

Curvas e superfícies implícitas: noções de geometrias diferencial e discreta

MARIA ANDRADE¹, ALLYSON CABRAL¹, VINÍCIUS MELLO², ADELAILSON PEIXOTO³ AND THOMAS LEWINER¹

Department of Mathematics — PUC — Rio de Janeiro — Brazil

Department of Mathematics — UFBA — Salvador — Brazil

Institute of Mathematics — UFAL — Maceió — Brazil

www.matmidia.mat.puc-rio.br/{maria,allyson,tomlew}.

<http://www.dmat.ufba.br/docentes/vinicius-moreira-mello>.

www.im.ufal.br/professores/adelailson.

Abstract. O estudo de curvas e superfícies permite introduzir, a alunos de graduação, conceitos simples das diversas geometrias, em particular diferencial e discreta. Através de exemplos palpáveis e com bases elementares de geometria descritiva, o aluno pode apreciar o potencial das ferramentas tanto diferenciais como numéricas, e o seu uso para criar expressões calculáveis que respeitem as invariâncias geométricas. Deste ponto de vista, o curso oferece ao aluno algumas noções úteis no seu currículo de pós-graduação em matemática tanto pura como aplicada, assim como expõe algumas linhas atuais da pesquisa matemática. Os objetivos deste curso é apresentar os fundamentos teóricos das geometrias euclidianas, diferenciais e discretas, com base nos casos de curvas e superfícies implícitas, além de mostrar algumas aplicações atuais como, por exemplo, a visualização de curvas e superfícies no computador. A teoria diferencial básica será introduzida a partir de noções intuitivas de imersão e submersão e das definições formais de curvaturas nos casos paramétrico e implícito, e seu cálculo usando ferramentas de geometria diferencial. Na parte discreta, mostraremos os problemas de definir invariantes discretos, e de aproximar os valores das derivadas. O final do curso ilustrará aplicações no computador e os temas de pesquisa atual que eles envolvem, em particular em torno do algoritmo Marching Cubes (artigo mais citado da história da computação gráfica) e dos problemas topológicos envolvidos na construção de representações discretas de superfícies implícitas.

Keywords: *Implicit curve. Implicit surface. Curvature. Isosurface. Marching Cubes.*



Figure 1: Superfície implícita colorida usando a informação de diferentes curvatura.

Sumário

1	Introdução	2	Fórmulas para Superfícies Implícitas	12
2	Curvas e Superfícies Paramétricas	3	Noções de Topologia de Superfícies	13
	Curvas Regulares	3	Interpolação e Derivação Discreta	16
	Comprimento de Arco e Curvatura	3	Interpolação Local	16
	Superfícies Regulares	5	Derivação de Sinais Interpolados	17
	Primeira Forma Fundamental	7	Visualização de Superfícies Implícitas	19
	Segunda Forma Fundamental	8	Problema da Invariância	21
	Cálculo das curvaturas	9	5 Representações Discretas Globais	23
3	Curvas e Superfícies Implícitas	11	Representações Discretas	23
	Derivação sobre uma Curva Implícita	11	Funções Multilineares por Partes	24
	Fórmulas para Curvas Implícitas	11	Marching Squares: Isocurvas	26
	Derivação numa Superfície Implícita	12	Marching Cubes: Isosuperfícies	27
			Problemas de Geometria e Topologia	29
		6	Conclusão	30
			Referências Bibliográficas	31

1 Introdução

A geometria é uma das disciplinas mais antigas da matemática, e se enriqueceu com muitas interações ao longo do tempo. Começando com problemas práticos de medição, do qual tirou o nome, a geometria descritiva foi formalizada muito cedo por Euclides. Os axiomas da geometria Euclidiana permitiram garantir a coerência entre as plantas desenhadas e a realização dos projetos de engenharia e da arquitetura da época. Depois, Arquimedes associou às definições construtivas de área e volume, métodos para calculá-los, abrindo ligações entre a geometria e o cálculo integral.

Por um lado, a geometria descritiva cresceu através das suas aplicações, a mais conhecida delas vindo das artes com a invenção da projeção perspectiva. Estas aplicações geraram grandes desafios, por exemplo definir medições de objetos tridimensionais diretamente a partir de uma projeção. Por outro lado, a introdução de coordenadas e equações de forma por Descartes e Fermat, junto com o desenvolvimento rápido do cálculo pela física, em particular o cálculo diferencial de Newton e Leibniz, permitiu estender os métodos de Arquimedes e Apolônio para o cálculo geométrico. As formalizações da geometria diferencial de Gauss e Riemann permitiram abrir a geometria para analisar problemas muito mais amplos, e até contribuir de volta à física, em particular na teoria da relatividade geral de Einstein.

Assim, o desenvolvimento espetacular do cálculo infinitesimal o tornou, até o século XX, a ferramenta predominante para tratar de problemas geométricos. Hoje, o computador é a principal ferramenta nas engenharias e em muitas ciências. Porém, ainda não é usado de forma extensiva em geometria, apesar do fundamento matemático da ciência da computação.

O presente livro descreve algumas abordagens diferenciais e discretas para curvas e superfícies implícitas. Os primeiros capítulos apresentam a formalização usual de geometria diferencial para superfícies implícitas. O leitor achará referências mais completas sobre a geometria diferencial de curvas e superfícies no livro de M. do Carmo [7]. Os últimos capítulos agrupam e discutem algumas das técnicas usadas para adaptar o cálculo diferencial em geometria para as capacidades discretas dos computadores. Eles tratam de temas independentes mas relacionados: o estudo da aproximação de quantidades geométricas locais e a construção de representações discretas de superfície. Para uma descrição mais completa do uso de superfícies discretas no computador, o leitor pode procurar no livro de L. Velho, L.H. de Figueiredo e J. Gomes [26].

O presente livro é destinado a alunos de graduação e procura ao mesmo tempo apresentar conceitos simples de geometria, e mostrar alguns problemas simples que são ainda temas de pesquisa atual. Em particular, o contraste entre a estrutura rigorosa dos primeiros capítulos e a apresentação mais prática dos últimos coloca em evidência o estado ainda iniciante da geometria discreta. O livro contém muitas sugestões de exercícios para ajudar a assimilar os conceitos

apresentados, e recomendamos fortemente que o leitor aproveite estas pausas na passividade da leitura para tornar este livro mais dinâmico. Atenderemos com muito prazer os leitores para tirar dúvidas, receber críticas ou sugestões por correio eletrônico.

*Maria, Allyson, Vinícius, Adailson e Thomas,
Rio de Janeiro, Maceió, Salvador e Aracaju. Janeiro de
2011.*

2 Curvas e Superfícies Paramétricas

A geometria diferencial de curvas e superfícies tem sido estudada desde as idéias iniciais que levaram ao cálculo, isto é desde tempos anteriores a Newton e Leibniz. A geometria diferencial Euclidiana clássica é o estudo de invariantes diferenciais em relação ao grupo dos movimentos rígidos. Neste sentido, vamos estudar propriedades geométricas de curvas em $\mathbb{R}^2$, tais como vetores tangente e normal, comprimento de arco e a curvatura da curva. Já em superfícies contidas em $\mathbb{R}^3$ apresentaremos os conceitos de plano tangente, vetor normal, primeira e segunda formas fundamentais e curvaturas.

A principal referência utilizada ao longo deste capítulo foi o livro de M. do Carmo [7].

(a) Curvas Regulares

Uma curva é dita diferenciável se ela pode ser descrita, ou parametrizada, por funções diferenciáveis. Porém, isto não implica que o desenho ou traço da curva seja suave. A suavidade das curvas ocorre nos casos regulares.

Nesta seção, vamos estudar curvas diferenciáveis em $\mathbb{R}^2$, mais precisamente curvas regulares. Além disso, conceituaremos alguns tipos de curvas como por exemplo: simples, periódica e fechada.

Definição 2.1. Uma curva diferenciável parametrizada é uma aplicação diferenciável $\alpha : I \rightarrow \mathbb{R}^2$, onde I é um intervalo real.

A palavra diferenciável na definição acima significa que α é uma aplicação que leva cada $t \in I$ em um ponto $\alpha(t) = (x(t), y(t)) \in \mathbb{R}^2$ tal que as funções reais $x(t)$, $y(t)$ são diferenciáveis.

O conjunto imagem C da aplicação α , dada por

$$C = \{(x(t), y(t)), t \in I\}$$

é chamado traço de α . A aplicação α é dita uma parametrização de C e denotaremos t o parâmetro da curva α .

Caso $I = (a, b)$, então os pontos limites $\alpha(a)$ e $\alpha(b)$, caso existam, são chamados pontos inicial e final de α , respectivamente. Se $\alpha(a) = \alpha(b)$ dizemos que α é uma curva fechada. Uma curva $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ é dita periódica se existe um número real $\rho > 0$, tal que

$$\alpha(t) = \alpha(t + \rho), \forall t \in \mathbb{R}.$$

O menor número ρ_0 tal que a equação acima é satisfeita é chamado de período.

A curva $\alpha : I \rightarrow \mathbb{R}^2$ é dita simples se a aplicação α for injetiva, isto é, se $\alpha(t_1) = \alpha(t_2)$, com $t_1, t_2 \in I$, então $t_1 = t_2$. Em outras palavras, não tem auto-intersecções.

Seja P um ponto em uma curva C . Dentre todas as retas passando por P existe uma reta que melhor aproxima a curva, tal reta é chamada de *reta tangente*. O vetor $\alpha'(t) = (x'(t), y'(t))$ é chamado de vetor tangente da curva α em t .

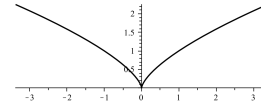


Figura 2: Curva com cúspide.

Exemplo 2.2. A aplicação $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $\alpha(t) = (t^3, t^2)$ com $t \in \mathbb{R}$ é uma curva suave diferenciável parametrizada cujo traço está esboçado na figura 2. Note que $\alpha'(0) = (0, 0)$, isto é o vetor tangente é nulo para $t = 0$.

Exemplo 2.3. A aplicação $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $\alpha(t) = (t, |t|)$, $t \in \mathbb{R}$, não é uma curva diferenciável parametrizada, pois $|t|$ não é diferenciável em $t = 0$.

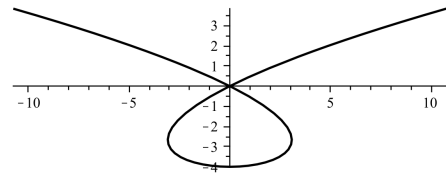


Figura 3: Curva com auto-interseção.

Exemplo 2.4. A aplicação α dada por $\alpha(t) = (t^3 - 4t, t^2 - 4)$, $t \in \mathbb{R}$, é uma curva diferenciável parametrizada não injetiva, pois temos $\alpha(2) = \alpha(-2) = (0, 0)$ (ver Figura 3).

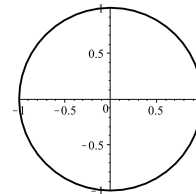


Figura 4: Círculo de raio 1.

Exemplo 2.5. A curva dada pela aplicação $\alpha : (0, 4\pi) \rightarrow \mathbb{R}^2$ com $\alpha(t) = (\cos(t), \sin(t))$ é uma curva diferenciável, periódica e fechada, mas não é injetiva, pois $\alpha(\pi) = \alpha(3\pi) = (-1, 0)$ (ver figura 4).

Definição 2.6. Seja $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ uma curva diferenciável parametrizada. Dizemos que α é **regular** se $\alpha'(t) \neq 0$, $\forall t \in I$. Caso contrário, o ponto $\alpha(t)$ tal que $\alpha'(t) = 0$ é **singular**.

(b) Comprimento de Arco e Curvatura

Consideraremos a partir de agora apenas curvas regulares.

Comprimento de arco

Nesta subseção mostraremos como calcular o comprimento de arco entre dois pontos de uma curva.

Definição 2.7. *Seja $t_0 \in I$. O comprimento de arco de uma curva parametrizada regular $\alpha : I \rightarrow \mathbb{R}^2$, dada por $\alpha(t) = (x(t), y(t))$ a partir do ponto t_0 , é definido por*

$$L(t) = \int_{t_0}^t \|\alpha'(u)\| du = \int_{t_0}^t \sqrt{x'(u)^2 + y'(u)^2} du.$$

Como $\alpha'(t) \neq 0$, temos que L é uma função diferenciável e $\frac{dL}{dt} = \|\alpha'(t)\|$. Se $\|\alpha'(t)\| = 1$, dizemos que α está parametrizada pelo comprimento de arco, então, neste caso,

$$L(t) = \int_{t_0}^t dt = t - t_0.$$

Exercício 2.8. *Mostre que o comprimento de arco está determinado de forma única a menos de uma constante.*

Exemplo 2.9. [Círculo de raio r] *Seja a curva $\alpha : [0, 2\pi) \rightarrow \mathbb{R}^2$ dada por*

$$\alpha(t) = (r \cos(t), r \sin(t)),$$

cujo traço é um círculo de raio r e centro na origem $(0, 0)$. Observemos que,

$$\alpha'(t) = (-r \sin(t), r \cos(t)) \text{ e } \alpha(t) \cdot \alpha'(t) = 0 \quad \forall t.$$

Isto significa que o vetor tangente é perpendicular ao raio. Notemos que $\alpha'(t) \neq 0$, $\forall t \in [0, 2\pi)$, logo podemos calcular o comprimento de arco do círculo que é dado por

$$L(\alpha) = \int_0^{2\pi} \|\alpha'(t)\| dt = \int_0^{2\pi} r dt = 2\pi r.$$

Dada uma curva $\alpha : I \rightarrow \mathbb{R}^2$ com parâmetro t , podemos reparametrizar a curva aplicando outro intervalo sobre I e usando a composição como uma nova curva. Mais precisamente, seja $h : J \rightarrow I$ diferenciável, onde $J \subset \mathbb{R}$ é um intervalo aberto real, então a reparametrização de α é

$$\beta = \alpha \circ h : J \rightarrow \mathbb{R}^2, \quad \beta(s) = \alpha(h(s)), \quad h(s) = t.$$

É fácil verificar, usando a regra da cadeia, que $\beta'(s) = \alpha'(h(s)) \cdot \frac{dh(s)}{ds}$.

Teorema 2.10. *Toda curva regular pode ser reparametrizada para obter velocidade unitária.*

Demonstração Seja α uma curva regular definida em I . O comprimento de arco é definido por

$$s(t) = \int_{t=a}^t \|\alpha'(u)\| du,$$

pelo teorema fundamental do cálculo temos que

$$\frac{ds}{dt} = \|\alpha'(t)\| > 0.$$

Usando o teorema do valor médio, segue que s é estritamente crescente em I . Portanto, é injetiva. Logo, s tem inversa na sua imagem, a qual denotaremos por $t(s)$ e suas respectivas derivadas estão relacionadas da seguinte forma

$$\frac{dt}{ds}(s) = \frac{1}{ds/dt(t(s))} > 0.$$

Seja $\beta(s) = \alpha(t(s))$. Então

$$|\beta'(s)| = \|\alpha'(t(s))\| \left| \frac{dt}{ds}(s) \right| = 1.$$

□

Curvatura

A curvatura indica o quanto a curva muda a direção. Seja $\alpha : I \rightarrow \mathbb{R}^2$ uma curva parametrizada pelo comprimento de arco. Podemos expressar essa direção como uma base positiva de $\mathbb{R}^2$ a partir de elementos geométricos da curva.

Vamos denotar por $\mathbf{t}(s)$ o vetor *tangente* $\alpha'(s)$, ou seja $\mathbf{t} : I \rightarrow \mathbb{R}^2$ é um vetor diferenciável e $\|\mathbf{t}\| = 1$. Existem somente dois vetores ortogonais a $\mathbf{t}$. Definimos então $\mathbf{n}(s) = J\mathbf{t}(s)$, onde $J : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ é a rotação de 90 graus no sentido anti-horário. O vetor $\mathbf{n}(s)$ é chamado de *normal* da curva $\alpha(s)$ em s . Com esta escolha temos que $\mathbf{n} : I \rightarrow \mathbb{R}^2$ é diferenciável e satisfaz:

$$\|\mathbf{n}(s)\| = 1, \langle \mathbf{t}(s), \mathbf{n}(s) \rangle = 0 \text{ e } \det(\mathbf{t}(s), \mathbf{n}(s)) = 1, \forall s \in I.$$

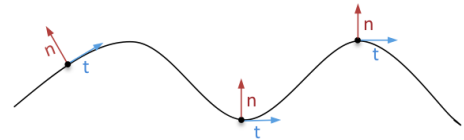


Figura 5: Vetores normais e tangentes.

Uma maneira de medir o quanto a curva muda de direção é observando como a base $\{\mathbf{t}(s), \mathbf{n}(s)\}$ associada a cada ponto da curva varia quando nos movemos ao longo da curva. Esta mudança pode ser controlada pelo significado das derivadas de $\mathbf{t}'(s)$ e $\mathbf{n}'(s)$. Diferenciando as expressões:

$$\|\mathbf{t}\|^2 = \|\mathbf{n}\|^2 = 1 \text{ e } \langle \mathbf{t}(s), \mathbf{n}(s) \rangle = 0,$$

obtemos

$$\begin{aligned} \langle \mathbf{t}'(s), \mathbf{t}(s) \rangle &= \langle \mathbf{n}'(s), \mathbf{n}(s) \rangle = 0 \text{ e} \\ \langle \mathbf{t}'(s), \mathbf{n}(s) \rangle &+ \langle \mathbf{t}(s), \mathbf{n}'(s) \rangle = 0. \end{aligned}$$

Portanto, o vetor $\mathbf{t}'(s)$ está na direção de $\mathbf{n}(s)$, isto é, existe uma função diferenciável $\kappa : I \rightarrow \mathbb{R}^2$, tal que $\mathbf{t}'(s) = \kappa(s) \mathbf{n}(s)$ ou ainda $\kappa(s) = \langle \mathbf{t}'(s), \mathbf{n}(s) \rangle$. O número

real $\kappa(s)$ é chamado de *curvatura* da curva α em $s \in I$. Então, temos

$$\mathbf{t}'(s) = \kappa(s) \mathbf{n}(s) \text{ e } \mathbf{n}'(s) = -\kappa(s) \mathbf{t}(s).$$

Notemos que $|\kappa(s)| = \|\mathbf{t}'(s)\| = \|\alpha''(s)\|$ que é o valor absoluto da aceleração da curva α . Como $\|\alpha'(s)\| = 1, \forall s \in I$, esta aceleração é centrípeta e não tangencial. Por outro lado,

$$\kappa(s) = \langle \mathbf{t}'(s), \mathbf{n}(s) \rangle = \langle \alpha''(s), J\alpha'(s) \rangle = \det(\alpha'(s), \alpha''(s)),$$

e então o sinal de $\kappa(s)$ informa sobre a orientação da base formada pelo vetor velocidade e a aceleração da curva. Isto é, se $\kappa(s) > 0$ então a curva muda sua direção no sentido anti-horário e se $\kappa(s) < 0$ no sentido horário.

Uma maneira simples de ver o que acabamos de falar é o seguinte, suponhamos que estamos viajando de Itabaiana para Aracaju pela BR-101 que a representaremos pela curva C , fixemos um sentido no qual a curva C é percorrida. Neste caso dizemos que a curva é orientada, então podemos colocar um sinal na curvatura para indicar se estamos dobrando à direita (-) ou à esquerda (+) (veja a Figura 6).

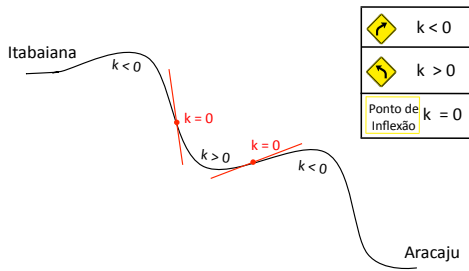


Figura 6: Sinais da curvatura.

Podemos ainda dar outra interpretação geométrica à curvatura da curva C . Seja P um ponto em C e seja r a reta tangente da curva em P . Existe um círculo que tangencia a reta r em P e que melhor aproxima a curva. Tal círculo é chamado de *círculo osculador*. Neste ponto a curvatura é dada como o inverso do raio, ou seja, quanto maior o raio do círculo no ponto P menor será a curvatura nesse ponto.

Exemplo 2.11 (Círculo). Seja $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ uma parametrização do círculo dado por

$$\alpha(t) = c + r \left(\cos\left(\frac{t}{r}\right), \sin\left(\frac{t}{r}\right) \right)$$

cujos centro é o ponto $c \in \mathbb{R}^2$ e raio $r > 0$. Se tomarmos $t_0 = 0$, temos

$$s(t) = \int_0^t \|\alpha'(u)\| du = rt, t \in (0, 2\pi].$$

Assim, uma reparametrização para α é $\beta(s) = c + r \left(\cos\left(\frac{s}{r}\right), \sin\left(\frac{s}{r}\right) \right)$, daí

$$\begin{aligned} \mathbf{t}'(s) &= \beta''(s) = -\frac{1}{r} \left(\cos\left(\frac{s}{r}\right), \sin\left(\frac{s}{r}\right) \right), \\ \mathbf{n}(s) &= J\beta'(s) = \left(-\cos\left(\frac{s}{r}\right), -\sin\left(\frac{s}{r}\right) \right), \end{aligned}$$

então $\kappa(s) = 1/r, \forall s \in \mathbb{R}$.

Exercício 2.12. Seja $\alpha : I \rightarrow \mathbb{R}^2$ uma curva regular definida por $\alpha(t) = (x(t), y(t))$, não necessariamente parametrizada pelo comprimento de arco. Mostre que a curvatura de α em $t \in I$ é dada por

$$\kappa = \frac{x'(t)y''(t) - x''(t)y'(t)}{\sqrt{(x'^2 + y'^2)^3}}.$$

Exercício 2.13. Seja a elipse definida por $\alpha(t) = (a\cos(t), b\sin(t))$, com $t \in I \subset \mathbb{R}$. Mostre que $\kappa = ab/(a^2\sin^2(t) + b^2\cos^2(t))^{3/2}$.

(c) Superfícies Regulares

Vemos exemplos de superfícies todos os dias, como pneus, balões, bolas de futebol, latas, por exemplo. Uma superfície regular em $\mathbb{R}^3$ é obtida tomando pedaços do plano, deformando-os e colando-os entre si de tal forma que a figura resultante não tenha pontas, arestas ou auto-interseções. A definição seguinte descreve essa propriedade de uma maneira mais formal.

Definição 2.14. Um subconjunto $S \subset \mathbb{R}^3$ é uma superfície regular se, para cada $p \in S$, existe uma vizinhança V de p em $\mathbb{R}^3$ e uma aplicação $\sigma : U \rightarrow V \cap S$ de um aberto U de $\mathbb{R}^2$ sobre $V \cap S$ tal que:

- σ é diferenciável.
- σ é um homeomorfismo. Como σ é contínua pela condição 1, isto significa que σ tem inversa $\sigma^{-1} : V \cap S \rightarrow U$ que é contínua.
- Para todo $q \in U$ a diferencial $d\sigma_q : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ é injetiva.

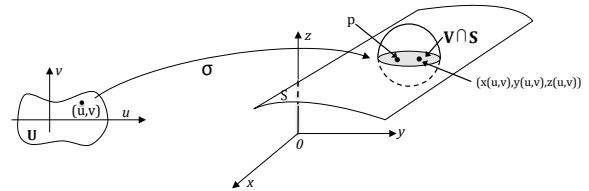


Figura 7: Definição de superfície.

Escrevendo σ em coordenadas, $\sigma(u, v) = \{x(u, v), y(u, v), z(u, v)\}$, podemos dizer que σ é diferenciável é equivalente a dizer que as funções x, y e z são diferenciáveis.

A última condição da definição equivale a $\sigma_u \times \sigma_v \neq 0$, ou seja, os vetores $\sigma_u = \frac{\partial \sigma}{\partial u}$ e $\sigma_v = \frac{\partial \sigma}{\partial v}$ são linearmente independentes.

A aplicação σ é chamada de *parametrização* em p . Ela tem o mesmo papel que a parametrização da curva α para superfícies, porém a expressão de σ pode variar de região para região permitindo mais tipos de superfícies (ver topologia no final do capítulo 3). A vizinhança $V \cap S$ de p é chamada *vizinhança coordenada* e as variáveis u, v são denominadas *coordenadas locais* de S .

Exemplo 2.15 (Plano). Seja $S = \{(x, y, z) \in \mathbb{R}^3 | ax + by + cz = d\}$ com $(a, b, c) \neq (0, 0, 0)$. Se $c \neq 0$, podemos reescrever a superfície como $S = \{(x, y, z) \in \mathbb{R}^3 | z = Ax + By + C\}$. Definimos a aplicação diferenciável $\sigma(u, v) = (u, v, Au + Bv + C)$. Temos que σ é diferenciável, σ é um homeomorfismo, cuja aplicação inversa é $\sigma^{-1}(x, y, z) = (x, y)$ e suas derivadas parciais são $\sigma_u = (1, 0, A)$ e $\sigma_v = (0, 1, B)$ são linearmente independentes em cada ponto. Logo, S é uma superfície diferenciável regular. Esta superfície pode ser apresentada por uma única parametrização.

Exercício 2.16. Seja $U \subset \mathbb{R}^2$ um conjunto aberto e seja $f : U \rightarrow \mathbb{R}$ aplicação diferenciável. O gráfico de f é o seguinte conjunto

$$S = \{(x, y, z) \in \mathbb{R}^3 | (x, y) \in U, z = f(x, y)\}.$$

Mostre que o gráfico é uma superfície regular.

Mudança de parâmetros

É interessante sabermos descrever analiticamente uma aplicação $f : S \rightarrow \mathbb{R}$ sobre uma superfície regular. Uma forma natural de pensarmos sobre o significado seria escolher para cada $p \in S$ uma parametrização, mas o problema é que tal ponto pode pertencer a duas ou mais parametrizações e precisa garantir que o valor de $f(p)$ seja o mesmo em todas. Então, é importante mostrarmos que isso não depende do sistema de coordenadas escolhido. Neste sentido, podemos enunciar o seguinte resultado:

Proposição 2.17 (Mudança de parâmetros). Seja p um ponto de uma superfície regular S e sejam $\sigma : U \subset \mathbb{R}^2 \rightarrow S$ e $\tau : V \subset \mathbb{R}^2 \rightarrow S$ duas parametrizações de S tais que $p \in \sigma(U) \cap \tau(V) = W$. Então, a mudança de coordenadas $h = \sigma^{-1} \circ \tau : \tau^{-1}(W) \rightarrow \sigma^{-1}(W)$ é um difeomorfismo: h é diferenciável e tem inversa diferenciável h^{-1} .

Demonstração Exercício. $\square$

Definição 2.18. Seja $f : S \rightarrow \mathbb{R}$ uma função definida em um subconjunto aberto V de uma superfície regular S . Então, f é diferenciável em $p \in V$ se para alguma parametrização $\sigma : U \subset \mathbb{R}^2 \rightarrow S$, com $p \in \sigma(U) \subset V$, a composição $f \circ \sigma : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ é diferenciável em $\sigma^{-1}(p)$. A função f é diferenciável em V se é diferenciável em todos os pontos de V .

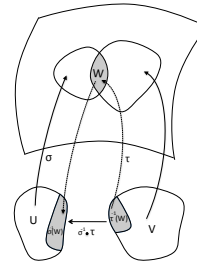


Figura 8: Mudança de Parâmetros

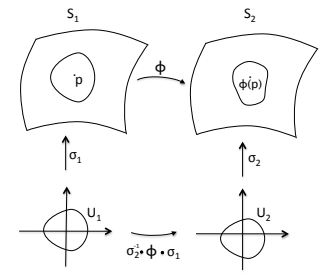


Figura 9: Aplicação diferenciável entre duas superfícies

Exemplo 2.19. Seja $f(p)$ uma função que calcula o quadrado da distância de um ponto p da superfície S a um ponto fixo $p_0 \in \mathbb{R}^3$. Temos que $f(p) = \|p - p_0\|^2$, $p \in S$, é uma função diferenciável.

Definição 2.20. Sejam S_1 e S_2 duas superfícies regulares. Dizemos que uma aplicação $\phi : V_1 \subset S_1 \rightarrow S_2$ é diferenciável em $p \in V_1$, se, dadas as parametrizações

$$\sigma_1 : U_1 \subset \mathbb{R}^2 \rightarrow S_1 \quad \text{e} \quad \sigma_2 : U_2 \subset \mathbb{R}^2 \rightarrow S_2,$$

onde $p \in \sigma_1(U_1)$ e $\phi(\sigma_1) \subset \sigma_2(U_2)$, então a aplicação

$$\sigma_2^{-1} \circ \phi \circ \sigma_1 : U_1 \rightarrow U_2$$

é diferenciável em $\sigma_1^{-1}(p)$.

Exemplo 2.21. Seja $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $\phi(x, y, z) = (xa, yb, zc)$, onde a, b e c são números reais não-nulos. Temos que ϕ é diferenciável e que a restrição $\phi|_{S^2}$ é uma aplicação da esfera

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$$

sobre o elipsóide

$$E^2 = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \right\}.$$

Plano tangente

Podemos também considerar curvas desenhadas sobre S . Mostraremos que, para cada $p \in S$, o conjunto de vetores tangentes às curvas parametrizadas de S , passando por p , constituem um plano, o qual denotaremos por $T_p S$ (plano tangente à superfície em p).

Definição 2.22. Uma curva parametrizada diferenciável $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$ é uma aplicação diferenciável $\alpha(t) = (x(t), y(t), z(t))$. Uma curva α é desenhada sobre S se para todo $t \in I$, existem funções $u(t), v(t)$ diferenciáveis, tais que $\alpha(t) = \sigma(u(t), v(t))$.

Definição 2.23. Fixado $(u_0, v_0) \in U \subset \mathbb{R}^2$, as curvas

$$\begin{aligned}\alpha(t) &= \sigma(u(t), v_0), \\ \beta(t) &= \sigma(u_0, v(t)),\end{aligned}$$

são chamadas as curvas coordenadas de σ em $(u_0, v_0) \in U$.

Proposição 2.24. Seja $\sigma : U \subset \mathbb{R}^2 \rightarrow S$ uma parametrização de uma superfície regular S e seja $q \in U$. O subespaço vetorial de dimensão 2, $d\sigma_q(\mathbb{R}^2) \subset \mathbb{R}^3$ coincide com o conjunto de vetores tangentes às curvas desenhadas sobre S passando em p .

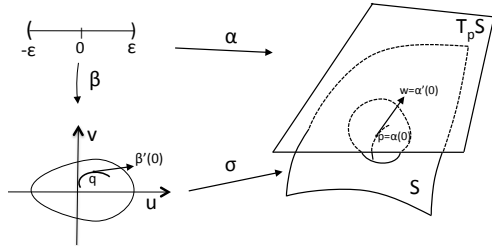


Figura 10: Plano tangente.

Demonstração A demonstração é deixada como exercício, mas a ideia está na figura 10.) $\square$

Ou seja, o plano tangente de S em p pode ser visto como

$$T_p(S) = \{v, v \text{ é tangente a } S \text{ em } p\}.$$

As coordenadas de um vetor $w \in T_p S$ na base associada a uma parametrização σ são determinadas da seguinte forma. Seja $\beta : (-\epsilon, \epsilon) \rightarrow U$ uma curva em U dada por $\beta(t) = (u(t), v(t))$ e seja $\alpha : (-\epsilon, \epsilon) \rightarrow S$ definida por $\alpha(t) = \sigma \circ \beta(t)$, com $\beta(0) = q = \sigma^{-1}(p)$. Então

$$\begin{aligned}\alpha'(0) &= \frac{d}{dt}(\sigma \circ \beta) \\ &= \sigma_u(q)u'(0) + \sigma_v(q)v'(0)\end{aligned}$$

Assim, na base $\{\sigma_u(q), \sigma_v(q)\}$ w tem coordenada $u'(0), v'(0)$.

Vale notar que a noção de plano tangente é transportada (preservada) por aplicações diferenciáveis. Sejam S_1 e S_2 duas superfícies regulares e seja $\phi : V \subset S_1 \rightarrow S_2$ aplicação diferenciável. Seja $p \in V$, sabemos que todo vetor $v \in T_p S_1$ é o vetor velocidade $\alpha'(0)$ de uma curva diferenciável $\alpha : (-\epsilon, \epsilon) \rightarrow V$ com $\alpha(0) = p$. A curva $\beta = \phi \circ \alpha$ é tal que $\beta(0) = \phi(p)$ e, portanto, $\beta'(0)$ é um vetor tangente em $T_{\phi(p)} S_2$.

Exercício 2.25. Dado w , como acima, mostre que o vetor $\beta'(0)$ não depende da escolha de α e que a aplicação $d\phi_p : T_p S_1 \rightarrow T_{\phi(p)} S_2$ definida por $d\phi_p(w) = \beta'(0)$ é linear.

(d) Primeira Forma Fundamental

Nesta seção, estudaremos a primeira forma fundamental, objeto que permite medir a área de regiões e o comprimento de curvas em superfícies.

O produto interno de $\mathbb{R}^3 \supset S$, induz em cada plano tangente $T_p S$ de uma superfície regular S um produto interno, que indicaremos por $\langle \cdot, \cdot \rangle : T_p S \times T_p S \rightarrow \mathbb{R}$ que associa a cada par de vetores $(w_1, w_2) \in (T_p S)^2$ um número real. A este produto interno, que é uma forma bilinear simétrica, corresponde uma forma quadrática $I_p : T_p S \rightarrow \mathbb{R}$ dada por $I_p(w) = \langle w, w \rangle_p = |w|^2 \geq 0$.

Definição 2.26. A forma quadrática $I_p : T_p S \times T_p S \rightarrow \mathbb{R}$ é chamada de primeira forma fundamental da superfície $S \subset \mathbb{R}^3$ em $p \in S$.

Vamos expressar I_p em termos da base $\{\sigma_u, \sigma_v\}$ de $T_p S$. Seja $w \in T_p S$, então existe uma curva diferenciável $\alpha : (-\epsilon, \epsilon) \rightarrow S$ dada por $\alpha(t) = \sigma(u(t), v(t))$ tal que $\alpha(0) = p = \sigma(u_0, v_0)$ e $\alpha'(0) = w$. Obtemos os coeficientes da primeira forma fundamental:

$$\begin{aligned}I_p(\alpha'(0)) &= \langle \alpha'(0), \alpha'(0) \rangle_p \\ &= \langle \sigma_u u' + \sigma_v v', \sigma_u u' + \sigma_v v' \rangle_p \\ &= \langle \sigma_u, \sigma_u \rangle_p u'^2 + 2\langle \sigma_u, \sigma_v \rangle_p u'v' + \langle \sigma_v, \sigma_v \rangle_p v'^2 \\ &= Eu'^2 + 2Fu'v' + Gv'^2,\end{aligned}$$

onde os valores das funções são calculadas em $t = 0$ e os coeficientes: $E(u_0, v_0) = \langle \sigma_u, \sigma_u \rangle$, $F(u_0, v_0) = \langle \sigma_u, \sigma_v \rangle$ e $G(u_0, v_0) = \langle \sigma_v, \sigma_v \rangle$ são os coeficientes da primeira forma fundamental na base $\{\sigma_u, \sigma_v\}$ de $T_p S$. Fazendo p variar na vizinhança coordenada correspondente a $\sigma(u, v)$, obtemos que as funções $E(u, v)$, $F(u, v)$ e $G(u, v)$ são diferenciáveis nessa vizinhança.

Exemplo 2.27. Seja $P \subset \mathbb{R}^3$ um plano passando por $p_0 = (x_0, y_0, z_0)$ e contendo os vetores ortonormais w_1 e w_2 . Então uma parametrização para P é

$$\mathbf{x}(u, v) = p_0 + uw_1 + vw_2.$$

Os coeficientes da primeira forma fundamental são, para esse plano: $E = 1$, $F = 0$ e $G = 1$.

Exercício 2.28. Considere a esfera $\mathbb{S}^2(r)$ centrada na origem e de raio $r > 0$. Sejam $V = \{(\theta, \phi); 0 < \theta < \pi, 0 < \phi < 2\pi\}$ e $\sigma : V \rightarrow \mathbb{R}^3$ dada por

$$\sigma(\theta, \phi) = (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta).$$

Mostre que σ é uma parametrização para $\mathbb{S}^2(r)$, e que os coeficientes da primeira forma fundamental são:

$$E = r^2, \quad F = 0 \quad e \quad G = r^2 \sin^2 \theta.$$

Comprimento de curva sobre uma superfície

Agora mostraremos como a primeira forma fundamental está relacionada com o comprimento de arco de curvas. Seja $\alpha : J \subset \mathbb{R} \rightarrow S$ curva parametrizada, vimos que o comprimento de arco é

$$s(t) = \int_0^t \|\alpha'(t)\| dt = \int_0^t \sqrt{I(\alpha'(t))} dt.$$

Se $\alpha(t) = \sigma(u(t), v(t))$ está contida em uma vizinhança coordenada correspondente à parametrização $\sigma(u, v)$, podemos calcular o comprimento de arco de α entre $a \leq t \leq b$ por

$$s(t) = \int_a^b \sqrt{E(u')^2 + 2Fu'v' + Gv'^2} dt.$$

Área de uma região em uma superfície

Nesta subseção, mostraremos a relação entre os coeficientes da primeira forma fundamental e a área de uma região em uma superfície.

Seja $\sigma : U \subset \mathbb{R}^2 \rightarrow S$ uma parametrização da superfície S . A função:

$$\|\sigma_u \times \sigma_v\| = \sqrt{EG - F^2}$$

representa a área do paralelogramo de lados σ_u e σ_v . Integrando esta a área infinitesimal $\|\sigma_u \times \sigma_v\| du dv$ sobre a superfície obtemos a área dela.

Definição 2.29. Seja $R \subset S$ uma região limitada de uma superfície regular contida na imagem da parametrização $\sigma : U \subset \mathbb{R}^2 \rightarrow S$. Então a área de R é dada por

$$A(R) = \iint_{\sigma^{-1}(R)} \|\sigma_u \times \sigma_v\| du dv.$$

Exercício 2.30. Mostre que $A(R)$ não depende da escolha da parametrização.

(e) Segunda Forma Fundamental

Nesta seção, estudaremos a segunda forma fundamental, objeto que permite introduzir as curvaturas, em particular as curvaturas Gaussiana e média, as quais dão informações sobre o comportamento local da superfície.

Aplicação normal de Gauss

A seguir, vamos estudar a aplicação normal de Gauss. A variação desta dá origem ao conceito de curvatura. Seja $\sigma : U \subset \mathbb{R}^2 \rightarrow S$ uma parametrização da superfície regular S . Podemos escolher para cada $p \in \sigma(U)$, um vetor normal unitário, dado por

$$\mathbf{N}(p) = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}(p), \quad p \in \sigma(U).$$

A aplicação $\mathbf{N} : \sigma(U) \rightarrow \mathbb{S}^2$ que toma seus valores na esfera unitária $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 + z^2 = 1\}$ é chamada *aplicação normal de Gauss*.

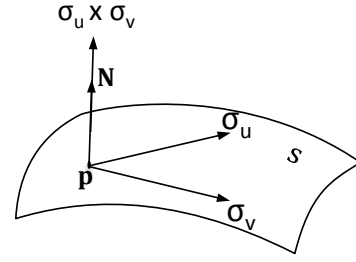


Figura 11: Vetor normal.

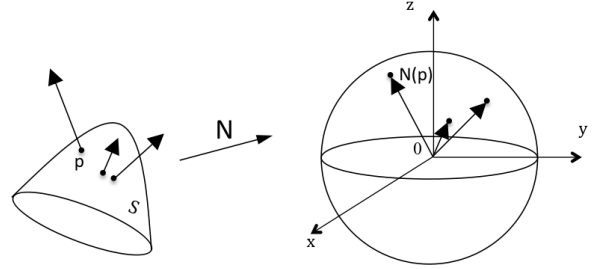


Figura 12: A aplicação de Gauss.

A sua derivada $d\mathbf{N}_p : T_p S \rightarrow T_{\mathbf{N}(p)} S$ é uma aplicação linear. Como $T_p S$ e $T_{\mathbf{N}(p)} S$ são paralelos (perpendiculares ao mesmo normal) podemos identificá-los, assim

$$d\mathbf{N}_p : T_p S \rightarrow T_p S.$$

Exemplo 2.31. Seja o plano $P = \{(x, y, z) \in \mathbb{R}^3 / ax + by + cz + d = 0\}$. O vetor normal $\mathbf{N}$ no ponto $p \in P$, dado por $\mathbf{N} = (a, b, c) / \sqrt{a^2 + b^2 + c^2}$, é constante, logo $d\mathbf{N}_p = 0$.

Exercício 2.32. Mostre que a aplicação $d\mathbf{N}_p$ é auto-adjunta, isto é $\langle d\mathbf{N}_p v, w \rangle = \langle v, d\mathbf{N}_p w \rangle$.

Definição 2.33. Seja $p \in S$ e seja $d\mathbf{N}_p : T_p S \rightarrow T_p S$ a diferencial da aplicação de Gauss. O determinante de $d\mathbf{N}_p$ é chamado curvatura Gaussiana K de S em p . A metade do traço de $d\mathbf{N}_p$ é chamado de curvatura média H de S em p .

Definição 2.34. A forma quadrática $\mathcal{I}_p : T_p S \times T_p S \rightarrow \mathbb{R}$, definida em $T_p S$ por $\mathcal{I}_p(v) = -\langle d\mathbf{N}_p(v), v \rangle$ é chamada segunda forma fundamental de S em p .

Vamos escrever a expressão da segunda forma fundamental na base $\{\sigma_u, \sigma_v\}$. Seja $\alpha(t) = \sigma(u(t), v(t))$ uma curva parametrizada em S , com $\sigma(0) = p$. O vetor tangente a $\alpha(t)$ em p é $\alpha'(t) = \sigma_u u' + \sigma_v v'$ e $d\mathbf{N}(\alpha') = \mathbf{N}'(u(t), v(t)) = \mathbf{N}_u u' + \mathbf{N}_v v'$. Portanto,

$$\begin{aligned} \mathcal{I}_p(\alpha') &= -\langle d\mathbf{N}_p(\alpha'), \alpha' \rangle \\ &= -\langle \mathbf{N}_u u' + \mathbf{N}_v v', \sigma_u u' + \sigma_v v' \rangle \\ &= e u'^2 + 2f u' v' + g v'^2, \end{aligned}$$

onde $e = -\langle \mathbf{N}_u, \sigma_u \rangle$, $f = -\langle \mathbf{N}_v, \sigma_u \rangle$ e $g = -\langle \mathbf{N}_v, \sigma_v \rangle$. Os coeficientes e , f e g são chamados coeficientes da segunda forma fundamental.

Curvatura

Daremos agora uma interpretação geométrica da segunda forma fundamental.

Definição 2.35. Seja C uma curva regular na superfície S passando por $p \in S$, κ a curvatura de C em p , e $\cos(\theta) = \langle \mathbf{n}, \mathbf{N} \rangle$, onde $\mathbf{n}$ é o vetor normal a C e $\mathbf{N}$ é o vetor normal a S em p . O número $\kappa_n = \kappa \cos(\theta)$ é chamado a curvatura normal de $C \subset S$ em p .

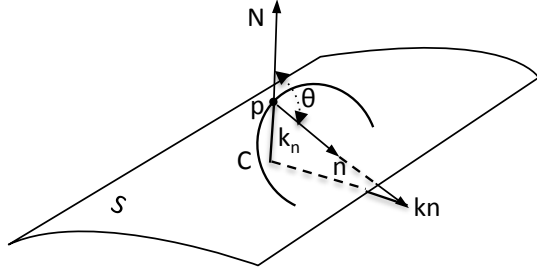


Figura 13: Curvatura normal.

A curvatura normal κ_n é o comprimento da projeção do vetor $\kappa \mathbf{n}$ sobre o normal à superfície em p , com sinal dado pela orientação $\mathbf{N}$ de S em p .

A curvatura normal de C não depende da orientação de C , mas troca de sinal com uma mudança de orientação da superfície.

Podemos dar uma interpretação geométrica da segunda forma fundamental $\mathcal{I}_p$ utilizando a curvatura normal. De fato, Seja $C \subset S$ uma curva parametrizada pelo comprimento de arco $\alpha(s)$, com $\alpha(0) = p$. Se indicarmos por $\mathbf{N}(s)$ a restrição do vetor normal $\mathbf{N}$ à curva $\alpha(s)$, teremos $\mathcal{I}_p(\alpha'(0)) = \kappa_n(p)$.

Exercício 2.36. Com a notação acima, prove que $\mathcal{I}_p(\alpha'(0)) = \kappa_n(p)$.

Como $d\mathbf{N}_p$ é uma aplicação auto-adjunta sabemos pelo teorema espectral que existe uma base $\{e_1, e_2\}$ ortonormal de $T_p S$ tal que

$$d\mathbf{N}_p(e_1) = -k_1 e_1 \text{ e } d\mathbf{N}_p(e_2) = -k_2 e_2.$$

Definição 2.37. O máximo k_1 da curvatura e o mínimo k_2 da curvatura são chamadas curvaturas principais em p ; as direções correspondentes, isto é, as direções dadas pelos autovalores e_1 e e_2 são chamadas direções principais em p .

Em termos das curvaturas principais, as curvaturas Gaussiana e média são, respectivamente, dadas por

$$K = k_1 k_2 \text{ e } H = \frac{1}{2}(k_1 + k_2).$$

As curvaturas principais podem ser interpretadas da seguinte maneira: dado $p \in S$, existe um plano tangente e o

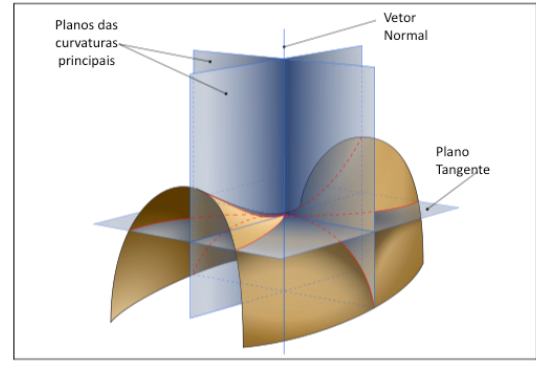


Figura 14: Interpretação das curvaturas principais (© wikipedia).

vetor normal $\mathbf{N}$, definidos em p . Considerando todos os planos que passam por p e contêm $\mathbf{N}$, as intersecções destes planos com a superfície nos dão uma família de curvas. As curvaturas máxima e mínima dessa família são as curvaturas principais da superfície.

(f) Cálculo das curvaturas

Vamos calcular as fórmulas explícitas para as curvaturas Gaussiana e média de uma superfície parametrizada regular, em função dos coeficientes da primeira e da segunda forma fundamental. Vimos na subseção anterior que $d_p \mathbf{N} : T_p S \rightarrow T_p S$, assim podemos escrever $\mathbf{N}_u, \mathbf{N}_v$ na base $\{\sigma_u, \sigma_v\}$ do plano tangente $T_p S$, ou seja, existem números reais $(a_{i,j})_{1 \leq i,j \leq 2}$ tais que

$$\begin{aligned} \mathbf{N}_u &= a_{11}\sigma_u + a_{12}\sigma_v, \\ \mathbf{N}_v &= a_{21}\sigma_u + a_{22}\sigma_v. \end{aligned} \quad (1)$$

Vamos encontrar os coeficientes $a_{i,j}$ em termos da base $\{\sigma_u, \sigma_v, \mathbf{N}\}$.

Tomando o produto interno de cada uma das igualdades da expressão (1) por σ_u e σ_v , obtemos

$$\begin{aligned} \langle \sigma_u, \mathbf{N}_u \rangle &= a_{11}\langle \sigma_u, \sigma_u \rangle + a_{12}\langle \sigma_u, \sigma_v \rangle, \\ \langle \sigma_v, \mathbf{N}_v \rangle &= a_{21}\langle \sigma_v, \sigma_u \rangle + a_{22}\langle \sigma_v, \sigma_v \rangle, \\ \langle \sigma_u, \mathbf{N}_v \rangle &= a_{21}\langle \sigma_u, \sigma_u \rangle + a_{22}\langle \sigma_v, \sigma_u \rangle, \\ \langle \sigma_v, \mathbf{N}_u \rangle &= a_{11}\langle \sigma_v, \sigma_u \rangle + a_{12}\langle \sigma_v, \sigma_v \rangle. \end{aligned} \quad (2)$$

Como $\langle \sigma_u, \mathbf{N} \rangle = 0 = \langle \sigma_v, \mathbf{N} \rangle$, temos

$$\begin{aligned} \langle \sigma_u, \mathbf{N}_u \rangle &= -\langle \sigma_{uu}, \mathbf{N} \rangle, \\ \langle \sigma_v, \mathbf{N}_v \rangle &= -\langle \sigma_{vv}, \mathbf{N} \rangle, \\ \langle \sigma_u, \mathbf{N}_v \rangle &= -\langle \sigma_{uv}, \mathbf{N} \rangle. \end{aligned}$$

Assim,

$$\begin{aligned} e &= -\langle \mathbf{N}_u, \sigma_u \rangle = \langle \mathbf{N}, \sigma_{uu} \rangle, \\ f &= -\langle \mathbf{N}_v, \sigma_v \rangle = \langle \mathbf{N}, \sigma_{vv} \rangle = \langle \mathbf{N}, \sigma_{vu} \rangle = -\langle \mathbf{N}_u, \sigma_v \rangle, \\ g &= -\langle \mathbf{N}_v, \sigma_u \rangle = \langle \mathbf{N}, \sigma_{uv} \rangle. \end{aligned}$$

Vamos obter, agora, os coeficientes $(a_{i,j})_{1 \leq i,j \leq 2}$ em termos de e , f e g . Usando as equações (2), segue que

$$\begin{aligned} -f &= \langle \mathbf{N}_u, \sigma_v \rangle = a_{11} \langle \sigma_u, \sigma_v \rangle + a_{21} \langle \sigma_v, \sigma_v \rangle = a_{11}F + a_{21}G, \\ -f &= \langle \mathbf{N}_v, \sigma_u \rangle = a_{12} \langle \sigma_u, \sigma_u \rangle + a_{22} \langle \sigma_u, \sigma_v \rangle = a_{12}E + a_{22}F, \\ -e &= \langle \mathbf{N}_u, \sigma_u \rangle = a_{11} \langle \sigma_u, \sigma_u \rangle + a_{21} \langle \sigma_v, \sigma_u \rangle = a_{11}E + a_{21}F, \\ -g &= \langle \mathbf{N}_v, \sigma_v \rangle = a_{12} \langle \sigma_u, \sigma_v \rangle + a_{22} \langle \sigma_v, \sigma_v \rangle = a_{12}F + a_{22}G, \end{aligned}$$

onde $E = \langle \sigma_u, \sigma_u \rangle$, $F = \langle \sigma_v, \sigma_u \rangle$ e $G = \langle \sigma_v, \sigma_v \rangle$ são os coeficientes da primeira forma fundamental. Em termos de matrizes

$$-\begin{pmatrix} e & f \\ f & g \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Em particular temos

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1}.$$

Como

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} = \frac{1}{EG - F^2} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix},$$

segue que

$$\begin{aligned} a_{11} &= \frac{fF - eG}{EG - F^2}, \\ a_{12} &= \frac{gF - fG}{EG - F^2}, \\ a_{21} &= \frac{eF - fE}{EG - F^2}, \\ a_{22} &= \frac{fF - gE}{EG - F^2}. \end{aligned}$$

Daí, obtemos

$$K(p) = \det(a_{i,j}) = \frac{eg - f^2}{EG - F^2} = \frac{\det(\mathcal{I}_p)}{\det(I_p)}.$$

Calculando a curvatura, vimos anteriormente que k_1 e k_2 satisfazem

$$d\mathbf{N}(v) = -k(v) = -kI(v), \text{ para algum } v \in T_p S - \{0\},$$

onde $I : T_p S \rightarrow T_p S$ é a aplicação identidade. Por definição de autovalores temos que a aplicação linear $d\mathbf{N} = kI$ não é invertível. Logo,

$$\det \begin{pmatrix} a_{11} + k & a_{12} \\ a_{21} & a_{22} + k \end{pmatrix} = 0.$$

Isto é equivalente a:

$$k^2 + k \underbrace{(a_{11} + a_{22})}_{\text{tr}(A)} + \underbrace{(a_{11}a_{22} - a_{12}a_{21})}_{\det(A)} = 0,$$

aqui $A = (a_{i,j})_{1 \leq i,j \leq 2}$.

Como k_1 e k_2 são as raízes da equação acima, obtemos:

$$H = \frac{k_1 + k_2}{2} = -\frac{1}{2} \text{tr}(A) = -\frac{1}{2} \frac{eG - 2fF + gE}{EG - F^2}.$$

Exemplo 2.38 (Plano). Vimos anteriormente que $d\mathbf{N} = 0$. Logo, $K = H = 0$.

Exemplo 2.39 (Esfera). Consideremos a parametrização definida no exercício 2.28. Temos que o vetor normal em cada ponto é

$$\mathbf{N} = (\sin(\theta)\cos(\psi), \sin(\theta)\sin(\psi), \cos(\theta)).$$

Verificamos que os coeficientes da segunda forma fundamental são: $e = -r$, $f = 0$ e $g = -r\sin^2(\theta)$. Logo, as curvaturas Gaussiana e média, respectivamente, são $K = r^{-2}$ e $H = r^{-1}$.

Exercício 2.40. Verifique os cálculos do exemplo anterior.

3 Curvas e Superfícies Implícitas

Neste capítulo estudaremos as curvas e superfícies dadas por funções implícitas. Encontraremos os elementos geométricos do capítulo anterior a partir de funções implícitas. Além disso, daremos algumas noções de topologia de curvas e superfícies.

A noção de submersão nos permite definir a superfície como um conjunto de pontos que satisfazem certas propriedades. Por exemplo, a esfera é o conjunto de pontos que são equidistantes a um ponto, o centro da esfera. O teorema de Sard nos diz que quase sempre tal definição gera uma superfície regular, ou seja, a medida total de probabilidade de ocorrer tal evento é 1.

Teorema 3.1. [Teorema de Sard] *O conjunto dos valores críticos de f tem medida total, ou seja, a probabilidade de um valor ser crítico para f é nula.*

Utilizamos neste capítulo algumas referências, a saber: o artigo de R. Goldman [10], onde se encontra uma vasta exposição sobre o cálculo de curvatura em superfícies implícitas, e o livro de M. Armstrong [2], na parte de topologia.

(a) Derivação sobre uma Curva Implícita

Nesta seção vamos definir o vetor gradiente e exibiremos o teorema da função implícita, o qual garante que localmente uma curva pode ser vista como um gráfico de função real.

Seja $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ função diferenciável. O vetor gradiente, o qual denotaremos por ∇f , é uma aplicação diferenciável definida por

$$\begin{aligned} \nabla f : U &\longrightarrow \mathbb{R}^3 \\ p &\longmapsto \nabla f(p) = (f_x(p), f_y(p), f_z(p)), \end{aligned}$$

onde f_x, f_y e f_z são as derivadas de f com relação a x, y e z .

O teorema da função implícita diz que localmente uma superfície pode ser vista como um gráfico (caso particular da superfície paramétrica). Ele permite encontrar propriedades geométricas como vetores tangente, normal e curvaturas de superfícies implícitas a partir do capítulo anterior. Enunciaremos o teorema para $\mathbb{R}^3$, mas o resultado vale para o $\mathbb{R}^n$.

Teorema 3.2. [Teorema da função implícita] *Sejam $U \subset \mathbb{R}^3$ um aberto, $p = (x_0, y_0, z_0) \in U$, $a \in \mathbb{R}$ e seja $f : U \rightarrow \mathbb{R}$ função diferenciável. Suponhamos que $f(p) = a$ e $f_z(p) \neq 0$. Então existem uma vizinhança V de (x_0, y_0) em $\mathbb{R}^2$ e uma vizinhança W de z_0 em $\mathbb{R}$ tal que $V \times W \subset U$ e uma função diferenciável $g : V \rightarrow W$ com $g(x_0, y_0) = z_0$ tal que:*

$$\{p \in V \times W \mid f(p) = a\} = \{(x, y, g(x, y)) \in \mathbb{R}^3 \mid (x, y) \in V\}.$$

Em outras palavras, se $f_z(p) \neq 0$, podemos utilizar a equação $f(x, y, z) = a$ para expressar z como uma função de x e y , em uma certa vizinhança de p .

Em cálculo, a regra da cadeia é uma fórmula para derivada da função composta de duas funções. A regra da cadeia

afirma que: $(f \circ g)'(x) = (f(g(x)))' = f'(g(x))g'(x)$, e é aplicada para funções de mais de uma variável. Considere a função $f(x, y)$, e as funções $x(t)$ e $y(t)$. A derivada da composta $f(t) = f(x(t), y(t))$ é dada por

$$f'(t) = f_x(x(t), y(t)) \cdot x'(t) + f_y(x(t), y(t)) \cdot y'(t).$$

(b) Fórmulas para Curvas Implícitas

Nesta seção encontraremos os vetores tangente $\mathbf{t}$ e normal $\mathbf{n}$ e a curvatura κ de uma curva implícita $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. A principal ideia para encontrarmos elementos geométricos é usarmos o teorema da função implícita o qual garante que localmente a curva pode ser vista como um gráfico. Curvas implícitas podem ser descritas pela equação $f(x, y) = a$, onde $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Desse ponto de vista, uma curva implícita é o conjunto de pontos $C = \{(x, y) \in U \mid f(x, y) = a\}$ que satisfazem a equação. Vamos supor que o vetor gradiente $\nabla f = (f_x, f_y)$ é não-nulo. Podemos supor sem perda de generalidade que $f_y \neq 0$, então existe um intervalo aberto $J \subset \mathbb{R}$ e uma função $g : J \rightarrow \mathbb{R}$ tal que a curva localmente é dada como um gráfico $\{(x, g(x)); x \in J\}$. Usando a seção 2(b) temos as fórmulas dos elementos $\mathbf{t}, \mathbf{n}$ e κ , mas observemos que tais fórmulas são dadas em função de g

$$\begin{aligned} \mathbf{t} &= (1 + g_x^2)^{-1/2}(1, g_x), \\ \mathbf{n} &= (1 + g_x^2)^{-1/2}(-g_x, 1), \\ \kappa &= \frac{g_{xx}}{(1 + g_x^2)^{3/2}}, \end{aligned} \quad (3)$$

assim devemos buscar informações das derivadas da g com as da f . Um cálculo relativamente simples mostra que:

$$\begin{aligned} g_x &= -\frac{f_x}{f_y}, \\ g_{xx} &= \frac{-f_{xx}}{f_y} + \frac{2f_{xy}f_x}{f_y^2} - \frac{f_{yy}f_x^2}{f_y^3}. \end{aligned} \quad (4)$$

Portanto,

$$\begin{aligned} \mathbf{t} &= (f_x^2 + f_y^2)^{-1/2}(f_y, -f_x), \\ \mathbf{n} &= (f_x^2 + f_y^2)^{-1/2}(f_x, f_y), \\ \kappa &= -\frac{f_{xx}f_x^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2}{(f_x^2 + f_y^2)^{3/2}}. \end{aligned} \quad (5)$$

Exercício 3.3. Verifique as expressões (3), (4) e (5).

Exercício 3.4. Considere o círculo de raio r dado pela seguinte função $f(x, y) = x^2 + y^2 - r^2$. Mostre que

$$\mathbf{t} = r^{-1}(y, -x), \quad \mathbf{n} = r^{-1}(x, y), \quad \kappa = r^{-1}.$$

Exercício 3.5. Mostre que $\nabla f \neq 0$ equivale à curva ser regular.

(c) Derivação numa Superfície Implícita

Nesta seção, estudaremos as superfícies implícitas e veremos que o plano tangente em um ponto é dado pelo núcleo da função que define a superfície neste ponto. Além disso, escreveremos explicitamente o vetor normal $\mathbf{N}$ e as curvaturas Gaussiana e média da superfície. Para isso serão necessárias algumas definições e o teorema da função implícita.

Definição 3.6. *Seja $U \subset \mathbb{R}^3$ um conjunto aberto e seja $f : U \rightarrow \mathbb{R}$ aplicação diferenciável e seja $a \in \mathbb{R}$. Dizemos que a é um valor regular de f se, para cada $p \in U$ com $f(p) = a$, temos $(df)_p \neq 0$, ou equivalentemente, $\nabla f(p) \neq 0$.*

O próximo resultado mostra que a pré-imagem de um valor regular é uma superfície regular.

Proposição 3.7. *Seja $a \in \mathbb{R}$ um valor regular de uma função diferenciável, onde U é um conjunto aberto de $\mathbb{R}^3$. Se $S = f^{-1}(a)$ é um conjunto não-vazio, então S é uma superfície.*

Demonstração Basta utilizar o teorema da função implícita e o fato que todo gráfico é uma superfície regular. $\square$

Exemplo 3.8. *O elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ é uma superfície regular. De fato, é o conjunto $f^{-1}(0)$, onde*

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$$

é uma função diferenciável e 0 é um valor regular da função f , pois, $f_x = 2x/a^2$, $f_y = 2y/b^2$, $f_z = 2z/c^2$ se anulam simultaneamente apenas no ponto $(0, 0, 0)$, que não pertence a $f^{-1}(0)$. Em particular, temos que a esfera é uma superfície regular, basta tomar $a = b = c = 1$.

Exercício 3.9. *Mostre que o paraboloide $z = x^2 + y^2$ é uma superfície regular.*

Plano Tangente Seja $U \subset \mathbb{R}^3$ um aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável tal que a seja um valor regular de f pertencendo a sua imagem, temos que $S = f^{-1}(a)$ é uma superfície, com plano tangente

$$T_p S = \ker((df)_p : \mathbb{R}^3 \rightarrow \mathbb{R}), \forall p \in S.$$

De fato, se $v \in T_p S$, então existe uma curva $\alpha : (-\epsilon, \epsilon) \rightarrow S$ tal que $\alpha(0) = p$ e $\alpha'(0) = v$. Portanto, $(f \circ \alpha)(t) = a, \forall t$, e por diferenciação em $t = 0$,

$$(df)_p(v) = (f \circ \alpha)'(0) = 0.$$

Portanto, v pertence ao núcleo de $(df)_p$. Como $T_p S$ e $\ker((df)_p)$ são subespaços lineares de dimensão dois e um está contido no outro, temos o resultado.

Vetor Normal O espaço tangente à superfície $S = f^{-1}(a)$, onde a é um valor regular de f , no ponto p é o complementar ortogonal da reta gerada $\nabla f(p)$. Seja $v \in T_p S$, então existe uma curva $\alpha : (-\epsilon, \epsilon) \rightarrow S$ tal que $\alpha(0) = p$ e $\alpha'(0) = v$, logo $f \circ \alpha(t) = a$, temos que

$$\langle \nabla f(p), \alpha'(0) \rangle = (f \circ \alpha)'(0) = 0,$$

que mostra que $\nabla f(p)$ é ortogonal a $T_p S$.

Consideremos a superfície $S = \{(x, y, z) \in \mathbb{R}^3 / f(x, y, z) = 0\}$, onde f é uma função suave. Seja $p \in S$ tal que $\nabla f(p) \neq 0$, então podemos supor, sem perda de generalidade, que $f_z(p) \neq 0$. Isso implica, pelo teorema da função implícita, que existe uma função $g : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que a equação $z = g(x, y)$ descreve a superfície S em uma vizinhança de p , V_p , ou seja, nessa vizinhança S é parametrizada como um gráfico $\mathcal{G} = \{(x, y, g(x, y)) / (x, y) \in U\}$. Calcularemos os coeficientes da primeira forma fundamental em $p = (x, y, g(x, y))$, os quais denotaremos por E, F, G .

Usando a hipótese $f_z(p) \neq 0$, obtemos:

$$\begin{aligned} g_x(x, y) &= -\frac{f_x(p)}{f_z(p)}, \\ g_y(x, y) &= -\frac{f_y(p)}{f_z(p)}, \\ g_{xx}(x, y) &= -\frac{f_{xx}}{f_z} + 2\frac{f_{xz}f_x}{f_z^2} - \frac{f_{zz}f_x^2}{f_z^3}, \\ g_{xy}(x, y) &= -\frac{f_{xy}}{f_z} + \frac{f_{yz}f_x + f_y f_{xz}}{f_z^2} - \frac{f_y f_{zz}f_x}{f_z^3}, \\ g_{yy}(x, y) &= -\frac{f_{yy}}{f_z} + 2\frac{f_{yz}f_y}{f_z^2} - \frac{f_{zz}f_y^2}{f_z^3}. \end{aligned} \quad (6)$$

(d) Fórmulas para Superfícies Implícitas

Dos resultados anteriores, podemos deduzir os coeficientes da primeira forma fundamental:

$$E = \frac{f_z^2 + f_x^2}{f_z^2}, \quad F = \frac{f_x f_y}{f_z^2}, \quad G = \frac{f_z^2 + f_y^2}{f_z^2}. \quad (7)$$

O vetor normal é

$$\mathbf{N} = \frac{(-g_x, -g_y, 1)}{\sqrt{g_x^2 + g_y^2 + 1}} = (f_z^2 + f_x^2 + f_y^2)^{-1/2} (f_x, f_y, f_z). \quad (8)$$

Exercício 3.10. *Verifique as expressões (6), (7) e (8).*

Os coeficientes da segunda forma fundamental e, f e g são

$$\begin{aligned} e &= \langle \mathbf{N}, (0, 0, g_{xx}) \rangle \\ &= -\frac{f_{xx}f_z^2 - 2f_{xz}f_x f_z + f_{zz}f_x^2}{f_z^2 \sqrt{f_z^2 + f_x^2 + f_y^2}} = \frac{\det(A_1)}{f_z^2 |\nabla f|}, \\ f &= \langle \mathbf{N}, (0, 0, g_{xy}) \rangle \\ &= -\frac{f_{xy}f_z^2 - f_z f_{yz}f_x - f_z f_y f_{xz} + f_y f_{zz}f_x}{f_z^2 \sqrt{f_z^2 + f_x^2 + f_y^2}} = \frac{\det(A_2)}{f_z^2 |\nabla f|}, \end{aligned}$$

$$\begin{aligned}
g &= \langle N, (0, 0, g_{yy}) \rangle = -\frac{f_{yy}f_z^2 - 2f_{yz}f_yf_z + f_{zz}f_y^2}{f_z^2\sqrt{f_z^2 + f_x^2 + f_y^2}} \\
&= \frac{\det(A_3)}{f_z^2|\nabla f|},
\end{aligned}$$

onde

$$\begin{aligned}
A_1 &= \begin{vmatrix} f_{xx} & f_{xz} & f_x \\ f_{xz} & f_{zz} & f_z \\ f_x & f_z & 0 \end{vmatrix}, \\
A_2 &= \begin{vmatrix} f_{xy} & f_{yz} & f_y \\ f_{xz} & f_{zz} & f_z \\ f_x & f_z & 0 \end{vmatrix} e \\
A_3 &= \begin{vmatrix} f_{yy} & f_{yz} & f_y \\ f_{yz} & f_{zz} & f_z \\ f_y & f_z & 0 \end{vmatrix}.
\end{aligned}$$

As curvaturas Gaussiana e médias em p são dadas por:

$$\begin{aligned}
K &= \frac{\det(A_1)\det(A_3) - \det(A_2)^2}{f_z^2|\nabla f|^4} \\
&= \frac{(f_{zz}f_{yy} - f_{yz}^2)f_x^2 + (-2f_{xy}f_{zz} + 2f_{xz}f_{yz})f_yf_x}{(f_z^2 + f_x^2 + f_y^2)^2} \\
&+ \frac{2(-f_{xz}f_{yy} + f_{xy}f_{yz})f_xf_z + (f_{xx}f_{zz} - f_{xz}^2)f_y^2}{(f_z^2 + f_x^2 + f_y^2)^2} \\
&+ \frac{-2(-f_{xz}f_{xy} + f_{xx}f_{yz})f_yf_z + (f_{xx}f_{yy} - f_{xy}^2)f_z^2}{(f_z^2 + f_x^2 + f_y^2)^2}, \\
H &= \frac{1}{2|\nabla f|^3} \left(\det(A_1) \left(1 + \frac{f_y^2}{f_z^2} \right) - 2\det(A_2) \left(\frac{f_xf_y}{f_z^2} \right) + \right. \\
&\quad \left. + \det(A_3) \left(1 + \frac{f_x^2}{f_z^2} \right) \right) \\
&= \frac{(f_y^2 + f_z^2)f_{xx} + (f_x^2 + f_z^2)f_{yy} + (f_x^2 + f_y^2)f_{zz}}{2|\nabla f|^3} \\
&- \frac{(f_xf_yf_{xy} + f_xf_zf_{xz} + f_yf_zf_{yz})}{|\nabla f|^3}.
\end{aligned}$$

Observação. Notemos que ao permutarmos as variáveis x , y e z nas expressões de K e H obtemos o mesmo resultado. Isto decorre da invariância das curvaturas.

Exercício 3.11. Verifique as expressões de e , f , g , K e H .

Exercício 3.12. Faça um programa utilizando o Maple para calcular os vetores tangente e normal, as curvaturas Gaussiana e média nos casos em que a superfície é dada na forma paramétrica e implícita.

(e) Noções de Topologia de Superfícies

Mostraremos algumas definições básicas sobre a topologia de superfície, em particular o teorema da classificação de superfícies limitadas, sem bordo e conexas. Restringimos ao caso de superfície implícita considerando apenas superfícies sem bordo (o bordo seria o que chamamos de “furo” ou fronteira da superfície).

Definição 3.13. Se X é um conjunto, uma topologia em X é um conjunto $\tau \subset \mathcal{P}(X)$ (conjunto das partes) tal que

- $\emptyset \in \tau$, $X \in \tau$,
- Toda família enumerável $(A_j)_{j \in J}$ de elementos de τ deve verificar $\bigcup_{j \in J} A_j \in \tau$,
- Toda família finita $(A_j)_{j \in J}$ de elementos de τ deve verificar $\bigcap_{j \in J} A_j \in \tau$.

Os elementos de τ são chamados de abertos.

Um espaço topológico é um par ordenado (X, τ) sendo X um conjunto e τ uma topologia em X .

Exemplo 3.14. O espaço $\mathbb{R}$ com a topologia formada pela união de intervalos abertos é um espaço topológico.

Os espaços topológicos permitem definir formalmente conceitos como conectividade.

Definição 3.15. Uma cisão de um conjunto X é uma decomposição $X = A \cup B$, onde $A \cap B = \emptyset$ e os conjuntos A e B são abertos em X . Uma cisão é dita trivial quando A ou B é vazio.

Exemplo 3.16 (Cisão não-trivial). Na topologia de $\mathbb{R}$ definida pelas uniões de abertos, considere o conjunto $\mathbb{R} - \{0\} = (-\infty, 0) \cup (0, \infty)$.

Definição 3.17. Um conjunto chama-se conexo quando não admite outra cisão além da trivial. Assim, se X é conexo, $X = A \cup B$, com A e B disjuntos e abertos em X , então $A = \emptyset$ ou $B = \emptyset$.

Exemplo 3.18. Os conjuntos $\mathbb{R}^n$ e $\emptyset$ são conjuntos conexos, já o conjunto $\mathbb{R} - \{0\}$ é desconexo, pois admite uma cisão não-trivial.

Definição 3.19. Seja $x \in X \subset \mathbb{R}^n$. A componente conexa do ponto x no conjunto X é a reunião C_x de todos os subconjuntos conexos de X que contém o ponto x .

Exemplo 3.20. A componente conexa de 1 em $\mathbb{R} - \{0\}$ é $(0, \infty)$.

Orientação de Superfícies Dada uma superfície regular S cuja parametrização é $\sigma(u, v)$, temos que os vetores σ_u e σ_v são linearmente independentes, logo $\{\sigma_u, \sigma_v, \mathbf{N}\}$ é uma base positiva orientada para $\mathbb{R}^3$. Isto quer dizer que a matriz cujas colunas são esses vetores tem determinante positivo.

Definição 3.21. Dizemos que a superfície S é orientável se for possível escolher, para cada $p \in S$, uma orientação $T_p S$ que varie continuamente com p , ou seja se existir uma aplicação contínua $\mathbf{N} : S \rightarrow \mathbb{S}^2$ tal que, para cada p , $\mathbf{N}(p)$ seja ortogonal a $T_p S$. A tal campo de vetores normais $\mathbf{N}$ chamamos orientação de S .

Exemplo 3.22. As superfícies implícitas são orientáveis, pois o campo $\mathbf{N}(p) = \frac{\nabla f(p)}{\|\nabla f(p)\|}$ é uma orientação de $f^{-1}(a)$.

Exemplo 3.23. Uma superfície que é um gráfico de uma função diferenciável é uma superfície orientável: é um caso particular de superfícies que podem ser cobertas por uma única parametrização.

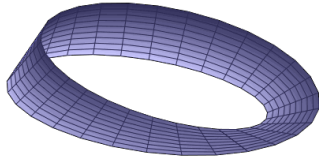


Figura 15: Faixa de Möbius.

Exemplo 3.24. A faixa de Möbius $\mathbb{M}$ é a superfície obtida quando colamos as duas extremidades de um retângulo alongado de papel de modo a fazer coincidir os vértices opostos. Tal superfície é não-orientável.

A topologia estuda as propriedades do espaço topológico que permanecem inalteradas sobre certas transformações. As deformações decorrentes destas transformações não podem envolver cortes ou colagens para serem topologicamente válidas. A transformação que vamos estudar é conhecida como homeomorfismo, ou seja, é uma aplicação entre espaços topológicos que é bijetiva, contínua e cuja inversa também é contínua. O espaço topológico que vamos trabalhar é o $\mathbb{R}^n$. É bom entender o conceito de função contínua, a saber

Definição 3.25. Dizemos que $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua, se para qualquer aberto $U \subset \mathbb{R}^m$ a pré-imagem $f^{-1}(U)$ é um aberto de $\mathbb{R}^m$.

Algumas superfícies bidimensionais como a esfera, o cilindro e o toro moram no $\mathbb{R}^3$, como está ilustrado na figura 16. Um exemplo mais complicado é a Garrafa de Klein pois qualquer tentativa de representá-la em três dimensões terá auto-interseção. Podemos construir um modelo da Garrafa de Klein da seguinte forma. Primeiro construímos um retângulo com orientação, em seguida colamos dois dos lados desse retângulo transformando-o em um cilindro e então

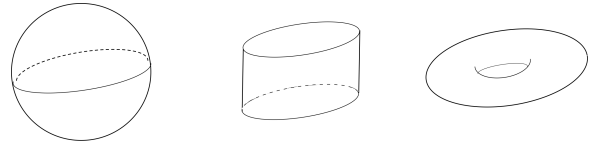


Figura 16: Superfícies: esfera, cilindro e toro.

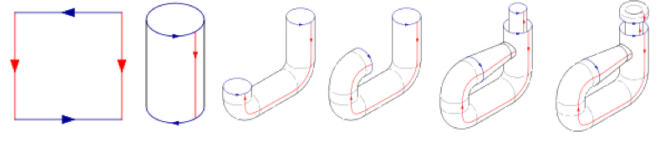


Figura 17: Construção da garrafa de Klein (© wikipedia).

as extremidades do cilindro são identificadas nas direções opostas. Para fazer isso o cilindro precisa ser dobrado em torno de uma extremidade e empurrado pelo lado como mostra a Figura 17. A garrafa de Klein pode ser representada em um espaço 4-dimensional sem auto-interseção.

Definição 3.26. Dizemos que duas superfícies M e N são topologicamente equivalentes se existe $h : M \rightarrow N$ homeomorfismo.

Exemplo 3.27. Seja $\mathbb{S}^2$ a esfera. Se a apertamos um pouco obtemos uma outra superfície que é topologicamente equivalente à esfera (ver Figura 18).

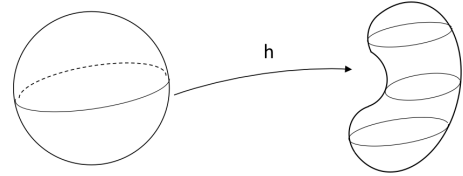


Figura 18: Superfícies topologicamente equivalentes.

Vamos considerar superfícies limitadas, sem bordo e conexas, e classificar essas superfícies para saber se são topologicamente equivalentes. Vimos na Figura 18 um exemplo de superfícies topologicamente equivalentes. Vamos construir alguns exemplos de superfícies que são limitadas, sem bordo e conexas.

Exemplo 3.28 (Esfera com uma alça). Consideramos a esfera $\mathbb{S}^2$ e tiramos dois discos disjuntos e então adicionamos um cilindro identificando seus bordos circulares com os bordos da esfera sem os discos (ver a Figura 19). Este processo é o que chamamos de adicionar alças. Repetindo o processo obteremos uma esfera com duas, três ou qualquer número finito de alças. Observemos que uma esfera com uma alça é homeomorfa a um toro.

Exemplo 3.29 (Plano Projetivo). Começamos com uma esfera $\mathbb{S}^2$ e retiramos um disco e adicionamos a faixa de Möbius neste lugar. O resultado é uma superfície fechada chamada plano projetivo denotada por $\mathbb{P}^2$ (ver a Figura 20). Um fato importante é que $\mathbb{P}^2$ mora em $\mathbb{R}^4$.

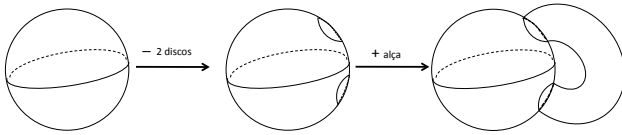


Figura 19: Esfera com uma alça.

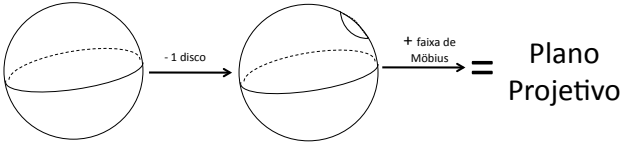


Figura 20: Construção plano projetivo.

Exemplo 3.30 (Garrafa de Klein). Começamos com uma esfera $\mathbb{S}^2$ e retiramos dois discos. Esta nova superfície é homeomorfa ao cilindro. Agora adicionamos uma faixa de Möbius em cada bordo (círculo) do cilindro. O resultado é uma superfície fechada chamada garrafa de Klein (ver a Figura 21).

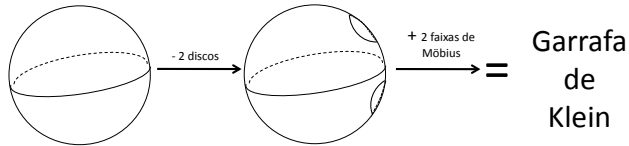


Figura 21: Construção da garrafa de Klein.

Teorema 3.31 (Classificação de superfícies). Qualquer superfície limitada, sem bordo e conexa é homeomorfa a:

- Esfera.
- Esfera com um número finito g de alças.
- Esfera com um número finito g de discos removidos e substituídos por g faixas de Möbius.

Demonstração A prova deste resultado pode ser encontrada no livro de M. Armstrong [2]. $\square$

As superfícies nas duas primeiras famílias são orientáveis. A esfera pode ser vista como um 0 toro e no último item temos que as superfícies são não orientáveis, em particular o plano projetivo e a garrafa de Klein são não orientáveis. O número g de alças é chamado **genus** da superfície.

Superfícies com várias componentes conexas são classificadas pela classe de cada uma de suas componentes conexas e assim, resta classificar superfícies conexas.

A partir do teorema da classificação é possível definir outro invariante topológico, ou seja, ao fazermos deformações na superfície essa propriedade topológica não é alterada.

Definição 3.32. A característica de Euler χ para superfícies limitadas, sem bordo e conexas é definida por

- $\chi = 2$, se a superfície é uma esfera.

- $\chi = 2 - 2g$, se a superfície é uma esfera com um número finito g de alças.
- $\chi = 2 - g$, se a superfície é uma esfera com um número finito g de discos removidos e substituídos por g faixas de Möbius.

Exercício 3.33. Mostre que a característica de Euler da garrafa de Klein é 0.

4 Interpolação e Derivação Discreta

As fórmulas para calcular as curvaturas de superfícies implícitas são calculáveis no computador, dado os valores das derivadas da função implícita f . Quando a função implícita é dada por uma fórmula algébrica, é possível usar cálculo simbólico para obter estas derivadas. Porém, nos casos usuais, a função é apenas medida (amostrada) como um sinal em alguns pontos do espaço, geralmente os vértices de uma grade (reticulado) regular. Neste capítulo estudaremos como, a partir do sinal discreto, encontrar uma aproximação contínua. Além disso, veremos como calcular derivadas a partir de dados interpolados e com isso extrair informações como normais e valores de curvaturas.

(a) Interpolação Local

Em muitas aplicações computacionais, desejamos obter uma função contínua $f : \mathbb{R} \rightarrow \mathbb{R}$ a partir de um sinal discreto $g[k]$, de tal modo que $f(k) = g[k]$ para todo $k \in \mathbb{Z}$. Esse problema é denominado *problema de interpolação de sinais discretos*. Nestes casos, podemos empregar interpolação linear, de modo que $f(t) = (1 - t)g[i] + tg[i + 1]$, onde $i = \lfloor x \rfloor$ denota a parte inteira e $t = \{x\}$ a parte fracionária de x . Interpolação linear é uma solução muito eficiente para o problema de interpolação. Entretanto, em alguns problemas, como no cálculo de curvaturas, mais uma condição é necessária, a de que f seja de classe C^2 . Neste caso, deve-se recorrer a outro método de interpolação, pois em geral as funções obtidas por interpolação linear não são diferenciáveis nos pontos $k \in \mathbb{Z}$.

Interpolação polinomial

Uma abordagem geral para o problema da interpolação é considerar que a função f pertence a um certo espaço de funções que possua as propriedades desejadas. Em aplicações de Computação Gráfica, os espaços mais usados são os *espaços das funções splines de grau n* ,

$S^n = \{f \in C^{n-1}(\mathbb{R}) \mid f|_{[k, k+1]} \text{ é um polinômio de grau } n\}$,

ou seja, S^n é o espaço das funções de classe C^{n-1} que são polinomiais por partes nos intervalos $[k, k + 1]$ da reta.

Exercício 4.1. Verifique que para cada sinal discreto $g[k]$ existe uma única função em S^1 que satisfaz à condição de interpolação, exatamente a função obtida por interpolação linear.

Nos modelos tradicionais de visualização espera-se apenas que a superfície seja regular, em outras palavras, que f definida em $\mathbb{R}^3$ seja de classe C^1 , dessa forma podem ser definidos planos tangentes em cada um dos seus pontos. Esse fato é muito importante pois garante que vetores normais possam ser calculados sobre esses pontos para serem utilizados, por exemplo, no processo de iluminação da cena. Incluiremos mais uma condição, a de que f seja de classe C^2 , ou seja, que suas derivadas até a segunda ordem sejam contínuas. Essa garantia é fundamental para o cálculo da curvatura sobre pontos em superfícies descritas por f .

Entretanto, para que f seja de classe C^2 , devemos buscá-la no espaço S^3 , conhecido como *espaço das splines cúbicas*. Nas próximas seções, mostraremos como resolver o problema de interpolação para *splines* cúbicas, ou seja, dado um sinal $g[k]$, determinar a única função $f \in S^3$ tal que $f(k) = g[k]$, para todo $k \in \mathbb{Z}$. Além disso, será apresentado um método para obtenção de valores de curvatura sobre superfícies.

Interpolação por splines

Para uma boa compreensão do problema de interpolação cúbica, é necessária uma prévia caracterização dos espaços S^n . Analisando o método de interpolação linear, é fácil ver que se $f \in S^1$ então

$$f(x) = \sum_{k \in \mathbb{Z}} c[k] B_1(x - k), \quad (9)$$

onde $B_1(x)$ é a *função B-spline de grau 1* exibida na Figura 22(a) e $c[k] = g[k]$. O prefixo B antes da palavra *spline* significa que as translações inteiras da função B_1 formam uma base do espaço vetorial S^1 . Note que como B_1 possui suporte compacto, para cada x o somatório em (9) é finito.

É possível mostrar, que para cada espaço S^n , com $n > 1$, existe uma função B_n , chamada *função B-spline de grau n* , tal que o conjunto $\{B_n(\cdot - k)\}_{k \in \mathbb{Z}}$ é uma base do espaço S^n , ou seja, cada função $f \in S^n$ pode ser escrita da forma

$$f(x) = \sum_{k \in \mathbb{Z}} c[k] B_n(x - k), \quad (10)$$

para coeficientes $c[k]$ unicamente determinados por f . O gráfico das funções B_n para $n = 0, \dots, 3$ pode ser visto na Figura 22.

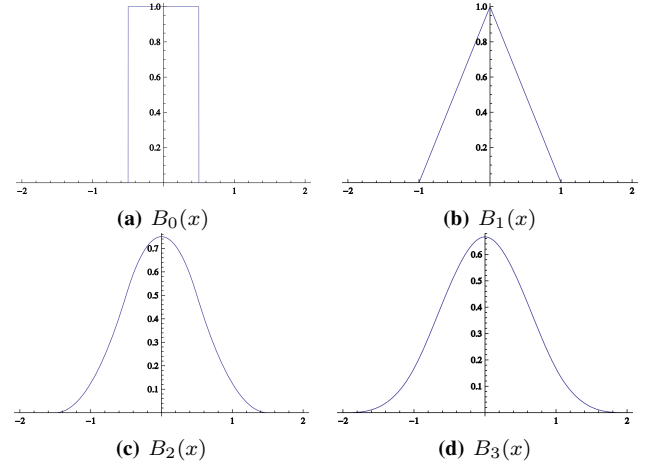


Figura 22: Funções B-spline.

Exercício 4.2. Existem várias fórmulas recursivas para as funções B_n . Mostre que $B_{n+1} = B_n \star B_0$, onde B_0 é a função característica do intervalo $(-\frac{1}{2}, \frac{1}{2}]$, conhecida como função box, e a operação $f \star g$ é o produto de convolução entre duas funções,

$$f \star g(x) = \int_{-\infty}^{\infty} f(t)g(x - t)dt.$$

Aplicando a definição recursiva, concluímos que

$$B_3(x) = \begin{cases} 0 & , \text{ se } |x| \geq 2 \\ \frac{1}{6} \cdot (2 - |x|)^3 & , \text{ se } 1 \leq |x| < 2 \\ \frac{2}{3} - \frac{1}{2}|x|^2 \cdot (2 - |x|) & , \text{ se } |x| < 1 \end{cases} \quad (11)$$

Como o suporte da função B_3 tem largura 4, em geral cada ponto x da função $f(x)$ sofre a influência de quatro funções de base $B_3(\cdot - k)$, mais exatamente para $k = i - 1, i, i + 1, i + 2$ onde $i = \lfloor x \rfloor$. Mas se x é inteiro, apenas três funções de base contribuem para o valor de $f(x)$, aquelas com $k = i - 1, i, i + 1$, onde $i = x$. Portanto, temos que para todo $k \in \mathbb{Z}$

$$\begin{aligned} f(k) &= c[k-1]B_3(k - (k-1)) + \\ &\quad c[k]B_3(k - k) + c[k+1]B_3(k - (k+1)) \\ &= c[k-1]B_3(1) + c[k]B_3(0) + c[k+1]B_3(-1) \\ &= \frac{1}{6}c[k-1] + \frac{4}{6}c[k] + \frac{1}{6}c[k+1]. \end{aligned}$$

Voltando ao problema de interpolação, o problema de encontrar a função $f \in S^3$ que interpola o sinal discreto $g[k]$ resume-se a determinar coeficientes $c[k]$ tais que

$$g[k] = \frac{1}{6}c[k-1] + \frac{4}{6}c[k] + \frac{1}{6}c[k+1] \quad (12)$$

para todo $k \in \mathbb{Z}$. Note que os coeficientes $c[k]$ podem ser interpretados como um sinal discreto de modo que a equação (12) pode ser colocada na forma de uma *convolução discreta*

$$g[k] = \left(\begin{bmatrix} \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{bmatrix} \star c \right)[k], \quad \text{onde} \quad (g \star h)[k] = \sum_{i \in \mathbb{Z}} g[i]h[k-i].$$

Considerando que o produto de convolução é associativo, obtemos:

$$c[k] = \left(\begin{bmatrix} \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{bmatrix}^{-1} \star g \right)[k],$$

e o problema se reduz a encontrar a inversa com respeito à convolução do sinal discreto $\begin{bmatrix} \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{bmatrix}$.

Existem algumas técnicas que podem ser empregadas com esse intuito, uma delas é a *Transformada Z* [23, 1], mas nesse caso em particular é fácil verificar diretamente que

$$h[k] = \frac{-6z}{1 - z^2} z^{|k|},$$

com $z = -2 + \sqrt{3}$, é a inversa de $\begin{bmatrix} \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{bmatrix}$, basta usar a definição de convolução discreta e o fato que $z^2 + 1 = -4z$.

(b) Derivação de Sinais Interpolados

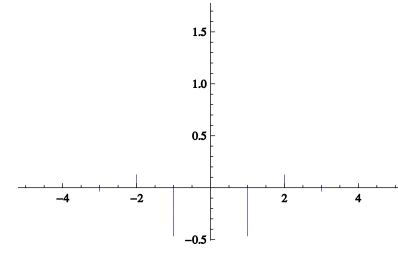


Figura 23: Gráfico do sinal $h[k]$.

Derivadas de funções splines

Agora que sabemos como avaliar a função f utilizando *B-Splines* cúbicas, podemos analisar o problema de calcular as derivadas desta função utilizando os coeficientes obtidos.

Encontrar as derivadas da função f se torna uma tarefa simples quando a caracterizamos como

$$f(x) = \sum_{k \in \mathbb{Z}} c[k]B_3(x - k),$$

pois, pela linearidade da derivação,

$$f'(x) = \sum_{k \in \mathbb{Z}} c[k]B_3'(x - k) \text{ e } f''(x) = \sum_{k \in \mathbb{Z}} c[k]B_3''(x - k).$$

As derivadas de primeira e segunda ordem da função B_3 são descritas pelas equações 13 e 14.

$$B_3'(x) = \begin{cases} 0 & , \text{ se } |x| \geq 2 \\ \frac{1}{2}x^2 + 2x + 2 & , \text{ se } -2 < x \leq -1 \\ -\frac{3}{2}x^2 - 2x & , \text{ se } -1 < x \leq 0 \\ \frac{3}{2}x^2 - 2x & , \text{ se } 0 < x \leq 1 \\ -\frac{1}{2}x^2 + 2x - 2 & , \text{ se } 1 < x < 2 \end{cases} \quad (13)$$

$$B_3''(x) = \begin{cases} 0 & , \text{ se } |x| \geq 2 \\ x + 2 & , \text{ se } -2 < x \leq -1 \\ -3x - 2 & , \text{ se } -1 < x \leq 0 \\ 3x - 2 & , \text{ se } 0 < x \leq 1 \\ -x + 2 & , \text{ se } 1 < x < 2 \end{cases} \quad (14)$$

Estas derivadas estão representadas na Figura 24.

Caso tridimensional

Os passos descritos nas seções anteriores para o cálculo de interpolação utilizando funções *B-spline* cúbicas podem ser facilmente estendidos para o caso tridimensional. Neste contexto, f passa a ser avaliada a partir de um produto tensorial.

No caso geral temos então o seguinte

$$f(x, y, z) = \sum_{i, j, k \in \mathbb{Z}} c[i, j, k]B_3(x - i)B_3(y - j)B_3(z - k).$$

(15)

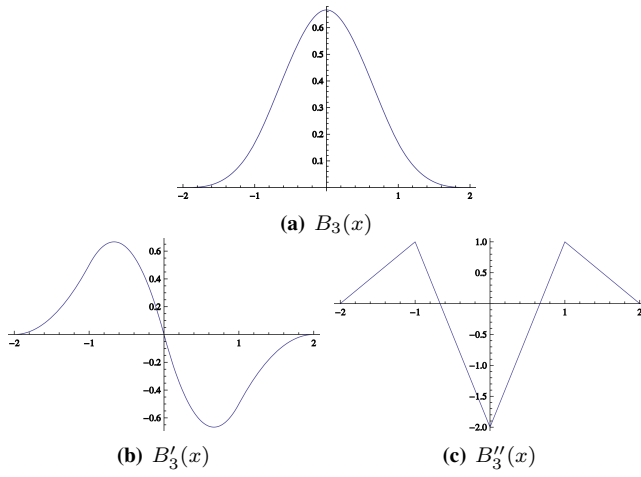


Figura 24: Derivadas da função spline cúbica.

Podemos ainda calcular as derivadas parciais de primeira e segunda ordem de f da seguinte forma

$$\begin{aligned}
 f_x(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B'_3(x-i) B_3(y-j) B_3(z-k), \\
 f_y(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B_3(x-i) B'_3(y-j) B_3(z-k), \\
 f_z(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B_3(x-i) B_3(y-j) B'_3(z-k), \\
 f_{xx}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B''_3(x-i) B_3(y-j) B_3(z-k), \\
 f_{xy}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B'_3(x-i) B'_3(y-j) B_3(z-k), \\
 f_{xz}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B'_3(x-i) B_3(y-j) B'_3(z-k), \\
 f_{yy}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B_3(x-i) B''_3(y-j) B_3(z-k), \\
 f_{yz}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B_3(x-i) B'_3(y-j) B'_3(z-k), \\
 f_{zz}(x, y, z) &= \sum_{i,j,k \in \mathbb{Z}} c[i, j, k] B_3(x-i) B_3(y-j) B''_3(z-k).
 \end{aligned}$$

Pela natureza das equações apresentadas, percebe-se facilmente que $f_{xy} = f_{yx}$, $f_{xz} = f_{zx}$ e $f_{yz} = f_{zy}$.

A Figura 25 apresenta as derivadas de primeira e segunda ordem da função f para o caso bidimensional.

Curvatura da interpolação por splines

Dadas as derivadas da interpolação por splines f vistas na seção anterior, podemos usar diretamente as fórmulas para superfícies implícitas vistas no capítulo 3 [10]. Lembramos que a normal da superfície implícita definida por f no ponto (x, y, z) é definida como $\mathbf{N} = \frac{\nabla f}{\|\nabla f\|}$, onde $\nabla f = (f_x, f_y, f_z)$. As informações sobre a curvatura estão

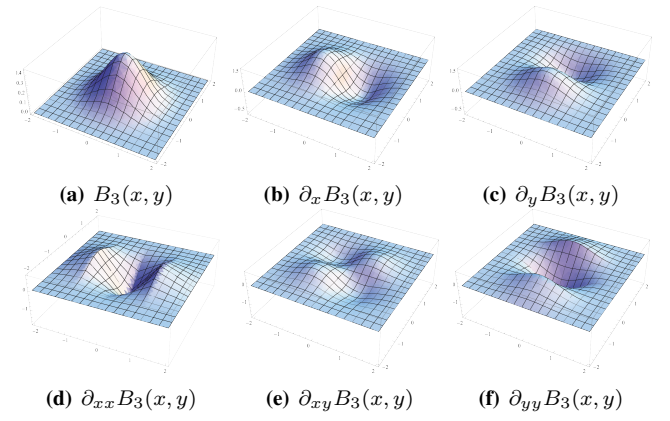


Figura 25: Funções B-spline bidimensionais e suas derivadas parciais de primeira e segunda ordem.

em $d\mathbf{N}$ que é em uma matriz 3×3 e pode ser expandida:

$$\begin{aligned}
 d\mathbf{N} &= - \left(\frac{d(\nabla f)}{\|\nabla f\|} - \frac{\nabla f \cdot \nabla(\|\nabla f\|)}{\|\nabla f\|^2} \right) \\
 &= - \frac{1}{\|\nabla f\|} \left(\mathbf{H} - \frac{\nabla f \cdot \nabla(\nabla f^T \cdot \nabla f)}{2\|\nabla f\|(\nabla f^T \cdot \nabla f)^{1/2}} \right) \\
 &= - \frac{1}{\|\nabla f\|} \left(\mathbf{H} - \frac{\nabla f \cdot (2\nabla f^T \cdot \mathbf{H})}{2\|\nabla f\|^2} \right) \\
 &= - \frac{1}{\|\nabla f\|} \left(\mathbf{I} - \frac{\nabla f \cdot \nabla f^T}{\|\nabla f\|^2} \mathbf{H} \right) \\
 &= - \frac{1}{\|\nabla f\|} (\mathbf{I} - \mathbf{N} \cdot \mathbf{N}^T) \mathbf{H}.
 \end{aligned}$$

No resultado acima, $\mathbf{I}$ é a matriz identidade e $\mathbf{H}$ a matriz Hessiana definida por

$$\mathbf{H} = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}.$$

O produto $\mathbf{N}\mathbf{N}^T$ é um operador linear que projeta vetores sobre a reta suportada por $\mathbf{N}$. O operador $\mathbf{I} - \mathbf{N}\mathbf{N}^T$ projeta sobre o complemento ortogonal desse subespaço, que vem a ser o plano tangente à isosuperfície em p . Definindo $\mathbf{P} = \mathbf{I} - \mathbf{N}\mathbf{N}^T$ temos

$$d\mathbf{N} = - \frac{1}{\|\nabla f\|} \mathbf{P} \cdot \mathbf{H}. \quad (16)$$

A equação (16) mostra que a variação do vetor normal da superfície $d\mathbf{N}$ é a projeção (por $\mathbf{P}$) da matriz Hessiana sobre o plano tangente re-escalada pelo fator escalar $1/\|\nabla f\|$. A composição com a projeção $\mathbf{P}$ significa que $d\mathbf{N}$ admite $\mathbf{N}$ como um autovetor associado ao autovalor 0. Existe então uma base ortonormal $\{v_1, v_2, \mathbf{N}\}$ de $\mathbb{R}^3$ onde v_1 e v_2 são as direções principais de curvatura associados aos autovalores k_1 e k_2 . Nessa base, $d\mathbf{N}$ é diagonal:

$$d\mathbf{N} \approx \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Como estas quantidades são invariantes por mudança de base, temos que o traço de $d\mathbf{N}$ é $\text{tr}(d\mathbf{N}) = k_1 + k_2$ e sua norma de Frobenius é $|d\mathbf{N}|_F = \sqrt{\text{tr}(d\mathbf{N}d\mathbf{N}^T)} = \sqrt{k_1^2 + k_2^2}$. Podemos achar a diferença das curvaturas principais a partir de $2|d\mathbf{N}|_F^2 - \text{tr}(d\mathbf{N})^2 = (k_1 - k_2)^2$ e deduzir k_1 e k_2 :

$$k_{1,2} = \frac{\text{tr}(d\mathbf{N}) \pm \sqrt{2|d\mathbf{N}|_F^2 - \text{tr}(d\mathbf{N})^2}}{2}.$$

Lembramos que a curvatura gaussiana é definida como $K = k_1 \cdot k_2$ e a curvatura média como $H = (k_1 + k_2)/2$.

As Figuras 26 and 27 apresentam um exemplo da aplicação da teoria que estamos estudando na visualização de curvaturas em superfícies detalhada na próxima seção.

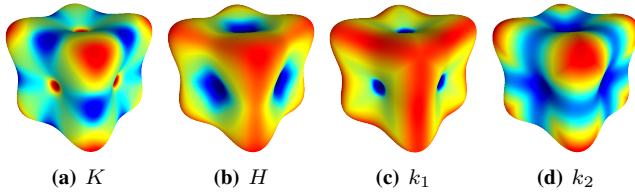


Figura 26: Exemplo de visualização baseada em curvatura. Vermelho indica curvaturas positivas e em azul aparecem as curvaturas negativas. Na imagem estão representadas a curvatura Gaussiana, a curvatura média e as curvaturas principais k_1 e k_2 .

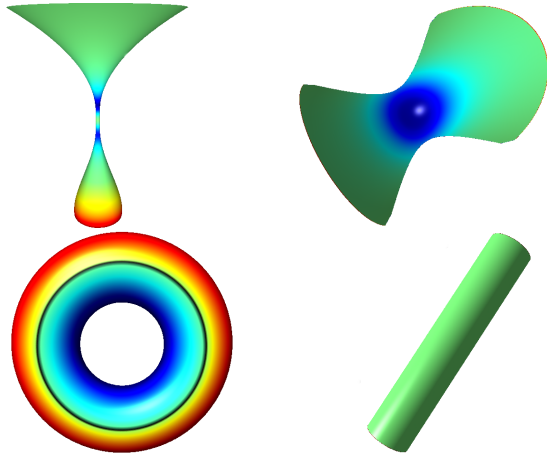


Figura 27: Exemplo de visualização baseada em curvatura em diferentes superfícies. Vermelho indica curvaturas positivas, verde curvatura nula e azul curvaturas negativas.

(c) Visualização de Superfícies Implícitas

Nesta seção, vamos descrever brevemente como as figuras de superfícies deste capítulo foram geradas. Empregamos um algoritmo simples e direto para a visualização de superfícies implícitas: o lançamento de raios (*ray casting* em inglês) [29]. Além do lançamento de raios básico para superfícies, descreveremos também o lançamento de raios para volumes.

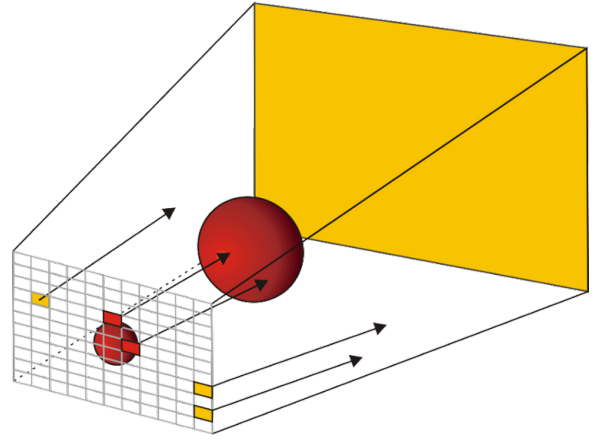


Figura 28: Lançamento de Raios.

Lançamento de Raios para Superfícies

Dada uma superfície implícita $S = f^{-1}(a)$, o primeiro passo do algoritmo consiste em considerar um *grid* regular, que representa a imagem resultante do algoritmo, onde cada elemento é chamado de *pixel*. Para cada *pixel* P_{ij} é traçado um raio $r(t)$ que parte do centro de projeção o e passa pelo ponto P_{ij} , ou seja, $r(t) = P_{ij} + t(P_{ij} - o)$. Conhecendo a equação do raio e a equação da superfície, podemos calcular os pontos de interseção entre o raio r e a superfície S fazendo $f(r(t)) = a$. Note que chegamos assim a uma equação na variável t . Se a equação não possui solução, temos que o raio não intersecta a superfície e portanto o *pixel* P_{ij} pode ser colorido com uma cor de fundo pré-determinada. Caso contrário, calculamos o menor valor positivo t_1 que satisfaz a equação acima, de modo que $p = r(t_1)$ é a primeira interseção do raio com a superfície. A Figura 28 ilustra o algoritmo de lançamento de raios.

Resta agora decidir que cor devemos aplicar ao *pixel* P_{ij} correspondente. Isso depende de dois fatores: do *modelo de iluminação*, ou seja, de como o processo de iluminação é simulado no computador para gerar um efeito de sombreamento mais ou menos realista, e de *atributos* associados à superfície S , tais como cor, textura e outras propriedades físicas que se deseje simular.

Um modelo de iluminação prático e fácil de implementar, embora não acurado fisicamente, é o modelo de *Blinn-Phong* [28]. A ideia básica é calcular um fator I_p que representa a intensidade de luz proveniente da fonte de luz presente na cena que é refletida pela superfície S no ponto p na direção do raio r .

A Figura 29 ilustra os componentes do modelo de iluminação. O primeiro fator considerado é a reflexão difusa que é provocada pela absorção e irradiação uniformemente distribuída da luz sobre o objeto iluminado. Esse efeito é modelado pela Lei de Lambert, descrita por:

$$I_p = k_d(\mathbf{N} \cdot \mathbf{L}), \quad (17)$$

onde k_d é a *constante de reflexão difusa* determinada pelo material que constitui o objeto, $\mathbf{N}$ é o vetor normal à su-

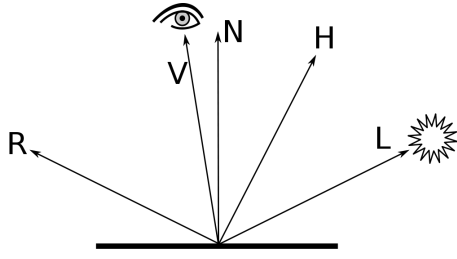


Figura 29: Modelo de Iluminação de Blinn-Phong (© wikipedia).

perfície S no ponto p e $\mathbf{L}$ é o vetor que representa a direção da fonte de luz.

Aliada à componente do parágrafo anterior, para simplificar o modelo, a contribuição da energia luminosa provocada pela reflexão da luz por outros objetos contidos na cena é considerada em um novo elemento chamado de luz ambiente. Dessa maneira, o modelo pode ser novamente descrito como

$$I_p = k_a + k_d(\mathbf{N} \cdot \mathbf{L}), \quad (18)$$

onde k_a é a constante de reflexão do ambiente.

Finalmente, podemos considerar a componente especular que simula propriedades de reflexão da superfície e portanto depende em grande medida da posição do observador relativamente ao ponto p , ou seja, depende do vetor $\mathbf{V}$ que liga o ponto p ao observador (exatamente o vetor diretor do raio r , no sentido oposto). A componente especular é responsável pelo efeito de brilho intenso em uma determinada região da superfície conhecido como *highlight*. Chegamos então à seguinte expressão:

$$I_p = k_a + k_d(\mathbf{N} \cdot \mathbf{L}) + k_s(\mathbf{N} \cdot \mathbf{H})^\alpha, \quad (19)$$

onde k_s é a constante de reflexão especular, α é o expoente de reflexão especular e $\mathbf{H}$ é o vetor que indica a direção de *highlight* máximo, definido por

$$\mathbf{H} = \frac{\mathbf{L} + \mathbf{V}}{\|\mathbf{L} + \mathbf{V}\|}.$$

Encontrada a intensidade de luz refletida no ponto p , a cor $\mathbf{C}_{ij}$ do pixel P_{ij} é apenas o produto da intensidade I_p pela cor $\mathbf{C}$ da superfície:

$$\mathbf{C}_{ij} = I_p \mathbf{C}.$$

Note que a cor de um pixel normalmente é representada por um vetor $\mathbf{C} = (r, g, b)$ que representa a porcentagem das cores primárias vermelho (r), verde (g) e azul (b).

Lançamento de Raios para Volumes

Em muitas aplicações é necessário visualizar o interior de volumes sólidos do espaço, mas o algoritmo de lançamento de raios clássico permite visualizar apenas superfícies. Veremos agora como estender o algoritmo descrito na subseção anterior para visualização de um volume descrito por uma função escalar $f : [0, 1]^3 \rightarrow [0, 1]$. A ideia é que podemos pensar na função f como uma maneira de classificar os pontos do cubo $[0, 1]^3$ atribuindo um valor de trans-

parência a cada um deles. O processo de visualização de volumes que utiliza esses valores de transparência é chamado de *Visualização Volumétrica*.

Uma maneira de interpretar a visualização volumétrica é considerar que

$$[0, 1]^3 = \cup_{a \in [0, 1]} f^{-1}(a),$$

ou seja, o cubo $[0, 1]^3$ é a união de todas as superfícies de nível da função f . Assim, podemos aplicar as técnicas de visualização de superfícies implícitas para cada um dos níveis $a \in [0, 1]$, integrando-as de alguma forma. É possível combinar os atributos contidos em cada uma dessas superfícies de maneira a visualizar partes diferentes do volume simultaneamente através de um efeito de transparência ilustrado na Figura 30.

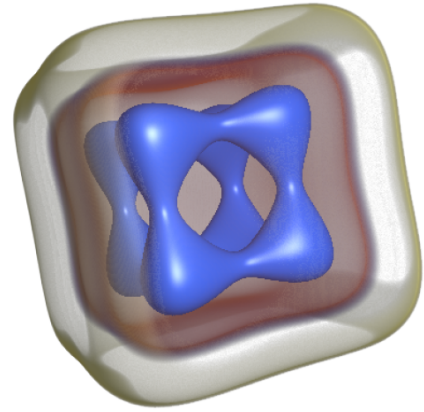


Figura 30: Visualização de volumes com efeito de transparência.

O segredo para realizar essa integração é considerar que existe um atributo τ , chamado *transparência*, associado a cada ponto $p \in [0, 1]^3$. Este atributo normalmente é considerado constante em cada superfície $S_a = f^{-1}(a)$, sendo portanto uma função de a , ou seja,

$$\tau(p) = \gamma(f(p)),$$

onde $\gamma : [0, 1] \rightarrow [0, 1]$ é denominada *função de transparência*. Assim, alterando o valor da função de transferência no ponto a , podemos tornar uma certa superfície S_a mais ou menos transparente.

Note que a transparência tem um caráter multiplicativo. Por exemplo, se um vidro tem 50% de transparência, isso significa que metade da luz que incide em um lado passa para o outro lado. Se justapomos dois destes vidros, temos que apenas metade da metade da luz passa, ou seja, $25\% = 50\% \times 50\%$. Como em certos casos é mais conveniente somar que multiplicar, vamos introduzir a grandeza ω , denominada *opacidade*, de modo que

$$\omega(p) = -\ln \tau(p).$$

Assim, somar opacidades é equivalente a multiplicar transparências, pelas propriedades do logaritmo. Note ainda que

quando a transparência vale 100% a opacidade vale 0 e quando a transparência vale 0%, a opacidade é infinita.

Durante o lançamento de raios, um mesmo raio pode intersectar diferentes superfícies de nível, com diferentes valores de transparência e cor. No algoritmo descrito anteriormente, considerávamos apenas a primeira interseção. O problema então é: como compor esses valores para gerar o resultado final?

A ideia é simplesmente integrar a contribuição de intensidade e cor ao longo do raio r , considerando a atenuação causada pela transparência acumulada, ou seja,

$$\mathbf{C}_{ij} = \int_{t_A}^{t_B} e^{-\int_{t_A}^t \omega(r(s)) ds} I_{r(t)} \mathbf{C}(r(t)) dt, \quad (20)$$

onde t_A e t_B são os parâmetros dos pontos de entrada e saída do raio no cubo $[0, 1]^3$. Note o aparecimento da função $\mathbf{C}(p)$ que dá a cor no ponto p .

Analogamente ao caso da transparência, normalmente consideramos que a cor $\mathbf{C}(p)$ é constante em cada superfície de nível, de modo que

$$\mathbf{C}(p) = (\gamma_r(f(p)), \gamma_g(f(p)), \gamma_b(f(p))),$$

onde γ_r, γ_g e γ_b são as funções de transferência dos componentes de cor, as quais podemos controlar para escolher as cores finais de cada superfície de nível.

A equação (20) pode ser justificada fisicamente [27], mas é mais fácil pensar nela intuitivamente. Simplesmente estamos integrando todas as contribuições de cor de todas as superfícies de nível ao longo do raio, ponderando cada uma delas com um fator da forma

$$\exp\left(-\int_{t_A}^t \omega(r(s)) ds\right)$$

que mede a transparência total acumulada de $r(t_A)$ até $r(t)$.

Na prática, a integral da equação (20) pode ser aproximada por uma soma de Riemann sobre uma partição finita de maneira bastante eficiente. Além disso, os computadores mais modernos possuem placas aceleradoras gráficas dedicadas (as GPU's) que permitem que os cálculos relativos a vários raios diferentes sejam efetuados simultaneamente, o que aumenta ainda mais a eficiência do processo, de modo que mesmo volumes de grandes dimensões podem ser visualizados em tempo real. Na nossa implementação, utilizamos o *framework* de visualização volumétrica VOREEN [30], ao qual adicionamos módulos para cálculos de interpolação e local e cálculo de curvaturas, descritas nas seções anteriores.

Visualização da Curvatura

Finalmente, vamos explicar como incorporar a informação de curvatura na visualização de superfícies e volumes. Vamos denotar por $k(p)$ uma das curvaturas estudadas no capítulo 3 (K, H, k_1 e k_2) no ponto p da superfície S_a , com $a = f(p)$. Queremos modificar a função de cor $\mathbf{C}(p)$ de modo que a cor dependa da curvatura $k(p)$, ao contrário das subseções anteriores onde se usava o nível a para definir a cor da superfície.

Uma pequena dificuldade inicial é que os valores de $k(p)$ não são conhecidos *a priori*, podendo assumir qualquer valor real, ao passo que os valores das componentes de cor estão limitados a valores no intervalo $[0, 1]$. A função

$$n_\lambda(x) = \begin{cases} \frac{e^{\lambda x}}{2} & , \text{ se } x < 0 \\ 1 - \frac{e^{-\lambda x}}{2} & , \text{ se } x \geq 0 \end{cases},$$

exibida na Figura 31, mapeia o intervalo $(-\infty, +\infty)$ no intervalo $(0, 1)$, onde λ é um parâmetro que controla a inclinação do gráfico. Podemos assim definir a *curvatura normalizada* $k_n(p) = n_\lambda(k(p))$ e calcular a cor $\mathbf{C}(p)$ fazendo

$$\mathbf{C}(p) = (\gamma_r(k_n(p)), \gamma_g(k_n(p)), \gamma_b(k_n(p))).$$

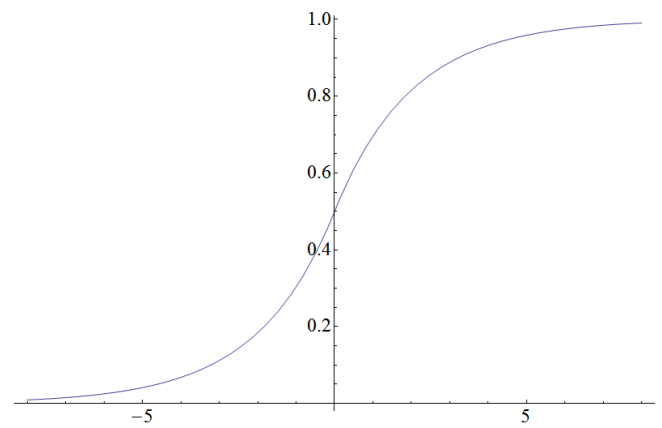


Figura 31: Gráfico da função $n_{0,5}(x)$.

(d) Problema da Invariância

Uma propriedade desejável ao estudarmos propriedades de geometria Euclidiana, como, por exemplo, vetores tangentes e normais ou curvaturas, é a invariância por movimentos rígidos (rotações e translações). Isto é importante porque podemos identificar um mesmo objeto (curvas ou superfícies) em posições distintas do espaço, pois ele terá as mesmas propriedades geométricas. De fato, a propriedade essencial para usar uma medida em física ou nas engenharias (e no dia a dia) é a invariância, e isso confere às medidas usuais (área e curvatura por exemplo) o seu significado mais útil.

Um exemplo simples para visualizar essa invariância é o caso da esfera. Façamos uma rotação e uma translação na esfera, temos que as curvaturas Gaussiana e média da superfície não se alteram. Formalmente, dizemos que uma propriedade geométrica $\mathcal{A}$ em um objeto $\mathcal{O}$ é invariante por movimentos rígidos T se, para todo $p \in \mathcal{O}$, termos $\mathcal{A}(T(p)) = \mathcal{A}(p)$.

Exercício 4.3. Mostre que as curvaturas Gaussiana e média são invariantes por movimentos rígidos. Dê um exemplo

mostrando que a curvatura Gaussiana não é invariante por escalonamento (mesmo uniforme).

Vamos supor de agora por diante que estamos trabalhando com transformações lineares. Os vetores tangentes σ_u e σ_v são naturalmente contravariantes sobre qualquer transformação linear A , ou seja:

$$\sigma_{\{u,v\}}(A(p)) = A\sigma_{\{u,v\}}(p).$$

Isto é uma consequência direta da regra da cadeia.

Relações de ortogonalidade não são preservadas sobre algumas transformações lineares A , portanto o vetor normal $\mathbf{N}$ não é invariante por movimentos rígidos. Entretanto, a direção do normal é covariante (se $\langle \mathbf{N}, \sigma_{\{u,v\}} \rangle = 0$, então $\langle A^{-T}\mathbf{N}, A\sigma_{\{u,v\}} \rangle = 0$), isto é, se aplicarmos uma transformação rígida A na superfície S teremos que o novo vetor normal é $A^{-T}\mathbf{N}$.

No caso discreto, surgem alguns problemas quando vamos obter as propriedades geométricas, como vetores tangentes e normais ou curvaturas, invariantes por movimentos rígidos. Uma das dificuldades é a boa estimativa das derivadas em cada ponto da grade e feito isso teremos que interpolar os valores das derivadas para obtermos uma boa representação dessas propriedades geométricas. Ao aplicarmos uma transformação rígida em nossa amostra de dados, o sinal $g[i, j, k]$ não é transformado por este movimento rígido pois aparecem pontos de amostragem que não são necessariamente combinações sistemáticas de amostras anteriores. Por isso, muitas vezes perdemos a invariância das propriedades geométricas. Isto torna-se algo interessante de estudar visto que a propriedade de invariância é importante em várias aplicações, como reconhecimento e reconstrução de curvas ou superfícies.

5 Representações Discretas Globais

Com as ferramentas de interpolação do capítulo anterior, podemos definir a geometria local de curvas e superfícies implícitas definidas por dados discretos. Isto é uma maneira de relacionar parte das definições diferenciais (capítulo 3) com os dados discretos. Porém, lembramos que as superfícies tem propriedades geométricas locais, como curvatura, mas devem ter uma consistência global para poder expressar, por exemplo, a área total ou o número de componentes conexas.

Este capítulo introduz representações e mecanismos de geometria discreta para expressar curvas e superfícies de forma global. De uma certa forma, esses mecanismos criam uma base para mudanças de parâmetros entre as regiões onde pode-se aplicar o teorema da função implícita e usar os cálculos do capítulo anterior. Para isto, introduzimos representações de curvas e superfícies discretas, essencialmente curvas poligonais e malhas triangulares, e mecanismos de gerar essas representações a partir de um sinal discreto.

O estudo global das superfícies requer frequentemente, até na Geometria Diferencial, abordagens menos fundamentadas no cálculo e mais nos processos construtivos. No caso discreto, as construções são expressas como algoritmos para adaptar-se ao computador. Os procedimentos são infelizmente formalizados como condição apenas suficiente (“deste jeito funciona”) mais do que de forma necessária (“existe apenas uma solução com hipóteses mínimas”). Isto é principalmente devido à juventude destas áreas que precisam ainda ser melhor formalizadas. Mas do lado didático, têm a vantagem de envolver conceitos de várias áreas da Matemática: desde a Combinatória (complexos celulares), até Topologia e Geometria.

(a) Representações Discretas

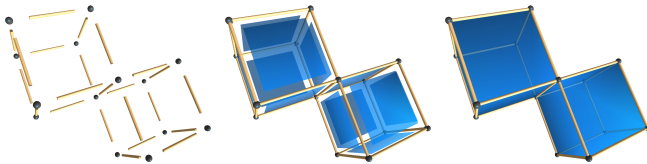


Figura 32: Um complexo celular com vértices, arestas e faces.

Complexo celular

A representação discreta usual de curvas e superfícies é similar à definição de superfície regular (seção 2.4): colando pedaços de reta ou de plano, eventualmente deformados. A diferença é que cada um dos pedaços não se recobrem, mas são postos lado a lado, colados pela fronteira deles. Cada pedaço é chamado de *célula*, tendo dimensão 0, 1 ou 2 se corresponder a um ponto, um pedaço de reta ou um pedaço de plano deformado.

Mais precisamente usando a noção de homeomorfismo vista no capítulo 3 [15]:

Definição 5.1 (Célula). Uma célula T de dimensão d é a imagem de um aberto limitado U de $\mathbb{R}^d$ por uma aplicação ϕ bijetiva, contínua, de inversa contínua e estendível ao bordo de U . O bordo de T é a imagem do bordo de U por ϕ .

A aplicação ϕ modela a deformação de cada pedaço T . A superfície discreta é um conjunto de tais células justapostas, coladas no bordo delas (ver Figura 32). O bordo comum a duas células tem que ser uma célula de dimensão menor, a saber

Definição 5.2 (Complexo celular). Um complexo celular C é um conjunto de células disjuntas tais que o bordo de uma célula seja a união de células de dimensão menores.

Complexo simplicial

Essa definição de complexo celular (finito mergulhado) permite representar a grande maioria das superfícies diferenciais de forma exata. Porém, apesar de discreta, esta representação nem sempre pode ser usada no computador, principalmente por causa da deformação ϕ e da eventual complexidade do bordo. Por isso é comum usar um caso particular de complexos celulares, mais simples, chamada de *complexos simpliciais* [5] (ver Figura 33). Neste caso, as células são simplexes, que generalizam pontos, segmentos de reta, triângulos, tetraedros, pentachoron, etc.

Definição 5.3 (Simplexo). Um simplexo T de dimensão d é o fecho convexo de $d + 1$ pontos em posição geral.

O fecho convexo é o menor conjunto convexo contendo os pontos. Por exemplo, o fecho convexo de três pontos no plano é o triângulo tendo esses pontos como vértices. Os pontos precisam ser em posição geral (não tendo três pontos alinhados, quatro pontos coplanares) para evitar que o simplexo degenera.

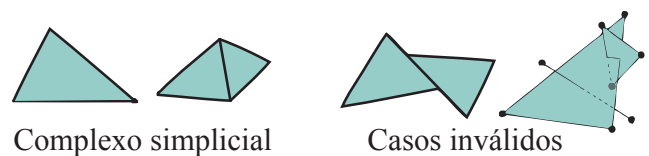


Figura 33: Um complexo simplicial e casos não válidos

Os simplexes de dimensão 0, 1 e 2 são respectivamente chamados de vértices, arestas e triângulos. Observe que as faces de dimensão $d - 1$ de um simplexo de dimensão d são simplexes de dimensão $d - 1$, pois são os fechos convexas de d dos $d + 1$ pontos.

Exercício 5.4. Verifique que um simplexo de dimensão d tem $\binom{d}{k}$ faces de dimensão k .

Propriedade de variedade local

Um *complexo simplicial* é um complexo celular onde cada célula é um simplexo. Para um complexo simplicial corresponder a uma curva (contínua), é necessário garantir que a vizinhança de cada ponto possa ser parecida com um segmento de reta. Chamamos esta propriedade de variedade local. Para isso, não pode haver simplexos de dimensão maior ou igual a 2, nem vértices isolados (não contidos no bordo de outra célula). Cada ponto de uma aresta verifica a propriedade de variedade local (as células são disjuntas). Para os vértices, eles podem ser considerados no meio de um segmento de reta se ele pertencer exatamente ao bordo de duas arestas. Neste caso, chamamos o complexo simplicial de *curva poligonal*.

Para um complexo simplicial corresponder a uma superfície (não necessariamente suave), não pode haver simplexos de dimensão maior ou igual a 3, nem vértices ou arestas isolados (não contidos no bordo de outra célula). Cada ponto de um triângulo verifica a propriedade de variedade local. Para a aresta, ela pode ser considerada no meio de um pedaço de plano se ela pertencer exatamente ao bordo de dois triângulos.

Esta propriedade será fundamental para os algoritmos de *Marching Cubes*. Um vértice verifica a propriedade de variedade local se a união das arestas e dos triângulos em volta dele é conexa. Neste caso, chamamos o complexo simplicial de *malha triangular* ou de *superfície discreta triangulada*.

Exercício 5.5. Verifique que a condição de variedade local para um vértice é válida apenas se as arestas em volta verificarem a condição de variedade local.

A característica de Euler χ vista no final do capítulo 3 pode ser expressa de forma algébrica numa superfície discreta como

$$\chi = \#_2 - \#_1 + \#_0,$$

onde $\#_2$ é o número de triângulos, $\#_1$ o número de arestas e $\#_0$ o número de vértices [9, 2].

Exercício 5.6. Verifique em exemplos que operações que não mudam a topologia, como dividir um triângulo em três inserindo o baricentro, não mudam χ . Verifique também que tem operações que não mudam o valor de χ , mas modificam a topologia, como acrescentar uma componente conexa orientável de genus 1.

(b) Funções Multilineares por Partes

A opção apresentada no capítulo anterior para construir uma representação discreta de curva ou de superfície a partir de um sinal discreto $g[i, j]$ ou $g[i, j, k]$ passa por interpolar o sinal g por uma função contínua (ou suave) f , e considerar a superfície implícita definida por f . orém, as interpolações por splines podem gerar funções implícitas muito complexas, tipicamente polinomiais por partes em várias variáveis e de grau alto. Isso dificulta a resolução

numérica da equação $f(x, y, z) = 0$, e por consequência a construção da superfície no computador. Nesta seção, veremos uma interpolação diferente, chamada de multilinear, que gera uma função f contínua (além de C^∞ por partes), e permite uma construção mais direta da superfície discreta no computador.

Interpolação linear por partes

Dado um sinal unidimensional $g : \mathbb{Z} \rightarrow \mathbb{R}$, geramos uma interpolação $f_i : [0, 1] \rightarrow \mathbb{R}$ dentro de cada intervalo entre $g[i]$ e $g[i + 1]$, de modo que $f_i(0) = g[i]$ e $f_i(1) = g[i + 1]$. Assim, a interpolação global $f : \mathbb{R} \rightarrow \mathbb{R}$ de g é deduzida por: $f(x) = f_i(x - i)$ com i igual a parte inteira de x : $i = \lfloor x \rfloor$.

Exercício 5.7. Mostre que as restrições $f_i(0) = g[i]$ e $f_i(1) = g[i + 1]$ garantem que a interpolação f seja contínua.

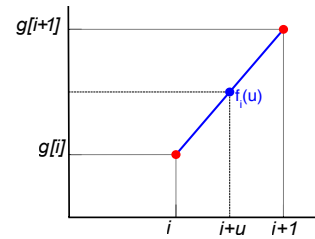


Figura 34: Interpolação linear de dois pontos do sinal.

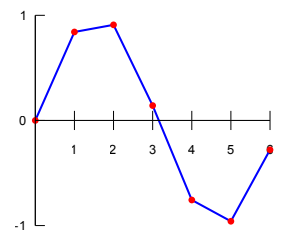


Figura 35: Interpolação linear por partes.

A interpolação f é dita *linear por partes* se cada interpolação f_i tem um segmento de reta como gráfico, sendo o único verificando as restrições acima (ver Figura 34):

$$f_i(u) = (1 - u) \cdot g[i] + u \cdot g[i + 1]. \quad (21)$$

Exercício 5.8. Mostre que a formulação acima verifica as condições $f_i(0) = g[i]$ e $f_i(1) = g[i + 1]$.

Juntado as expressões acima, obtemos a expressão completa (ver Figura 35)

Definição 5.9 (Interpolação linear por partes). Dada um sinal discreto $g : \mathbb{Z} \rightarrow \mathbb{R}$, a sua interpolação linear por partes é dada por

$$f(x) = (1 - x + \lfloor x \rfloor) \cdot g[\lfloor x \rfloor] + (x - \lfloor x \rfloor) \cdot g[\lfloor x \rfloor + 1], \quad \forall x \in \mathbb{R}.$$

Solução de $f(x) = 0$

Esta formulação da interpolação facilita muito o processo de resolução de $f(x) = 0$. Primeiro, observamos que se $g[i]$ e $g[i + 1]$ são ambos positivos, $f_i(u) = 0$ não tem solução para $u \in [0, 1]$. De fato, o gráfico de f_i é, por definição, um segmento de reta começando e terminando acima do eixo horizontal $y = 0$, portanto não a corta. Similarmente, não tem solução se $g[i]$ e $g[i + 1]$ são ambos negativos. Por outro

lado, o gráfico de f_i cortará uma vez o eixo horizontal se $g[i]$ e $g[i+1]$ tem sinais opostos, ou se um dos dois valores é nulo. O ponto u_0 de corte, solução de $f_i(u) = 0$, é obtido a partir da expressão (21):

$$u_0 = \frac{g[i]}{g[i] - g[i+1]}. \quad (22)$$

Exercício 5.10. Sabendo que $g[i]$ e $g[i+1]$ tem sinais opostos, verifique que $u_0 \in [0, 1]$.

Interpolação bilinear por partes

Quando o sinal $g[i, j]$ é bidimensional, podemos estender a interpolação acima. Cada parte será agora formada pelo quadrado entre $g[i, j]$, $g[i+1, j]$, $g[i, j+1]$ e $g[i+1, j+1]$. A interpolação local $f_{i,j} : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ tem que respeitar as restrições

$$\begin{aligned} f_{i,j}(0, 0) &= g[i, j] & , & \quad f_{i,j}(1, 0) = g[i+1, j] , \\ f_{i,j}(0, 1) &= g[i, j+1] & e & \quad f_{i,j}(1, 1) = g[i+1, j+1] . \end{aligned}$$

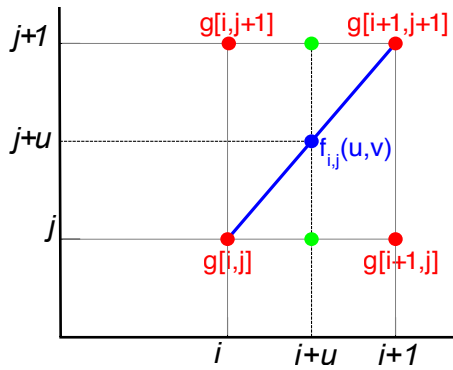


Figura 36: Interpolação bilinear de quatro pontos do sinal.

A interpolação bilinear por partes é obtida definindo $f_{i,j}$ como a única quádrlica que respeite as restrições acima (ver Figura 36)

$$\begin{aligned} f_{i,j}(u, v) &= (1-u)(1-v) \cdot g[i, j] \\ &+ u(1-v) \cdot g[i+1, j] \\ &+ (1-u)v \cdot g[i, j+1] \\ &+ uv \cdot g[i+1, j+1]. \end{aligned} \quad (23)$$

Exercício 5.11. Mostre que a formulação (23) verifica as restrições de igualdade nas extremidades do quadrado.

A propriedade essencial dessa interpolação é que $f_{i,j}$ é uma interpolação linear ao longo de qualquer reta paralela a um eixo. Mais precisamente, se fixarmos um valor real a em $[0, 1]$ para a coordenada u , $f_{i,j}$ tem a expressão da interpolação linear entre os valores $f_{i,j}(a, 0)$ e $f_{i,j}(a, 1)$ (e similarmente se fixar o valor de v). Isto permite calcular a interpolação bilinear a partir de duas interpolações lineares (ver Figura 36).

Em particular, a interpolação bilinear restrita ao longo de uma aresta do quadrado, por exemplo entre $g[i, j]$ e $g[i+1, j]$

é a interpolação linear (expressão (21)) entre $g[i, j]$ e $g[i+1, j]$. Como a aresta entre $g[i, j]$ e $g[i+1, j]$ é comum ao quadrado entre $g[i, j]$, $g[i+1, j]$, $g[i, j+1]$ e $g[i+1, j+1]$ e ao quadrado entre $g[i, j-1]$, $g[i+1, j-1]$, $g[i, j]$ e $g[i+1, j]$, o valor de $f_{i,j}$ e de $f_{i,j-1}$ é igual na aresta. Portanto, isto garante que a interpolação bilinear por partes obtida a partir dos $f_{i,j}$ é contínua passando de um quadrado ao vizinho (através das arestas).

Outra consequência importante é que a solução de $f(x, y) = 0$ ao longo das arestas do reticulado $\mathbb{Z}^2$ é dada pela expressão da interpolação linear (22).

Interpolação trilinear por partes

Quando o sinal $g[i, j, k]$ é tridimensional, podemos seguir a construção anterior. Cada parte será agora formada pelo cubo de diagonal $g[i, j, k]$, $g[i+1, j+1, k+1]$. A interpolação local no cubo $f_{i,j,k} : [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ tem que respeitar as restrições $\{f_{i,j,k}(o_u, o_v, o_w) = g[i + o_u, j + o_v, k + o_w]\}$ para $o_{uvw} \in \{0, 1\}$.

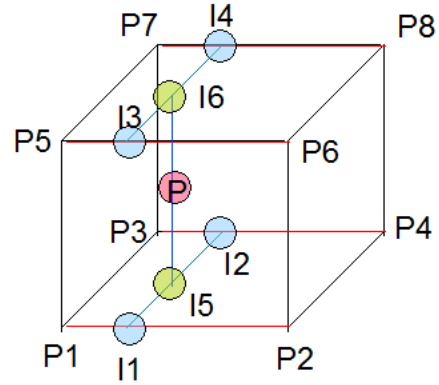


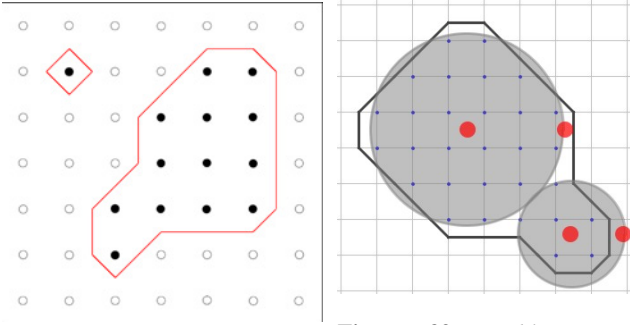
Figura 37: Interpolação trilinear dentro de um cubo pode ser feita por combinações de várias interpolações lineares ao longo dos eixos.

A interpolação trilinear por partes é obtida definindo $f_{i,j,k}$ como a única cúbica que respeite as restrições acima

$$\begin{aligned} f_{i,j,k}(u, v, w) &= (1-u)(1-v)(1-w) \cdot g[i, j, k] \\ &+ u(1-v)(1-w) \cdot g[i+1, j, k] \\ &+ (1-u)v(1-w) \cdot g[i, j+1, k] \\ &+ uv(1-w) \cdot g[i+1, j+1, k] \\ &+ (1-u)(1-v)w \cdot g[i, j, k+1] \\ &+ u(1-v)w \cdot g[i+1, j, k+1] \\ &+ (1-u)vw \cdot g[i, j+1, k+1] \\ &+ uvw \cdot g[i+1, j+1, k+1]. \end{aligned} \quad (24)$$

Essa interpolação conserva a propriedade de ser bilinear em qualquer plano paralelo aos eixos, e linear em qualquer reta paralela a um eixo, garantindo a continuidade da interpolação e a simplicidade da solução ao longo das arestas de cada cubo (ver Figura 37).

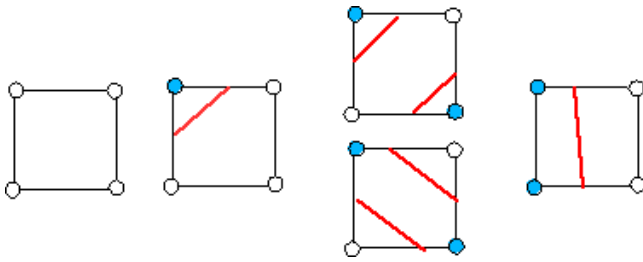
Exercício 5.12. Desenhe no computador várias curvas implícitas com valores arbitrários para $g[i + o_u, j + o_v, k + o_w]$.

(c) **Marching Squares: Isocurvas****Figura 38:** Exemplo de aplicação do Marching Squares.**Figura 39:** Problemas topológicos gerados pela má resolução das ambiguidades.

Um algoritmo bastante usado para construir representações discretas de curvas a partir de sinais discretos bidimensionais é o *Marching Squares* [14]. A base do algoritmo é de construir os vértices da curva poligonal resolvendo $f(x, y) = 0$ ao longo das arestas de cada quadrado, usando a solução simples do caso linear (expressão (22)). Em cada quadrado da interpolação bilinear por partes são criadas arestas ligando estes vértices (ver Figura 38). Na maioria dos casos, um quadrado terá nenhum ou dois vértices, e assim será criada zero ou uma aresta ligando os dois vértices (ver Figura 40).

Exercício 5.13. Supondo que $\forall (i, j) \in \mathbb{Z}^2, g[i, j] \neq 0$, mostre que em cada quadrado da interpolação bilinear por partes poderá haver apenas 0, 2 ou 4 vértices.

Exercício 5.14. Verifique que o número de soluções de $f(x, y) = 0$ pode ser infinito se deixar a condição $\forall (i, j) \in \mathbb{Z}^2, g[i, j] \neq 0$.

**Figura 40:** Casos do Marching Squares.**Ambiguidades**

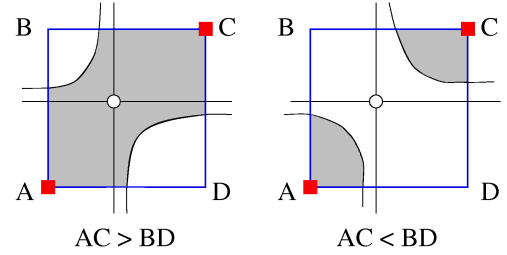
No caso bidimensional, o único caso ambíguo (se $g[i, j] \neq 0$) é quando tiver 4 vértices (soluções de $f(x, y) = 0$) no mesmo quadrado. Este caso acontece quando um par de cantos diagonalmente opostos do quadrado tem valores positivos, e o outro par tem valores negativos. A interpolação bilinear $f_{i,j}(u, v)$ parametriza assim um parabolóide hiperbólico com o ponto (u_s, v_s) correspondendo à sela pertencente ao quadrado. A curva implícita $f_{i,j}(u, v) = 0$ é

então uma hipérbole, com seu centro pertencente ao quadrado. A orientação dessa hipérbole, dada pelo determinante

$$g[i, j] \cdot g[i + 1, j + 1] - g[i + 1, j] \cdot g[i, j + 1]$$

permite escolher como ligar os quatro pontos [20].

Exercício 5.15. Verifique que o ponto (u_s, v_s) , onde o gradiente de $f_{i,j}$ é nulo, pertence ao quadrado, e que o sinal de $f_{i,j}(u_s, v_s)$ é relacionado ao determinante $g[i, j] \cdot g[i + 1, j + 1] - g[i + 1, j] \cdot g[i, j + 1]$.

**Figura 41:** Caso ambíguo do Marching Squares.

Uma maneira de formular este problema é considerar a superfície $S_- = \{(x, y) \in \mathbb{R}^2, f(x, y) < 0\}$, e olhar se S_- intersectada com o interior do quadrado é conexa ou não (ver Figura 41). Note que estas ambiguidades são importantes para garantir a coerência do resultado do *Marching Squares*, em particular em relação à topologia da curva gerada (ver Figura 39).

Propriedades e limitações

O procedimento acima permite uma implementação eficaz no computador respeitando a base da interpolação bilinear por partes. Considerando que $g[i, j] \neq 0$, pode-se garantir que o conjunto dos vértices e das arestas geradas é uma curva poligonal com a propriedade de variedade local, exceto eventualmente na fronteira do reticulado quando limitado.

Porém, desenvolver o algoritmo para incluir valores nulos do sinal é mais delicado, e pode não garantir a construção de uma variedade. A condição $g[i, j] \neq 0$ corresponde ao critério de valor regular para as curvas implícitas.

Exercício 5.16. Mostre, com um exemplo, que o critério $g[i, j] \neq 0$ não é necessário para garantir que o Marching Squares crie uma curva discreta.

Um problema mais delicado é a coerência entre as interpolações por splines, permitindo cálculos de curvaturas por derivação, e bilinear, permitindo a representação geométrica da curva. Em particular perto dos casos ambíguos, o sinal da curvatura obtido pela interpolação por splines pode não bater com a concavidade da curva poligonal gerada.

Alternativas

No caso do sinal discreto não vir num reticulado equivalente a $\mathbb{Z}^2$, mas em pedaços formados por triângulos, é possível usar da interpolação baricêntrica em vez de bilinear. Essa interpolação não tem caso ambíguo, pois tem no máximo dois vértices solução de $f = 0$ nas arestas por triângulo. É possível usar esta técnica dividindo cada quadrado em dois triângulos, mas isto resolve as ambiguidades arbitrariamente, e não a partir dos dados.

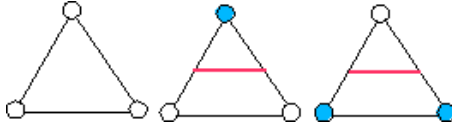


Figura 42: Casos do Marching Triangle.

Exercício 5.17. Mostre que, no caso de interpolação baricêntrica dentro de um triângulo com $g \neq 0$, pode ter apenas dois casos: nenhum vértice ao longo das arestas do triângulo, ou exatamente dois vértices.

É possível gerar no computador uma quantidade importante de pontos na curva com interpolações mais complexas (como splines por exemplo) até ter a impressão de uma curva contínua, ou até poder ligar os vértices vizinhos sem ambiguidade. Porém, é um processo bem mais custoso numericamente, e delicado de estender para mais dimensões.

(d) Marching Cubes: Isosuperfícies

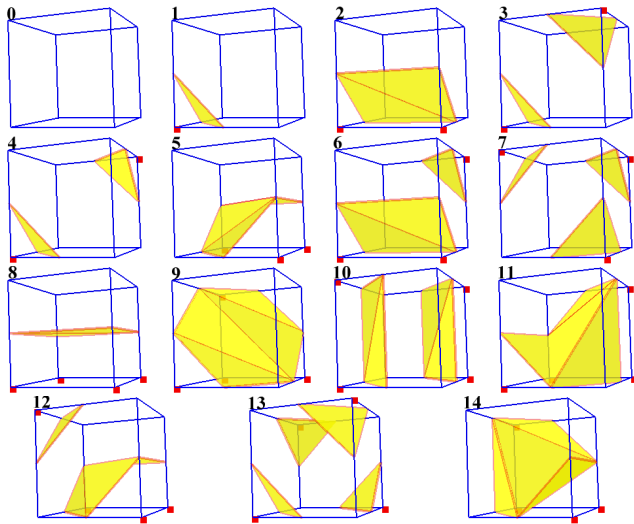


Figura 43: Casos básicos do algoritmo Marching Cubes.

O princípio do caso tridimensional é similar, porém a complexidade de criar triângulos ligando vértices nas arestas de um cubo é bem maior que no caso do quadrado. O algoritmo *Marching Cubes* [14] usa a interpolação trilinear e calcula num primeiro tempo os eventuais vértices em cada aresta do reticulado usando de novo a solução do caso linear.

Geração de triângulos

Supondo de novo que $g[i, j, k] \neq 0$, temos no máximo um vértice por aresta. Considerando um cubo por vez, temos que gerar triângulos ligando os vértices, com as restrições de gerar uma malha triangular válida. Em particular, os triângulos não podem se intersectar, e eles tem que se justapor corretamente com os triângulos vizinhos, em particular com os triângulos dos cubos vizinhos.

Existem casos simples, por exemplo, quando um canto do cubo tem valor de g positivo e todos os outros negativos. Neste caso, existem três arestas do cubo saindo do canto positivo, indo para cantos negativos, portanto apenas três arestas do cubo vão ter vértices. Com apenas três vértices, dá para criar apenas um triângulo e o caso está resolvido (ver Figura 43).

Ambiguidades

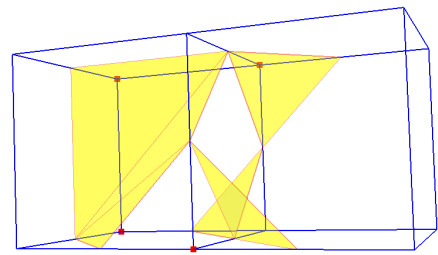


Figura 44: As faces ambíguas tem que ser resolvidas da mesma maneira nos dois cubos adjacentes, se não gera um complexo simplicial válido, como aqui usando diretamente a tabela base com os casos 12 e 3.

Porém, pode haver configurações dos sinais dos cantos do cubo de tal forma que faces do cubo seriam ambíguas para o *Marching Squares*. Neste caso, é importante tomar cuidado o algoritmo crie triângulos em cada cubo vizinho desta face de modo que as arestas batam (ver Figura 44). O mais simples é usar o mesmo teste do determinante que o *Marching Squares*. O custo deste teste é que deve-se prever mais configurações ainda, pois cada configuração de sinal deve ser subdividida por configurações de faces ambíguas [17].

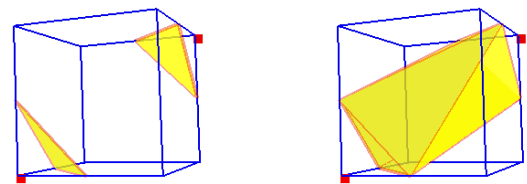


Figura 45: Ambiguidade interna a um cubo no caso 4, deixando as faces ambíguas com a mesma configuração.

Além disso, a interpolação trilinear pode gerar topologias diferentes com as mesmas restrições às faces do cubo [8, 13, 21] (ver Figura 45). Similarmente ao caso bidimensional, uma maneira de entender este problema consiste em

considerar o sólido interior à superfície $S_- = \{(x, y, z) \in \mathbb{R}^3, f(x, y, z) < 0\}$. A questão da topologia da interpolação trilinear dentro do cubo pode ser formulada da seguinte maneira: o sólido S_- intersectado com o interior do cubo é conexo ou não. Isto pode ser verificado olhando cortes horizontais do cubo, que são quadrados, e ver o teste do determinante em cada corte. O sólido S_- será conexo no cubo se e somente se este teste sempre marcar S_- como conexo no corte (ver Figura 46).

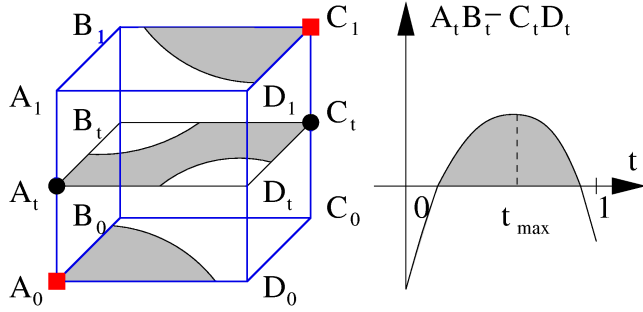


Figura 46: Resolução da ambiguidade interna com a interpolação trilinear.

Exercício 5.18 (difícil). *Demonstre esta última equivalência.* (Dica: use o fato que $f_{i,j,k}$ é linear ao longo de qualquer reta paralela a um eixo, e portanto muda de sinal no máximo uma vez.)

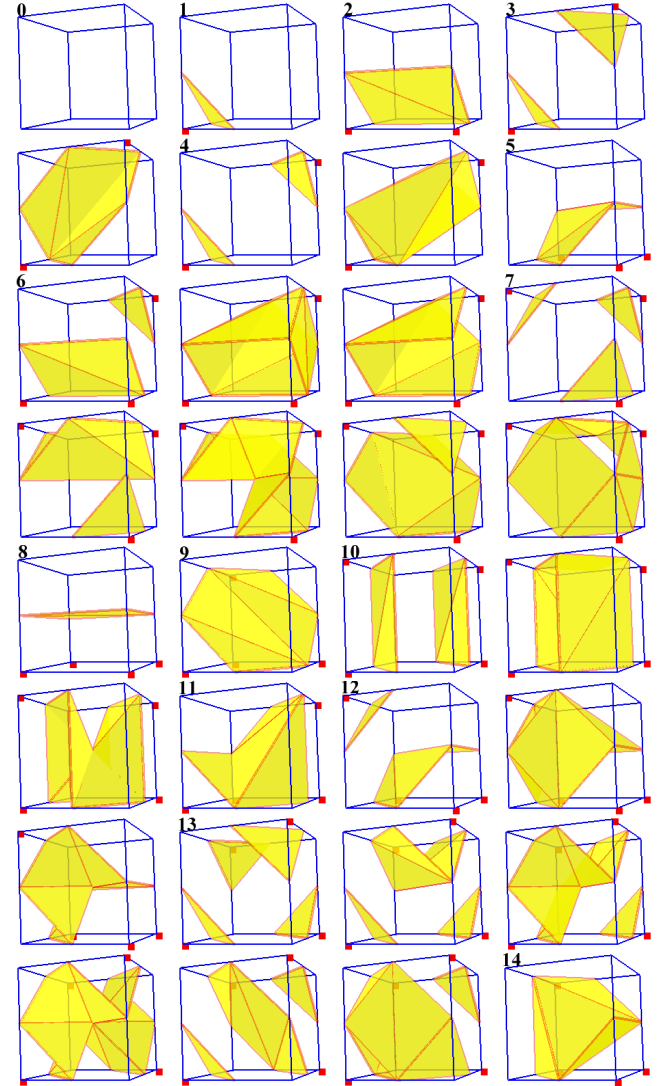


Figura 47: Casos do Marching Cubes com ambiguidades.

Alternativas

Contando que os sinais podem ser ou positivo ou negativo em cada um dos 8 cantos do cubo, temos $2^8 = 256$ configurações básicas, redutíveis à 15 casos se tirar casos equivalentes por rotação ou simetria. Porém, se incluirmos os sub-casos para acomodar as ambiguidades, isso gera 33 casos bases derivados em mais de 730 por simetria [11] (ver Figura 47). Esta complexidade tornou a busca por alternativas ao *Marching Cubes* original [19].

Similarmente ao caso bidimensional, pode-se dividir cada cubo em tetraedros e usar interpolação baricêntrica, o que gera apenas 3 casos distintos [4, 25]. Porém, a topologia da superfície gerada não depende mais apenas do sinal discreto g mas também da escolha (arbitrária) da decomposição dos cubos em tetraedros.

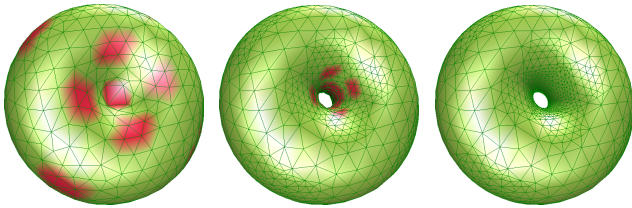


Figura 48: Reticulado adaptado para garantir a topologia do resultado: as partes vermelhas indicam que a aritmética exata não validou a região.

No caso que uma função diferenciável g tiver originada o sinal $g(g[i, j, k] = g(i \cdot \delta x, j \cdot \delta y, k \cdot \delta z))$ e que g for conhecida, pode-se também refinar o reticulado até que em cada cubo a topologia da superfície seja simples [24, 12, 22]. Em teoria, este procedimento pode gerar um reticulado infinitamente denso (por exemplo numa superfície com uma alça arbitrariamente pequena), o que não é viável no computador. Porém, com técnicas de aritmética exata, é possível isolar estes casos (ver Figura 48).

(e) Problemas de Geometria e Topologia

As abordagens detalhadas anteriormente não têm um formalismo único: estudamos interpolação por splines para o estudo local e por funções multilineares para criar globalmente uma variedade discreta. Idealmente, usar-se-ia uma única interpolação para tudo, abrindo o caminho para estudar categorias de interpolação com as suas respectivas propriedades. Esta tendência faz parte da pesquisa atual, mas está só começando!

Existem já alguns elementos simples para desenvolver teorias com este objetivo. A abordagem mais antiga é de garantir convergência das construções: se refinarmos infinitamente o reticulado e se o sinal g convergir (localmente ou uniformemente) para uma função implícita diferenciável, as curvaturas calculadas por splines convergem para as curvaturas da superfície implícita diferenciais? A topologia gerada pelo *Marching Cubes* vai corresponder à topologia da superfície suave? A resposta é *a priori* positiva, apesar de requerer por

enquanto condições de regularidade sobre g que não são necessárias no caso diferencial [16, 6, 18].

Um outro problema é a invariância da superfície gerada. Porém, o processo de amostragem ao longo do reticulado não é invariante por movimentos rígidos, pois privilegia as direções paralelas aos eixos. Isto dificulta a análise de invariantes no caso implícito discreto.

Finalmente, uma abordagem recente e promissora consiste em preservar as relações entre a topologia e a geometria. Por exemplo, o teorema de Gauss-Bonnet estipula que a integral da curvatura Gaussiana numa superfície sem bordo é igual à $2\pi\chi$, onde χ é a característica de Euler. Isto é válido nos complexos celulares definindo a curvatura Gaussiana como o déficit angular em cada vértice v : $2\pi - \sum \beta_i$ onde β_i são os ângulos em v dos triângulos tendo v na fronteira [3]. Porém, usando a estimativa da curvatura por splines e a característica de Euler dada pelo complexo simplicial resultando do *Marching Cubes* com interpolação trilinear, esta relação não vale mais. Pode assim servir de critério para construir uma teoria de interpolação mais coerente.

6 Conclusão

Vimos no decorrer do livro que o caminho de adaptar as ferramentas de geometria diferencial ao mundo discreto do computador é delicado, gerando ainda resultados aproximados, pouco formalizados e até um pouco frustrantes. Existem tentativas de formalização na linha de geometria diferencial discreta, com alguns sucessos, mas ainda é uma área de pesquisa atual. Uma das razões das frustrações da geometria discreta vem de se comparar teorias antigas, como a geometria diferencial, tendo diversas interpretações e intuições, além de formalização elegante, com tentativas recentes de reconstruir resultados similares, sem a base familiar do cálculo diferencial.

Um outro elemento que dificulta a criação de teorias discretas é o hábito do cálculo infinitesimal. Olhando o caminho desde Euclides, o cálculo infinitesimal começou apenas como ferramenta para modelar e medir formas. Esta ferramenta é claramente inadequada para o computador, onde a combinatória e a álgebra têm mais aplicações. Por exemplo, pode ser mais certo e elegante tentar definir o comprimento de uma curva diretamente a partir da geometria descritiva do que recorrer à aproximações de conceitos infinitesimais no computador.

Devemos expandir nossa intuição no uso das ferramentas computacionais, apesar de já termos confiança e hábito de usar derivadas e uma certa desconfiança dos erros numéricos no computador. É uma área que apenas começa... Esperamos poder incluir alguns dos leitores no rol dos pesquisadores que contribuirão nesta área!

Agradecimentos Os autores agradecem o apoio do CNPq da CAPES e da FAPERJ.

Referências

- [1] M. Unser, A. Aldroubi and M. Eden. B-spline signal processing. ii. efficiency design and applications. *IEEE Transactions on Signal Processing*, 41:834–848, 1993.
- [2] M. A. Armstrong. *Basic topology*. McGraw-Hill, London, 1979.
- [3] T. Banchoff. Critical points and curvature for embedded polyhedra. *Differential Geometry*, 1:257–268, 1967.
- [4] J. Bloomenthal. An implicit surface polygonizer. In P. Heckbert, editor, *Graphics Gems IV*, pages 324–349. Academic Press, Boston, 1994.
- [5] J.-D. Boissonnat and M. Yvinec. *Algorithmic Geometry*. Cambridge University Press, 1998.
- [6] V. Borrelli, F. Cazals and J.-M. M.-M. Morvan. On the angular defect of triangulations and the pointwise approximation of curvatures. *Computer Aided Geometric Design*, 20(6):319 – 341, 2003.
- [7] M. do Carmo. *Geometria Diferencial de Curvas e Superfícies*. IMPA, Rio de Janeiro, 2005.
- [8] E. V. Chernyaev. Marching cubes 33: Construction of topologically correct isosurfaces. Technical Report CERN CN 95-17, CERN, 1995.
- [9] F. H. Croom. *Basic concepts of algebraic topology*. Springer-Verlag, New York, 1978.
- [10] R. Goldman. Curvature formulas for implicit curves and surfaces. *Computer Aided Geometric Design*, 22(7):632–658, 2005.
- [11] T. Lewiner, H. L. Lopes, A. W. Vieira and G. Tavares. Efficient implementation of Marching Cubes’ cases with topological guarantees. *Journal of Graphics Tools*, 8(2):1–15, 2003.
- [12] H. Lopes, J. B. Oliveira and L. H. de Figueiredo. Robust adaptive polygonal approximation of implicit curves. *Computers & Graphics*, 26(6):841–852, 2002.
- [13] A. Lopes and K. Brodlie. Improving the robustness and accuracy of the marching cubes algorithms for isosurfacing. *IEEE Transactions on Visualization and Computer Graphics*, 9:16–29, 2003.
- [14] W. E. Lorensen and H. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Proceedings of the 14th ACM Siggraph annual conference on Computer graphics and interactive techniques*, volume 21, pages 163–169, 1987.
- [15] A. T. Lundell and S. Weingram. *The topology of CW complexes*. Van Nostrand Reinhold, New York, 1969.
- [16] M. Meyer, M. Desbrun, P. Schröder and A. Barr. Discrete differential–geometry operators for triangulated 2–manifolds. In H.-C. Hege and K. Polthier, editors, *Mathematical Visualization III*. Springer, Berlin, 2002.
- [17] C. Montani, R. Scateni and R. Scopigno. A modified lookup table for implicit disambiguation of Marching Cubes. *The Visual Computer*, 10(6):353–355, Dec. 1994.
- [18] J.-M. Morvan and B. Thibert. Smooth surface and triangular mesh : comparison of the area, the normals and the unfolding. In C. Hoffman and W. Bronsvort, editors, *Solid Modeling and Applications*, pages 147–158. ACM, 2002.
- [19] T. S. Newman and H. Yi. A survey of the marching cubes algorithm. *Computers & Graphics*, 30(5):854–879, Oct. 2006.
- [20] G. M. Nielson and B. Hamann. The asymptotic decider: resolving the ambiguity in marching cubes. In *Visualization*, pages 83–91. IEEE, 1991.
- [21] G. M. Nielson. On Marching Cubes. *Transactions on Visualization and Computer Graphics*, 9(3):283–297, 2003.
- [22] A. Paiva, H. Lopes, T. Lewiner and L. H. de Figueiredo. Robust adaptive meshes for implicit surfaces. In *Sibgrapi 2006 (XIX Brazilian Symposium on Computer Graphics and Image Processing)*, pages 205–212, Manaus, AM, 2006. IEEE.
- [23] M. Unser, A. Aldroubi, M. Eden and L. Fellow. B-spline signal processing: Part i-theory. *IEEE Transactions on Signal Processing*, 41:821–833, 1993.
- [24] A. van Gelder and J. Wilhelms. Topological considerations in isosurface generation. *ACM Transactions on Graphics*, 13(4):337–375, 1994.
- [25] L. Velho. Simple and efficient polygonization of implicit surfaces. *Journal of Graphics Tools*, 1(2):5–25, 1996.
- [26] L. Velho, L. H. de Figueiredo and J. Gomes. *Implicit Objects in Computer Graphics*. Springer-Verlag, 1998.
- [27] A. C. Paiva, R. B. Seixas and M. Gattas. *Introdução à Visualização Volumétrica*. Pontifícia Universidade Católica do Rio de Janeiro, 1999.
- [28] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *Siggraph*, volume 11, pages 192–198, 1977.
- [29] A. S. Glassner. *An Introduction to Ray Tracing*. Morgan Kaufmann, 2002.
- [30] J. Meyer-Spradow, T. Ropinski, J. Mensmann and K. Hinrichs. Voreen: A rapid-prototyping environment for ray-casting-based volume visualizations. *IEEE Computer Graphics and Applications*, 29(6):6–13, 2009.