5. Rastergrafik-Algorithmen

5.1. Grundlegende 2D-Zeichenalgorithmen

Problem:
virtuelles Zeichenblatt mit stetigen Koordinaten (float *x*, float *y*)
→ reales Rasterdisplay mit diskreten Koordinaten (int *X*, int *Y*)

Nicht-ganzzahligen Punkten (*x*, *y*) werden ganzzahlige *Alias-Punkte* (*X*, *Y*) zugeordnet.
Dabei: Ungenauigkeiten (und im Extremfall sehr unschöne Effekte).

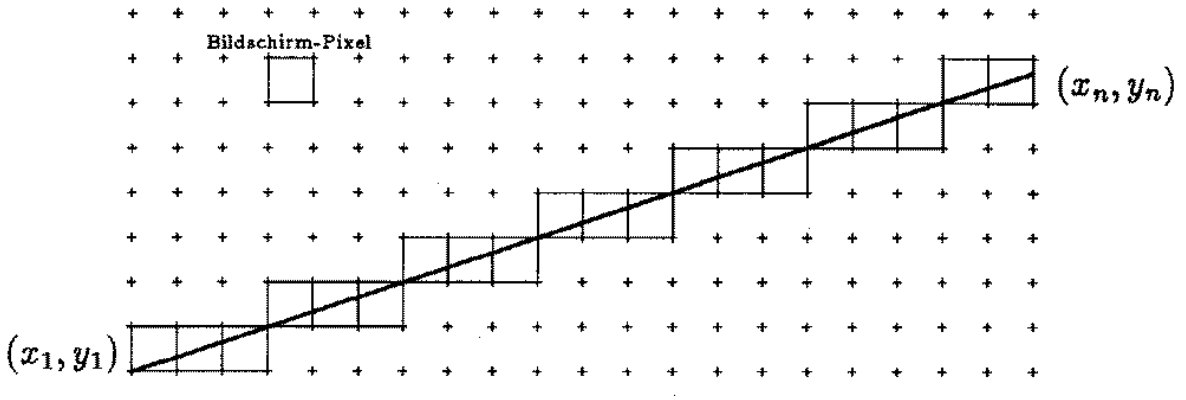
Ziel 1: Linien sollen ihrem idealen Verlauf möglichst nahe kommen
Ziel 2: Ausgleich möglicher, verbleibender Störeffekte.

Zunächst zu Ziel 1:
Raster-Konvertierungen für Geraden(-stücke), Kreise, Ellipsen...

Zunächst: *Linien* (*Geradenstücke*) mit Dicke von 1 Pixel darstellen

⇒ für Gerade mit Steigung zwischen –1 und 1:
setze in jeder Spalte genau 1 Pixel.

Für andere Geraden: Symmetrie ausnutzen (Koordinaten vertauschen)!



(aus Rauber 1993)

Abb.: Die gesetzten Pixel sind als Quadrate dargestellt.

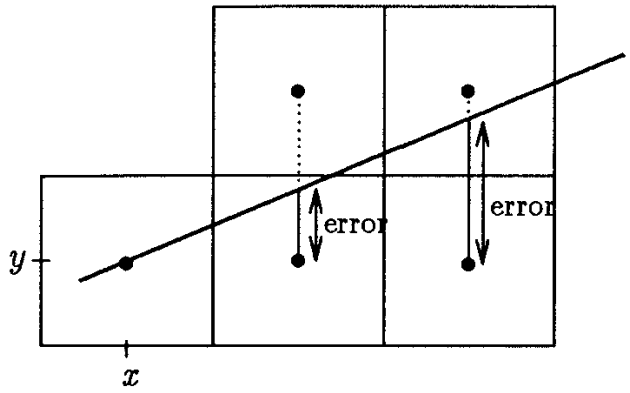
Naiver Algorithmus:
aus *x*-Werten von x_1 bis x_n die *y*-Werte nach Geradengleichung $y = mx+b$ berechnen und auf nächstgelegene Integer-Zahl runden.

Nachteile:
- Floating-Point-Arithmetik (langsamer als Integer-Arithmetik)
- Multiplikationen
- Rundungsoperationen

möglichst einfache Operationen: wichtig auch für evtl. Hardware-Realisierung!

schrittweise Verbesserung:

- *y* inkrementieren statt durch Multiplikation berechnen
- Rundungsoperation durch Variable *error* und Vergleichsabfrage ersetzen (*error* liefert Abstand der idealen Geraden vom Pixel):



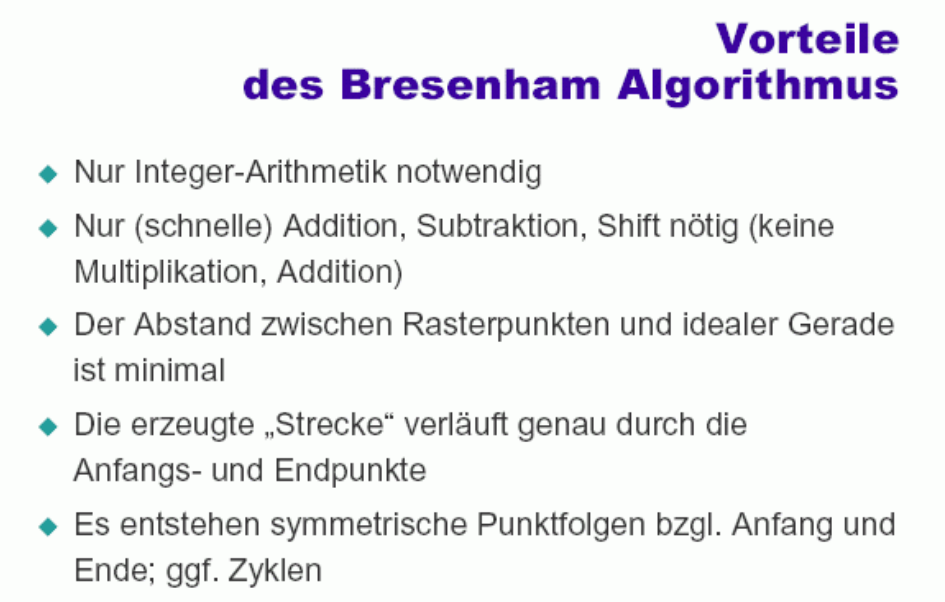
wenn *error* >= 0.5 ist, setze Pixel um 1 höher und dekrementiere *error* um 1. Damit optimale Nähe der gesetzen Pixel zur idealen Geraden gewährleistet.

Immer noch floating point-Operationen
⇒ weitere Verbesserung: alles mit $dx = x_n - x_1$ multiplizieren, dann Rechnung nur noch im Integer-Bereich!

Zusätzlich wird error mit $-dx/2$ (bzw. –int(dx/2)) statt mit 0 initialisiert (⇒ Vergleich nur noch mit 0 statt mit 0,5).

Ergebnis:
Bresenham-Algorithmus zur Linienrasterung (1965)

```
void Linie (int x1, int y1, int xn, int yn,
            int value)
/* Anfangspunkt (x1, y1), Endpunkt (xn, yn),
   Steigung zwischen 0 und 1 angenommen */
{
int error, x, y, dx, dy;
dx = xn - x1; dy = yn - y1;
error = -dx/2; y = y1;
for (x = x1; x <= xn; x++)
   {
   set_pixel(x, y, value);
   error += dy;
   if (error >= 0)
      {
      y++;
      error -= dx;
      }
   }
}
```

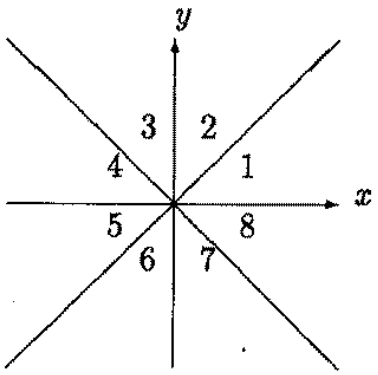


(Krömker 2001)

In obiger Form nur für 1. Oktanden.
Verallgemeinerung: Ausnutzung der Symmetrie.

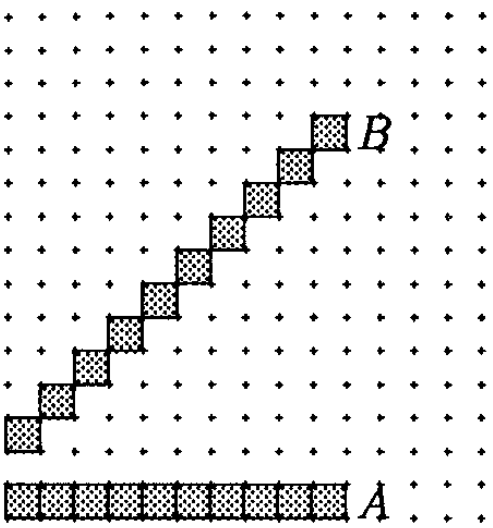
"*Treibende Achse*" = Achse, deren Werte mit fester Schrittweite 1 durchlaufen werden (im obigen Fall: *x*-Achse).

Fallunterscheidung nach dem Oktanden der Geraden:



Oktand	treibende Achse	andere Achse
1	x	inkrementiert
2	y	inkrementiert
3	y	dekrementiert
4	x	dekrementiert
5	x	inkrementiert
6	y	inkrementiert
7	y	dekrementiert
8	x	dekrementiert

Nachteil:
Linien unterschiedlicher Steigung erscheinen unterschiedlich hell.



Linie *B* ist $\sqrt{2}$ -mal länger als Linie *A*, hat aber dieselbe Zahl von Pixeln $\Rightarrow$ erscheint schwächer!

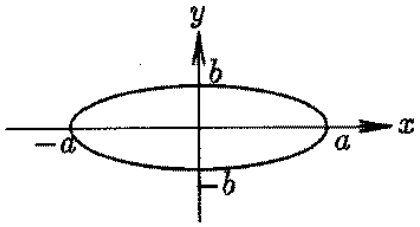
Abhilfe: Intensität (Parameter **`value`**) mit der Steigung der Linie erhöhen.

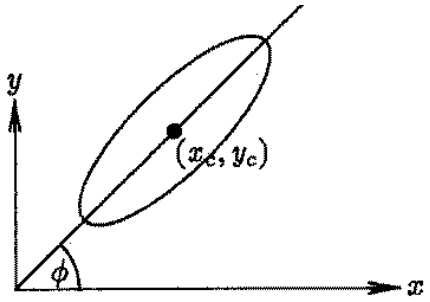
Rasterung von Ellipsen

Parametermethode (auch: trigonometrische Methode):
Ellipsengleichung in Parameterform (Parameter θ = Winkel im Mittelpunkt):

$$x = a \cos \theta, \quad y = b \sin \theta.$$

θ durchläuft die Werte von 0 bis 360° bzw. bis 2π.
a und *b* sind die Abschnitte auf den Hauptachsen. Hier Hauptachsen = Koordinatenachsen.





Um den Winkel ϕ gedrehte Ellipse mit Mittelpunkt (x_c, y_c):

$x = a \cos \phi \cos \theta - b \sin \phi \sin \theta + x_c = A \cos \theta - B \sin \theta + x_c$
$y = a \sin \phi \cos \theta + b \cos \phi \sin \theta + y_c = C \cos \theta + D \sin \theta + y_c$

mit $A = a \cos \phi$, $B = b \sin \phi$, $C = a \sin \phi$, $D = b \cos \phi$

(durch Anwendung einer Rotation und Translation auf die ungedrehte Ellipse, vgl. Transformationsmatrizen (Kap. 4) bzw. Lineare Algebra).

- Durchlaufe äquidistante Parameterwerte
- verbinde die zugehörigen Ellipsenpunkte durch Linien (Geradensegmente).

n = Anzahl gewünschter Geradensegmente $\Rightarrow$
Schrittweite **`incr`** = $2\pi/n$.

Zu jedem $\theta_i = \theta_{i-1}$ + **`incr`** berechne zugehörigen Punkt (x_i, y_i),
zeichne Linie von (x_{i-1}, y_{i-1}) nach (x_i, y_i).
Dazu: Merken des vorangegangenen Punktes.

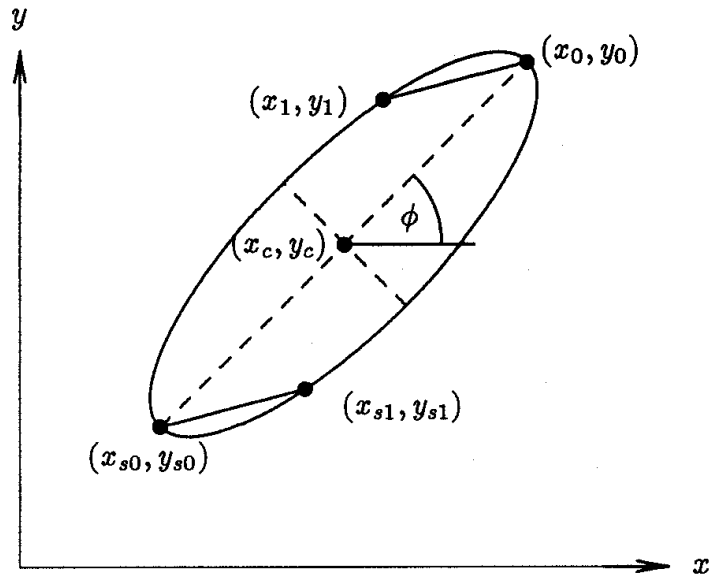
Implementation:

```
void ellipse(float a, float b, float phi,
          float xc, float yc, int n, int value)
{
float cosphi, sinphi, theta, pi, costheta,
   sintheta, A, B, C, D, x0, x1, y0, y1, incr;
int i;
pi = 4*arctan(1);
cosphi = sintab[pi/2 + phi];
sinphi = sintab[phi];
A = a*cosphi; B = b*sinphi;
C = a*sinphi; D = b*cosphi;
x0 = A + xc; y0 = C + yc;
theta = 0; incr = 2*pi/n;
for (j=1; j <= n; j++)
   {
   theta += incr;
   costheta = sintab[pi/2 + theta];
   sintheta = sintab[theta];
   x1 = A*costheta - B*sintheta + xc;
   y1 = C*costheta + D*sintheta + yc;
   Draw_line(x0, y0, x1, y1, value);
   x0 = x1; y0 = y1;
   }
}
```

Kosinusberechnung wird auf Sinus zurückgeführt.

Sinusberechnung wird hier durch Lookup-Table ersetzt.
Wegen sin x = sin(π–x) und sin x = –sin(x–π) braucht man nur die Werte zwischen 0 und $\pi/2$ zu speichern!

Verbesserung des Algorithmus (Halbierung der Laufzeit):
Ausnutzung der Symmetrie der Ellipse.



Nachteil der Parametermethode:
Feste Unterteilung der Parameterwerte $\Rightarrow$ Ungenauigkeit an Stellen starker Krümmung.

Abhilfe: nichtlineare Einteilung der Parameterwerte (**`incr`** abhängig von der Krümmung).

anderer Ansatz:
Polynom-Methode

Ausgangspunkt jetzt:
kartesische Achsenabschnittsform der Ellipsengleichung (Polynom 2. Grades).
Ungedrehte Ellipse:

$$(x/a)^2 + (y/b)^2 = 1$$

Vorgehen:
Auflösen nach y,
x durchläuft alle in Frage kommenden Werte in Einerschritten,
2 y-Werte werden jeweils berechnet.

Nachteil: hoher Rechenaufwand (Quadratwurzeln!).

alternativer Ansatz:
Scangeraden-Methode
(auch für andere Arten von Kurven)

Prinzip: horizontale Scangerade wird über die (ideale) Ellipse geschoben, für jede y-Position werden die Schnittpunkte bestimmt und im Raster approximiert.

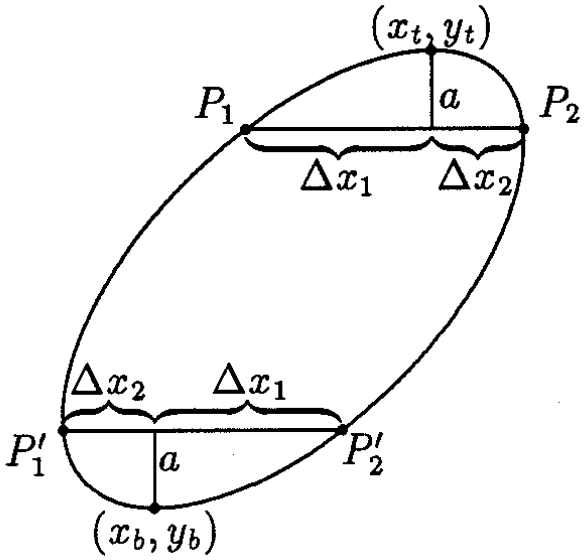
Rechnung geht wieder aus von der Achsenabschnittsform (s.o.)
Für gedrehte und verschobene Ellipsen: x und y entsprechend substituieren durch $x \cos \phi - y \sin \phi + x_c$ bzw. $x \sin \phi + y \cos \phi + y_c$.
Scangeraden-Gleichung: $y = y_i$ (= konst.)

Schnittpunkt-Berechnung: Einsetzen von $y = y_i$ in die Ellipsengleichung und Auflösen nach x (quadratische Gleichung; genaue Rechnung s. Rauber 1993, S. 31 ff.). Ergebnis: zwei x-Werte für die Schnittpunkte.

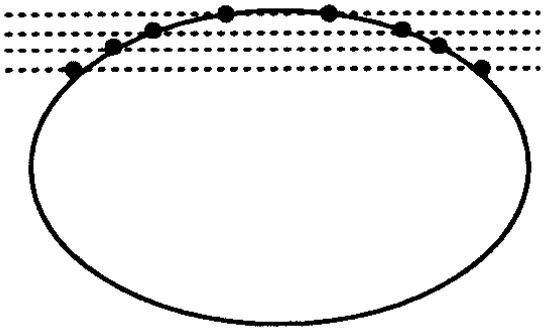
Berechnung nur nötig für solche Scangeraden, die die Ellipse schneiden:
Berechnung der oberen und unteren Ausdehnung y_t und y_b der Ellipse durch

$$y_{t,b} = \pm \sqrt{(b^2 \cos^2 \phi + a^2 \sin^2 \phi)}.$$

Ferner: Ausnutzung der Symmetrie:

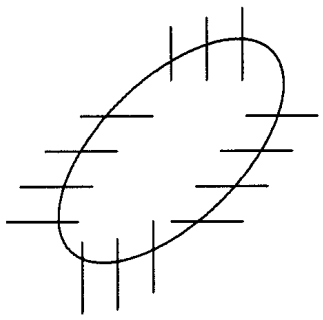


Nachteil der Methode:
In jeder Bildschirmzeile, die die Ellipse trifft (außer der oberen und unteren Tangente), werden genau 2 Pixel gesetzt
$\Rightarrow$ in Bereichen mit Steigung nahe 0 können "Lücken" auftreten



Abhilfe:

- überprüfe *x*-Abstand zum zuletzt gesetzten Pixel, wenn > 1: setze die dazwischenliegenden Pixel der Scangeraden
- oder: verwende vertikale Scangeraden in Bereichen mit Steigung zwischen –1 und 1:



Für letztere Vorgehensweise benötigt man die Parameterwerte mit Tangentensteigung +1 und –1:

bei $\theta_1 = \arctan(-b / (a \tan(\pi/4 - \phi)))$, $\theta_2 = \theta_1 + \pi$
ist die Steigung +1,
bei $\theta_3 = \arctan(-b / (a \tan(\pi/4 - \phi)))$, $\theta_4 = \theta_3 + \pi$ ist sie –1.
(Herleitung bei Rauber 1993.)

Vorteil der Scangeraden-Methode:
Direkt auch zum Füllen von Ellipsen geeignet (verbinde die Schnittpunkte jeweils durch horizontales Scan-geraden-Stück).

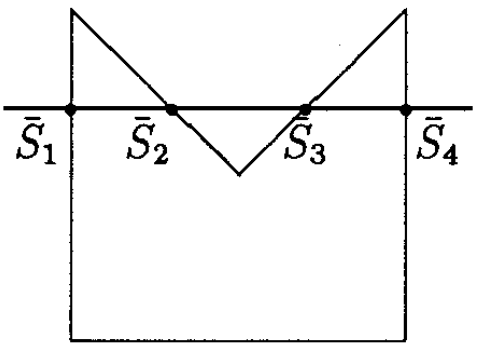
Kreise:
Spezialfälle der Ellipsen.
Für Kreise können (theoretisch) weitere Symmetrien ausgenutzt werden.
Beachte aber:
Für viele Monitore sind Höhe und Breite der Pixel ungleich
$\Rightarrow$ Kreise müssen als "echte" Ellipsen (mit $a \neq b$) konstruiert werden.

Füllen von Polygonen und geschlossenen Kurven

Die Scangeraden-Methode

Voraussetzung: geschlossenes Polygon *P* liegt in geometrischer Beschreibung vor.

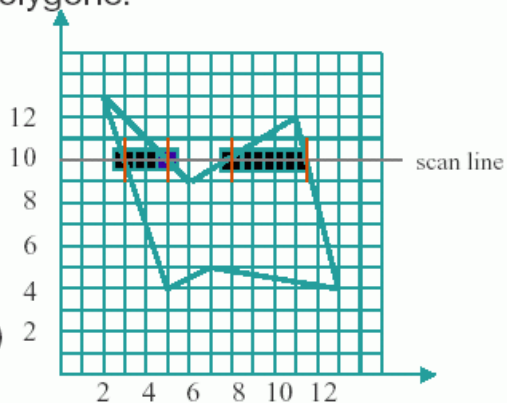
- Scangerade parallel zur *x*-Achse wird von unten nach oben über *P* geschoben,
- für jede Lage der Scangeraden werden die Schnittpunkte mit dem Polygon berechnet.
- Für konvexe Polygone gibt es genau 2 Schnittpunkte oder keinen,
- für konkave Polygone gibt es eine gerade Zahl von Schnittpunkten, die nach steigendem *x* geordnet werden.



1. **Finde Schnittpunkt der scan line mit allen Kanten des Polygons**
2. **Sortiere Schnittpunkte nach wachsender x-Koordinate**
3. **Fülle alle Pixel zwischen Paaren aufeinanderfolgender Schnittpunkte die im Inneren des Polygons liegen.**

(Dieser Algorithmus behandelt beliebige Polygone:
Regel von der ungeraden Parität:
Parität ist am Anfang 0 und wird mit jedem Schnittpunkt um eins inkrementiert.
Pixel wird gesetzt falls Parität ungerade.)

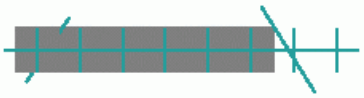
Anm.: Alle Polygoneckpunkte sind auf Integerwerte gerundet.
Bei allen konvexen Polygonen (Dreiecken) ensteht nur ein Span



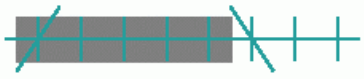
Es sind Sonderfälle zu beachten:

Scan line Algorithmus
Behandlung von Sonderfällen

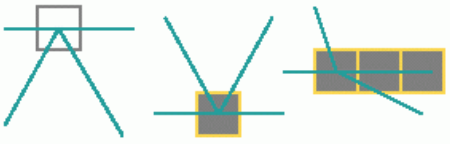
Schnitt Floatingpoint-x-Wert
Innen sein➔ abrunden
außen sein ➔ aufrunden



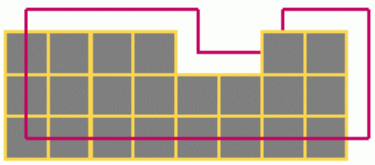
Schnitt Integer-x-Wert
linker Endpunkt eines Spans: Innen
Rechter Endpunkt eines Spans: Außen



gemeinsamer Eckpunkt (Vertex) von Nachbarpolygonen
y_{max}-Vertex einer Kante wird **nicht** zur Berechnung der Parität gezählt



horizontale Kanten
Pixel horizontaler Kanten werden zur Berechnung der Parität nicht gezählt und fallen für weitere Berechnungen heraus



(Krömker 2001)

Nach der Regel für den y_{max}-Knoten einer Kante werden in diesem Beisp. für die Scangerade y=0 die Kanten *a* und *b*, für die Scangerade y=2 die Kanten *c* und *d*, für die Scangerade y=4 aber keine Kante geschnitten:

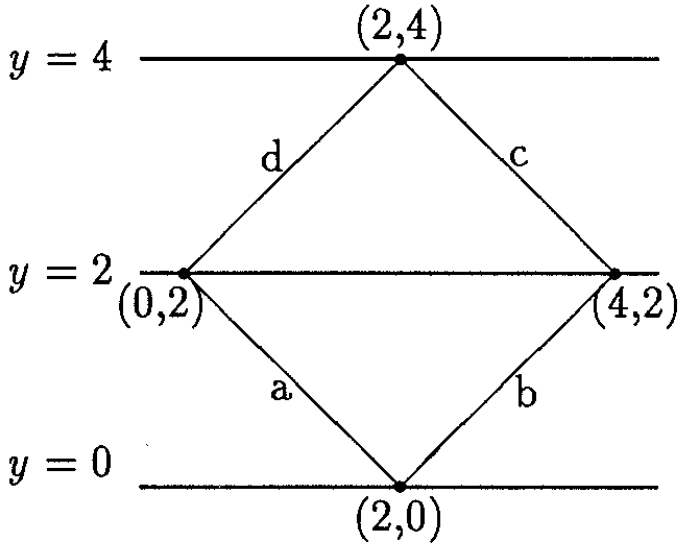
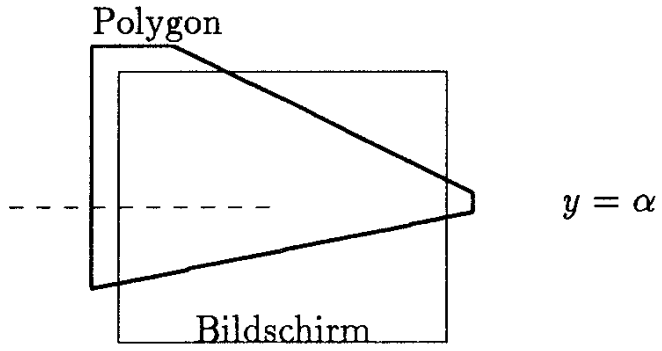


Abb. (*) (aus Rauber 1993)

Vorsicht beim Füllen von Polygonen, die sich über den Bildschirmrand erstrecken:



Wenn nur die sichtbaren Polygonkanten übergeben werden, kann z.B. für die Scangerade $y = \alpha$ nicht entschieden werden, ob die entspr. Bildschirmzeile innerhalb oder außerhalb des Polygons liegt.

Abhilfe: Beim Clipping (siehe später) müssen die Teile des Fensterrandes, die im Polygon liegen, als zusätzliche Kanten an den Scangeraden-Algorithmus übergeben werden.

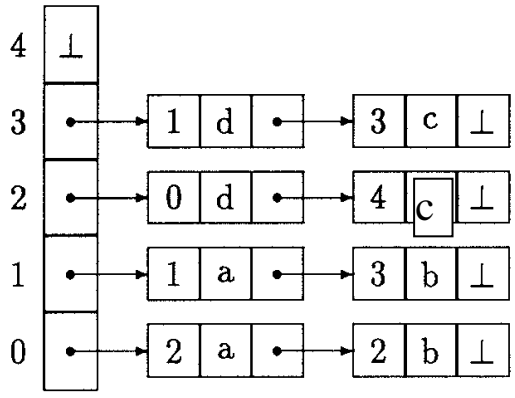
Genaueres Vorgehen beim Einfärben:

naiver Algorithmus:
äußere Schleife: Scangerade verschieben
innere Schleife: Schnittpunkte mit allen Polygonkanten berechnen

effizienter:
äußere Schleife über die Polygonkanten *e*
innere Schleife: berechne für e alle Schnittpunkte mit allen Scangeraden zwischen $y_{min}(e)$ und $y_{max}(e)$ und lege diese sortiert in einer (globalen) Datenstruktur ab.
Danach durchlaufe die Datenstruktur und färbe die entspr. Segmente für jede Scangerade.

Vorteil: Bei Berechnung der Schnittpunkte kann ein inkrementeller Algorithmus verwendet werden (Variante des Bresenham-Algorithmus).
In der Datenstruktur brauchen für jeden Schnittpunkt nur die *x*-Werte separat abgespeichert zu werden, der *y*-Wert gilt für die gesamte Scangerade.

Dieses Beispiel bezieht sich auf obige Abb. (*):



Der Übersichtlichkeit halber sind hier noch die Labels der Punkte mit angegeben worden (a,b,c,d) – diese braucht man in der Praxis nicht.

Diese Datenstruktur könnte in C etwa so realisiert werden:

```
struct scanline
    {
    struct xpixlist *list;
    struct scanline *next;
    int ypix;
    }
struct xpixlist
    {
    int xpix;
    struct xpixlist *next;
    }
```

Es gibt eine noch platzsparendere Variante, die einen *bucket sort*-Algorithmus zur Erstellung einer Kantenliste ausnutzt (s. Foley et al. 1990).

Für beliebige Kurven als Grenzen der Füllregion muss nur die Berechnung der Schnittpunkte geändert werden.

Die Saatfüll-Methode

wird benutzt, wo umgebende Polygone bereits auf dem Bildschirm dargestellt sind (interaktive Grafik)
keine geometrische Repräsentation des Polygons nötig!

stattdessen: Pixel der Grenzlinie müssen durch spezifische Grenzfarbe erkennbar sein.

Grundidee:
Setze ein Pixel im zu füllenden Bereich
setze dessen Nachbarpixel auf die Füllfarbe (sofern diese nicht die Grenzfarbe haben)
wiederhole dies, bis kein Pixel mehr neu gefärbt werden kann.

Beachte:
Die Nachbarschaft darf keine Diagonal-Pixel enthalten, da sonst vom Bresenham-Algorithmus gerasterte Geraden "undicht" sind:

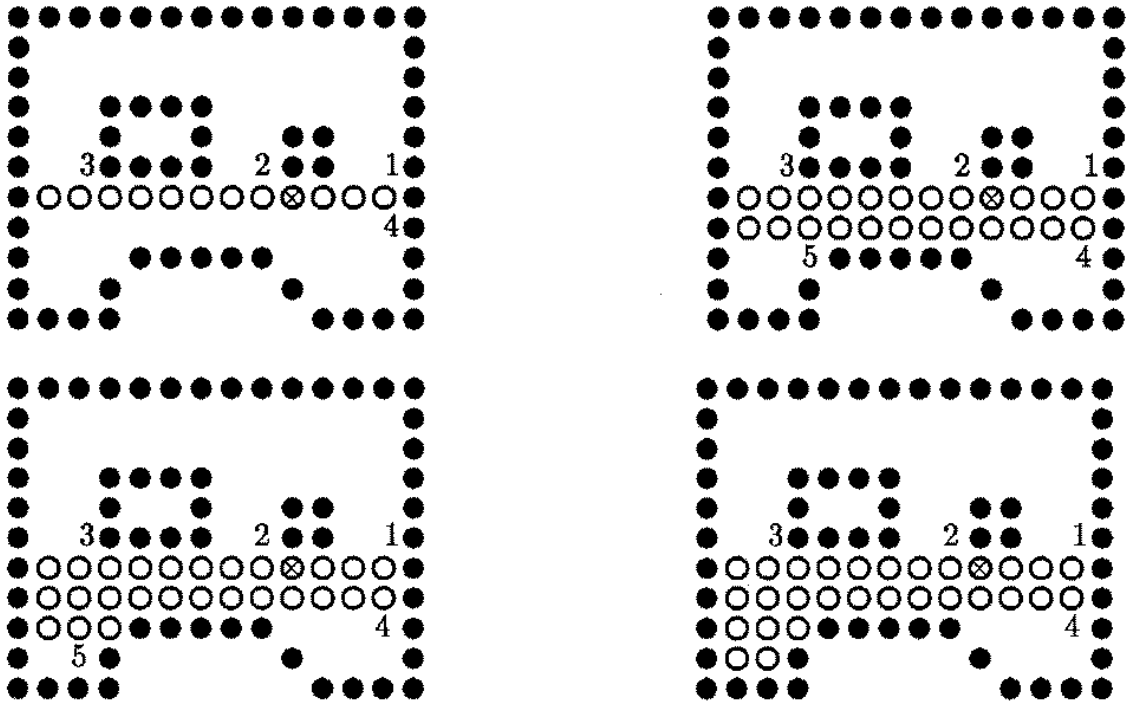
Implementierung:

naiver Ansatz: rekursiver Aufruf;
ungünstig wegen Beanspruchung des Rekursionsstacks

effizienter:
"Horizontalisierung" des Füllens; 2 Schleifen: innere erzeugt "Pixelläufe" in horiz. Richtung, äußere iteriert dies nach oben und unten.

Beachte: Auch hier ist (außer bei konvexen Polygonen) ein Stack erforderlich, der das Auftreten zusätzlicher Pixelläufe verwaltet.

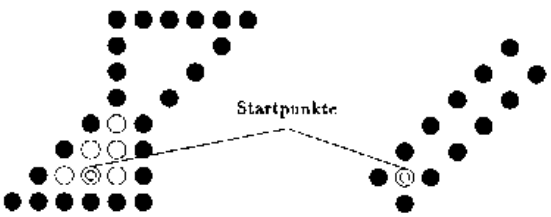
Beispiel:



(aus Rauber 1993)

Hier entstehen bei 1, 2, 3, 4 und 5 jeweils neue Pixelläufe. Jeder Pixellauf wird durch sein rechtestes Pixel repräsentiert. Der Algorithmus terminiert, wenn der Stack, der die Pixelläufe verwaltet, leer ist.

Nachteil:
Bei "engen" Polygonen kann der Saatfüll-Algorithmus vorzeitig stoppen. Abhilfe: Zweite Saat setzen.



(aus Fellner 1992)

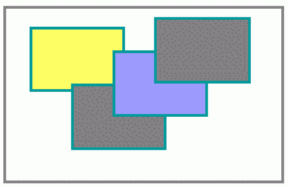
Ausschnittsbildung (Windowing und Klipping)

Um von der gesamten vorhandenen Bildinformation nur einen Ausschnitt darzustellen wird ein üblicherweise rechteckiges Fenster (Window) definiert, dessen Ränder den interessierenden Bildausschnitt begrenzen.

Um ein fehlerfreies Bild zu erhalten, muß die außerhalb des Fensters liegende Bildinformation vor der Bildausgabe abgeschnitten werden (Clipping = Klippen).

In der Regel wird das Fenster in Weltkoordinaten spezifiziert und mit "World-coordinate-window" oder einfach Window" bezeichnet.

Viewport



- Bildschirm in verschiedene Darstellungs-flächen (Viewports) aufteilen.
- Ohne Clipping würden die Inhalte der einzelnenViewports sich gegenseitig beeinflussen, also falsche Bilder erzeugen
- Auf dem Bildschirm wird als Koordinatensystem das Bildschirmkoordintensystem verwendet und das Ausgabefenster in diesem angegeben und Viewport genannt.
- Die Transformation zwischen dem Weltkoordinatensystem und dem Bildschirmkoordinatensystem heißt Window-Viewport-Transformation.

(Krömker 2001)

Clipping von Linien

häufige Aufgabe: Abschneiden von Geradensegmenten an einem Polygon (häufig an einem Rechteck). *Clipping*, *Clippen*.

einfachste Variante:
für jedes Pixel während der Bildschirm-Darstellung prüfen, ob es innerhalb des Clipping-Polygons liegt.

- sinnvoll, wenn Schnittpunktberechnung zu aufwändig
- oder wenn das zu clippende Objekt sehr klein ist (nur wenige Pixel zu prüfen).

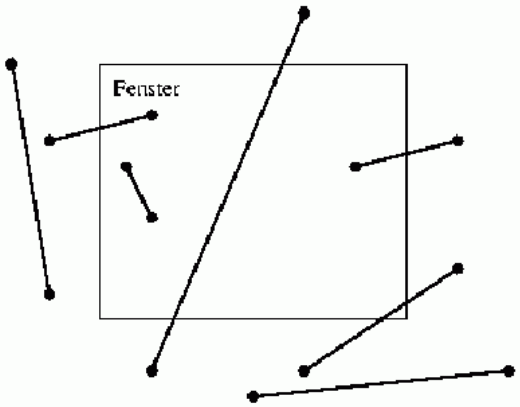
Analytische Methode:

Schnittpunkte berechnen
Fallunterscheidungen schon bei einfachsten Objekten erforderlich.

Clippen eines Geradensegments *s* an einem rechteckigen Fenster:

Clipping von Liniensegmenten

Beim Clipping von Liniensegmenten an einem rechteckigen Fenster (allgemeiner konvexen Objekt) entsteht entweder **kein** oder (bei Unterteilung einer Geraden in sichtbare und unsichtbare Teile) **ein** sichtbarer Teil. Vektoren, deren beide Endpunkte oberhalb, unterhalb, rechts oder links des Fensters liegen, sind völlig unsichtbar. Können diese Vektoren einfach aussortiert werden, so ist eine erhebliche Beschleunigung des Clipping zu erwarten. Der Cohen-Sutherland-Algorithmus (s.u.) nutzt diese Eigenschaft.



Direkte Berechnung, wenn zu clippende Linie und Fensterkante beide in Parameterform:

Endpunkte von s: P_1 und P_2
Parametergleichung der Geraden, auf der s liegt:
$g = P_1 + t(P_2 - P_1)$, t beliebig
Für s selbst: $t \in [0; 1]$.

Fenster-Eckpunkte: $A = (x_{min}, y_{min})$, $B = (x_{max}, y_{min})$,
$C = (x_{max}, y_{max})$, $D = (x_{min}, y_{max})$.
Gleichung der Geraden durch die Fenster-Eckpunkte A, B:
$f = A + u(B - A)$. (analog für BC, CD, DA.)
Schnittpunkt aus Gleichsetzen der Geradengleichungen:
$P_1 + t(P_2 - P_1) = A + u(B - A)$
(entspr. 2 Gleichungen mit 2 Unbek.) $\Rightarrow$ Lösungspaar t, u.

$0 \le t \le 1 \Leftrightarrow$ Schnittpunkt liegt auf dem gegebenen Segment
$0 \le u \le 1 \Leftrightarrow$ Schnittpunkt liegt auf der Fensterkante.
Fall A: beide Endpunkte von s liegen innerhalb des Fensters
$\Rightarrow$ s liegt ganz innerhalb des Fensters
Fall B: ein Endpkt. liegt innerhalb, einer außerhalb
$\Rightarrow$ (nur) ein Schnittpunkt zu bestimmen
Fall C: beide Entpkte liegen außerhalb des Fensters
$\Rightarrow$ s **kann** durch das Fenster verlaufen; beide Schnittpunkte zu bestimmen.

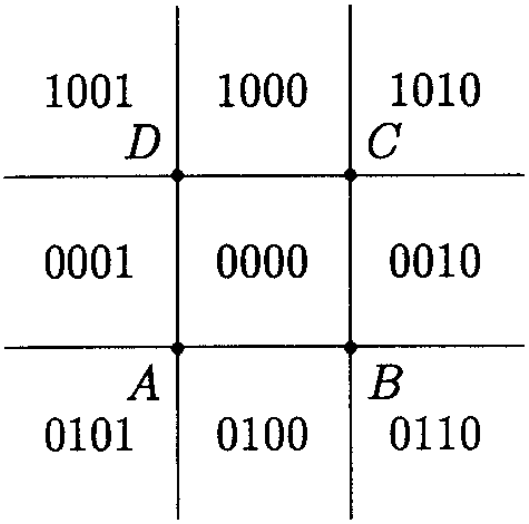
Ziel: Zahl der Schnittpunktberechnungen vermindern!

Beispiel: Wenn beide Endpunkte von s links des Fensters liegen, kann s nicht durch das Fenster verlaufen.
Ebenso: wenn beide rechts, beide oberhalb, beide unterhalb.

Systematisierung:

Algorithmus von Cohen und Sutherland

Kennzeichnung der 9 Bereiche in Bezug auf das Fenster durch 4 Bits. Bit 0 gesetzt für Bereiche links des Fensters, 1 für rechts, 2 für unten, 3 für oben: ("*outcodes*")

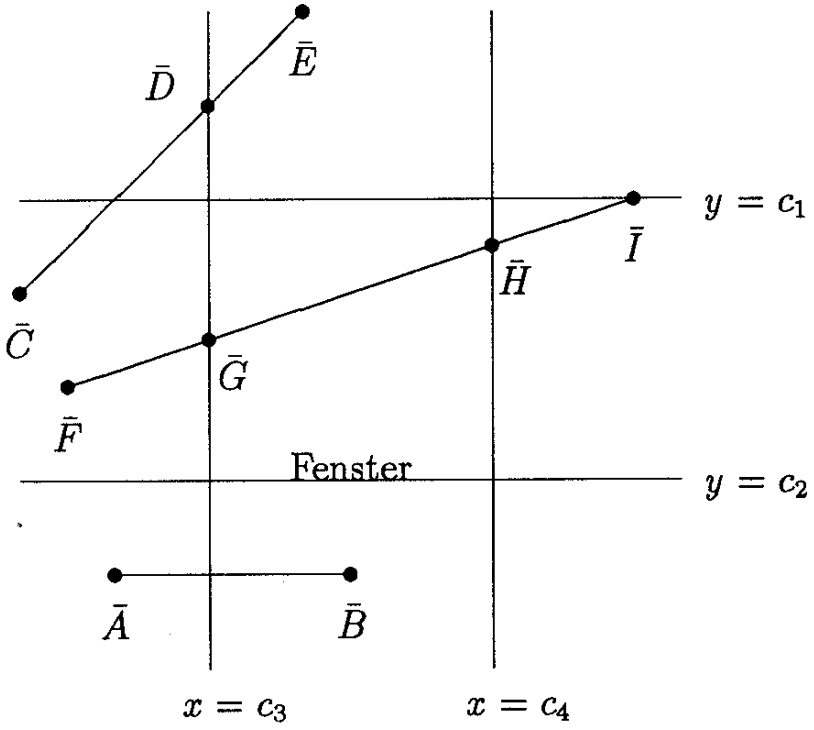


⇒ Bitmuster von horizontal oder vertikal benachbarten Bereichen unterscheiden sich in genau 1 Bitposition.

Ordne den Endpunkten von *s* ihre Bitcodes *c*1 und *c*2 zu.
*c*1 | *c*2 = 0 (bitweises Oder) ⇔ *c*1 und *c*2 sind beide 0000 ⇔ *s* liegt ganz im Fenster.

*c*1 & *c*2 ≠ 0 (bitweises Und) ⇔ *c*1 und *c*2 haben in mindestens einer Stelle gemeinsame Bits ⇔ *s* liegt auf einer Seite des Fensters, kann nicht durch das Fenster gehen!

in allen übrigen Fällen: Teile *s* mittels einer der Fenstergeraden in zwei Teile (hier Schnittpunktberechnung erforderlich!), wiederhole das Verfahren für beide Teile (einer muss ganz außen liegen!). Mit welcher Fenstergerade schneiden: entscheide mittels des Bitcodes z.B. von *c*1 (wenn *c*1 ≠ 0, sonst *c*2).
Wenn Bit 3 gesetzt: entspr. Punkt liegt ganz oberhalb des Fensters; wähle die obere Begrenzungsgerade. Analog für die anderen Bits.



(aus Rauber 1993)

3 Beispielfälle für die Anwendung des Algorithmus

Programmskizze:

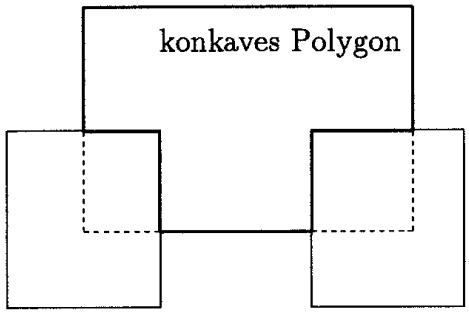
```
void CS_clip (float x1, float y1, float x2,
              float y2, float xmin, float xmax,
              float ymin, float ymax,
              int value)
{
int c1, c2; float xs, ys;
c1 = code(x1, y1);  c2 = code(x2, y2);
if (c1 | c2 == 0x0)
   Draw_Line(x1, y1, x2, y2, value);
else
   if (c1 & c2 != 0x0)
      return;
   else
      {
      intersect(xs, ys, x1, y1, x2, y2,
                xmin, xmax, ymin, ymax);
      if (is_outside(x1, y1))
         CS_clip(xs, ys, x2, y2, xmin, xmax,
                 ymin, ymax, value);
      else
         CS_clip(x1, y1, xs, ys, xmin, xmax,
                 ymin, ymax, value);
      }
}
```

`intersect` berechnet in den Rückgabeparametern **`xs`** und **`ys`** einen Schnittpunkt zwischen dem Geradensegment und einer der Fenstergeraden, die anhand der 4-Bit-Zahl der Endpunkte des Geradensegments ausgewählt wird.
`is_outside` berechnet, ob der Punkt (**`x1`**, **`y1`**) bzgl. derselben Fenstergeraden außerhalb des Fensters liegt.

(nach Rauber 1993)

Clipping bzgl. beliebiger (auch konkaver) Polygone:

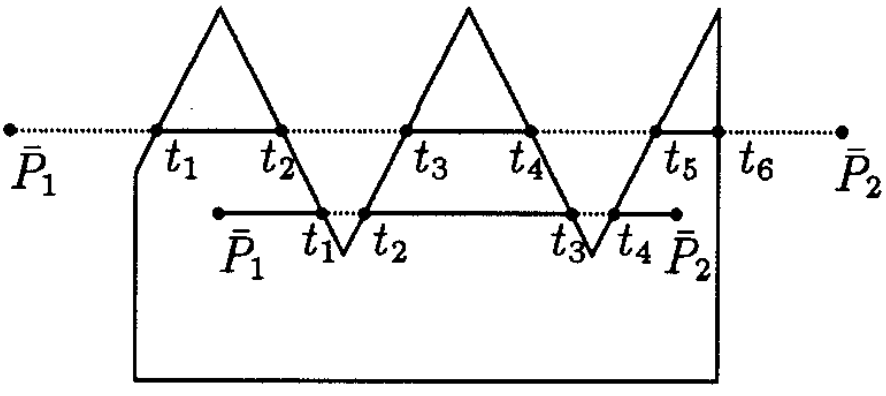
wichtig z.B. bei Systemen, die das Überlappen mehrerer Fenster auf einem Bildschirm erlauben!



Vorgehensweise:
Bestimme alle *n* Schnittpunkte des Geradensegments *s* mit dem Clipping-Polygon *P*.
Ordne diese nach aufsteigenden Parameterwerten.
Entscheide, welche von den *n*+1 Segmenten von *s* zwischen den Schnittpunkten im Inneren von *P* liegen ("sichtbar sind"):

Liegt der Endpunkt P_1 von *s* innerhalb von *P*?
Teststrahl von P_1 aus, zähle Anzahl der Schnittpunkte mit *P*, Paritätsprüfung (ungerade $\Rightarrow$ innerhalb).

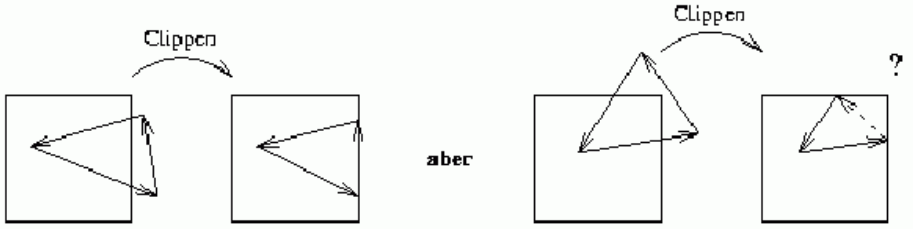
Wenn P_1 innerhalb von *P*: das bei P_1 beginnende und danach jedes zweite Parameterintervall liegen in *P*.
Sonst genau die hierzu komplementären.



(aus Rauber 1993).

Polygon-Clipping

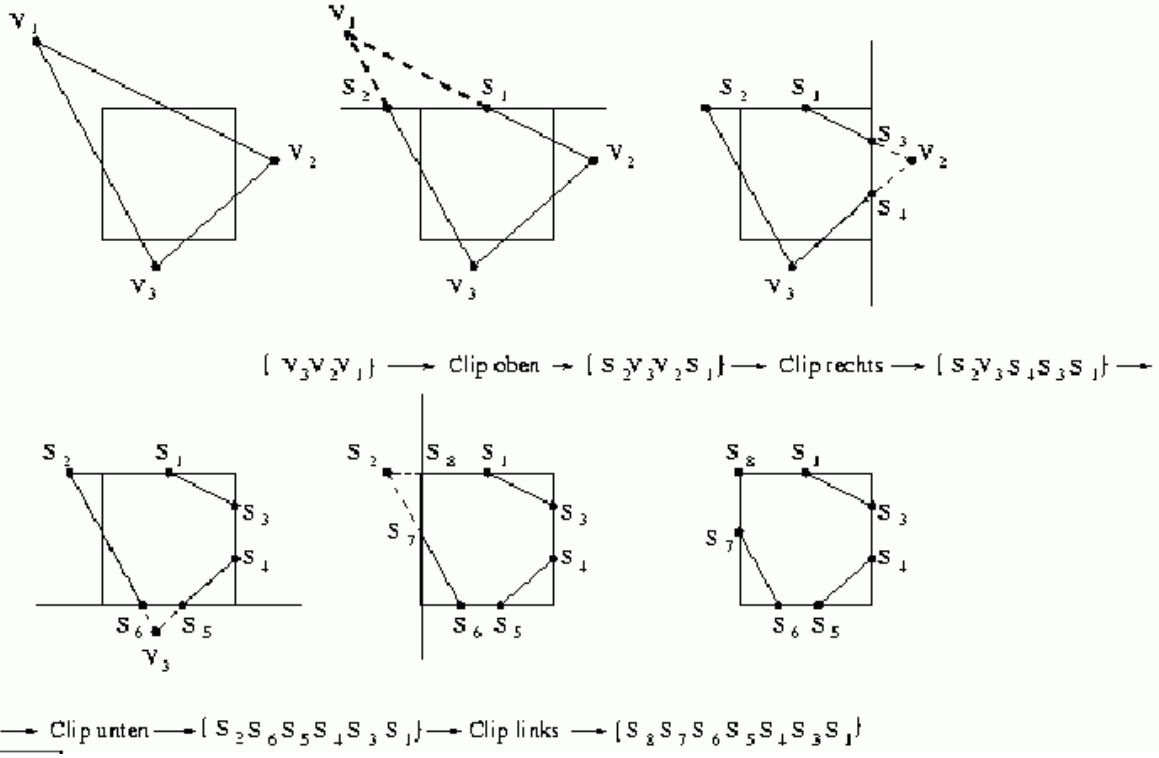
Ein Klipping-Algorithmus für geschlossene Polygone muß als Ergebnis des Klippingvorgangs wieder geschlossene Polygone liefern. Dies ist nur durch richtige Einbeziehung von Teilen der Fensterbegrenzung in das „geklippte" Polygon möglich.



Probleme entstehen, wenn das zu „klippende" Polygon Ecken des Fensters umschließt.

Verschiedene Methoden wurden vorgeschlagen, um diese Sonderfälle behandeln zu können. Die einfachste Methode ist die Umkehrung der Schleifenschachtelung: Das gesamte Polygon wird zunächst an einer Fen ergrenze geklippt, anschließend wird an der nächsten Fenstergrenze geklippt etc Dies ist die wesentliche Idee des Sutherland-Hodgman-Algorithmus.

Sutherland-Hodgeman-Algorithmus



(Krömker 2001)

5.2 Graustufendarstellung

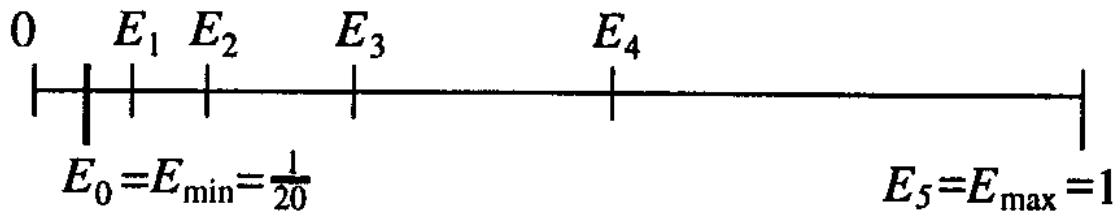
Diskrepanz zwischen (physikalischer) Luminanz und (wahrgenommener) Helligkeit von Graustufen:

Wahrnehmung reagiert in Relation zur empfangenen Lichtenergie in etwa logarithmisch

(+ weitere Einflüsse durch Monitor bzw. Art des Papiers)

⇒ *Transferfunktion* erforderlich, die aus theoret. Grauwert den adäquaten technischen Helligkeitswert berechnet

z.B. für 6 Graustufen E_0 bis E_5:



Die Transferfunktion kann an Eigenschaften der Hardware angepasst sein und tabellarisch (durch Stützstellen und Interpolation) spezifiziert sein.

Kleinste realisierbare Intensität E_{min} (i. allg. > 0): abhängig vom Ausgabemedium.

Bildschirme	0,005–0,025
Foto-Abzüge	0,01
Dias	0,001
Zeitungspapier	0,1

(nach Bungartz et al. 1996)

Sinnvolle Anzahl von Graustufen: $n < 1 - \frac{\ln(E_{\min})}{\ln(1{,}01)}$

Darstellung der Intensitätsstufen:

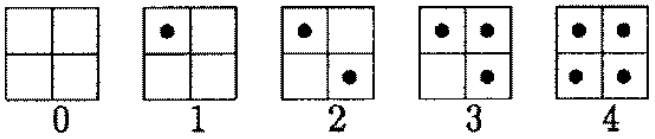
- auf Hardwareseite mehrere Intensitäten pro Pixel

oder

- *Halbtonverfahren*:

Einteilung des Bildschirms in Makropixel
Wahrnehmung fasst nahe zusammenliegende Pixel-Werte zu Graustufe zusammen

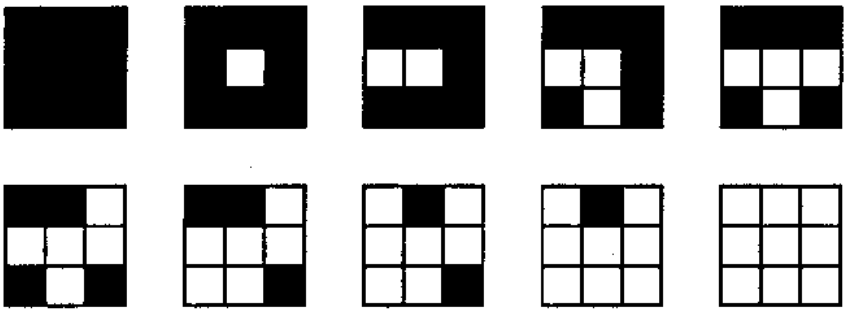
Beispiel: 2 x 2 - Halbtonmatrix



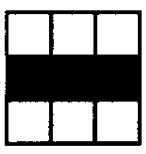
$$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$$

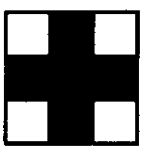
3 x 3:

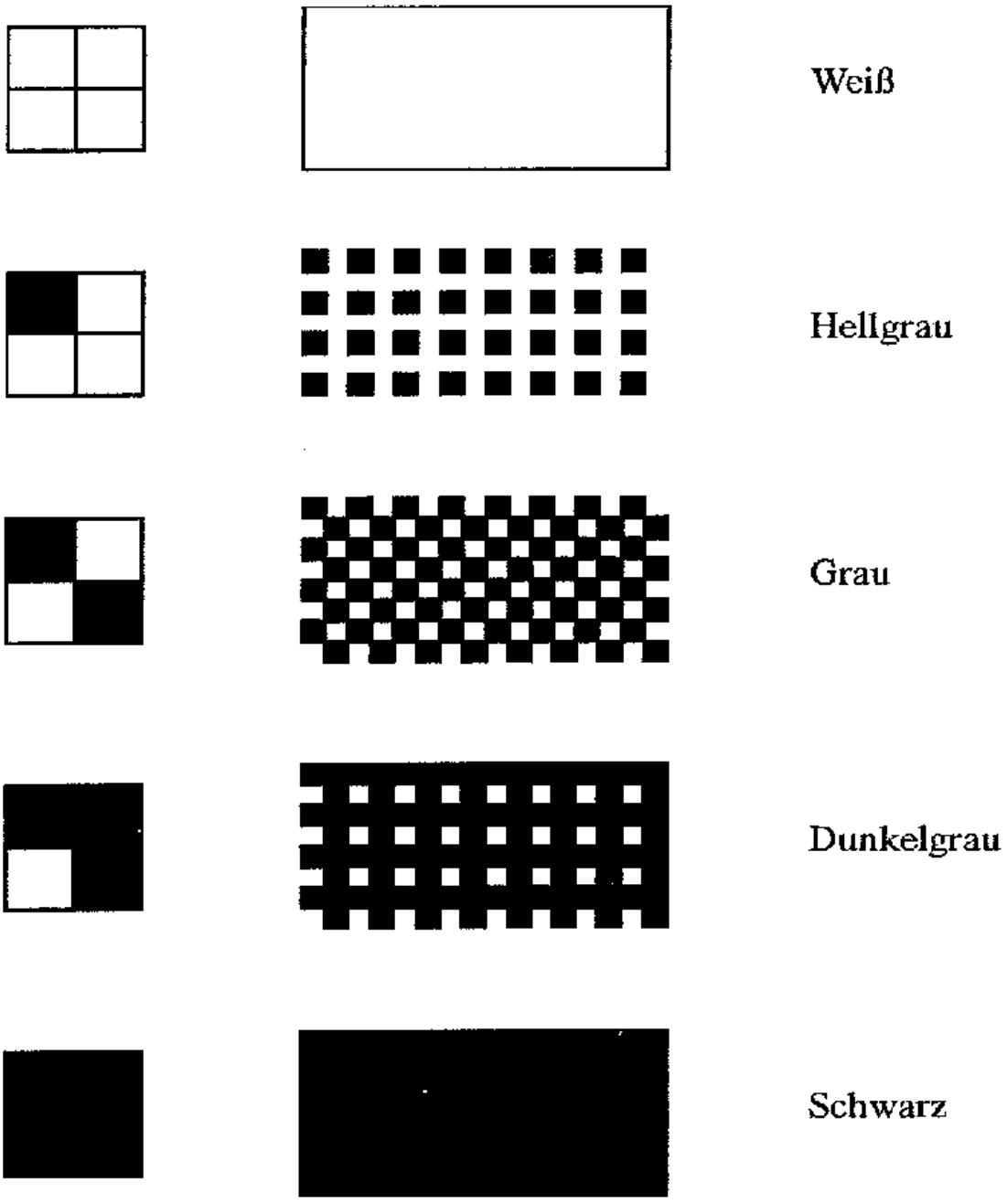
$$\begin{pmatrix} 6 & 8 & 4 \\ 1 & 0 & 3 \\ 5 & 2 & 7 \end{pmatrix} \triangleq$$



Wahl der Matrix ist nicht beliebig:
(a) Vermeidung optischer Artefakte (Linienmuster)
(b) hardwareabhängig: Vermeidung des Setzens benachbarter Pixel

a) 

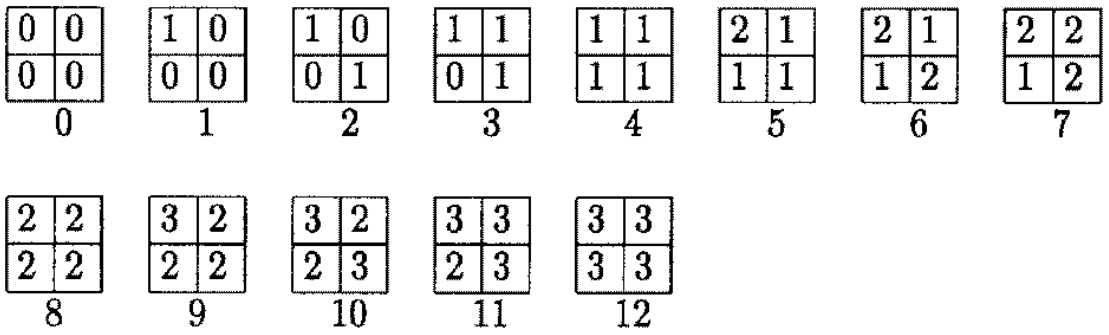
b) 



Vermeidung regelmäßiger Muster:

- Wechsel zwischen mehreren Halbtonmatrizen
- Zufallsauswahl der gesetzten Pixel

Verfahren der Halbtonsimulation auch für Ausgabegeräte mit mehr als 2 darstellbaren Intensitätsstufen anwendbar
z.B. bei 4 Intensitätsstufen 0–3:



Dither-Verfahren:

Intensitätswert wird nicht für Makro-Pixel berechnet, sondern für jedes einzelne Pixel. Darstellung in Abhängigkeit von der Position in einer Dither-Matrix (deren Kopien periodisch den Bildschirm überdecken).

```
for (y=0; y < rows; y++)
   {
   for (x=0; x < columns; x++)
      {
      i = x mod n; j = y mod n;
      if (intensity(x, y) > dither[i, j])
         set_pixel(x, y, 1);
      }
   }
```

Einträge in der Dither-Matrix = Schwellenwerte für die Bildschirmpixel

Wahl der Dither-Matrix:
Vermeidung artifizieller Muster in Bereichen konstanter Intensität

Beispiele:

2	6	4
5	0	1
8	3	7

0	8	2	10
12	4	14	6
3	11	1	9
15	7	13	5

21	10	17	14	23
15	2	6	4	9
20	5	0	1	18
12	8	3	7	13
24	18	19	11	22

(aus Rauber 1993)

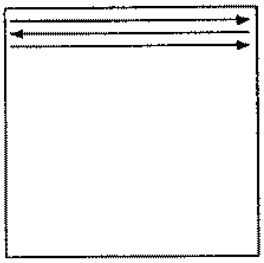
Vorteil gegenüber Halbton:
bei Vergrößerung von *n* erhöht sich Anzahl der darstellbaren Graustufen ohne merkliche Änderung der räumlichen Auflösung
(bis zu Schwellenwert für *n*: bei Auflösung 1024 x 1024 ca. $n = 10$ bis 16).

Fehlerverteilungsverfahren:

Grundidee: Bei Darstellung des Intensitätswerts in einem Pixel, für den nur diskrete Intensitätswerte möglich sind, wird bei Rundung auf den nächstgelegenen möglichen Wert ein Fehler gemacht. Man versucht, diesen Fehler in den Nachbarpixeln auszugleichen.

Verschiedene Varianten: unterscheiden sich in Reihenfolge der Pixel und Auswahl der Pixel für die Fehler-(ausgleichs-) Verteilung.

Floyd-Steinberg-Verfahren: 3 benachbarte Pixel, Zickzack-Abarbeitung.



Lauf nach rechts

(x,y)	$\frac{3}{8}$
$\frac{3}{8}$	$\frac{1}{4}$

Lauf nach links

$\frac{3}{8}$	(x,y)
$\frac{1}{4}$	$\frac{3}{8}$

Vorteil:

- Verf. liefert oft gute Ergebnisse (besser als Halbton u. Dither)

Nachteile:

- sequentiell
- in ungünstigen Fällen Akkumulation des Fehlers, "Geisterbilder"

Vermeidung dieser Nachteile:

Fehlerdiffusions-Verfahren

Kombination von Dithering und Fehlerverteilung

Bildschirm wird von Kopien einer *Diffusionsmatrix* überdeckt (analog zum Dither-Verfahren) $\Rightarrow$ Einteilung der Pixel in n^2 Klassen

Pixel-Intensitätswerte werden klassenweise berechnet
Fehler wird auf die benachbarten Pixel verteilt, die noch nicht berechnet sind

Wahl der Diffusionsmatrix:
möglichst wenige Einträge, die keinen oder nur einen Nachbarn mit größerem Eintrag haben (weil dort der Fehler gar nicht oder nur an 1 Nachbarn verteilt werden kann).

Beispiele für Diffusionsmatrizen:

34	48	40	32	29	15	23	31
42	58	56	53	21	5	7	10
50	62	61	45	13	1	2	18
38	46	54	37	25	17	9	26
28	14	22	30	35	49	41	33
20	4	6	11	43	59	57	52
12	0	3	19	51	63	60	44
24	16	8	27	39	47	55	36

25	21	13	39	47	57	53	45
48	32	29	43	55	63	61	56
40	30	35	51	59	62	60	52
36	14	22	26	46	54	58	44
16	6	10	18	38	42	50	24
8	0	2	7	15	31	34	20
4	1	3	11	23	33	28	12
17	9	5	19	27	49	41	37

Diffusionsverf. besonders bei Druckeranwendungen dem Dithering und dem Fehlerverteilungsverf. überlegen. Parallelisierbar, keine Geisterbilder.

Auswahl der *Farbschattierungen* (Helligkeitsstufen) für die Grundfarben R, G, B:

- logarithmische Wahrnehmungsskala ähnlich wie bei Graustufen
- zusätzlich Gamma-Korrektur des Bildschirms zu beachten (Nichtlinearität in Luminanzantwort des Displays)

5.3 Belegung der Farbtabelle (Color Lookup Table)

Anzahl darstellbarer Farben in einem digitalen Bild ist mitunter beschränkt

Ziel: möglichst gute Annäherung der Farben des darzustellenden Bildes durch die Farben in der Tabelle

naiver Ansatz:
uniforme Quantisierung

feste Zuordnung unabh. vom Bild, 3 Bit für Rotstufe, 3 Bit für Grünstufe, 2 Bit für Blaustufe

Standardbelegung der Tabelle bei wenigen darzustellenden Farben, z.B. bei grafischen Benutzungsoberflächen

verbessertes Verfahren:
Popularitätsalgorithmus

Belegung der Farbtabelle mit den 256 häufigsten Farbwerten des darzustellenden Bildes
Für im Bild auftretenden Farbwert wird der *nächstliegende* Farbwert der Tabelle verwendet

Farbabstände:

- euklidische Metrik im RGB-System

$$d(c_1, c_2) = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2}$$

zur Vermeidung der Wurzelberechnung häufig stattdessen:

- *Manhattan-Metrik*

$$d_m(c_1, c_2) = | R_1 - R_2 | + | G_1 - G_2 | + | B_1 - B_2 |$$

Einsparen der aufwändigen Minimumsuche für die Abstände: Merken bereits aufgetretener Farbwerte und ihrer zugeordneten Tabelleneinträge in weiterer Tabelle.

Nachteil des Popularitätsalgorithmus:
farbliche Details in kleinen Bildbereichen können völlig falsch dargestellt werden!

Vermeidung:

Median-Schnitt-Algorithmus

unterteilt den RGB-Einheitswürfel sukzessive in Teilquader (Schnitte parallel zu Koordinatenebenen)

1. bestimme kleinsten Teilquader des Würfels, der alle Pixel des Bildes enthält
2. teile diesen so, dass in beiden Hälften (etwa) gleichviele Pixel enthalten sind ("Median-Schnitt")
3. kontrahiere beide Teilquader auf ihre Extrem-koordinaten (vgl. Schritt 1)
4. wende hierauf wieder Schritt 2 an...
   bis 256 Quader erzeugt sind
   oder kein Quader mehr geteilt werden kann

- berechne dann für jeden Teilquader Mischfarbe durch gewichtete Mittelwertbildung der enthaltenen Pixel-Werte
diese Farbe wird in die Farbtabelle eingetragen

Bildschirm-Darstellung: jedes Pixel erhält Farbwert des Quaders, in dem es liegt

Datenstruktur: BSP-Baum (*binary space partitioning*)
Bestimmung eines konkreten Tabellen-Wertes zu einem Farbwert des Bildes durch top-down-Lauf durch den BSP-Baum.

Vorteil: sehr gute Farbdarstellung
- noch weiter verbesserbar durch nachträgl. Anwendung des Floyd-Steinberg-Fehlerverteilungsverfahrens

Nachteil: hoher Speicherplatzbedarf, hohe Laufzeit
(alle Pixel werden im BSP-Baum gespeichert)

"Kompromisslösung":
Octree-Quantisierung

regelmäßige rekursive Unterteilung des RGB-Würfels in 8 gleichgroße Unter-Würfel

- Aufbau des Octrees, bis jeder Teilwürfel nur noch 1 Pixel enthält
- Reduktion von den Blättern her (Mittelwertbildung für je 8 Tochter-Würfel), bis der Octree genau 256 Blattknoten hat

(Aufbau und Reduktion können speicherplatzsparend auch verschränkt werden)

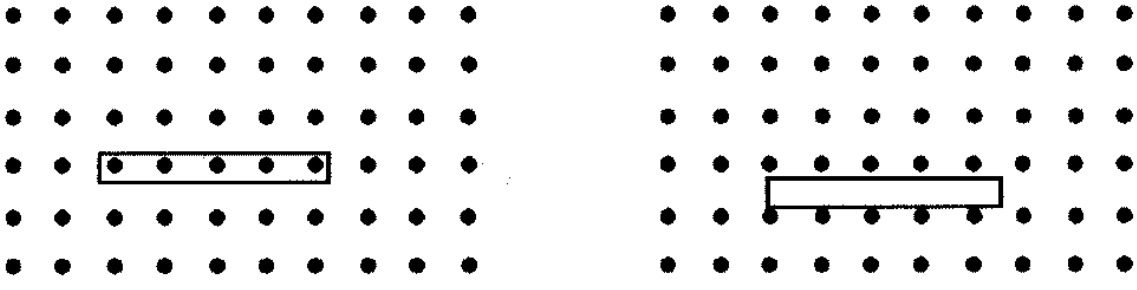
5.4 Antialiasing

Aliasing (Aliasierung):
Sammelbez. für Verfremdungseffekte, die durch die Rasterkonvertierung stetiger Objekte entstehen

- Lattenzaun-Problem: Objekt mit regelmäßigem, periodischem Aufbau ist nicht verträglich mit der Rasterung

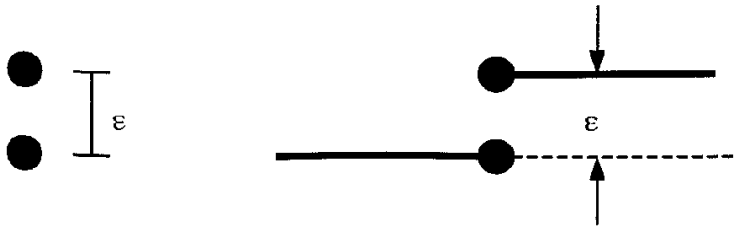
⇒ Teile verschwinden oder werden unregelmäßig dargestellt

- *Moiree-Effekte* bei Objekte mit regelmäßiger Textur, Schraffierung etc.: Vortäuschen großflächiger Muster
- dünne Objekte werden "verschluckt" oder verfremdet



- Treppenstufen-Effekte bei schrägen Linien und bei Kurven

Problem hierbei: Abstands-Auflösung des visuellen Systems ist bei Linien exakter als bei einzelnen Punkten
⇒ Treppenstufen sichtbar, auch wenn Pixel nicht unterscheidbar!



Antialiasing-Techniken:

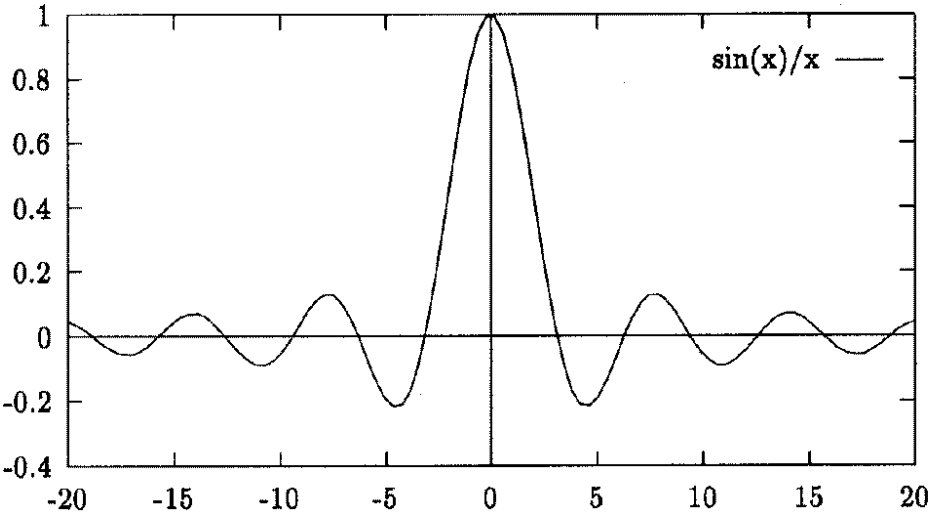
- *Prefiltering* (Vorfilter-Techniken): Vermeidung des Aliasing *vor* der Bildschirmdarstellung durch geeignete Bildmanipulation

- *Postfiltering* (Nachfiltern): Abschwächung des Aliasing durch Nachbehandlung des bereits gerasterten Bildes

- Modifikation der Zeichenalgorithmen

Prefiltering:
Abschneiden der höheren Frequenzen in der Fourierdarstellung (vgl. JPEG-Kompression)

rechnerisch durch Konvolution (Faltung) der Funktion mit einem Vielfachen der Funktion (sin *x*)/*x* (= der umgekehrten Fourier-Transformierten einer Rechteckfunktion):



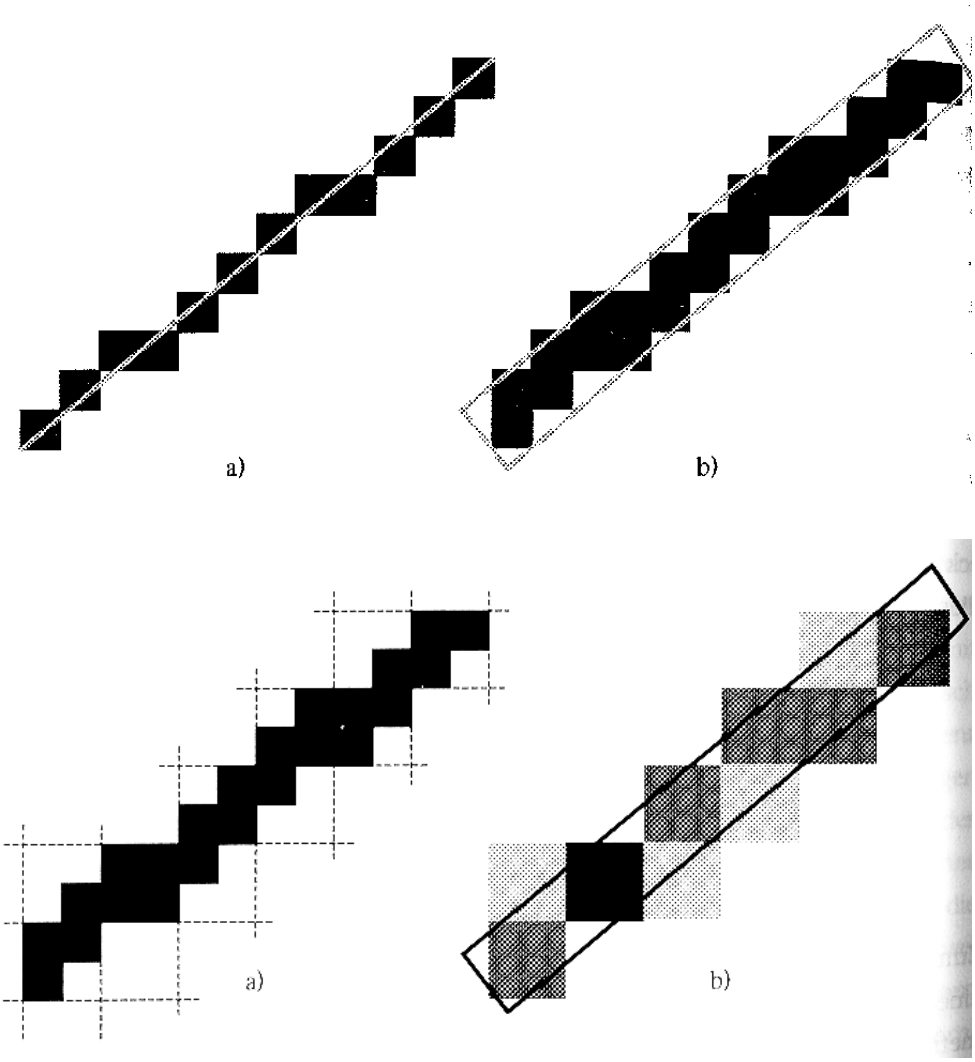
Abschneiden der Funktion erforderlich ⇒ nicht alle höheren Frequenzen verschwinden
dennoch sehr gute Qualitäten erreichbar
Nachteil: hoher Rechenaufwand
⇒ Verwendung anderer Filterfunktionen

oder: *heuristische Techniken*

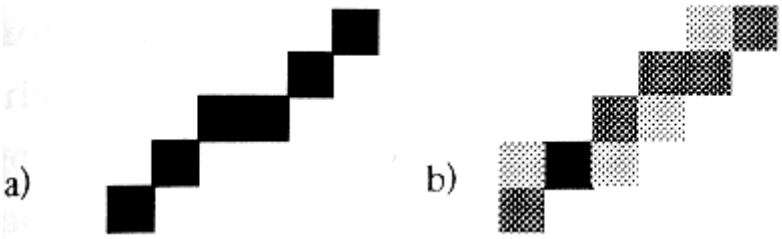
2 Techniken, die häufig für Linien angewandt werden:

1.

- Zeichnen in doppelter Auflösung
- anschließende Verdopplung
- anschließende Zuordnung eines Grauwertes je nach Bedeckungsgrad der 4x4-Makropixel (die den echten Pixeln entsprechen)



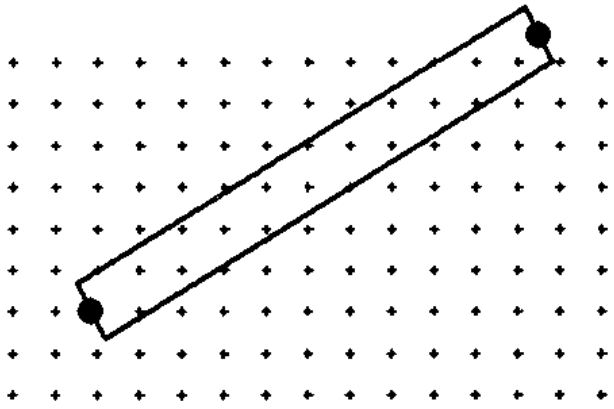
Ergebnis: "verschmierte" Linie, die in der Wahrnehmung aber "glatter" erscheint
Qualität: ca. entspr. doppelter Bildauflösung



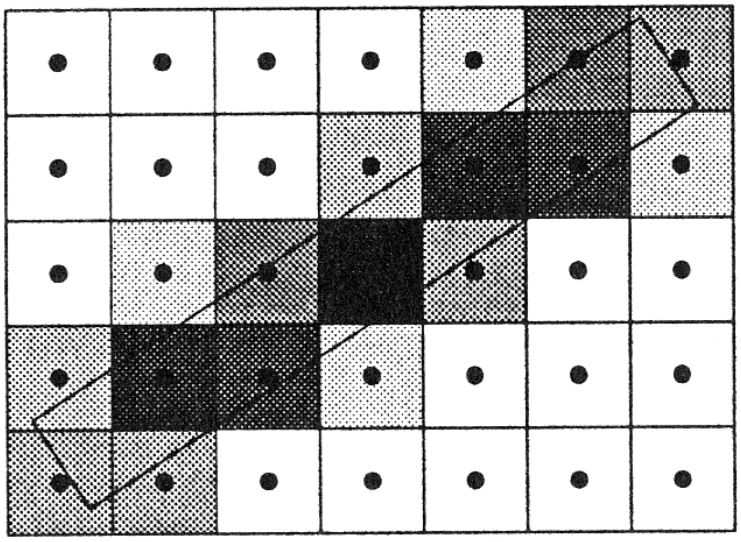
Gegenüberstellung der Rasterdarstellung einer Linie
(a) ohne Anti-Aliasing
(b) mit Anti-Aliasing
(c) mit doppelter Auflösung, ohne Anti-Aliasing:

2.
Area Sampling:

- Modellierung der Linie als rechteckiger Bereich



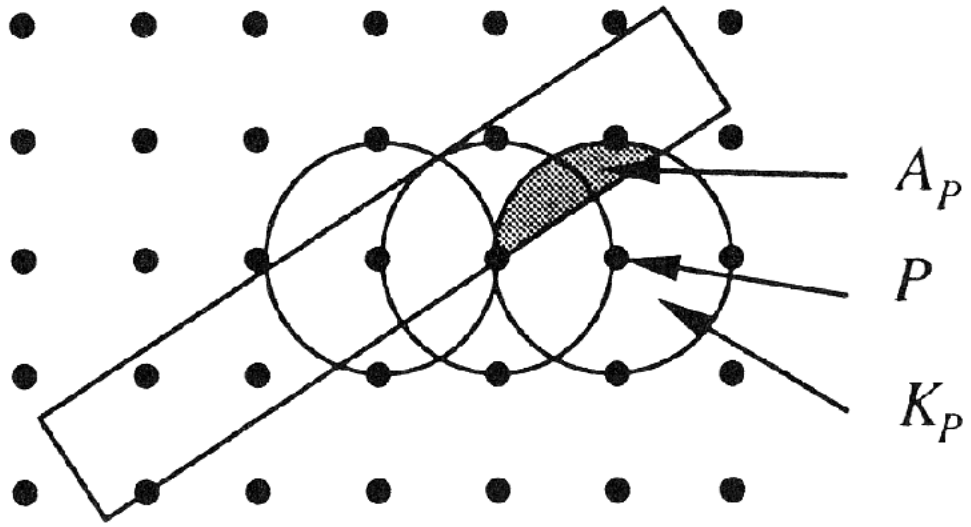
- Bestimmung der Graustufe eines Pixels nach dem Flächenanteil des Pixel(-Quadrats), den das Rechteck überstreicht (*unweighted area sampling*)



oder

- Berücksichtigung der Lage der überdeckten Fläche im Pixel bei der Festlegung der Graustufe:
  Weighted area sampling

den Pixeln werden überlappende Kreisscheiben zugeordnet:



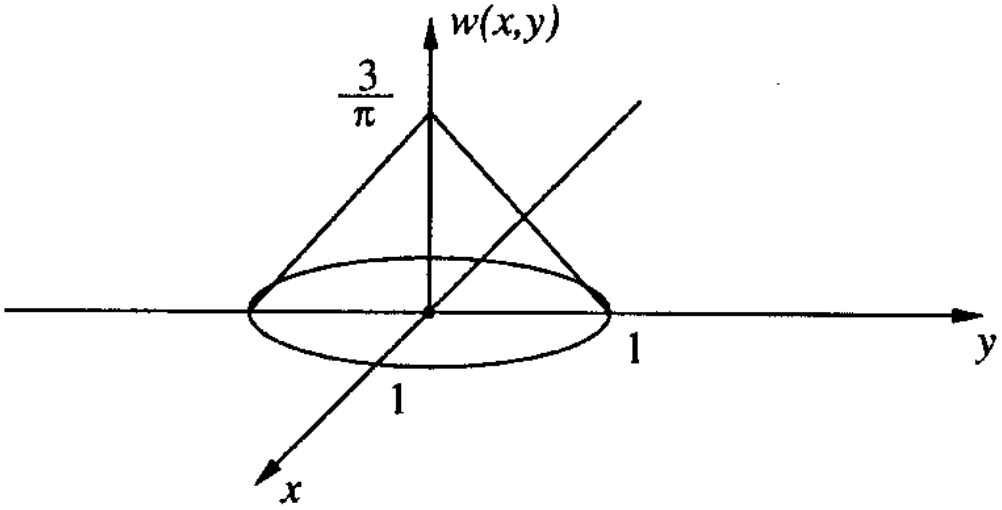
Pixel P, Kreisscheibe K_P, überdeckter Flächenteil A_P

Intensität des Pixels *P*:

$$I_P = \int_{A_P} w(x, y)dxdy$$

mit einer Gewichtsfunktion *w*, die $\int_{K_P} w(x, y)dxdy = 1$ erfüllt.

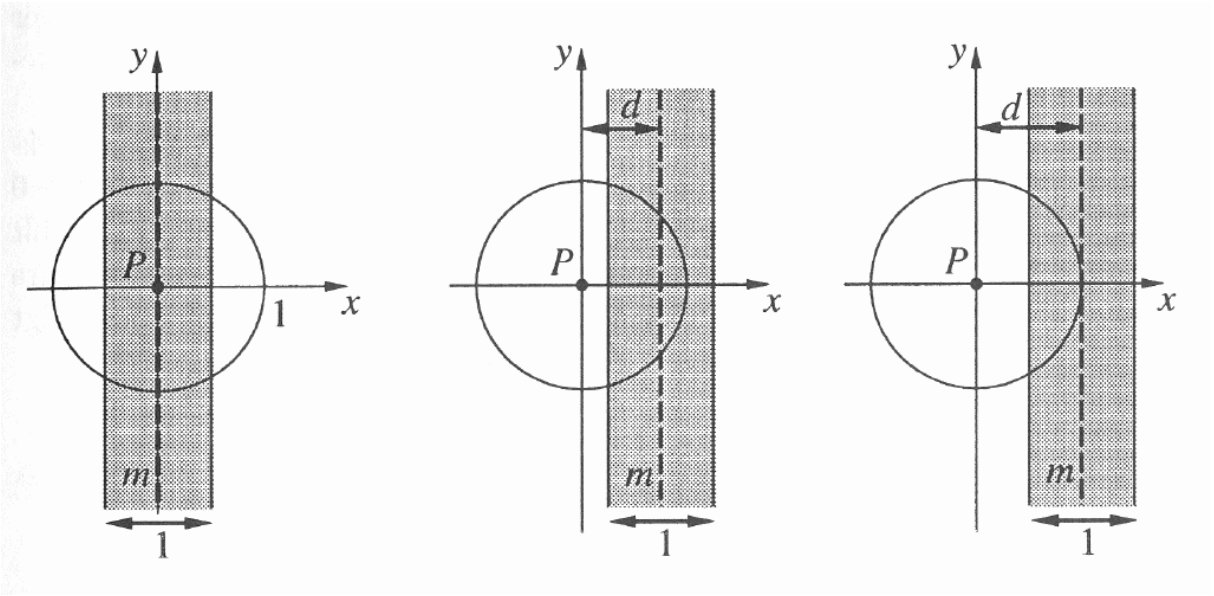
Beispielsweise *w* als Kegelfunktion:



Die Intensität hängt hier nur vom Abstand *d* der Pixelmitte zur Linien-Mittelachse ab ($0 \le d \le 1{,}5$).

Effiziente Implementation:

Algorithmus von Gupta und Sproull
Anstelle der Integralberechnung werden die Intensitätswerte in Abhängigkeit von *d* für 24 Werte von *d* ($d = i/16$, i=0;...;23) in einer Tabelle gespeichert. *d* kann beim Zeichnen inkrementell berechnet werden (vgl. Bresenham-Algorithmus).



Abhängigkeit der Intensität vom Abstand *d*.
Links: *d* = 0, Mitte: *d* = 0,75, rechts: *d* = 1,0.

Postfiltering-Techniken
(am häufigsten angewandte Antialiasing-Techniken)

Supersampling

Rasterung mit höherer Auflösung als der Bildschirm (Faktor 2 oder 4)
Intensitätswert eines Bildschirmpixels als Mittelwert der zugehörigen (4 oder 16) Pixel im errechneten Rasterbild

d.h.: Anwendung eines 2x2- oder 4x4-Filters auf das Rasterbild

$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

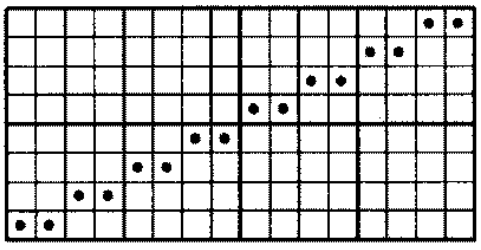
ungewichteter Filter

$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$\frac{2}{36}$	$\frac{4}{36}$	$\frac{4}{36}$	$\frac{2}{36}$
$\frac{2}{36}$	$\frac{4}{36}$	$\frac{4}{36}$	$\frac{2}{36}$
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Filter mit stärkerer Gewichtung des mittleren Pixelblocks

Nachteile:

- Schwäche des Supersamplings bei Bildern mit dünnen Linien



0	0	$\frac{4}{16}$	$\frac{4}{16}$
$\frac{4}{16}$	$\frac{4}{16}$	0	0

Ergebnis des Supersamplings einer Linie mit 4 x 4-Filter mit gleichmäßiger Gewichtsverteilung. Die resultierenden Bildschirmpixel zeigen nur 1/4 der max. Intensität und sind ebenso gestuft wie das Original.

- hoher Rechenzeit- und Speicherplatzaufwand

Abhilfe: *adaptives Supersampling*. Gröbere Rasterung in Bereichen konstanter Intensität.

Filter-Methode (ohne Supersampling)

Es wird mit der Auflösung des Bildschirms gearbeitet. Die endgültige Darstellung wird durch gewichtete Mittelung der Intensitäten benachbarter Pixel erzeugt: "Verschmieren" der Intensitäten der Pixel.

3 x 3 - Filter:

$\frac{1}{36}$	$\frac{1}{9}$	$\frac{1}{36}$
$\frac{1}{9}$	$\frac{4}{9}$	$\frac{1}{9}$
$\frac{1}{36}$	$\frac{1}{9}$	$\frac{1}{36}$

Für höhere Bildschirm-Auflösungen: 5 x 5 - Filter (hier: *Bartlett-Filter*):

$\frac{1}{81}$	$\frac{2}{81}$	$\frac{3}{81}$	$\frac{2}{81}$	$\frac{1}{81}$
$\frac{2}{81}$	$\frac{4}{81}$	$\frac{6}{81}$	$\frac{4}{81}$	$\frac{2}{81}$
$\frac{3}{81}$	$\frac{6}{81}$	$\frac{9}{81}$	$\frac{6}{81}$	$\frac{3}{81}$
$\frac{2}{81}$	$\frac{4}{81}$	$\frac{6}{81}$	$\frac{4}{81}$	$\frac{2}{81}$
$\frac{1}{81}$	$\frac{2}{81}$	$\frac{3}{81}$	$\frac{2}{81}$	$\frac{1}{81}$

Die Summe der Gewichtsfaktoren ergibt 1.
Der Filter wird "über die Pixel des Rasterbildes geschoben".

Vor der Anwendung: Replikation der Randzeilen und -spalten.