8.7.2023

Bézier Kurven

Lovis Rentsch
mathe@lovisrentsch.de

Inhaltsverzeichnis

1 Einleitung

Um Kurven im Internet darzustellen benutzt man Bézier Kurven. Denn sie erlauben es einem Rundungen darzustellen, die nicht als Funktion, etwa weil sie sich selbst schneiden, oder als Kreis dargestellt werden können. Außerdem brauchen Bézier Kurven nur sehr wenig Speicherplatz. Deutlich weniger, als eine vergleichbare Kurve als Rasterbild (PNG,JPG,…). Denn hier werden die Kurven nicht als viele Einzelteile beschrieben, wie es in pixelbasierten Formaten der Fall ist, sondern es werden nur einige wenige Eckpunkte gegeben, aus denen man für jeden Punkt die Stellung der Kurve, mit einer rechnerische sehr einfachen Funktion, bestimmen kann. Daher sind fast alle Formen im digitalen Raum Bézier Kurven. Außerdem stützen sie sich auf eine sehr faszinierende Mathematik.

2 Funktionsweise

Eine Bézier Kurve wird über sogenannte Kontrollpunkte definiert. Die Linie selbst wird dann immer aufs neue berechnet, um den gewünschten Ausschnitt zu zeigen. So ergibt sich die Skalierbarkeit, die es von den pixelbasierten Bildformaten unterscheidet. Auf dem Bildschirm werden die berechneten Kurven natürlich trotzdem als Pixel gezeichnet, aber es kann eben immer die maximale Auflösung berechnet werden.

Eine solche Kurve kann verschiedene Grade haben[1]. Eine lineare Bézier Kurve (ersten Grades) ist eine Linie, die die beiden Kontrollpunkte P_0 und P_1, den Anfang und den Endpunkt, verbindet. Wenn die Kurve nun gezeichnet, beziehungsweise berechnet werden soll, dann wendet man die Formel

$$P(t) = (1 - t)P_0 + tP_1$$

an. Wobei t ein Wert zwischen 0 und 1 ist. Und die Strecke zwischen den Beiden Punkten wird dann für t berechnet. Man kann also sagen, dass t der Prozentsatz, der Strecke $\overline{\mathrm{AB}}$ ist[2]. Die höheren Grade funktionieren genauso. Bei einer quadratischen Bézier Kurve hat man 3 Kontrollpunkte. Zwei davon sind die Anfangs und Endpunkte, die von der Kurve geschnitten werden (P_0 und P_2). Wenn man hier nun die Kurve zeichnen möchte „baut" man erst mal die linearen Kurven $\overline{P_0P_1}$ und $\overline{P_1P_2}$. Dann verbindet man die zwei entstehenden, von t abhängigen Punkte (a und b). Wenn man hier nun den gleichen Algorithmus anwendet bekommt man den finalen Punkt P, der (durch die Punkte a und b) von t abhängig ist (also $P(t)$). Man kann es also zusammenfassen, dass man so lange die ‚lerp'-Funktion anwendet, bis nur noch ein Punkt übrig ist. Wenn man die Bewegung dieses Punktes in der Spanne von 0 bis 1 der Variable t verfolgt, erhält man die Kurve[3]. Man nennt diesen Algorithmus nach seinem Erfinder den „DeCasteljau Alorithmus" [1]. Hier die Beispielrechnung für eine kubische Bézier Kurve.

interaktive Demonstration des lerps [2]

[1] ähnlich wie Polynome. Den Vergleich werde ich später noch ziehen

[2] dies wird häufig auch als 'lerp'-Funktion beschrieben

[3] Bézier auf Desmos

$$a = \text{lerp}(P_0, P_1, t)$$
$$b = \text{lerp}(P_1, P_2, t)$$
$$c = \text{lerp}(P_2, P_3, t)$$
$$d = \text{lerp}(a, b, t)$$
$$e = \text{lerp}(b, c, t)$$
$$P = \text{lerp}(d, e, t)$$

Es gibt aber auch noch andere Schreibweisen, die teilweise für Computer schneller zu berechnen sein. Zum Beispiel die Schreibweisen als Polynom.

$$\begin{gathered} P = P_0 + \\ t(-3P_0 + 3P_1) + \\ t^2(3P_0 - 6P_1 + 3P_2) + \\ t^3(-P_0 + 3P_1 - 3P_2 + P_2) \end{gathered}$$

2.1 Vektoren

Es gibt noch eine alternative Art sich die Ort-Zeit-Funktion vorzustellen. Man kann von den Kontrollpunkten. Dafür sollte man die „Bernstein“ Schreibweise betrachten:

$$\begin{gathered} P(t) = \\ P_0(-t^3 + 3t^2 - 3t + 1) + \\ P_1(3t^3 - 6t^2 + 3t) + \\ P_2(-3t^3 + 3t^2) + \\ P_3(t^3) \end{gathered}$$

Hier kann man sehen, dass die Punkte als unabhängige Funktionen betrachtet werden können, die in Abbildung 1 beschrieben werden. Man kann die Ergebnisse dieser Funktionen als Stärke des Einflusses, den der dazugehörige Kontrollpunkt an diesem Wert von t hat.

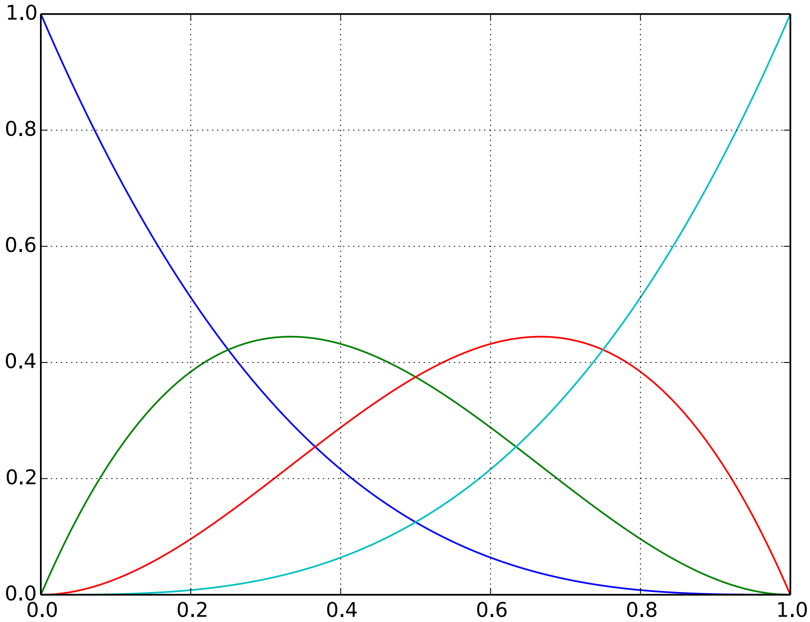

Abbildung 1: Einflüsse der Kontrollpunkte
Quelle,
interaktive Darstellung [2]

Am Anfang hat P_0 den Wert 1, denn die Kurve fängt in diesem Punkt an (der Punkt hat also alle Anziehungskräfte in sich vereint). Am Ende ist dies bei P_{Ende}[4] der Fall. Außerdem ist die Summe aller Werte immer 1, was auch für die Sichtweise der Anziehungskräfte, beziehungsweise des Einflusses auf die Kurve, spricht. Wenn man nun die Ortsvektoren der Punkte mit den Werten aus Abbildung 1 multipliziert und dann alle Ortsvektoren aller Punkte addiert bekommt man Punkt P[5] [3].

2.1.1 rationale Bézier Kurven

Bei einer normalen Bézier Kurve stammen die Faktoren, mit denen man den Vektor multiplizieren muss, aus der Entfernung zu $P(t)$. Aber man kann die Werte auch modifizieren (also die Graphen in Abbildung 1 anpassen). So wird einem noch mehr Freiheit bei der Gestaltung der Kurve gelassen ,ohne den Grad der Kurve zu erhöhen, was den Rechenaufwand wieder steigert[6] [4] [2] *(interaktive Darstellung)*.

2.2 splines

Als splines bezeichnet man eine Aneinanderreihung von mehreren Bézier Kurven. Sie mehrere Vorteile [4]. Eine Bézier Kurve hohen Grades ist schwer zu modifizieren. Das liegt daran, dass alle Punkte für jeden Wert, den t annehmen kann, einen Einfluss auf P haben. Darin liegt auch der Grund, warum es immer schwieriger wird Kurven hohen Grades zu berechnen. Man fasst dies unter dem Ausdruck, dass ein Punkt keine „lokale Kontrolle“ hat zusammen. Das ist auch an Abbildung 1 erkennbar, denn die Graphen haben alle immer einen Wert, der größer ist als 0 (außer bei $t = 0$ oder $t = 1$). Außerdem geht die Kurve auch nicht durch diese Punkte.

[4]Hier habe ich P mit 'Ende', und nicht mit einem Wert, gekennzeichnet, denn eine Bézier Kurve kann n-Grade haben

[5]ich gehe hierbei immer vom globalen Ursprung als Referenzpunkt aus

[6]mehr dazu unter 'splines'

Die schneidet nur den Anfangs- und den Endpunkt. Für die meisten Anwendungszwecke ist dies Unpraktisch, da man nicht einen Punkt verschieben kann, ohne dass die ganze Kurve sich verändert und außerdem muss man immer die Abstraktionsleistung erbringen die Punkte zu verschieben, die von der resultierenden Kurve nicht geschnitten werden. All das umgeht man (teilweise) mit der Aneinanderreihung von Bézier Kurven. In diesen Ketten geht man genauso vor, als ob es unabhängige Kurven wäre, was den Arbeitsaufwand nur linear steigen lässt, und nicht exponentiell, wie bei einer Kurve hohen Grades. t wird dann mit

$$0 \leq t \leq \text{Anzahl aneinandergereiter Kurven}$$

beschrieben (Abbildung 2).

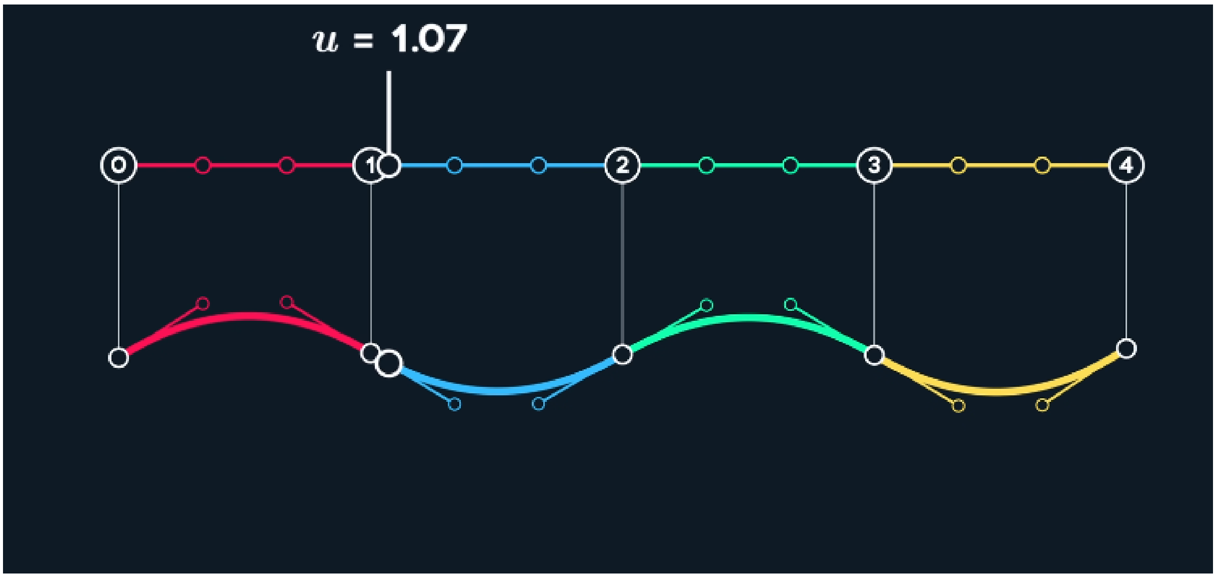

Abbildung 2: Kette aus kubischen Bézier Kurven (hier is $u = t$) [4]

Dafür werden meist kubische Kurven genutzt[7], denn sie geben einem die gewünschte Flexibilität bei der Gestaltung der Kurven, wobei gleichzeitig die Berechnung der Kurve noch einfach ist.

Erwähnenswert ist außerdem, dass es viele verschiedene Kurvenalgorithmen gibt, die alle verbunden werden können um ausgehend von Punkten Wege/Linien zu beschreiben. Bei den speziellen Anwendungen kann sich dann ein Algorithmus besser eignen, als ein anderer. Beispielsweise gibt es den „cardinal" Spline Algorithmus um Kurven zu beschrieben. Bei ihm hat man keine Kontrollpunkte, sondern man extrahiert aus der Position der Punkte zueinander die Ableitung an diesem Punkt. Dafür zeichnet man den Vektor von einem Punkt P_0 zu einem Punkt P_2 und legt diesen Vektor dann mit einem Faktor an P_1 als „Geschwindigkeit"[8] an. So erhält man eine Kurve, die durch alle Punkte geht, in der jeder Punkt lokale Kontrolle hat und die Kurve trotzdem einen weichen Verlauf hat[9]. Wenn die „cardinal" Spline einen Skalierungsfaktor von $\frac{1}{2}$ hat, nennt man sie eine „Catmull-Rom" Spline.

2.2.1 Kontinuität

Die Kontinuität ist ein Maß der Glätte der Kurve. Daher ist es sinnvoll die Ableitungen zu betrachten. Dabei geht es darum wie verbunden die Enden der Kurven einer Kurvenkette sind. Und das eben bei allen Ableitungen. Wenn die Kette nicht vollständig verbunden ist, so ist sie nicht kontinuierlich. Das hat zur Folge, dass auch keine „höhere" Kontinuität festgestellt

[7]also Kurven 3. Grades
[8]also als erste Ableitung
[9]mehr dazu unter Kontinuität

werden kann. C^0 wird also erreicht, wenn 2 Kurven mindestens einen gemeinsamen Punkt haben. Eine Kurve ist C^1 kontinuierlich, wenn die Ableitungen am Anfang und Ende der Kurve gleich sind ($A'(1) = B'(0)$). Da die Tangenten, die Strecken von den Punkten P_0 und P_3[10] und ihren nächsten Kontrollpunkten. Sie können als Maß der Änderungsrate betrachtet werden und müssen daher, um C^1 zu gewährleisten Gespiegelt sein. Das hat zur Folge, dass man mit diesem höheren Grad der Kontinuität auch etwas Kontrolle über das Verhalten der Kurve verliert, denn nun ist P_2 von der Position von P_4[11] abhängig, da die Tangenten um die Normale an P_3 achsensymetrisch sein müssen. Aber die Beschleunigung verändert sich immer noch instantan und nicht allmählich. Wenn das der Fall ist, dann ist die Kurve C^3 kontinuierlich. Jede Kurve, die C^3 kontinuierlich ist, ist immer C^∞ kontinuierlich. Allerdings hat man über die erste Kurve der Kette keinerlei Kontrolle mehr, wie die Kurve sich verhält, da alle Punkte von den ersten 4 Punkten abhängig sind. Man kann diese Kurve also auch beschrieben, wenn man die erste Kurve nimmt und den t Wert größer als 1 macht. Allgemein gilt die Kontinuität C^n:

$$A^i(t_{\text{end}}) = B^i(t_{\text{start}})$$
$$i \in \{0, ..., n\}$$

Eine einzelne Bézier Kurve ist immer C^∞ kontinuierlich. Die Einteilung ist daher nur für Splines, also Kurvenketten sinnvoll.

2.2.2 geometrische Kontinuität

[4]

Vorher ging es immer um die weichen Verläufe der Ableitungen. Das ist wichtig bei bestimmten Anwendungen, wie zum Beispiel Animationen. Aber es gibt auch noch eine andere Kontinuität, die man betrachten sollte, die beim Grafikdesign zum Beispiel wichtig ist. Man bezeichnet sie mit G^n. Gleich zu Anfang, sollte man anmerken, dass die Einschränkungen für die G Kontinuität weniger streng sind, als die für die C Kontinuität. Jede Kurve, die C^x kontinuierlich ist, ist auch immer G^x kontinuierlich[12]. Splines sind G^1 kontinuierlich, wenn die Tangenten an den Knotenpunkten parallel sind. Das hat zur Folge, dass die Oberfläche zwar Glatt ist, aber wenn diese Oberfläche aus einem reflektierendem Material besteht und man Licht auf sie scheint, wird das zurückgeworfene Licht verzerrt. Das liegt daran, dass die Änderungsrate der Oberfläche nicht kontinuierlich verläuft.[13] Auch hier kann man den Grad der Kontinuität erhöhen, auch wenn man dabei mehr und mehr Kontrolle über die Eigenschaften der Kurve verliert. Um den Grad nun zu erhöhen muss man ein neuen Parameter betrachten. Dabei handelt es sich um die *Rundung*.

[10] in kubischen, den üblichsten, Bézier Kurven

[11] dem ersten Punkt der Nachbarkurve

[12] mit Ausnahme einiger Sonderfälle, wie zum Beispiel Kurven in denen die erste Ableitung an einem Punkt 0 ist.

[13] Diese Verzerrung ist auch bei Körpern zu sehen, bei denen ein Kreisausschnitt an eine gerade Fläche anschließt

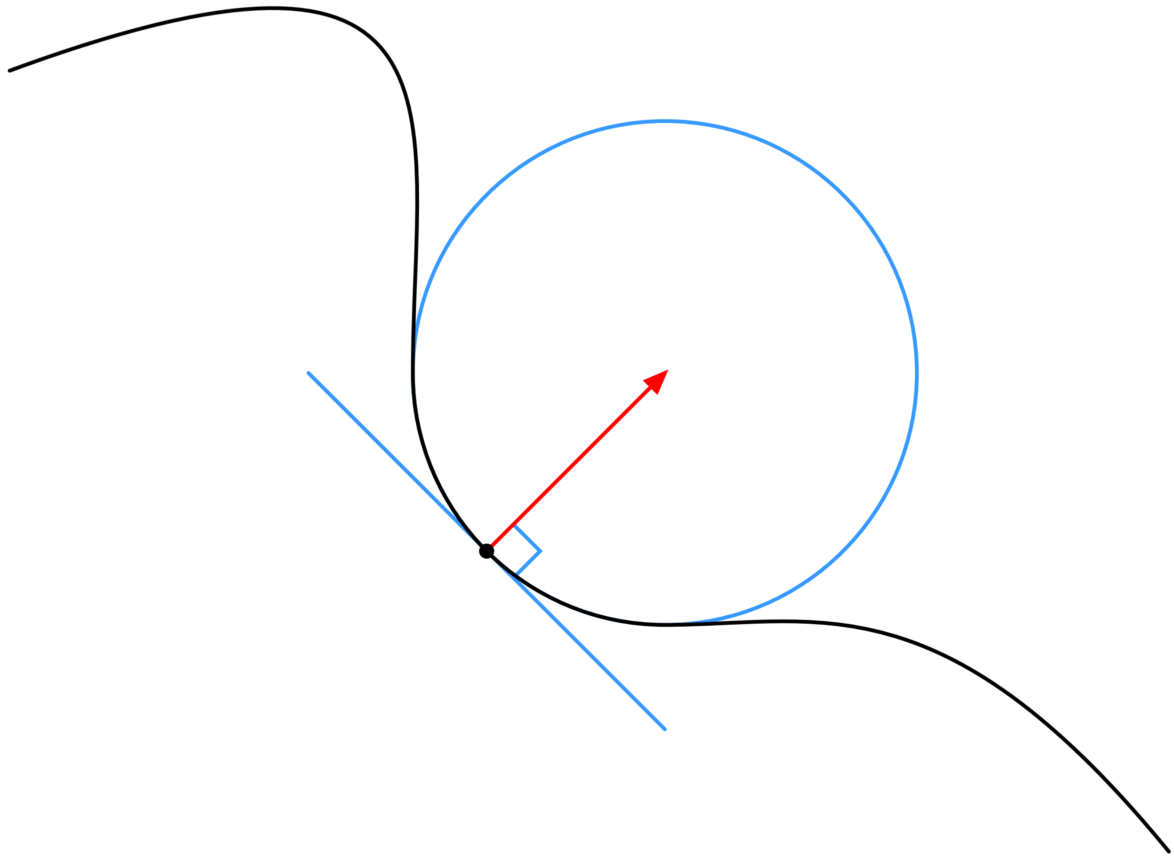

Abbildung 3: Die Krümmung im Punkt P
Quelle

Jeder Punkt auf einer beliebigen Kurve hat einen Kreis, der die Rundung an genau diesem Punkt beschreibt, wie in Abbildung 3 erkannbar ist. der Radius dieses Kreises ist

$$r = \frac{1}{k}$$

mit r als Rundungsfaktor. Mit diesem Wert kann man den sogenannten *Krümmungskamm zeichnen*. Dieser beschreibt den Radius des Krümmungskreises zu jedem Punkt. Wenn der Krümmungskamm eine geschlossene Linie bildet, dann ist die Kurvenkette G^2 kontinuierlich. Wenn die Ableitung des Krümmungskamms auch eine geschlossene Linie bildet, dann ist die Kurvenkette G^3 kontinuierlich.

2.3 Ableitungen

Bei Bézier Kurven ist die Ableitung immer auch eine Bézier Kurve, aber eine des nächst niedrigeren Grades. Die Ableitung ist nützlich, da sie immer tangential zu der zugehörigen Bézier Kurve ist. So kann man lokale Koordinaten bekommen, die immer um den gleichen Wert verschoben sind. In manchen Anwendungen ist das nützlich oder sogar zwangsweise erforderlich. Das ist immer der Fall, wenn man nicht nur einen Weg beschreiben möchte, sondern eine Linie, die eine bestimmte Dicke hat. Man kann also über die Ableitung die „Bewegungsrichtung“ und einen Normalvektor zu einem Beliebigen Punkt P der Kurve bekommen[14], über die man dann

[14] beziehungsweise für einen bestimmten Wert von t

die Stärke des Striches beschreiben kann, da man weiß wo „vorne“ und wo die „Seiten“ sind. Die Beschleunigung ist dann, wie oben bereits erwähnt nur in besonderen Fällen sinnvoll. [3]

2.4 Länge der Kurve

Man kann die Länge einer Bézier Kurve nicht berechnen. Man kann sie nur schätzen. Eine Möglichkeit das zu tun, ist die Kurve in kleine Teile zu zerteilen und dessen Länge zu bestimmen und die Längen zu addieren. Je mehr Teile wir haben, desto genauer ist die Schätzung, aber desto rechen aufwändiger ist es auch. Die summe wäre dann

$$\sum_{i=1}^{n} \left\| f\left(\frac{i}{n}\right) - f\left(\frac{i-1}{n}\right) \right\|$$

Um den zu einer bestimmten Entfernung von P_0 zugeordneten t-Wert zu bekommen, kann man

$$t = \frac{\text{Entferung}}{\text{Bogenänge}}$$

berechnen. Dann kann man das Berechnete t über die bekannten Funktionen einsetzen und kann so alle Parameter des Punktes einer bestimmten Entfernung vom Anfangspunkt bestimmen. Allerdings ist hier zu beachten, dass die t Werte nicht proportional zu der Entfernung des Punktes sind. Das liegt an den Ableitungen, die hier als Geschwindigkeiten zu den entsprechenden Werten stehen. Man muss also t als Zeit betrachten. Wenn man nun herausfinden möchte, wie weit (zeitlich gesehen) ein punkt von P_0 entfernt ist, muss man die Geschwindigkeit mit einbeziehen. t ist also nur dann zur Entfernung proportional, wenn der Graph der erstem Ableitung eine Kreisbahn ist. Eine Möglichkeit dies zu umgehen ist wieder die Linie in Abschnitte zu unterteilen, die alle gleich groß sind. Wenn man das berechnet, hat man eine Annäherung der Entfernung, nicht der Dauer des Erreichens dieses Punktes. [3]

interaktive Demonstration der Schätzung der Bogenlänge

3 Anwendungen

Wegen ihrer Vielseitigkeit finden Bézier Kurven in sehr vielen Bereichen Anwendung. Wie oben angesprochen basieren fast alle Formen auf Bildschirmen auf irgendeine Art und Weise auf Bézier Kurven. Andere Anwendungen sind zum Beispiel das designen von Plakaten oder Produkten. Hier ist es besonders wichtig, dass die Kurven nur von Punkten, und nicht von Bildgrößen abhängig sind, denn das macht sie vollkommen skalierbar. Mann muss nur alle Koordinaten mit dem gleichen Faktor multiplizieren und man bekommt die gleiche Kurve nur größer. Das hat auch etwas mit ihrem Ursprung in Vektoren, die ja auch skalierbar sind. In Arbeitsfeldern wie dem Produktdesign oder dem erstellen von vektorieller Kunst/Grafiken ist der G-Grad besonders wichtig, denn er gibt an, wie weich eine Kante/Oberfläche ist, und wie sie sich im Zusammenspiel mit Licht und Flüssigkeiten verhält. Wegen der Möglichkeit unendlich nah an das Bild „ran zu zoomen“ (es zu vergrößern / das Sichtfeld auf einen Ausschnitt zu begrenzen) ist es ideal für das Aufzeichnen wissenschaftlicher Daten, wie Messergebnisse in Graphen, bei denen es auch auf sehr kleine Schwankungen an kommt.

Ein weiterer Anwendungsbereich, den ich oben schon angesprochen habe sind Computersimulationen. Dazu zählen auch, aber nicht nur Computerspiele. Hier werden oft Wege definiert, die in ihrer Form flexibel oder auch vom Spieler einfach zu modifizieren sein müssen. Wege können auch zum Beispiel Kamerafahrten sein, bei denen es besonders wichtig ist, dass auf den C-Grad geachtet wird, denn sonst kann es zu störenden, abrupten Beschleunigungen kommen.

3.1 Kollisionsboxen

Ganz oft, vor allem in der Definition von Wegen oder Körpern in Spielen oder ähnlichen Simulationen , ist darauf zu achten, ob 2 Körper, die von Bézier Kurven definiert sind, kollidieren. Dabei nutzt man zuerst die vom Rechenaufwand günstigere Methode eine (möglichst kleine) Kollisionsbox um die Kurve zu zeichnen, bevor man überprüft, ob die Kurven kollidieren, denn wenn die Box nicht geschnitten wird, so kann die Kurve nicht geschnitten werden. Manchmal kann es reichen eine Box um die Kontrollpunkte zu zeichnen, da die Kurve immer zwischen den Punkten liegen muss. Aber häufig gibt es große Unterschiede zwischen der kleinen Box, die nur die Kurve beinhaltet, und der großen Box, die die Kontrollpunkte beinhaltet. Um die kleine Box zu bekommen, teilt man die „Bewegung" der Kurve in die X und Y-Bestandteile auf, und betrachtet diese unabhängig voneinander. Die lokalen Minima und Maxima dieser Funktionen sind die Schnittstellen der Kurve mit dem Rahmen der kleinen Box (Abbildung 4). Wenn eines der beiden lokalen Extrema nicht vorhanden ist, dann ist P_0 oder P_3 der Schnittpunkt mit der Box. Und auch wenn es ein Extremum gibt, so ist nicht unbedingt gegeben, dass nicht doch die Kontrollpunkte Teil der Box sind. Aber man sollte alle Punkte (der Startpunkt, der Endpunkt und die lokalen Extrema) betrachten. Das ist vom Rechenaufwand deutlich einfacher, als die genauen Schnittpunkte mit der Kurve zu berechnen. Auch hier ist die Ableitung wieder nützlich, denn die Extrema sind (wie immer) die Schnittpunkte der Ableitung mit den Achsen.

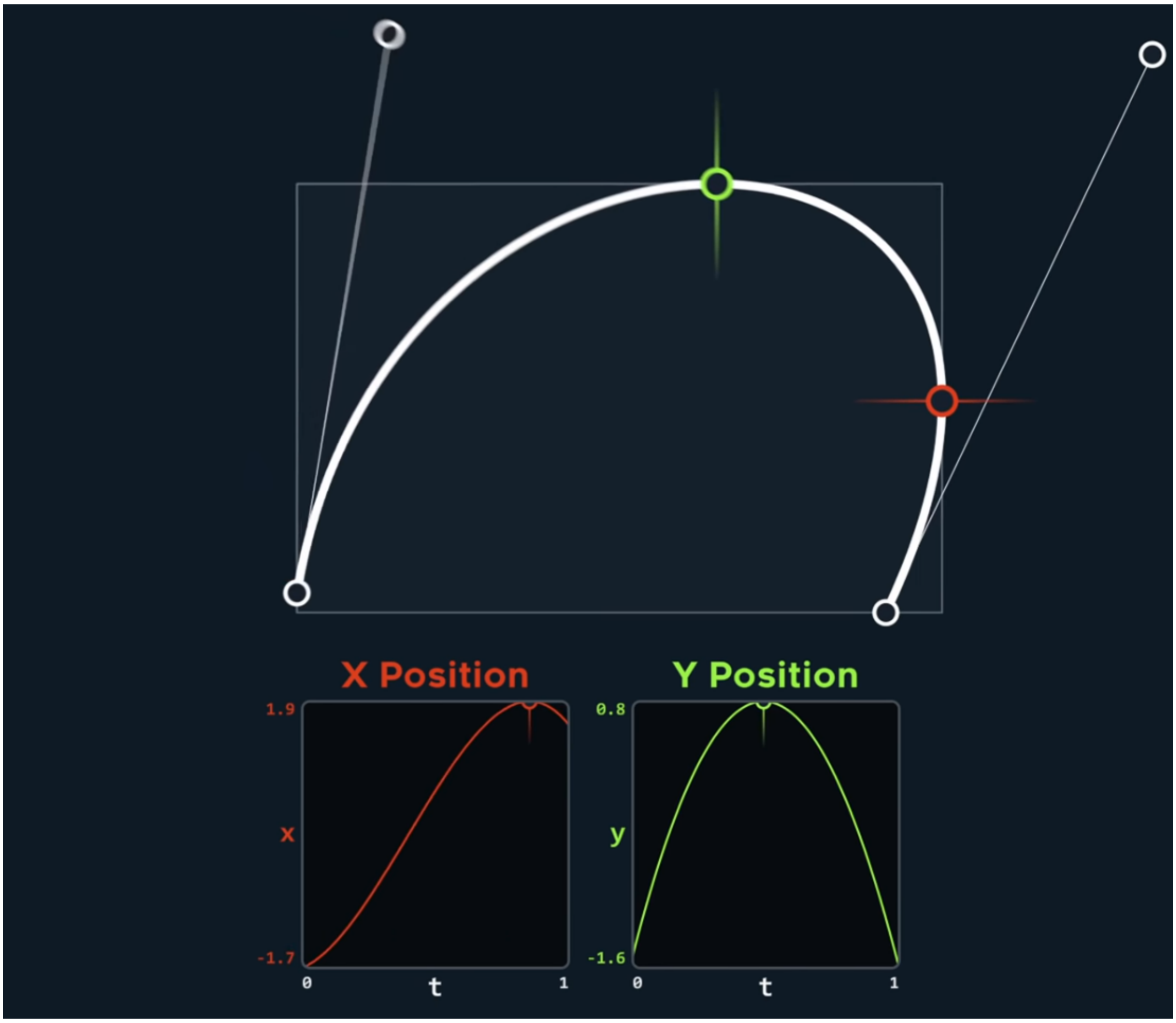

Abbildung 4: Die kleine Boundingbox [3]

Das hat nur den Vorteil, dass man die beiden Koordinaten nicht getrennt voneinander betrachten muss, sondern einfach sowohl die X-Achsen, als auch die Y-Achsen Schnittpunkte betrachten muss.

3.2 SVG

Ein weiterer sehr wichtiger Anwendungsbereich ist die Darstellung im SVG Format. Tatsächlich ist SVG der Standard für viele Anwendungen[15] , aber vor allem die Darstellung von Bildern/ Grafiken im Internet. Sie sind eine Alternative zu den gängigen Rasterbildformaten, die bei großen Bildschirmen pixelig wirken und außerdem die Lade-dauer der Webseite verlängern, da sie oft sehr große Dateien sind und SVG eben meist nicht. Außerdem erlaubt SVG es einem die Style-Elemente der Grafik direkt im HTML zu modifizieren, etwa mit CSS und eingebauten Klassen. Das geht, da SVG die Bézier Kurven in einen XML Format einbettet. So kann man Klassen erteilen und dann mit Stylekomponenten modifizieren. So kann man dann eben auch die Animationen umsetzen, denn man kann in CSS Events festlegen, die eine Transformation des Zielelements zur Folge hat, welches eben auch eine einzelne Form, Kurve oder eine Kurvengruppe sein kann.

Bézier Kurven kann man in SVG manuell definieren, indem man im `svg` Element das `d` Schlagwort benutzt. Dann gibt es verschiedene Befehle, mit denen man einen Virtuellen Stift benutzen kann um die Kontrollpunkte für die Kurve zu zeichnen. Mit `m` kann man den „Stift“ bewegen, ohne eine Spur zu hinterlassen. Mit `l` kann man eine gerade Linie zeichnen und mit `c` definiert man eine kubische Bézier Kurve. Als Argumente muss man dann noch die Punkte P_1,P_2 und P_3 angeben. Mit `q` kann man eine kubische Bézier Kurve definieren. Für alle diese Operationen gibt es die dazugehörigen Großbuchstaben, die anstatt lokaler Koordinaten die globalen/absoluten Werte als Referenzpunkt nehmen.

Beispielcode, in dem ich ein par Styles definieren und dann eine kubische Bézier Kurve:

```
<path
   d="M 50 50
      C 60 100, 75 50, 90 70"
   fill="none"
   stroke="blue"
   stroke-width="2"
   stroke-linejoin="round"
   stroke-linecap="round"
/>
```

Resultat in Abbildung 5 zu sehen.

[15] zum Beispiel auch die oben angesprochenen Graphen

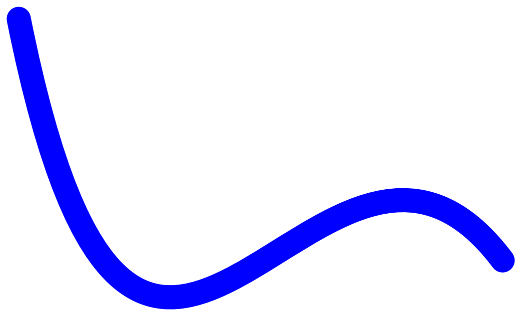

Abbildung 5: Die entstehende Kurve

4 Fazit

Bézier Kurven sind sehr vielfältig und haben viele wichtige Anwendungsbereiche. Außerdem sind sie die effizienteste Methode alle Kurven mathematisch und für Computer verständlich darzustellen. Die Kurven sind mit sehr interessanter Mathematik verbunden, die bei genauerer Betrachtung gar nicht so kompliziert ist und relativ viel mit dem Unterricht zu tun hat[16]. Daher hoffe ich, dass dieser Essay Ihnen die Bézier Kurven und ihre Funktionsweise etwas näher gebracht hat.

[16] Vektoren, Ableitungen, Normalvektoren,...

Bibliographie

[1] https://en.wikipedia.org/wiki/B%C3%A9zier_curve

[2] pomax, “A primer on bézier curves.” https://pomax.github.io/bezierinfo/

[3] Freya Holmér, “The beauty of bézier curves.”

[4] Freya Holmér, “The continuity of splines.”

---

5 Lizenz und Quellcode

Dieses Dokument ist unter CC BY-NC-SA veröffentlicht.

Der Quellcode ist *hier* zu finden.