CIV2802 – Sistemas Gráficos para Engenharia

2019.1

Geometria Computacional: Principais Algoritmos e Predicados

Luiz Fernando Martha

André Pereira

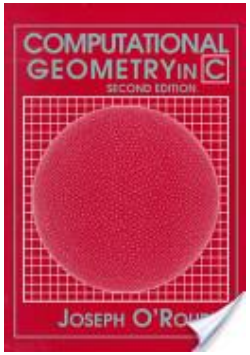
Universidade Federal Fluminense

Conteúdo

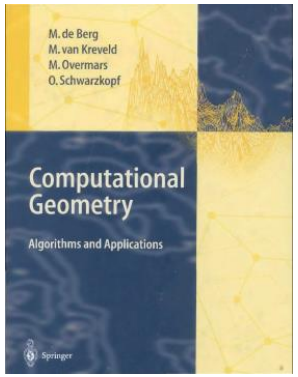
- Referência e fontes
- Introdução e escopo
- Necessidade de estruturas de dados
- Definições e notações gerais
- Área orientada de polígono
- Tesselagem de polígonos
- Predicados geométricos
- Ponto mais próximo em um segmento de reta
- Interseção de segmentos de reta
- Verificação de inclusão de ponto em polígono
- Aritmética exata e adaptativa

Referências e Fontes

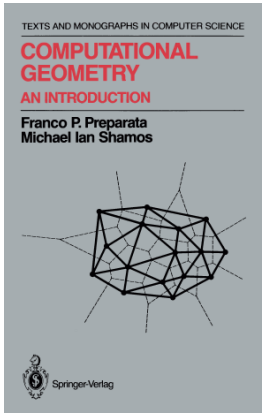
[OROURKE98]
Joseph O'Rourke
Computational Geometry in C
Cambridge University Press, 1998



[BERG97]
M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf
Computational Geometry - Algorithms and Applications
Springer, 1997

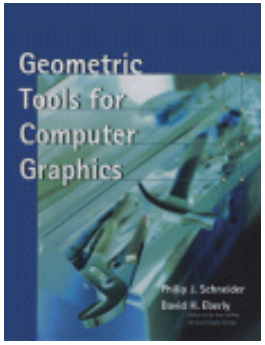


[PREPARATA85]
Franco P. Preparata, Michael Ian Shamos
Computational Geometry – An Introduction
Springer-Verlag, 1985

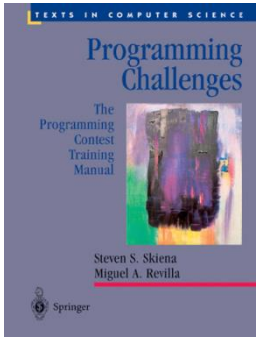


Referências e Fontes

[SCHNEIDER03]
Philip Schneider and David Eberly
Geometric Tools for Computer Graphics
Elsevier, 2003



[SKIENA02]
Steven S. Skiena, Miguel A. Revilla
Programming Challenges
Springer, 2002



[GHALI08]
Sherif Ghali
Introduction to Geometric Computing
Springer, 2008

[VINCE05]
John Vince
Geometry for Computer Graphics
Springer, 2005

Introdução e Escopo

Fonte: [OROURKE98]

Geometria Computacional é o estudo de algoritmos para resolver problemas geométricos em um computador. A ênfase nesse curso é no projeto de tais algoritmos, com de alguma forma menos atenção tomada na análise de desempenho. [OROURKE98]

Existem diversas áreas dentro de Geometria, e a que tem se tornado conhecida como Geometria Computacional, segundo [OROURKE98], é primariamente geometria discreta e combinatória. Portanto, polígonos tem um papel mais importante do que regiões com fronteiras curvas. A maioria dos trabalhos com curvas e superfícies contínuas são mais escopo da Modelagem Geométrica ou Modelagem de Sólidos, um campo com suas próprias conferências e textos, distintos da geometria computacional. Obviamente que existe substancial sobreposição, e não existe razão fundamental para que os campos sejam particionados dessa forma, eles parecem se fundir em alguma extensão. [OROURKE98]

Introdução e Escopo

Fonte: [BERG97]

A Geometria Computacional emergiu do campo de projeto e análise de algoritmos no final da década de 1970. Ela tem crescido como uma disciplina reconhecida com suas próprias revistas, conferências e uma grande comunidade de pesquisadores ativos. O sucesso do campo como uma disciplina de pesquisa pode por um lado ser explicado pela beleza dos problemas estudados e das soluções obtidas, e por outro lado, pelo domínio de diversas aplicações – computação gráfica, sistemas de informações geográficas (GIS), robótica, e outros – em que algoritmos geométricos tem um papel fundamental. [BERG97]

Para diversos problemas geométricos as primeiras soluções algorítmicas foram ou lentas ou difíceis de serem entendidas ou implementadas. Mais recentemente, uma grande quantidade de novas técnicas algorítmicas tem sido desenvolvidas que melhoram ou simplificam diversas das primeiras abordagens. A ideia desse curso é tentar fazer com que essas soluções algorítmicas modernas sejam acessíveis para uma grande audiência. [BERG97]

Introdução e Escopo

Fonte: [PREPARATA85]

Um grande número de áreas de aplicações tem sido a incubation bed da disciplina reconhecida atualmente como Geometria Computacional, já que tais aplicações fornecem inerentemente problemas geométricos que requerem o desenvolvimento de algoritmos eficientes. Estudos algorítmicos desses e outros problemas tem aparecido no século passado na literatura científica, com uma intensidade aumentando nas últimas três décadas. Porém, apenas recentemente, estudos sistemáticos de algoritmos geométricos tem sido assumidos, e um crescente número de pesquisadores tem sido atraído para essa disciplina, batizada como Geometria Computacional em um artigo de M. I. Shamos em 1975. Uma característica fundamental dessa disciplina é a percepção de que caracterizações clássicas de objetos geométricos são frequentemente não amenas para o projeto de algoritmos eficientes. Para tornar obvio essa inadequação, é necessário identificar os conceitos úteis e estabelecer suas propriedades, que irão conduzir para computações eficientes. De certa forma, a geometria computacional deve remodelar – sempre que necessário – a disciplina clássica em sua encarnação computacional.

Introdução e Escopo

Fonte: [SCHNEIDER03]

O campo da Geometria Computacional é muito vasto e é um dos que avançam mais rapidamente nos últimos tempos. Os tópicos gerais abordados neste curso são feixes convexos de conjuntos finitos de pontos, testes de ponto em polígono, tesselagem de polígonos (particionamento de polígonos em pedaços convexos ou triângulos) e interseções de segmentos de retas. [SCHNEIDER03]

A ênfase será nos algoritmos para implementar as várias ideias. Porém, é dada atenção especial para o problema de computação quando realizada por um sistema com números em formato de ponto flutuantes. Esses problemas ocorrem sempre que se precisa determinar quando pontos são colineares, coplanares e cocirculares. Isso é fácil de fazer quando o sistema computacional é baseado em aritmética de inteiros, mas bem problemático quando aritmética de pontos flutuantes é usada. [SCHNEIDER03]

Introdução e Escopo

Fonte: [SKIENA02]

Computação geométrica vai se tornando cada vez mais importante em aplicações como computação gráfica, robótica e projeto auxiliado por computador, isso tudo porque forma é uma propriedade inerente de objetos reais. Porém, objetos do mundo real não são feitos de linhas que vão até o infinito. Ao invés, a maioria dos programas de computador representam geometria como arranjos de segmentos de retas. Curvas ou formas arbitrarias fechadas podem ser representadas por coleções ordenadas de segmentos de retas ou polígonos.

Geometria computacional pode ser definida (para o propósito deste curso) como a geometria de segmentos de retas discretos e polígonos. É um assunto interessante e prazeroso, porém não tipicamente ensinado em cursos que requerem tal conhecimento. Isso dá ao estudante ambicioso que aprende um pouco de geometria computacional um estimulo e uma janela em uma área fascinante de algoritmos ainda hoje com pesquisas ativas. Livros excelentes sobre geometria computacional estão disponíveis na literatura, porém essa seção do curso deve ser suficiente para iniciar o estudande nesse assunto.

Necessidade de Estruturas de Dados

Algoritmos geométricos envolvem a manipulação de objetos que não são manipulados no nível da linguagem de máquina. O usuário tem que portanto organizar esses objetos complexos por meio de tipos de dados mais simples diretamente representáveis pelo computador. Essas organizações são universalmente referidas como **Estruturas de Dados**.

Os objetos complexos mais comuns encontrados no projeto de algoritmos geométricos são os conjuntos e as sequências (conjuntos ordenados). Estruturas de dados particularmente apropriadas para esses objetos combinatórios complexos são bem descritas na literatura padrão sobre algoritmos. Restringe-se aqui a revisar a classificação dessas estruturas de dados, junto com suas capacidades funcionais e desempenho computacional.

Seja S um conjunto representado em uma estrutura de dados e seja u um elemento arbitrário de um conjunto universal em que S é um subconjunto. As operações fundamentais que ocorrem na manipulação de conjuntos são:

1. **MEMBER**(u,S). Tem-se que $u \in S$? (Resposta SIM/NÃO.)
2. **INSERT**(u,S). Adiciona u em S.
3. **DELETE**(u,S). Remove u de S.

Fonte: [PREPARATA85]

Suponha agora que $\{S_1, S_2, \ldots, S_k\}$ é uma coleção de conjuntos (com interseção vazia por pares). Operações úteis nessa coleção são:

4. **FIND**(u). Reporta j, se u $\in S$.
5. **UNION**(S_i,S_j; S_k). Forma a união de S_i e S_j e chama esse novo conjunto de S_k.

Quando o conjunto universal é totalmente ordenado, as seguintes operações são muito importantes:

6. **MIN**(S). Reporta o elemento mínimo de S.
7. **SPLIT**(u,S). Particiona S em $\{S_1,S_2\}$, tal que $S_1 = \{v: v \in S \text{ e } v \leq u\}$ e $S_2 = S - S_1$.
8. **CONCATENATE**(S_1,S_2). Assumindo que, para arbitrário $u' \in S_1$ e $u'' \in S_2$ tem-se que $u' \leq u''$, forma o conjunto ordenado $S = S_1 \cup S_2$.

As Estruturas de Dados podem ser classificadas com base nas operações que elas supotam (sem se preocupar com eficiência). Assim, para conjuntos ordenados tem-se a seguinte tabela:

Data Structure	Supported Operations
Dictionary	MEMBER, INSERT, DELETE
Priority queue	MIN, INSERT, DELETE
Concatenable queue	INSERT, DELETE, SPLIT, CONCATENATE

Para eficiência, cada uma dessas estruturas de dados é normalmente realizada como uma árvore de busca binária balanceada pela altura. Com essa realização, cada uma das operações acima é executada em tempo proporcional ao logaritmo do número de elementos armazenados nessa estrutura de dados; o armazenamento é proporcional ao tamanho do conjunto.

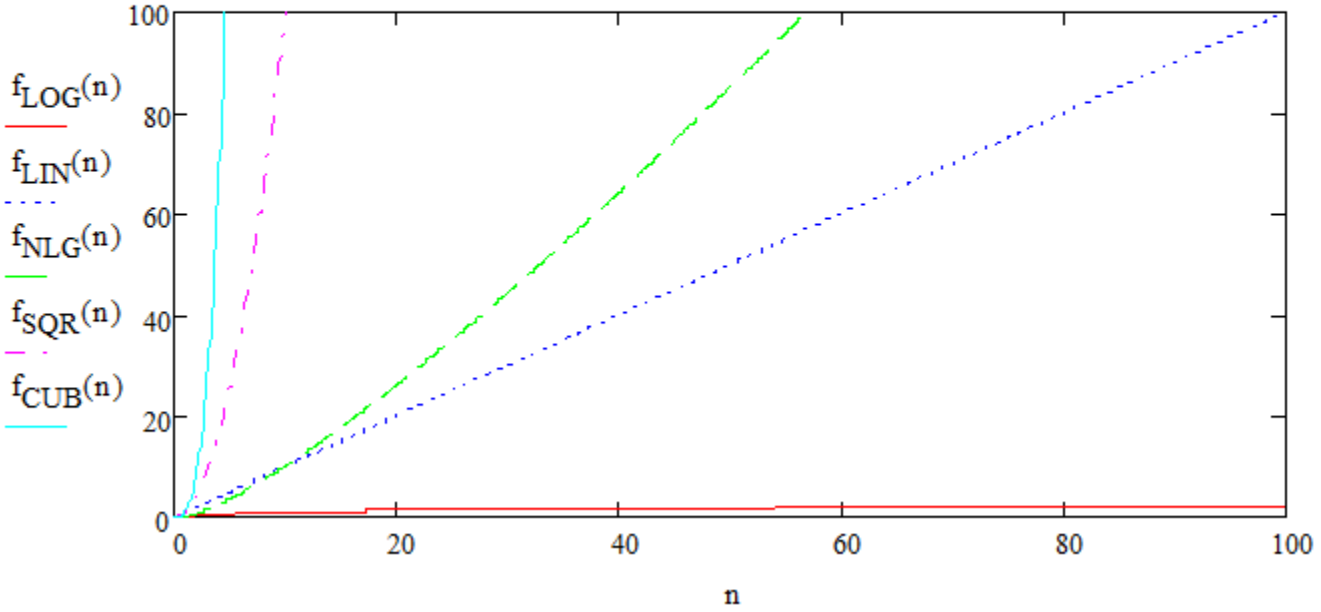
$f_{LOG}(n) := \log(n)$

$f_{LIN}(n) := n$

$f_{NLG}(n) := n \cdot \log(n)$

$f_{SQR}(n) := n^2$

$f_{CUB}(n) := n^3$



As estruturas de dados acima podem ser vistas abstratamente como uma coleção linear de elementos (uma **lista**), tal que inserções e remoções podem ser executadas uma posição arbitrária da coleção. Em alguns casos, alguns modos mais restritivos de acesso são adequados para algumas aplicações, com as subsequentes simplificações. Tais estruturas são: **Filas**, onde inserções ocorrem em um extremo e remoções no outro. **Pilhas**, onde ambas inserções e remoções ocorrem em um extremo (o topo da pilha). Claramente, um e dois ponteiros são tudo o que é necessário para gerenciar uma pilha ou uma fila, respectivamente.

As estruturas de dados padrões revisadas acima são usadas extensivamente em conjunto com os algoritmos da Geometria Computacional. Porém, a natureza dos problemas geométricos tem levado ao desenvolvimento de estruturas de dados não convencionais específicas.

Definições e Notações Gerais

Os objetos considerados normalmente em Geometria Computacional são normalmente conjuntos de pontos no espaço Euclidiano. Um sistema de coordenadas de referência é assumido, tal que cada ponto é representado por um vetor de coordenadas cartesianas da dimensão apropriada. Os objetos geométricos não consistem necessariamente de conjuntos finitos de pontos, mas tem que obedecer a convenção de ser finitamente especificado (tipicamente, como strings finitas de parâmetros). Então, considera-se além de pontos individuais, a linha reta contendo dois pontos dados, o segmento de linha reta definido pelos seus dois pontos extremos, o plano contendo três pontos dados, o polígono definido por uma (ordenada) sequência ou pontos, etc.

Essa seção não tem pretenção de fornecer definições formais dos conceitos geométricos usados neste curso; ela tem apenas os objetivos de revisar notações que são certamente conhecidas pelo leitor e de introduzir a notação adotada.

Por E^d denota-se o espaço euclidiano *d*-dimensional, isto é, o espaço das *d*-tuplas $(x_1,...,x_d)$ de números reais x_i, $i = 1,..., d$ com métrica $(\sum^d_{i=1} x_i^2)^{1/2}$.

Revisa-se agora a definição dos principais objetos considerados pela Geometria Computacional.

Fonte: [PREPARATA85]

Ponto. Uma d-tupla $(x_1,...,x_d)$ representa um ponto p de E^d; esse ponto também pode ser interpretado como um vetor com d componentes aplicado à origem de E^d, cuja extremidade livre é o ponto p.

Linha, **plano, variedade linear**. Dados dois pontos distintos q_1 e q_2 in E^d, a combinação linear

$$\alpha\, q_1 + (1 - \alpha)\, q_2 \qquad (\alpha \in \mathrm{R})$$

é uma **linha** de E^d. Mais geralmente, dados k pontos linearmente independentes $q_1, \ldots, q_k$ em E^d $(k \leq d)$, a combinação linear

$$\alpha_1\, q_1 + \alpha_2\, q_2 + \ldots + \alpha_{k-1} q_{k-1} + (1 - \alpha_1 - \ldots - \alpha_{k-1})\, q_k$$
$$(\alpha_j \in \mathrm{R}\,,\ j = 1, \ldots, k-1)$$

é um **variedade linear** de dimensão $(k - 1)$ in E^d.

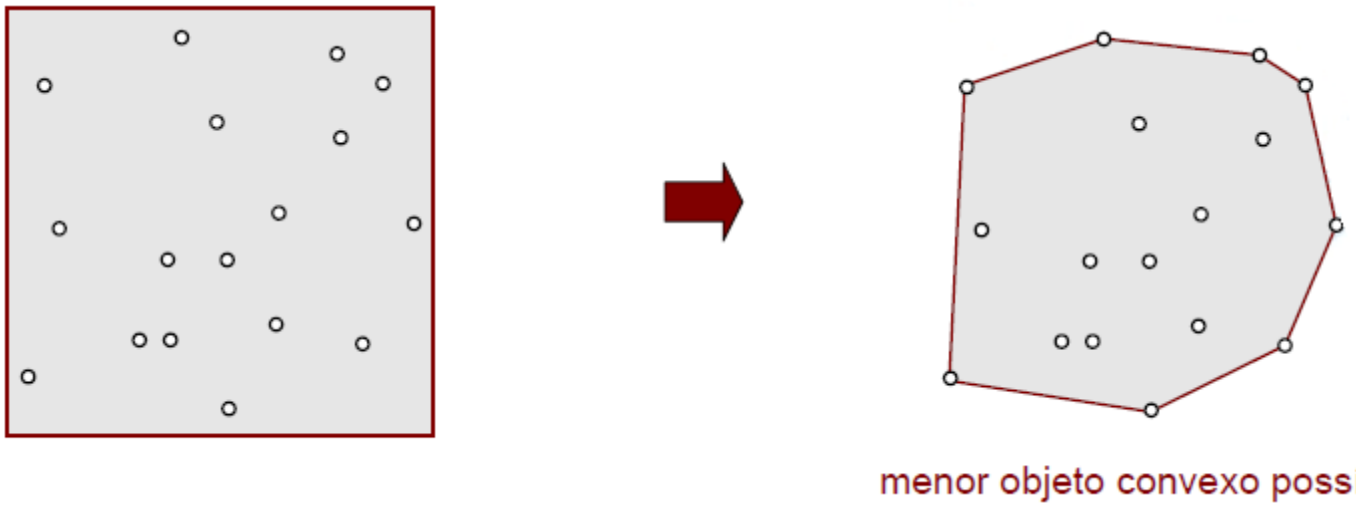
Segmento de reta. Dados dois pontos distintos q_1 e q_2 em E^d, se na expressão $\alpha q_1 + (1 - \alpha) q_2$ somarmos a condição $0 \leq \alpha \leq 1$, obtemos a combinação convexa de q_1 e q_2, ou seja,

$$\alpha\, q_1 + (1 - \alpha)\, q_2 \qquad (\alpha \in \mathrm{R}\,,\ 0 \leq \alpha \leq 1).$$

Esta combinação convexa descreve o segmento de reta que une os dois pontos q_1 e q_2. Normalmente esse segmento é denotado como $q_1 q_2$ (par não ordenado).

Conjunto convexo. Um domínio D em E^d é *convexo* se para qualquer dois pontos q_1 e q_2 em D, o segmento q_1q_2 está inteiramente contido em D. Pode ser mostrado que a interseção de domínios convexos é um domínio convexo.

Fecho convexo. O *fecho convexo* de um conjunto de pontos S em E^d é o limite do menor domínio convexo em E^d contendo S.



Polígono. Em E^2 um *polígono* é definido por um conjunto finito de segmentos de modo que cada segmento extremo é partilhado por exatamente duas arestas e nenhum subconjunto de arestas tem a mesma propriedade.

Os segmentos são as **arestas** e os seus extremos são os **vértices** do polígono. (Note que o número de vértices e arestas são idênticos.) Um polígono de n vértice é chamado um *n-gon*.

Um polígono é **simples** se não há nenhum par de arestas não consecutivas que compartilham um ponto. Um polígono simples particiona o plano em duas regiões disjuntas, o **interior** (limitado) e o **exterior** (ilimitado) que são separados por um polígono (teorema da curva de Jordan) . (Nota: na linguagem comum, o termo polígono é frequentemente usado para designar a união do contorno e do interior.)

Um polígono simples P é **convexo** se o seu interior é um conjunto convexo.

Um polígono simples P é em **forma de estrela**, se existe um ponto z não externo a P tal que, para todos os pontos p de P o segmento de linha zp fica totalmente dentro de P. (Assim, cada polígono convexo também é em forma de estrela.) O locus dos pontos z tendo a propriedade acima é o kernel do P. (Assim, um polígono convexo coincide com seu próprio kernel).

Grafo planar. Um grafo $G = (V, E)$ (conjunto de vértices V e de arestas E) é *planar* se ele pode ser incorporado no plano sem cruzamentos. Uma linha reta planar incorporada de um grafo planar determina uma partição do plano chamada *subdivisão planar* ou *mapa*. Deixe v, e, e f denotar, respectivamente, os números de vértices, arestas e regiões (incluindo a única região sem contorno) da subdivisão. Estes três parâmetros estão relacionados pela clássica *fórmula de Euler*

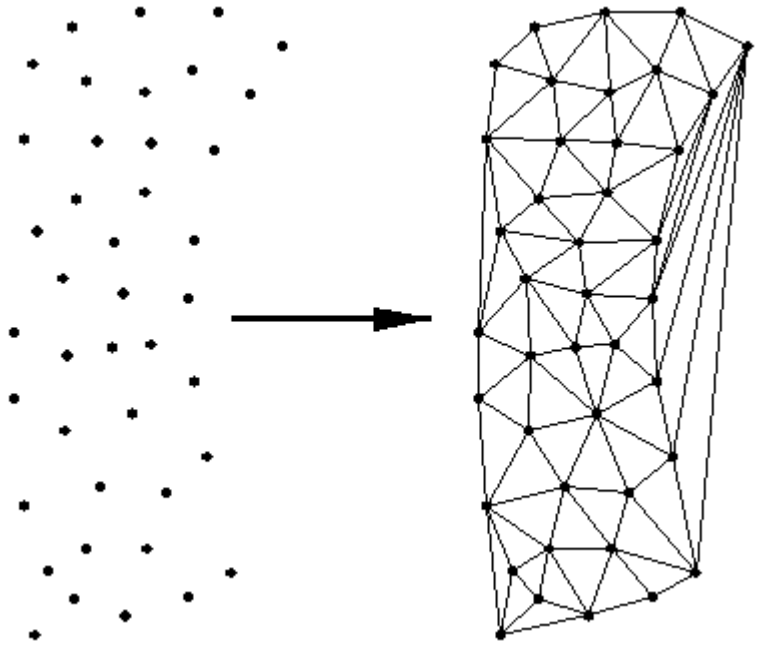
$$v - e + f = 2.$$

Se tivermos a propriedade adicional que cada vértice tem grau ≥ 3, então é um exercício simples provar as seguintes desigualdades

$$v \leq 2/3\, e, \qquad e \leq 3v - 6$$
$$e \leq 3f - 6, \qquad f \leq 2/3\, e$$
$$v \leq 2f - 4, \qquad f \leq 2v - 4$$

que mostram que v, e e f são proporcionais aos pares. (Observe que as três desigualdades mais à direita são incondicionalmente válidas.)

Triangulação. Uma subdivisão planar é uma *triangulação* se todas as suas regiões delimitadas forem triângulos. A *triangulação de um conjunto finito S* de pontos é um grafo planar em *S* com o número máximo de arestas (isto é equivalente a dizer que a triangulação de *S* é obtida juntando-se os pontos de *S* por segmentos de linhas retas que não se intersectam de modo que cada região interna ao fecho convexo se S é um triângulo).

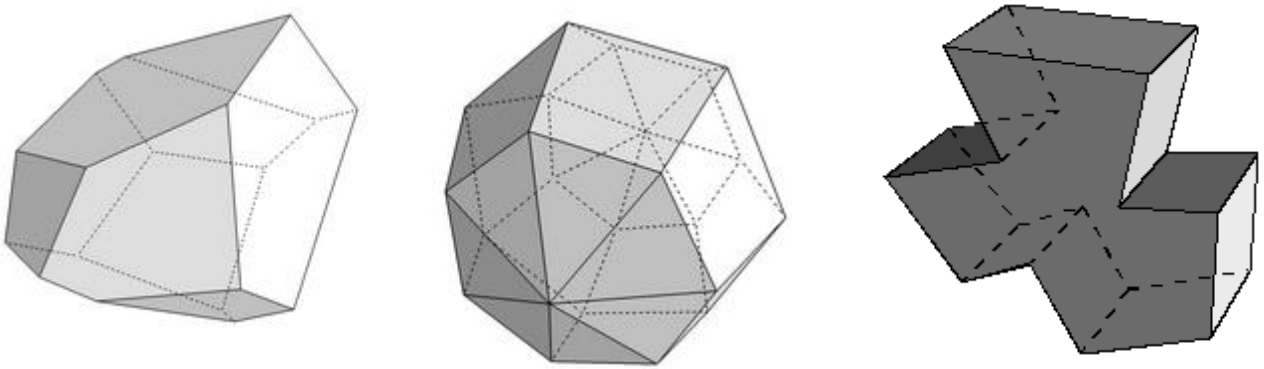


Poliedro. No E^3 um poliedro é definida por um conjunto finito de polígonos plano de tal modo que cada aresta de um polígono é partilhada por exatamente um outro polígono (polígonos adjacentes) e nenhum subconjunto de polígonos tem a mesma propriedade. Os vértices e as arestas dos polígonos são os vértices e as arestas do poliedro; os polígonos são as *faces* do poliedro

Um poliedro é *simples*, se não houver par de facetas não adjacentes que compartilham um ponto. Um poliedro simples particiona o espaço em dois domínios disjuntos, o *interior* (contornado) e no *exterior* (infinito). (Mais uma vez, na linguagem comum, o termo poliedro é frequentemente usado para designar a união do contorno e do interior.)

A superfície de um poliedro (de gênero zero) é isomorfo a uma subdivisão planar. Assim, o número *v*, *e*, e *f* de seus vértices, arestas e faces obedecem a fórmula de Euler.

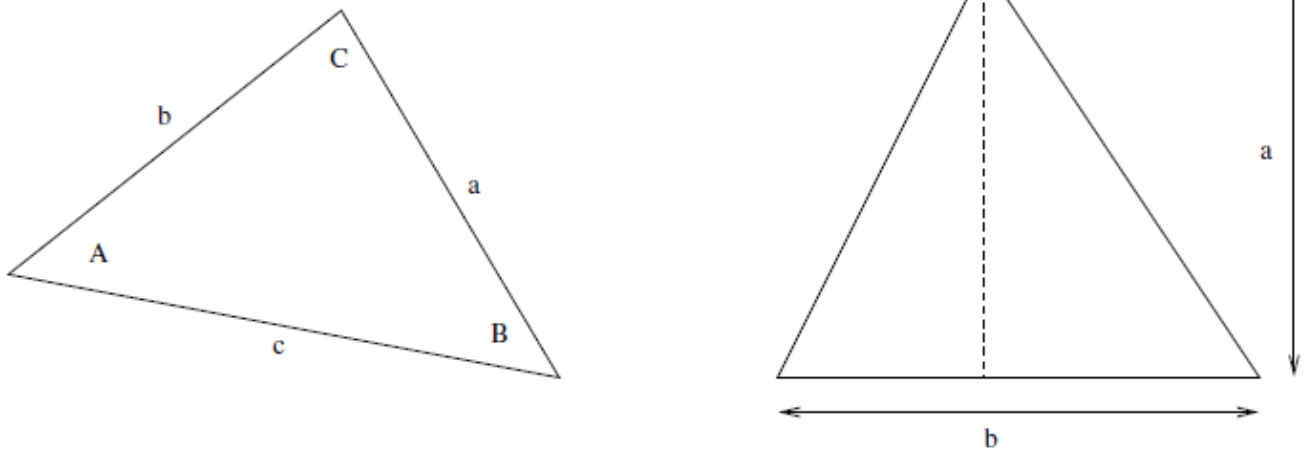
Um poliedro simples é *convexo* se o seu interior é um conjunto convexo.



Área Orientada de Polígono

Cálculo de Áreas Fonte: [SKIENA02]

Podemos calcular a área de um triângulo, a partir das coordenadas dos seus vértices, avaliando-se o produto vetorial definido a seguir. Notar que esse cálculo é facilmente implementável.



$$2 \cdot A(T) = \begin{vmatrix} a_x & a_y & 1 \\ b_x & b_y & 1 \\ c_x & c_y & 1 \end{vmatrix} = a_x b_y - a_y b_x + a_y c_x - a_x c_y + b_x c_y - c_x b_y$$

```
double signed_triangle_area(const Point* _A, const Point* _B, const Point* _C)
{
  return( (_A[X]*_B[Y] - _A[Y]*_B[X] + _A[Y]*_C[X]
         - _A[X]*_C[Y] + _B[X]*_C[Y] - _C[X]*_B[Y]) / 2.0 );
}
```

Cálculo de Áreas Fonte: [SKIENA02]

Podemos calcular a área de qualquer polígono triangular somando a área de todos os triângulos. Isto é fácil de implementar utilizando as rotinas já desenvolvidas que calculam a área de um triângulo.

No entanto, existe um algoritmo ainda mais esperto que se baseia na noção de áreas com sinais para triângulos, que foi utilizado na nossa rotina **`signed_triangle_area`**. Adequadamente somando as áreas com sinais dos triângulos definidos por um ponto arbitrário *p* com cada segmento do polígono *P* temos a área de *P*, porque os triângulos com áreas com sinais negativos cancelam a área fora do polígono. Isto se simplifica na computação da equação

$$A(P) = \frac{1}{2}\sum_{i=0}^{n-1}(x_i \cdot y_{i+1} - x_{i+1} \cdot y_i)$$

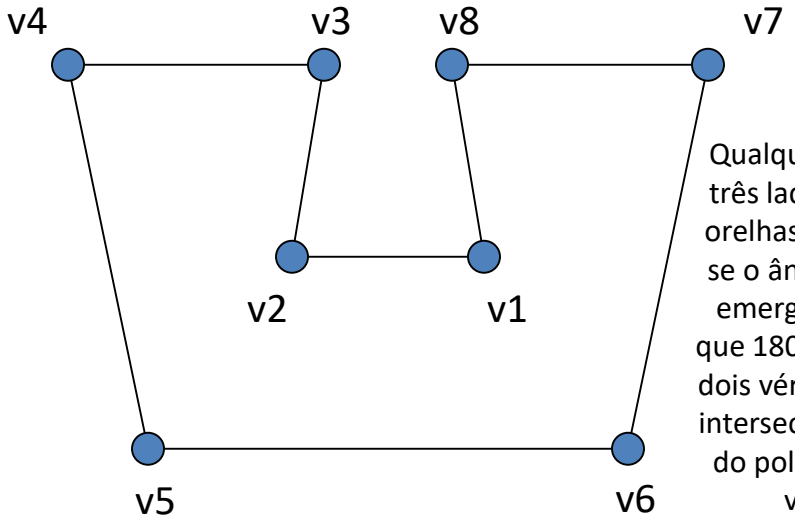
onde todos os índices são tomados modulo do número de vértices. Veja [OROURKE98] para uma exposição de porque isso funciona, mas que certamente leva a uma solução simples:

```
double area(polygon *_p)
{
  double total = 0.0; /* total area so far */
  int i, j;           /* counters */
  for (i=0; i<_p->n; i++) {
    j = (i+1) % _p->n;
    total += (_p->p[i][X]*_p->p[j][Y]) - (_p->p[j][X]*_p->p[i][Y]);
  }
  return(total / 2.0);
}
```

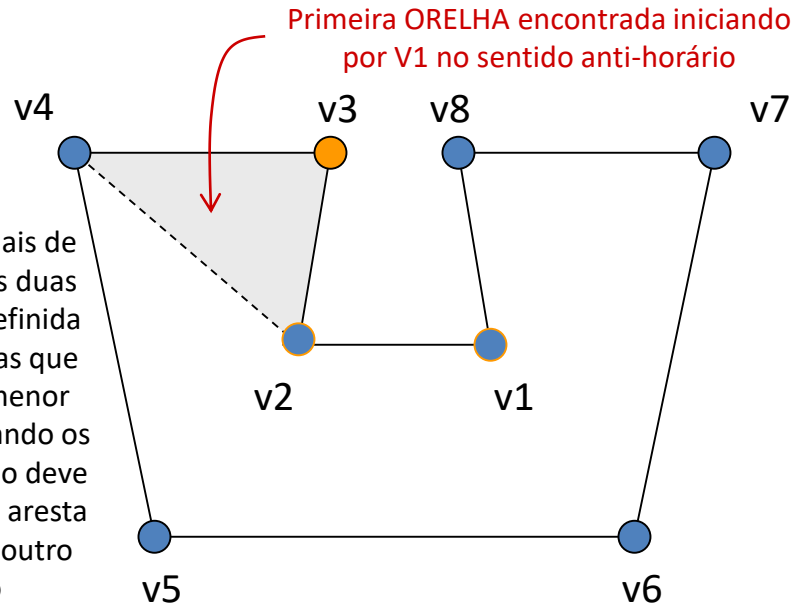
Algoritmos para Tesselagem de Polígonos

Como **tecer** uma face não convexa?

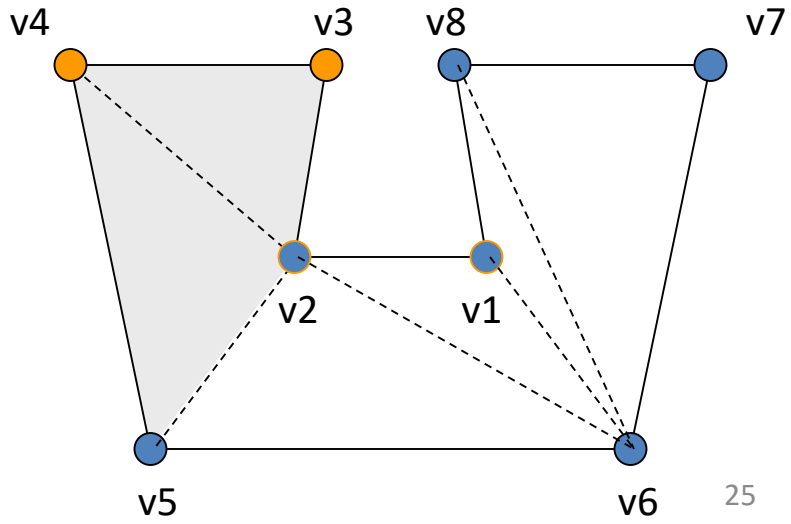
Solução de SKIENA & REVILLA, 2002, *Programming Challenges*, p.319



Qualquer polígono com mais de três lados tem pelo menos duas orelhas. Uma ORELHA é definida se o ângulo entre as arestas que emergem do vértice for menor que 180° e a corda conectando os dois vértices adjacentes não deve intersectar nenhuma outra aresta do polígono (i.e. nenhum outro vértice deve ficar no triângulo/orelha).



Encontrar ORELHA do polígono até remanescer um único triângulo.
POL = 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8
Defina duas listas com os vértices anteriores e posteriores:
L = 8 / 1 / 2 / 3 / 4 / 5 / 6 / 7
R = 2 / 3 / 4 / 5 / 6 / 7 / 8 / 1
Atualize a lista após o primeiro triângulo encontrado:
L = 8 / 1 / 2 / **2** / 4 / 5 / 6 / 7
R = 2 / **4** / 4 / 5 / 6 / 7 / 8 / 1
Atualize a lista após o próximo triângulo encontrado:
L = 8 / 1 / 2 / **2** / **2** / 5 / 6 / 7
R = 2 / **5** / 4 / 5 / 6 / 7 / 8 / 1
Repetir até que o número de triângulos seja menor que n-2



Predicados Geométricos

Introdução aos Predicados

Em qualquer algoritmo, existem funções que precisam ser avaliadas para definir que caminho seguir em sua sequência de execução. Em geral, essas funções retornam um valor Booleano (falso ou verdadeiro). Essas funções são chamadas de *predicados*.

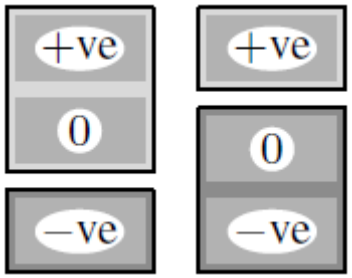
Em algoritmos de Geometria Computacional, o desenvolvimento de predicados geométricos robustos e precisos é de fundamental importância.

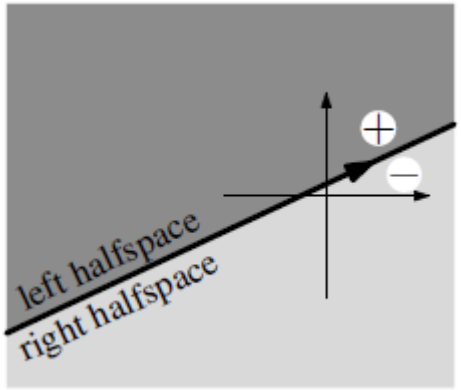
Tipo de Retorno de um Predicado

Geralmente pensa-se em predicados como funções com um tipo de retorno Booleano. O tipo booleano pode identificar se um contador atingiu algum limite, se uma tolerância predeterminada tem sido satisfeita, ou se o final de uma lista foi atingido. Tais predicados surgem na computação geométrica, mas um tipo adicional de teste é frequentemente necessário. Devido esse teste geométrico ter três resultados possíveis, refere-se a ele como um teste com ramificação ternária. Ainda mais frequente, tem-se o interesse em formar um predicado binário de três resultados possíveis.

A necessidade por três ramos em um teste pode ser vista quando considera-se uma linha orientada separando o plano. O plano é separado em pontos que se encontram no plano-médio positivo e pontos que se encontram no plano-médio negativo, tão bem como aqueles que se encontram na própria linha. Uma biblioteca geométrica oferecerá tal resposta ternária para os clientes, e o programador da aplicação irá decidir como os predicados devem ser formados.

Talvez uma dada aplicação só precise de uma resposta binária como, por exemplo, se um ponto está abaixo ou acima de um plano. Entretanto, testes geométricos devem ser oferecidos de tal maneira que também seja possível determinar se o ponto está exatamente sobre o plano, isto é, uma resposta ternária, se for possível.

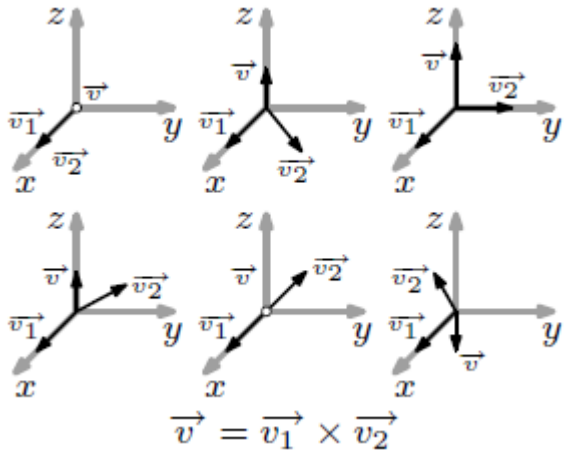




O Predicado Orientação no Plano

Determinar a orientação de um ponto em relação a linha definida por dois outros pontos é facilmente definida recorrendo a uma função que nos levará momentaneamente para uma terceira dimensão.

O Produto Vetorial. Existe mais de uma forma de definir o produto vetorial de dois vetores $\mathbf{v}_1$ e $\mathbf{v}_2$. Neste contexto, toma-se a visão clássica (em computação gráfica) que o produto vetorial $\mathbf{v} = \mathbf{v}_1 \times \mathbf{v}_2$ é um vetor que é simultaneamente ortogonal a $\mathbf{v}_1$ e $\mathbf{v}_2$, que obedece a regra da mão direita com relação aos dois vetores, e cuja magnitude é relacionada com as magnitudes dos dois vetores por $|\overrightarrow{v}| = |\overrightarrow{v_1}||\overrightarrow{v_2}| \sin\theta$



Para desenvolver uma intuição sobre produto vetoriais precisa-se apenas considerar como ele varia quando um dos dois vetores, por exemplo $\mathbf{v}_2$, move. Considere posicionar $\mathbf{v}_1$ tal que ele coincida com o eixo *x* positivo. Se $\mathbf{v}_2$ também coincide com o eixo *x*, o produto vetorial será o vetor zero. Isso é natural, pois os dois vetores não definem um plano, ou alternativamente, o paralelogramo que eles definem tem área zero. Agora considere que $\mathbf{v}_2$ gira em direção ao eixo *y*. A magnitude do produto vetorial aumenta até atingir um máximo quando $\mathbf{v}_1$ e $\mathbf{v}_2$ são ortogonais. Como $\mathbf{v}_2$ gira além de *y*, a magnitude de $\mathbf{v}$ retrai, atinge zero quando $\mathbf{v}_2 = -\mathbf{v}_1$, e quando $\mathbf{v}_2$ passa por $-x$, a direção de $\mathbf{v}$ é alinhada com o eixo $-z$.

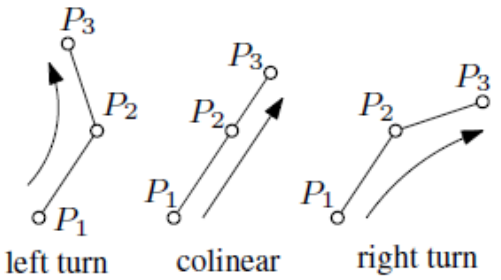
O Projeto do Predicado Orientação 2D

Alguém poderia argumentar que um predicado que reporta se três pontos são colineares seria necessário. Porém, ao invés de implementar tal predicado por si próprio, é mais conveniente implementar um predicado mais geral que determinará também colinearidade. Tal predicado de orientação no plano pode ter a seguinte assinatura

```
SIGN orient2d(const Point* _p1, const Point* _p2, const Point* _p3);
```

onde o tipo do retorno é definido como

```
enum SIGN {
  NEGATIVE = -1,
  ZERO = 0,
  POSITIVE = 1
};
```



Se necessário a implementação de vários predicados convenientes é agora simples. Os seguintes predicados binários delegam a requisição que eles recebem para a função `orient2d`, tal como

```
bool isLeftSide(const Point* _p1, const Point* _p2, const Point* _p3) {
  return orient2d(_p1,_p2,_p3) == POSITIVE;
}
bool areColinear(const Point* _p1, const Point* _p2, const Point* _p3) {
  return orient2d(_p1,_p2,_p3) == ZERO;
}
bool isRightSide(const Point* _p1, const Point* _p2, const Point* _p3) {
  return orient2d(_p1,_p2,_p3) == NEGATIVE;
}
```

Forma Matricial do Predicado Orientação 2D

Como a linha orientada P_2P_3 divide o plano em pontos encontrando-se sobre, a esquerda e a direita da linha, o sinal da expressão

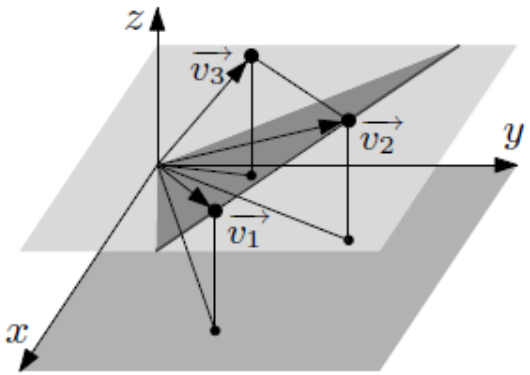
$$\overrightarrow{P_1P_2} \times \overrightarrow{P_2P_3}$$

identifica a localização do ponto P_3. Se o sinal for positivo, P_3 está a esquerda; se ele for zero, P_3 está sobre a linha; e se ele for negativo, P_3 está a direita da linha. O produto vetorial acima avalia o determinante

$$\begin{vmatrix} x_2 - x_1 & x_3 - x_2 \\ y_2 - y_1 & y_3 - y_2 \end{vmatrix},$$

que por um lado pode ser expandido no determinante 3x3

$$\begin{vmatrix} x_1 & x_2 - x_1 & x_3 - x_2 \\ y_1 & y_2 - y_1 & y_3 - y_2 \\ 1 & 0 & 0 \end{vmatrix},$$



onde os dois valores x_1 e y_1 podem ser arbitrariamente escolhido. Somando a primeira coluna com a segunda e o resultado da segunda com a terceira, obtém-se a equivalente forma homogênea

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}.$$

Interpretando essa expressão como três vetores em 3D ao invés de três pontos em 2D, o teste particular a ser usado irá depender se está testando a inclusão do ponto em questão em um plano-médio aberto ou fechado. Se deseja determinar se um ponto encontra-se no plano-médio esquerdo aberto, testa-se `orient2d(..)==POSITIVE`., e no plano-médio esquerdo fechado, testa-se `orient2d(..)==NEGATIVE`

O Predicado em qual Lado do Círculo

Da mesma forma que dois pontos definem naturalmente uma linha que separa o plano em duas regiões, além da linha separando elas, três pontos no plano P_1, P_2 e P_3 definem um círculo que divide o plano em duas regiões, além do próprio círculo.

Matrix Form of the Side of Circle Predicate

A circle with center (x_c, y_c) and radius r in the plane has the equation

$$(x - x_c)^2 + (y - y_c)^2 = r^2,$$

which expands to

$$(x^2 + y^2) - 2(xx_c + yy_c) + (x_c^2 + y_c^2 - r^2) = 0.$$

More generally,

$$A(x^2 + y^2) + Bx + Cy + D = 0$$

is the equation of a circle in the plane provided that $A \neq 0$.

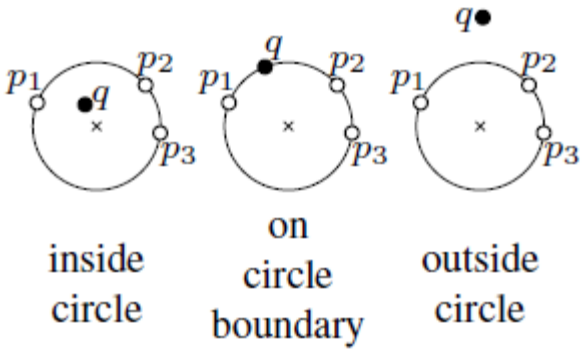
The equation above can be written as the determinant

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0,$$

where

$$A = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}, \qquad B = \begin{vmatrix} x_1^2 + y_1^2 & y_1 & 1 \\ x_2^2 + y_2^2 & y_2 & 1 \\ x_3^2 + y_3^2 & y_3 & 1 \end{vmatrix},$$

$$C = \begin{vmatrix} x_1^2 + y_1^2 & x_1 & 1 \\ x_2^2 + y_2^2 & x_2 & 1 \\ x_3^2 + y_3^2 & x_3 & 1 \end{vmatrix}, \qquad D = \begin{vmatrix} x_1^2 + y_1^2 & x_1 & y_1 \\ x_2^2 + y_2^2 & x_2 & y_2 \\ x_3^2 + y_3^2 & x_3 & y_3 \end{vmatrix}.$$



It is clear that the determinant in Eq. (2.1) vanishes if the point $P(x, y)$ coincides with any of the three points $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, or $P_3(x_3, y_3)$. Moreover, we know from § 2.2 that $A \neq 0$ if and only if the three given points are not colinear.

It is also clear that all the points lying either inside or outside the circle generate a positive determinant and that the points lying on the other side generate a negative determinant. Since exchanging any two rows in Eq. (2.1) would flip the sign of the determinant, the order of the three given points does matter. Clients of this predicate would likely rather not be careful in selecting a particular order for the three points and so it would be appropriate to take a small efficiency hit and compute the 3×3 determinant for the orientation of the three points in addition to computing the 4×4 determinant in Eq. (2.1). And so a point $P(x, y)$ can be classified with respect to a circle defined by three points by evaluating the following equation:

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{vmatrix} \times \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

$$= \text{side\_of\_circle}(P,P_1,P_2,P_3) \begin{cases} < 0 & \text{inside,} \\ = 0 & \text{on the circle boundary,} \\ > 0 & \text{outside.} \end{cases}$$

Ponto mais Próximo em um Segmento de Reta

PONTO MAIS PRÓXIMO EM UM SEGMENTO DE RETA UTILIZANDO PRODUTO INTERNO

Projeção do ponto C na reta AB:

$$C' = A + t_{C'}\left(B - A\right)$$

Ponto mais próximo de C no segmento AB:

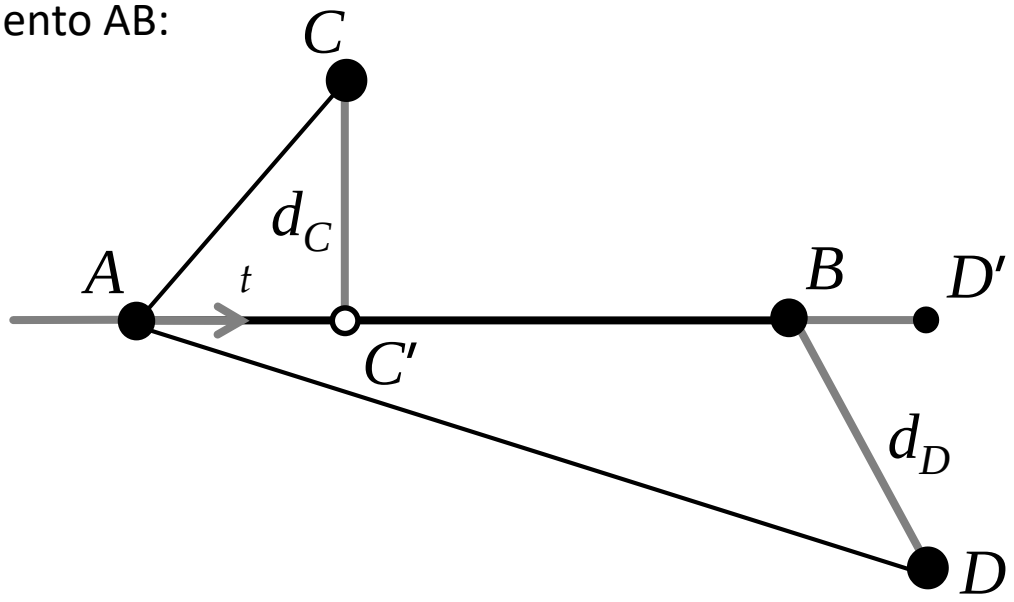
$$P = C'$$

Valor paramétrico do ponto C' no segmento AB:

$$t_{C'} = \frac{\overrightarrow{AB} \circ \overrightarrow{AC}}{\left|\overrightarrow{AB}\right|^2}$$

(produto interno)

$$0 < t_{C'} < 1$$



Projeção do ponto D na reta AB:

$$D' = A + t_{D'}\left(D - C\right)$$

Valor paramétrico do ponto D' no segmento AB:

$$t_{D'} = \frac{\overrightarrow{AB} \circ \overrightarrow{AD}}{\left|\overrightarrow{AB}\right|^2}$$

$$t_{D'} > 1$$

Ponto mais próximo de D no segmento AB:

$$P = B$$

Algoritmos para Interseção de Segmentos de Reta

Segmentos de Retas e Interseções Fonte: [SKIENA02]

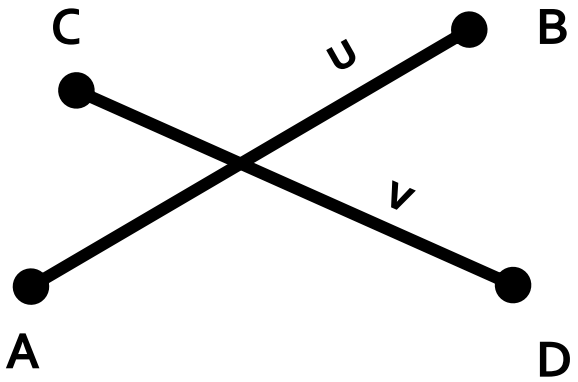
Um segmento de linha s é a porção de uma linha L que se encontra entre dois pontos dados , inclusive. Assim, os segmentos de linha são mais naturalmente representados por pares de pontos extremos.

A mais importante primitiva geométrica em segmentos, testar se um determinado par deles se cruzam, prova ser surpreendentemente complicada por causa de casos especiais difíceis que surgem. Dois segmentos podem estar em linhas paralelas , o que significa que não se cruzam em tudo. Um segmento pode intersectar a extremidade de um outro, ou os dois segmentos podem estar em cima um do outro de modo que eles se intersectem num segmento em vez de um único ponto.

Este problema de casos especiais geométricas , ou *degeneração* , complica seriamente o problema da construção de implementações robustas de algoritmos de geometria computacional. A degenerescência pode ser uma verdadeira dor no pescoço para lidar com eles. Leia qualquer especificação do problema com cuidado para ver se ele não promete linhas paralelas ou segmentos sobrepostos. Sem essas garantias , no entanto , é melhor programar defensivamente e lidar com eles.

A maneira correta de lidar com a degeneração é basear todos os cálculos em um pequeno número de primitivas geométricas cuidadosamente elaborados.

COMO TRATAR A INTERSEÇÃO DE SEGMENTOS DE RETA DE FORMA ROBUSTA E EFICIENTE?



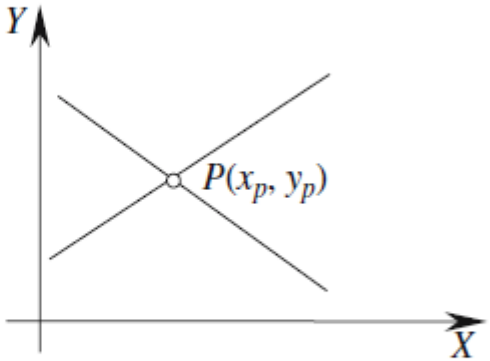
Fonte: Ricardo Marques

Fonte [VINCE05]

1.11.7 Point of intersection of two straight lines

General form of the line equation

Given $a_1x + b_1y + c_1 = 0$
$a_2x + b_2y + c_2 = 0$



then
$$\frac{x_P}{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}} = \frac{y_P}{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}} = \frac{-1}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

Intersect at
$$x_P = \frac{c_2b_1 - c_1b_2}{a_1b_2 - a_2b_1} \qquad y_P = \frac{a_2c_1 - a_1c_2}{a_1b_2 - a_2b_1}$$

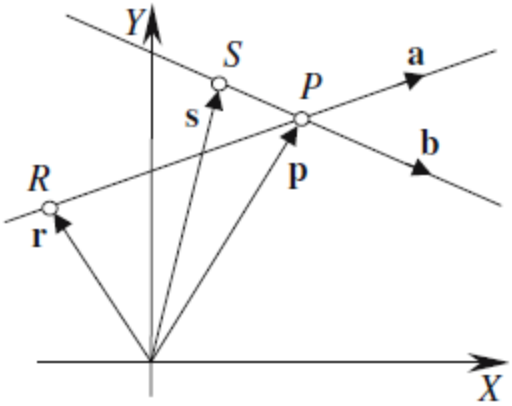
The lines are parallel if $a_1b_2 - a_2b_1 = 0$

Parametric form of the line equation

Given $\mathbf{p} = \mathbf{r} + \lambda\mathbf{a}$ $\mathbf{q} = \mathbf{s} + \varepsilon\mathbf{b}$

where $\mathbf{r} = x_R\mathbf{i} + y_R\mathbf{j}$ $\mathbf{s} = x_S\mathbf{i} + y_S\mathbf{j}$

and $\mathbf{a} = x_a\mathbf{i} + y_a\mathbf{j}$ $\mathbf{b} = x_b\mathbf{i} + y_b\mathbf{j}$



then $$\lambda = \frac{x_b(y_S - y_R) - y_b(x_S - x_R)}{x_b y_a - x_a y_b}$$

and $$\varepsilon = \frac{x_a(y_S - y_R) - y_a(x_S - x_R)}{x_b y_a - x_a y_b}$$

Point of intersection $x_P = x_R + \lambda x_a$ $y_p = y_R + \lambda y_a$

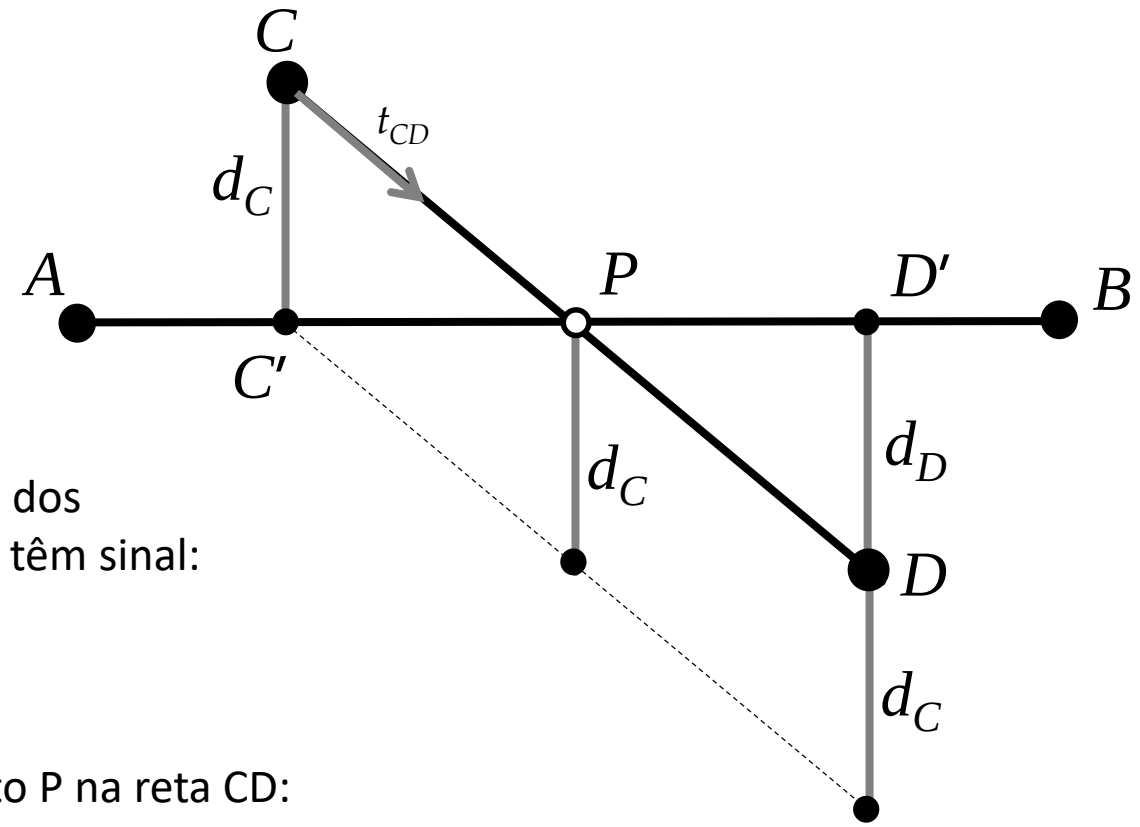
or $x_P = x_S + \varepsilon x_b$ $y_P = y_S + \varepsilon y_b$

The lines are parallel if $x_b y_a - x_a y_b = 0$

INTERSEÇÃO BASEADA NA REPRESENTAÇÃO PARAMÉTRICA DOS SEGMENTOS

$$P = A + t_{AB}(B - A)$$

$$P = C + t_{CD}(D - C)$$



Considere que as distâncias dos pontos C e D para a reta AB têm sinal:

$$d_C > 0 \qquad d_D < 0$$

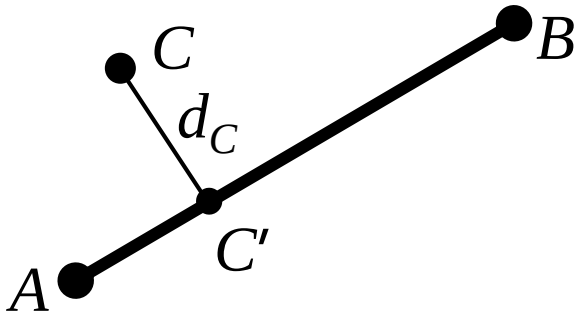
Valor paramétrico do ponto P na reta CD:

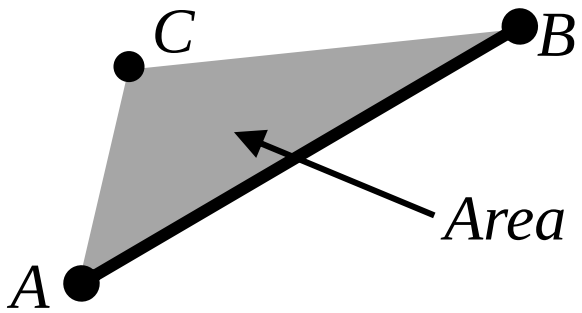
$$t_{CD} = \frac{|d_C|}{|d_C| + |d_D|} \qquad t_{CD} = \frac{d_C}{d_C - d_D} \qquad 0 \le t_{CD} \le 1$$

M. Gavrilova and J. G. Rokne. (2000) “Reliable line segment intersection testing”, CAD 32, 737–746

INTERSEÇÃO BASEADA NA REPRESENTAÇÃO PARAMÉTRICA DOS SEGMENTOS

Distância com sinal pode ser substituída por produto vetorial





$$Area = \frac{\left|\overrightarrow{AB}\right| \cdot d_C}{2}$$

$$Area = \frac{\overrightarrow{AB} \times \overrightarrow{AC}}{2}$$

$$Area = \frac{(B-A) \times (C-A)}{2}$$

Definição: (dobro da) área orientada do triângulo

$$orient2d(A,B,C) = (B-A) \times (C-A)$$

Portanto:

$$d_C = \frac{orient2d(A,B,C)}{\left|\overrightarrow{AB}\right|}$$

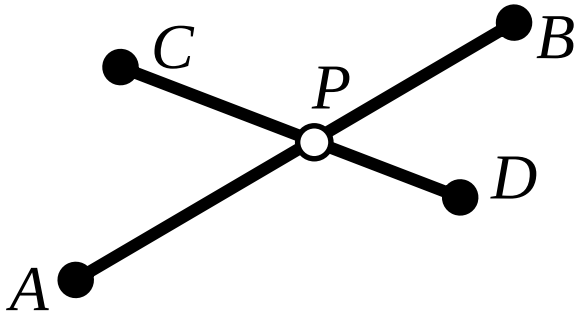
Analogamente:

$$d_D = \frac{orient2d(A,B,D)}{\left|\overrightarrow{AB}\right|}$$

Observe que os sinais das distâncias são resolvidos naturalmente.

INTERSEÇÃO BASEADA NA REPRESENTAÇÃO PARAMÉTRICA DOS SEGMENTOS

Valor paramétrico do ponto P na reta CD:



$$t_{CD} = \frac{d_C}{d_C - d_D}$$

$$d_C = \frac{orient2d(A,B,C)}{\left|\overrightarrow{AB}\right|} \qquad d_D = \frac{orient2d(A,B,D)}{\left|\overrightarrow{AB}\right|}$$

Portanto:

$$t_{CD} = \frac{orient2d(A,B,C)}{orient2d(A,B,C) - orient2d(A,B,D)} \qquad P = C + t_{CD}\left(D - C\right)$$

Analogamente:

$$t_{AB} = \frac{orient2d(C,D,A)}{orient2d(C,D,A) - orient2d(C,D,B)} \qquad P = A + t_{AB}\left(B - A\right)$$

Segment_Segment_Intersection(u, v):

if both endpoints of **u** are over **v** **then**
  return false
end if
$orient2d(C,D,A) > 0$ $orient2d(C,D,B) > 0$

if both endpoints of **u** are under **v** **then**
  return false
end if
$orient2d(C,D,A) < 0$ $orient2d(C,D,B) < 0$

if both endpoints of **v** are over **u** **then**
  return false
end if
$orient2d(A,B,C) > 0$ $orient2d(A,B,D) > 0$

if both endpoints of **v** are under **u** **then**
  return false
end if
$orient2d(A,B,C) < 0$ $orient2d(A,B,D) < 0$

if **u** and **v** are collinear **then**
  return false
end if
$orient2d(C,D,A) = 0$ $orient2d(C,D,B) = 0$

ou

$orient2d(A,B,C) = 0$ $orient2d(A,B,D) = 0$

~~**if** **u** and **v** are parallel **then**~~
  return false
~~**end if**~~
~~$orient2d(C,D,A) = orient2d(C,D,B)$~~

ou

~~$orient2d(A,B,C) = orient2d(A,B,D)$~~

if **u** touches **v** **then**
  return true
end if
(there are many cases)
$orient2d(C,D,B) = 0$ $orient2d(C,D,A) < 0$

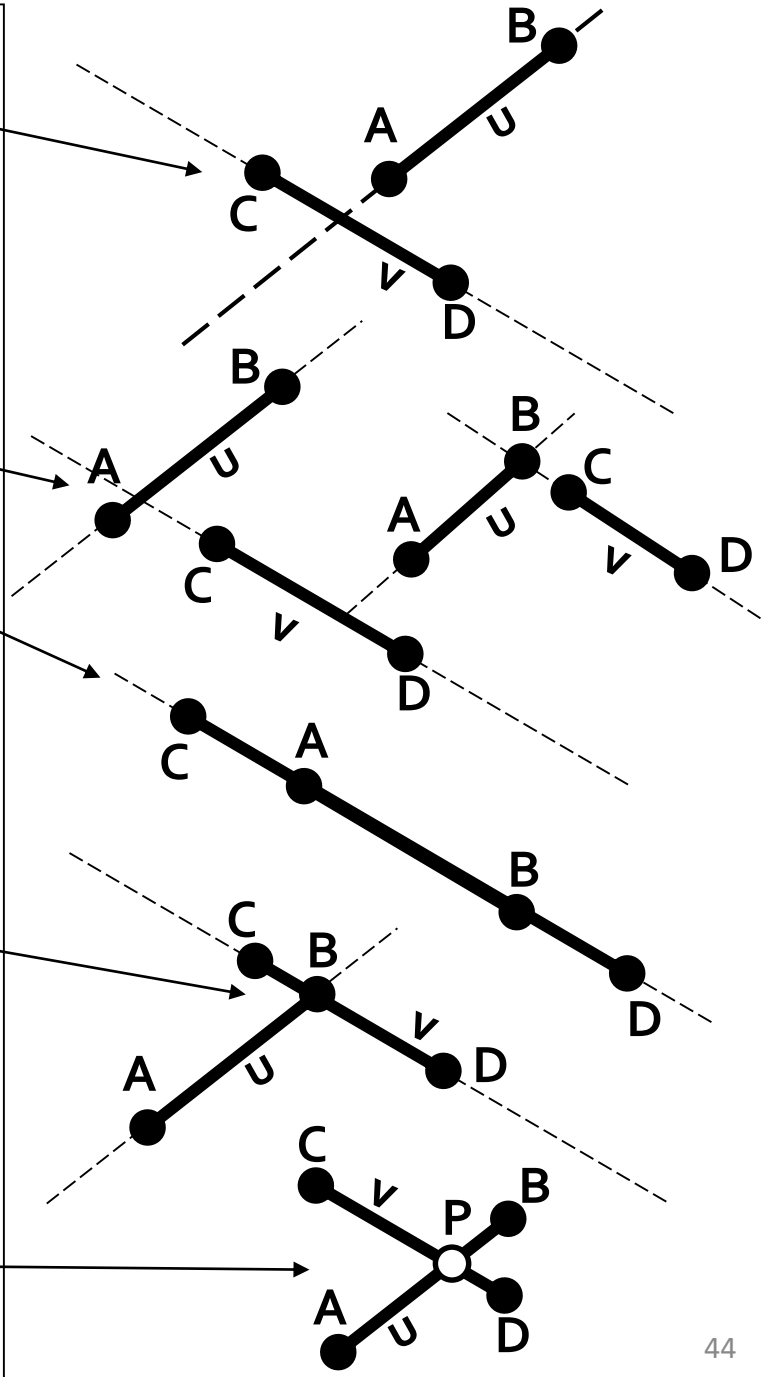
$P = B$

// When get to this point, there is an intersection point

$$t_{CD} = \frac{orient2d(A,B,C)}{orient2d(A,B,C) - orient2d(A,B,D)}$$

return true

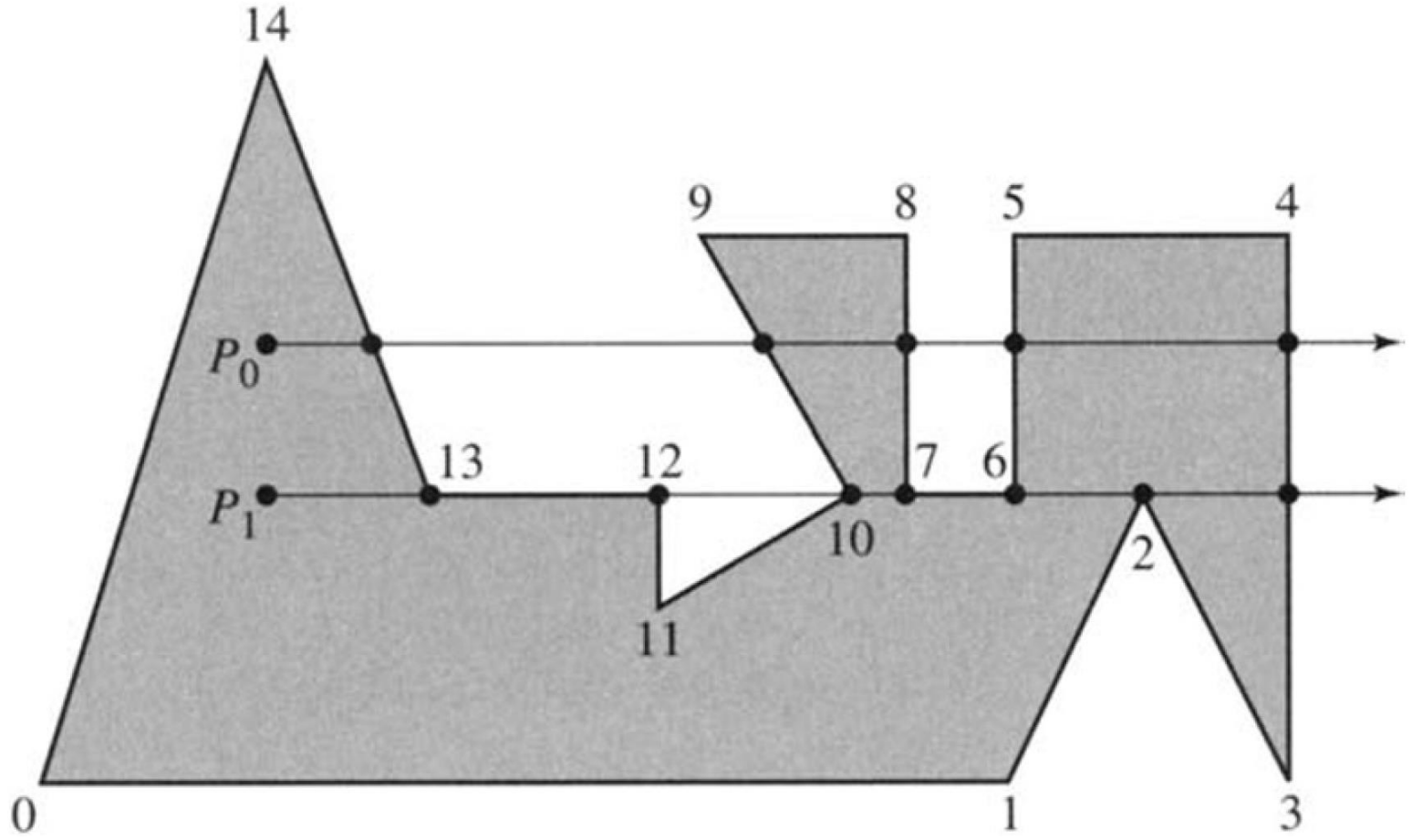
$$P = C + t_{CD}(D - C)$$



Algoritmo para verificação de ponto dentro de polígono

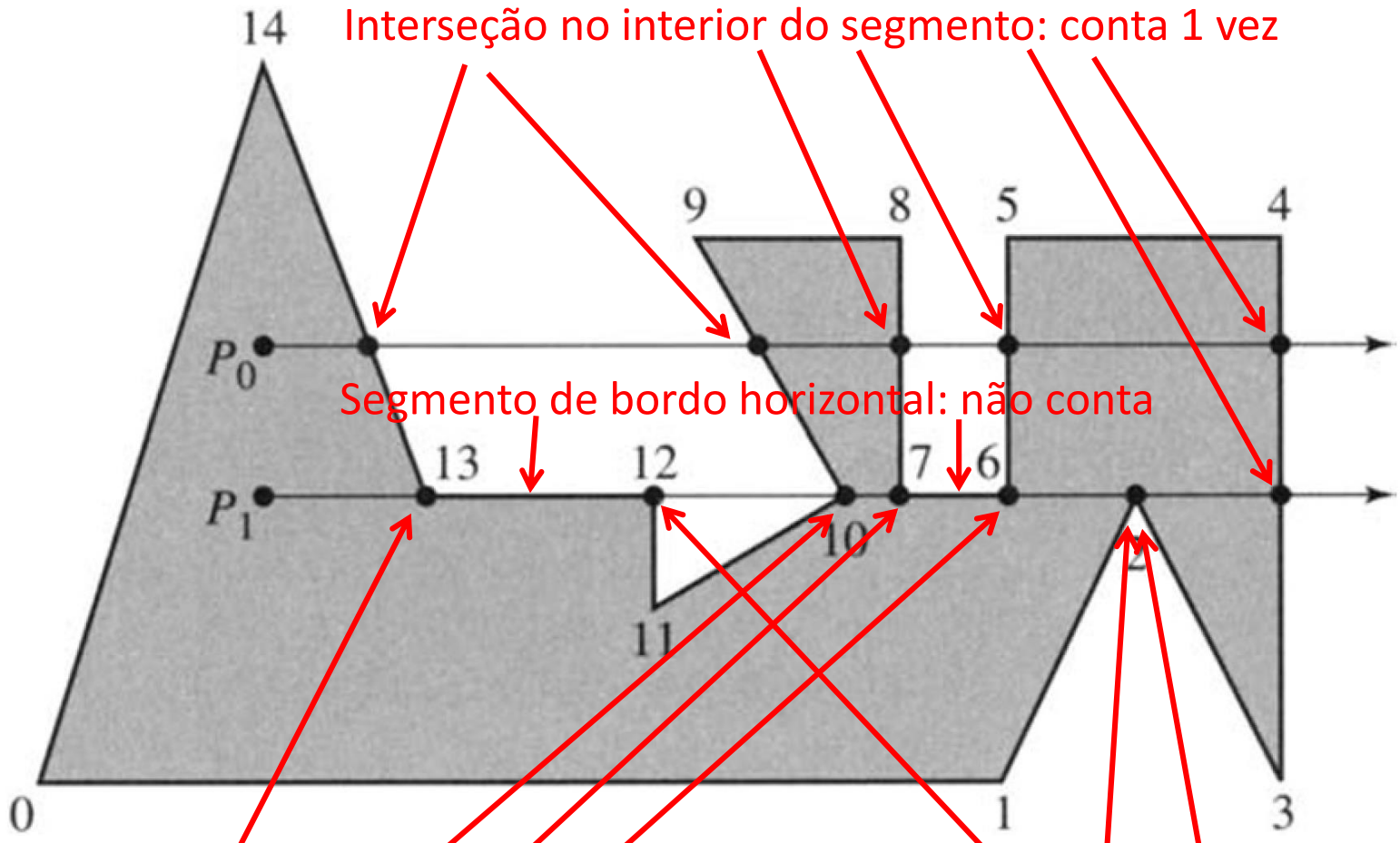
Algoritmo do raio (ou tiro)

Philip Schneider and David Eberly Geometric Tools for Computer Graphics, 2003, p.70



Uma semirreta (raio) que parte de qualquer ponto dentro de polígono em uma direção qualquer cortará as curvas no bordo do polígono um número impar de vezes. Se a semirreta cortar a fronteira do polígono um número par de vezes, o ponto está fora do polígono.

Critérios para contar interseções do raio com um segmento de bordo



Precisão Numérica
Aritmética Exata e Adaptativa

Por que utilizar Aritmética Exata?

Fonte: Ricardo Marques

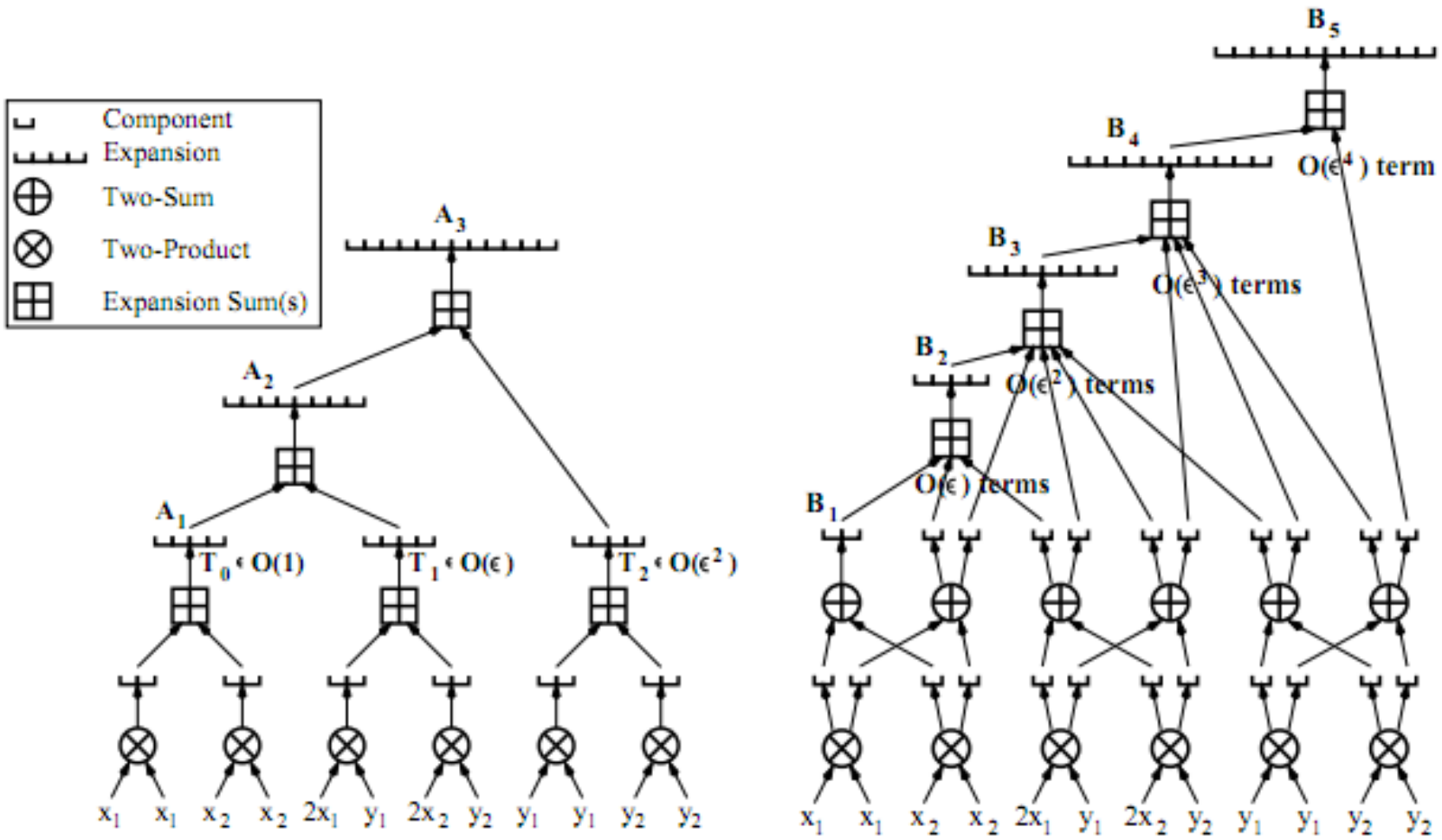
- Utilizar tolerâncias *hard-coded* não resolve!
  - Uma tolerância de 1e-07 pode ser suficiente para modelos de dimensões pequenas.
  - Mas se o modelo possuir dimensões de centenas de km, 1e-07 é inexpressivo. Nesse caso, 1e+00 é uma tolerância bem mais aceitável, representando um erro relativo de 1e-05, ou seja, 1 cm.
  - Ao mesmo tempo, uma tolerância de 1e+00 pode não faz sentido em modelos pequenos.
- Com Aritmética Exata, tolerâncias não são mais necessárias.

O que é Aritmética Exata?

Fonte: Ricardo Marques

- Aritmética Exata é uma técnica para se fazer cálculos com alto nível de precisão
  - 1e-08 é zero? -3.1415e-10 é zero?
- Evita erros de arredondamento:
  - 1e+08 + 1e-16 = 1e+08 ???
  - Nos operadores de aritmética exata, todo número (double) de entrada é quebrado em duas componentes (numéricas) não-sobrejacentes e com ordem de grandezas diferentes.
  - Ao se utilizar sucessivos operadores, componentes podem ser quebradas novamente. Ao fim, todas as componentes geradas são unidas, minimizando erro numérico.

O que é Aritmética Exata Adaptativa?



Numerical analysis is clearly the first place to look for answers about accuracy in a world of approximations. A lot is known about the accuracy of the output of a computation given the accuracy of the input. For example, to get precise answers with linear problems one would have to perform computations using four to five times the precision of the initial data. In the case of quadratic problems, one would need forty to fifty times the precision. Sometimes there may be guidelines that help one improve the accuracy of results. Given the problems with floating point arithmetic, one could try other types of arithmetic.

Bounded Rational Arithmetic. This is suggested in [Hoff89] and refers to restricting numbers to being rational numbers with denominators that are bounded by a given fixed integer. One can use the method of continued fractions to find the best approximation to a real by such rationals.

Infinite Precision Arithmetic. Of course, there are substantial costs involved in this.

"Exact" Arithmetic. This does not mean the same thing as infinite precision arithmetic. The approach is described in [Fort95]. The idea is to have a fixed but relatively small upper bound on the bit-length of arithmetic operations needed to compute geometric predicates. This means that one can do integer arithmetic. Although one does not get "exact" answers, they are reliable. It is claimed that boundary-based **faceted** modelers supporting regularized set operators can be implemented with minimal overhead (compared with floating point arithmetic). Exact arithmetic works well for linear objects but has problems with smooth ones. See also [Yu92] and [CuKM99].

Interval Analysis. See Chapter 18 for a discussion of this and also [HuPY96a] and [HuPY96b].

Just knowing the accuracy is not always enough if it is v
like. Geometric computations often involve many steps. Rather
accuracy only after data structures and algorithms have bee
perhaps also use accuracy as one criterion for choosing the
algorithms.

One cause for the problem indicated in Figure 5.48 is that
ent computations to establish a common fact. The question
segment intersected the edge of the cube was answered twice –
f and second by using the face **g**. The problem is in the redundancy in the representation of the edge and the fact that the intersection is determined from a collection of isolated computations. If one could represent geometry in a nonredundant way, then one would be able to eliminate quite a few inconsistency problems. Furthermore, the problem shown in Figure 5.48 would be resolved if, after one found the intersection with face **f**, one would check for intersections with all the faces adjacent to **f** and then resolve any inconsistencies.

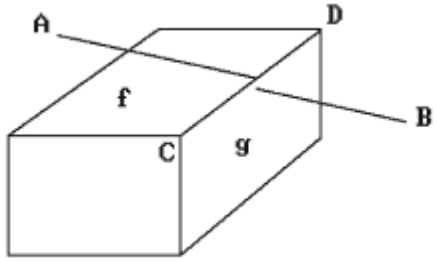


Figure 5.48. Intersection inconsistencies due to round-off errors.

Max K. Agoston
Computer Graphics and Geometric Modeling
Springer 2004

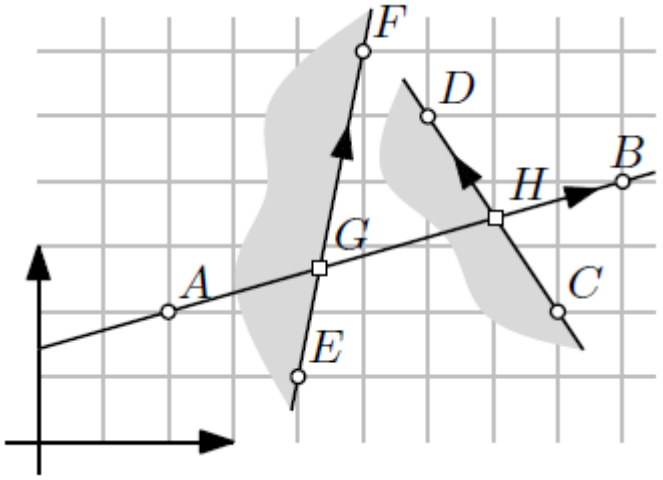
Sherif Ghali
Introduction to Geometric Computing
Springer
2008

But how do we know, when starting the design of a system, whether floating point numbers are adequate? The answer is sometimes easy. It is clear that interactive computer games, or systems that need to run in real time in general, cannot afford to use data types not provided by the hardware. This restricts the usable data types to int, long, float, and/or double. It is also clear that systems that perform Boolean operations on polygons or solids, such as the ones discussed in Chapter 28, will need to use an exact number type. In general, however, this is an important decision that needs to be made for each individual system. Genericity is a powerful device at our disposal to attempt to delay the choice of number type as long as possible, but to generate one executable or program, the various compromises have to be weighed and the decision has to be made.

At this time there is no silver bullet to determine whether to sacrifice efficiency and use an exact number type. A simple rule of thumb is to consider the compromise between speed and accuracy. If the system requirements suggest speed, then we have to sacrifice accuracy, and vice versa. The answer is of course easy if neither is required, but it is more often the case that both are.

Sherif Ghali
Introduction to Geometric Computing
Springer
2008

This theme is the topic of Chapter 7, but lest this issue appear to be of mere theoretical interest, an example is warranted. Consider clipping a segment AB in the plane by the two positive (i.e., left) halfspaces of CD and EF. In theory the resulting clipped segment does not depend on the order of the two clipping operations. The code below uses the type **float** to compute the final intersection point directly (Gd) by intersecting AB and EF, and also computes the ostensibly identical point indirectly (Gi) by first computing AH then intersecting AH and EF.



Sherif Ghali
Introduction to Geometric Computing
Springer
2008

```
int main()
{
   const Point_E2f A(2,2), B(9,5);
   const Point_E2f C(8,2), D(6,5);
   const Point_E2f E(4,1), F(5,6);

   const Segment_E2f AB(A,B), CD(C,D), EF(E,F);

   const Point_E2f Gd = intersection_of_lines(AB, EF);
   const Point_E2f H = intersection_of_lines(AB, CD);

   const Segment_E2f AH(A,H);
   const Point_E2f Gi = intersection_of_lines(AH, EF);

   // assert( Gd == Gi ); // fails

   print( Gd.x() );
   print( Gi.x() );

   print( Gd.y() );
   print( Gi.y() );
}
```

Sherif Ghali
Introduction to Geometric Computing
Springer
2008

After finding that the coordinates differ, we print the sign bit, the exponent, and the mantissa of the two x-coordinates then those of the y-coordinates (see § 7.5).

```
0 10000001 00011010000000000000000
0 10000001 00011010000000000000001
0 10000000 10000100000000000000000
0 10000000 10000100000000000000000
```

And so we see that the direct computation of the x-coordinate leads to a mantissa with a long trail of zeros, whereas the indirect computation leads to an ending least-significant bit of 1. This single-bit difference suffices for the equality operator to conclude that the two points are not equal. Performing the same computation using a combination of indirect steps only reduces the quality of the resulting floating point coordinates.