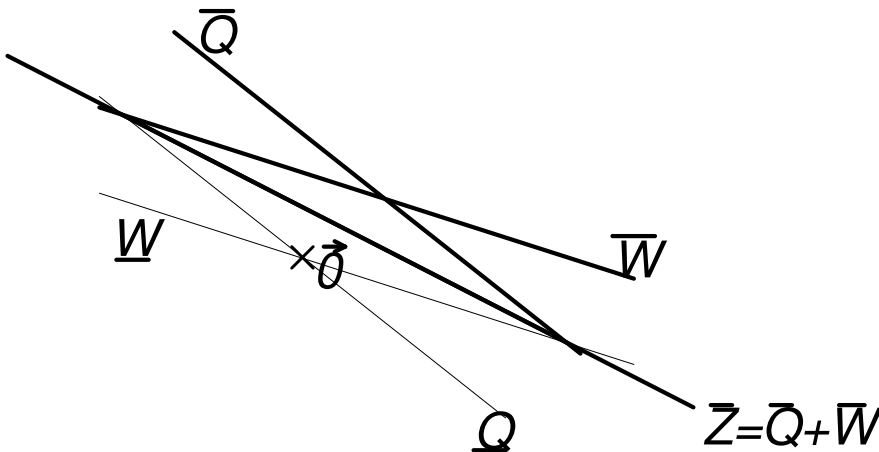


Fig. 2.8 Soma de vetores duais.

⁶ Esta definição apresenta dificuldades no caso em que os dois vetores sejam paralelos. Este caso pode ser resolvido escrevendo um dos vetores duais primeiramente como soma de dois outros e usar a associatividade da soma.

A multiplicação de um vetor dual por um número é definida da seguinte maneira: seja \bar{Q} um vetor dual e $\lambda \in \mathbb{R}$ um número. Para construir o vetor $\lambda\bar{Q}$ constrói-se uma reta R qualquer que passe pelo vetor zero e que intercepte o plano \bar{Q} . Nesta reta R define-se uma escala linear que tenha o valor zero no ponto do vetor zero e o valor um no ponto da intersecção de R com \bar{Q} . Nesta escala procura-se o ponto $1/\lambda$. O vetor dual $\lambda\bar{Q}$ é então o plano paralelo ao plano \bar{Q} e que passa pelo ponto $1/\lambda$ na escala em R .

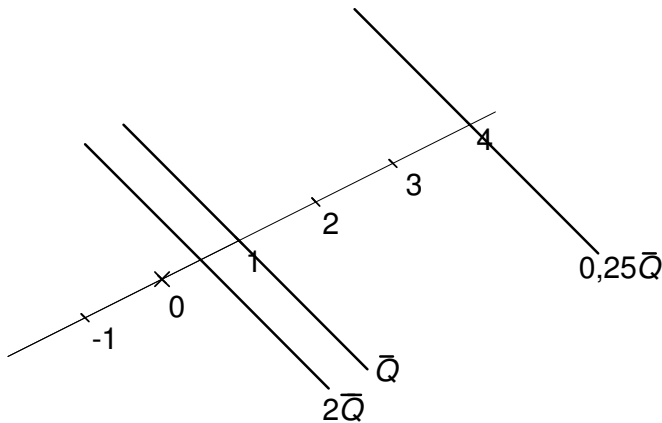
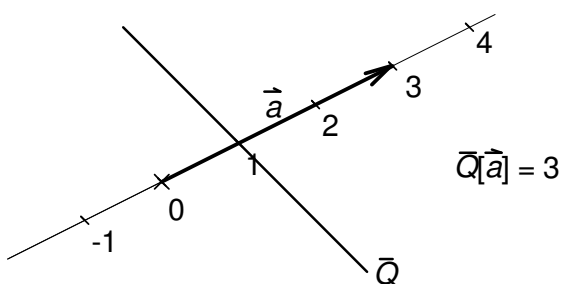


Fig. 2.9 Multiplicação de vetor dual \bar{Q} com números. A figura mostra os exemplos $2\bar{Q}$ e $\frac{1}{4}\bar{Q}$.

Com $\lambda=0$ obtém-se o vetor dual zero que seria um plano escondido no infinito. Pode-se mostrar que, com estas regras de soma e produto com números, o conjunto de vetores duais forma uma espaço linear que é chamado o espaço dual do espaço \mathfrak{X} e é costume escrevê-lo como \mathfrak{X}^* .

A propriedade mais importante dos vetores duais é a possibilidade de se definir a ação deles sobre vetores comuns: seja \bar{Q} um vetor dual e \vec{a} um vetor comum. Para formar a aplicação de \bar{Q} sobre \vec{a} , traça-se uma reta R na direção do vetor \vec{a} que passe pelo vetor zero. Depois define-se nesta reta uma escala linear da mesma forma como se fez para definir a multiplicação de um vetor dual por um número. Encostando a base do vetor \vec{a} no vetor zero, pode-se ler o valor de escala da ponta do vetor \vec{a} . Este número é a aplicação de \bar{Q} sobre \vec{a} . Escreveremos este número como $\bar{Q}[\vec{a}]$. No caso em que o vetor \vec{a} for paralelo ao plano \bar{Q} , a reta R não interceptaria o plano \bar{Q} em nenhum ponto. Ou poderíamos dizer ela intercepta \bar{Q} no infinito. Desta forma a unidade de escala na reta R



seria infinitamente grande e frente a esta unidade o vetor \vec{a} mediria zero. Neste caso temos então $\bar{Q}[\vec{a}] = 0$.

Fig. 2.10 Aplicação de um vetor dual num vetor.

A aplicação de \bar{Q} sobre vetores é uma operação linear, isto é $\bar{Q}[\alpha\vec{a} + \beta\vec{b}] = \alpha\bar{Q}[\vec{a}] + \beta\bar{Q}[\vec{b}]$. Pode-se mostrar que o conjunto de todos os mapeamentos lineares de \mathcal{T} em \mathbb{R} é idêntico ao espaço dual \mathcal{T}^* . Podemos ainda definir vetores duais dimensionais. Isto significa que o valor da aplicação do vetor dual num vetor não resulta num número, mas num valor com unidade. Por exemplo, para vetores \vec{a} que terminam no hiperplano \bar{Q} , a aplicação $\bar{Q}[\vec{a}]$ pode valer 1 Joule.

Você deve-se perguntar: por que não aprendemos de vetores duais antes do curso de termodinâmica? Isto tem uma explicação simples: No espaço euclidiano pode-se evitar o uso de vetores duais porque nestes espaços podemos descrever os mapeamentos lineares que mapeiam vetores em números com ajuda do produto escalar. Se existir um produto escalar no espaço \mathcal{T} podemos encontrar para cada vetor dual \bar{Q} de \mathcal{T}^* um vetor comum \vec{q} tal que para todos $\vec{a} \in \mathcal{T}$ vale $\bar{Q}[\vec{a}] = \vec{q} \cdot \vec{a}$. (Exercício: construa o vetor correspondente). O produto escalar é intimamente relacionado com a norma dos vetores. Tendo uma norma $\| \cdot \|$ no espaço \mathcal{T} que obedece a identidade de paralelograma ($\forall \vec{a}, \vec{b} \in \mathcal{T} : \|\vec{a} + \vec{b}\|^2 + \|\vec{a} - \vec{b}\|^2 = 2\|\vec{a}\|^2 + 2\|\vec{b}\|^2$), pode-se mostrar (Exercício: mostre-lo) que

$$\vec{a} \cdot \vec{b} \stackrel{\text{def.}}{=} \frac{1}{4} \left[\|\vec{a} + \vec{b}\|^2 - \|\vec{a} - \vec{b}\|^2 \right]$$

tem todas as propriedades de um produto escalar. Mas, na termodinâmica não existe a norma de vetores. A norma permite comparar comprimentos de vetores que apontam em direções diferentes. Na termodinâmica não podemos dizer se fizemos uma mudança de estado maior ou menor quando aumentamos a temperatura um tanto ou quando aumentamos o volume um outro tanto. Por esta razão não existe o produto escalar nos espaços tangentes dos espaços da termodinâmica e a introdução dos vetores duais se faz necessária.

2.4) Diferenciais e formas diferenciais

Bernhard Lesche

Da eletrostática conhecemos *campos vetoriais*, por exemplo, o campo elétrico \vec{E} . Este campo atribui a cada ponto (x, y, z) no espaço um vetor $\vec{E}(x, y, z)$. Da mesma forma poderíamos pensar em campos vetoriais definidos no espaço dos estados de um sistema termodinâmico. Um campo \vec{a} associaria a cada estado E um vetor $\vec{a}(E)$ que seria um elemento do espaço tangente deste ponto E . Mas resulta que campos vetoriais não são de grande utilidade na termodinâmica. Por outro lado, veremos que *campos de vetores duais* realmente são usados na termodinâmica. Este tipo de campo associaria a cada estado E um vetor dual $\bar{Q}(E)$ que seria um elemento do espaço dual do espaço tangente de E . Isto é $\bar{Q}(E) \in \mathcal{T}_E^*$.

Um campo de vetores duais é chamado uma *forma diferencial*.⁷ Um caso particular de forma diferencial são as diferenciais de funções.

⁷ Na verdade a definição de forma diferencial inclui ainda exigências de diferenciabilidade.

Seja F alguma função de estado de um sistema termodinâmico. Podemos, por exemplo, pensar no volume ou na energia interna de um gás ideal. O gás está num estado com pressão P e temperatura T . Agora submetemos o gás a um processo que o leva ao estado infinitesimalmente próximo com pressão $P + \delta P$ e temperatura $T + \delta T$. Conhecemos o valor de F no ponto (P, T) e queremos saber o valor de F no ponto $(P + \delta P, T + \delta T)$. Temos

$$F(P + \delta P, T + \delta T) - F(P, T) = \left(\frac{\partial F}{\partial P} \right)_T \delta P + \left(\frac{\partial F}{\partial T} \right)_P \delta T + \text{erro de ordem superior}$$

Como δP e δT são infinitesimais, podemos escrever

$$F(P + \delta P, T + \delta T) - F(P, T) = \left(\frac{\partial F}{\partial P} \right)_T \delta P + \left(\frac{\partial F}{\partial T} \right)_P \delta T \quad (2.4.1)$$

Como você pode ver, esta variação infinitesimal do valor da função depende linearmente do vetor $\vec{a} = \delta P \vec{e}_1 + \delta T \vec{e}_2$. Então com uma função F podemos definir um vetor dual em cada ponto que descreve este mapeamento linear. Este vetor dual definido em cada ponto é chamado a diferencial da função F e é escrito dF . Com esta notação temos

$$\delta F = F(P + \delta P, T + \delta T) - F(P, T) = dF(P, T)[\vec{a}] \quad (2.4.2)$$

onde $dF(P, T)[\vec{a}]$ é a aplicação do vetor dual $dF(P, T)$ sobre o vetor \vec{a} .

Note que $dF(P, T)$ não é uma grandeza infinitesimal nem é um número e nem depende do processo usado. $dF(P, T)$ é um vetor dual definido em cada ponto (P, T) do espaço de estados que depende unicamente da função F . Por outro lado $dF(P, T)[\vec{a}]$ é um valor infinitesimal que depende do processo usado através do vetor \vec{a} . A associação de pontos (P, T) e vetores duais $dF(P, T)$ define um campo de vetores duais dF que é chamado de *diferencial da função F* .

Como exemplo de diferencial vemos duas funções bem simples: a primeira é a função P e a segunda é a função T . Temos com $\vec{a} = \delta P \vec{e}_1 + \delta T \vec{e}_2$

$$\begin{aligned} dP[\vec{a}] &= \delta P \\ dT[\vec{a}] &= \delta T \end{aligned} \quad (2.4.3)$$

Com isso podemos combinar as equações (2.4.1), (2.4.2) e (2.4.3) e expressar dF como combinação linear de dP e dT .

$$dF = \left(\frac{\partial F}{\partial P} \right)_T dP + \left(\frac{\partial F}{\partial T} \right)_P dT \quad (2.4.4)$$

Portanto, dP e dT têm as propriedades de uma base. Para visualizar esta base graficamente notamos que as equações (2.4.3) podem ser escritas também na forma

$$dP[\vec{e}_1]=1, \quad dT[\vec{e}_2]=1, \quad dP[\vec{e}_2]=0, \quad dT[\vec{e}_1]=0 \quad (2.4.5)$$

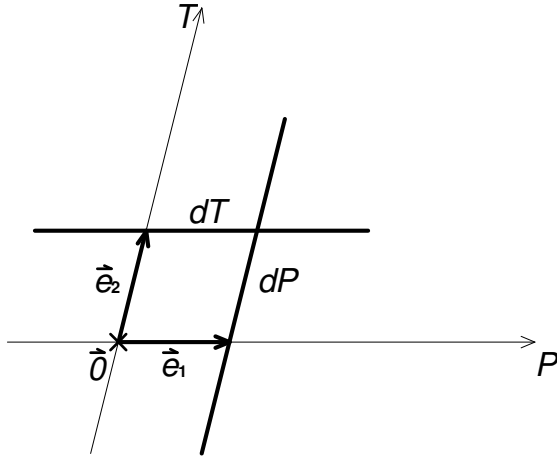


Fig. 2.11 As diferenciais dP e dT . As linhas de coordenadas P e T são propositadamente mostradas não ortogonais. No espaço de estados termodinâmicos não existe a noção de ortogonalidade.

No caso geral de um espaço de n dimensões, podemos escolher n coordenadas x_1, x_2, \dots, x_n e definir os vetores básicos dos espaços tangentes

$$\begin{aligned} \vec{e}_1 &= \frac{1}{\varepsilon} \overrightarrow{\delta x_1 = \varepsilon \setminus x_2, x_3, \dots, x_n} \\ \vec{e}_2 &= \frac{1}{\varepsilon} \overrightarrow{\delta x_2 = \varepsilon \setminus x_1, x_3, \dots, x_n} \\ &\cdot \\ &\cdot \\ \vec{e}_n &= \frac{1}{\varepsilon} \overrightarrow{\delta x_n = \varepsilon \setminus x_1, x_3, \dots, x_{n-1}} \end{aligned} \quad (2.4.6)$$

O vetor \vec{e}_k tem a direção da linha de coordenada k que é caracterizada pela constância de todas as outras coordenadas x_i ($i \neq k$). As diferenciais dx_k formam uma base do espaço dual do espaço tangente e temos em analogia com a equação (2.4.5)

$$dx_j[\vec{e}_k] = \delta_{jk} = \begin{cases} 1 & \text{para } j = k \\ 0 & \text{para } j \neq k \end{cases} \quad (2.4.7)$$

Note certa assimetria entre a base (\vec{e}_k) do espaço tangente e a base (dx_k) do espaço dual: um dx_k é bem definido se conhecermos apenas a coordenada x_k , mas para saber o que é \vec{e}_k temos que conhecer todas as coordenadas x_1, x_2, \dots, x_n .

A diferencial de uma função F seria

$$dF = \sum_{k=1}^n \left(\frac{\partial F}{\partial x_k} \right)_{x^j (j \neq k)} dx_k \quad (2.4.8)$$

e com a equação (2.4.7) temos

$$\boxed{dF[\bar{e}_k] = \left(\frac{\partial F}{\partial x_k} \right)_{x^j (j \neq k)}} \quad (2.4.9)$$

Podemos escrever uma *forma diferencial* geral (isto é um campo de vetores duais) como uma combinação linear dos vetores duais básicos dx_k

$$\bar{Q} = \sum_{j=1}^n q_j dx_j \quad (2.4.10)$$

onde os coeficientes q_j são funções de estado. Os matemáticos exigem ainda que estas funções sejam diferenciáveis para chamar o campo \bar{Q} uma forma diferencial. Nem toda forma diferencial é uma diferencial de uma função de estado. Pois sabemos (teorema de Clairaut e Schwarz⁸), que as derivadas parciais segundas (se elas foram funções contínuas) obedecem às relações

$$\frac{\partial^2 F}{\partial x_j \partial x_k} = \frac{\partial^2 F}{\partial x_k \partial x_j} \quad \text{para todo } k, j = 1, 2, \dots, n \quad (2.4.11)$$

Então para que uma forma diferencial (2.4.10) possa ser a diferencial de uma função é necessário que

$$\frac{\partial q_k}{\partial x_j} = \frac{\partial q_j}{\partial x_k} \quad \text{para todo } k, j = 1, 2, \dots, n \quad (2.4.12)$$

A diferença mais notável entre diferencial e forma diferencial que não é diferencial de uma função aparece quando integramos uma forma diferencial sobre um caminho. Podemos descrever um caminho C no espaço de estados como um ponto variável $C(\lambda)$ que depende de um parâmetro real entre 0 e 1. $C(0)$ seria o ponto inicial do caminho e $C(1)$ o ponto final. Vamos exigir que C seja diferenciável, isto é, as funções das coordenadas do ponto $C(\lambda)$ são funções diferenciáveis de λ . Se aumentarmos a partir do ponto $C(\lambda)$ o valor de λ por um acréscimo infinitesimal $\delta\lambda$, nos movemos para o ponto infinitesimalmente próximo $C(\lambda + \delta\lambda)$. O par de pontos $C(\lambda)$ e $C(\lambda + \delta\lambda)$ define um

⁸ Alexis Clairaut (1713-1765) e Hermann Schwarz (1843- 1021)

vetor $\overline{(C(\lambda), C(\lambda + \delta\lambda))}$ no espaço tangente do ponto $C(\lambda)$. Se temos uma forma diferencial \bar{Q} definida no espaço dos estados, podemos aplicar \bar{Q} no ponto $C(\lambda)$ sobre o vetor $\overline{(C(\lambda), C(\lambda + \delta\lambda))}$. O resultado é um valor infinitesimal $\bar{Q}(C(\lambda))\left[\overline{(C(\lambda), C(\lambda + \delta\lambda))}\right]$. Se repetirmos esta operação um número infinito de vezes até percorrer toda a curva C com avanços infinitesimais $\delta\lambda$ e se somarmos os valores infinitesimais resultantes $\bar{Q}(C(\lambda))\left[\overline{(C(\lambda), C(\lambda + \delta\lambda))}\right]$, obtemos a integral de \bar{Q} sobre o caminho C .

Esta descrição da integral $\int_C \bar{Q}$ é intuitiva, mas ela não serve para calcular uma integral.

Para escrever uma definição mais prática, vamos definir o vetor tangente da curva como

$$\vec{t}(\lambda) \stackrel{\text{def.}}{=} \lim_{\delta\lambda \rightarrow 0} \frac{\overline{(C(\lambda), C(\lambda + \delta\lambda))}}{\delta\lambda} \quad (2.4.13)$$

Com este vetor podemos definir a integral de \bar{Q} sobre o caminho C da seguinte forma:

$$\int_C \bar{Q} \stackrel{\text{def.}}{=} \int_0^1 \bar{Q}(C(\lambda))[\vec{t}(\lambda)] d\lambda \quad (2.4.14)$$

Em termos das coordenadas $x_k(\lambda) \equiv x_{k C(\lambda)}$ e com a representação de \bar{Q} na forma (2.4.10), esta definição da integral significa

$$\int_C \bar{Q} = \int_0^1 \left(\sum_{j=1}^n q_j \frac{dx_j(\lambda)}{d\lambda} \right) d\lambda \quad (2.4.15)$$

A integral pode ser definida também para caminhos que são constituídos por pedaços diferenciáveis. Neste caso o valor da integral é a soma dos valores dos pedaços diferenciáveis.

Se \bar{Q} for uma diferencial de uma função F , isto é, $\bar{Q} = dF$, sabemos que o valor da integral depende apenas dos pontos inicial e final do caminho

$$\int_C dF = F(C(1)) - F(C(0))$$

Mas se \bar{Q} não for a diferencial de uma função, a integral $\int_C \bar{Q}$ dependerá dos detalhes do caminho.

Veremos um exemplo: a forma diferencial do trabalho reversível⁹ num gás é $\bar{W} = -PdV$.
Escrevendo \bar{W} na forma (2.4.10) temos

$$\bar{W} = -PdV + 0 \cdot dT$$

Então a condição (2.4.12) não é satisfeita, pois $\left(\frac{\partial(-P)}{\partial T}\right)_V \neq 0$, mas $\left(\frac{\partial(0)}{\partial V}\right)_T = 0$.

Consequentemente \bar{W} não é a diferencial de uma função de estado. A integral de \bar{W} é o trabalho reversível fornecido para um sistema e este depende dos detalhes do processo. Mais tarde veremos que as máquinas térmicas de Carnot dependem desta propriedade do trabalho.

2.5) Fórmulas úteis para derivadas parciais

Bernhard Lesche

Sejam x_i (com $i = 1, \dots, n$) e \tilde{x}_i ($i = 1, \dots, n$) dois sistemas de coordenadas. Com a equação (2.4.8) podemos escrever a diferencial de uma função F nas coordenadas \tilde{x}

$$dF = \sum_{k=1}^n \left(\frac{\partial F}{\partial \tilde{x}_k} \right)_{x_j(j \neq k)} d\tilde{x}_k \quad (2.4.16)$$

Usando a (2.4.9) e aplicando aos os lados da equação (2.4.16) nos vetores básicos $\vec{e}_k = \frac{1}{\varepsilon} \overrightarrow{\delta x_k} = \varepsilon \backslash x^m; m \neq k$, obtemos a *regra de cadeia para derivadas parciais*:

$$\boxed{\left(\frac{\partial F}{\partial x_i} \right)_{x_k(k \neq i)} = \sum_{j=1}^n \left(\frac{\partial F}{\partial \tilde{x}_j} \right)_{x_k(j \neq k)} \left(\frac{\partial \tilde{x}_j}{\partial x_i} \right)_{x_k(k \neq i)}} \quad (2.4.17)$$

Imaginamos agora um espaço de duas dimensões. Sejam x, y e z funções neste espaço tal que cada um dos três pares de funções (x, y) , (x, z) e (y, z) serve como coordenadas. Mudando as coordenadas $(y, z) \rightarrow (x, z)$, podemos escrever com a regra da cadeia para qualquer função F

⁹ A definição deste conceito será explicada no capítulo 4.

$$(2.4.19) \quad \left(\frac{\partial F}{\partial x}\right)_z = \left(\frac{\partial F}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z + \left(\frac{\partial F}{\partial z}\right)_y \underbrace{\left(\frac{\partial z}{\partial x}\right)_z}_{=0} \quad (2.4.18)$$

Especialmente obtemos para $F = x$

$$\boxed{1 = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z} \quad (2.4.19)$$

Mudando as coordenadas $(x, z) \rightarrow (x, y)$, podemos escrever com a regra de cadeia:

$$\left(\frac{\partial F}{\partial x}\right)_y = \left(\frac{\partial F}{\partial x}\right)_z \underbrace{\left(\frac{\partial x}{\partial x}\right)_y}_{=1} + \left(\frac{\partial F}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \quad (2.4.20)$$

Especialmente obtemos para $F = y$

$$0 = \left(\frac{\partial y}{\partial x}\right)_z + \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \quad (2.4.21)$$

Combinando a (2.4.21) com a (2.4.19), obtemos

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial y}{\partial z}\right)_x = -1} \quad (2.4.22)$$

Agora vamos supor que tenhamos ainda uma quarta função u tal que o par (z, u) sirva também como um sistema de coordenadas. Mudando as coordenadas $(x, y) \rightarrow (y, z)$, podemos escrever com a regra de cadeia:

$$\left(\frac{\partial u}{\partial y}\right)_z = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial u}{\partial y}\right)_x \underbrace{\left(\frac{\partial y}{\partial y}\right)_z}_{=1} \quad (2.4.23)$$

Agora vamos botar $\left(\frac{\partial x}{\partial y}\right)_z$ em evidência:

$$\left(\frac{\partial u}{\partial y}\right)_z = \left[\left(\frac{\partial u}{\partial x}\right)_y + \frac{\left(\frac{\partial u}{\partial y}\right)_x}{\left(\frac{\partial x}{\partial y}\right)_z} \right] \left(\frac{\partial x}{\partial y}\right)_z \quad (2.4.24)$$

Usando a equação (2.4.19), podemos reescrever isso como

$$\left(\frac{\partial u}{\partial y}\right)_z = \left[\left(\frac{\partial u}{\partial x}\right)_y + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z \right] \left(\frac{\partial x}{\partial y}\right)_z \quad (2.4.25)$$

Usando novamente a regra de cadeia, podemos transformar a expressão em colchete numa derivada parcial de u :

$$\left[\left(\frac{\partial u}{\partial x}\right)_y + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z \right] = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial x}\right)_z + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z = \left(\frac{\partial u}{\partial x}\right)_z \quad (2.4.26)$$

Com isto a (2.4.25) toma a seguinte forma:

$$\frac{\left(\frac{\partial u}{\partial y}\right)_z}{\left(\frac{\partial u}{\partial x}\right)_z} = \left(\frac{\partial x}{\partial y}\right)_z \quad (2.4.27)$$

Utilizando finalmente a (2.4.19), obtemos:

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z = \frac{\left(\frac{\partial x}{\partial u}\right)_z}{\left(\frac{\partial y}{\partial u}\right)_z}} \quad (2.4.28)$$