Ein Weg zu Bézierkurven[1]

Die folgende Unterrichtseinheit hat den Kern des *Computer Aided Geometric Design* zum Inhalt, nämlich den Entwurf von Freiformkurven. Sie hat ein offenes Ende; entsprechend offen ist der Zeitbedarf.

In Bezug auf den Unterricht stellt die Einheit eine Verbindung von Analysis und Vektorgeometrie dar. Sie ist ohne Computer sinnlos; dynamische Geometrie–Software und Funktionsplotter sind beide zur Veranschaulichung, zum Experimentieren und zum Explorieren unabdingbar erforderlich.

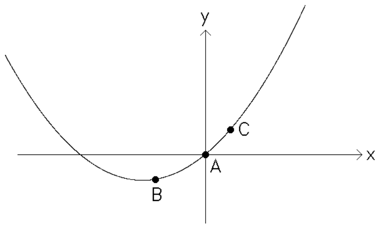
0. Einleitung

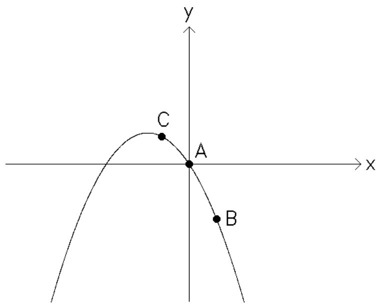
Kurven im Automobilbau werden auch von Designern entworfen, die in keinem Leistungskurs Mathematik waren. Sie zeichnen eine „schön" aussehende Kurve auf dem Papier (oder mit der Maus auf dem Bildschirm), und die Techniker stehen dann vor dem Problem, gemäß dieser Kurve ein Metallteil zu schneiden. Natürlich kann man aus der Designer–Kurve eine Holz– oder Metallvorlage basteln, aber das hat denn doch zu viele Nachteile: Die Vorlage ist schlecht zu transportieren, sie nutzt sich ab und ist dann nicht genau reproduzierbar. Besser wäre es, die Designer–Kurve mathematisch zu beschreiben. Da unser Designer seine wenigen Mathematikkenntnisse längst vergessen hat, kann man von ihm nicht verlangen, dass er zu seiner schönen Kurve geeignete Funktionsterme angibt.

Nun kann man sich damit behelfen, auf seiner Kurve Punkte genau auszumessen, und durch diese Punkte eine Kurve zu legen, von der man hofft, dass sie mit der Designer–Kurve gut übereinstimmt.

Dass dies nicht funktionieren kann, zeigen die folgenden Beispiele:

Durch 3 Punkte (etwa $A=\begin{pmatrix}0\\0\end{pmatrix}$; $B=\begin{pmatrix}-2\\-1\end{pmatrix}$; $C=\begin{pmatrix}1\\1\end{pmatrix}$) wird eine Parabel bestimmt, die in der üblichen Weise im Koordinatensystem liegt. Die Lösung (Ansatz: $f(x)=a\cdot x^2+b\cdot x$) lautet $f(x)=\frac{1}{6}\cdot x^2+\frac{5}{6}\cdot x$ (linkes Bild).





Aber hat der Designer das Koordinatensystem überhaupt mitbedacht? Dreht man das Koordinatensystem um A mit dem Winkel 90°, so erhält man die neuen Punkte $B=\begin{pmatrix}1\\-2\end{pmatrix}$; $C=\begin{pmatrix}-1\\1\end{pmatrix}$. Legt man nun durch A, B und C eine Parabel, so bekommt man $f(x)=-\frac{1}{2}\cdot x^2-\frac{3}{2}\cdot x$ (rechtes Bild).

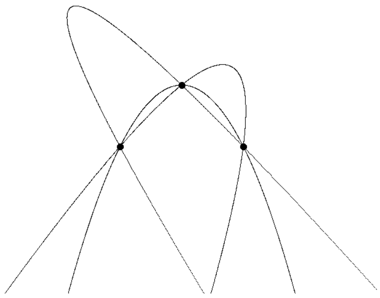
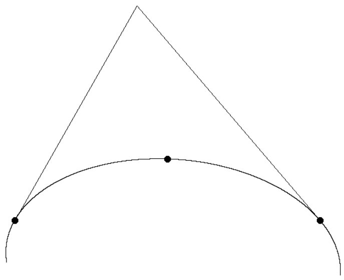
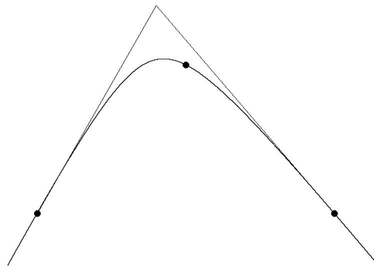
Die neue Parabel hat eine andere Öffnung als die erste. Auch die Scheitelpunkte gehen nicht durch Drehung auseinander hervor.

Allgemein kann man durch 3 (nicht kollineare) Punkte unendlich viele verschiedene Parabeln legen. Der Grund ist der folgende:

Die 3 Punkte seien A, B, C, und ein allgemeiner Kurvenpunkt sei $X(t)$. Gebilde mit dem allgemeinen Punkt $X(t)=U+t\cdot V$ sind bekanntlich Geraden; weiter unten wird gezeigt werden, dass Gebilde mit dem allgemeinen Punkt $X(t)=U+t\cdot V+t^2\cdot W$ Parabeln sind.

[1] Eine ausführlichere und in Teilen veränderte Fassung rschien in Mathematik in der Schule **38** (5), S. 303 - 314 (2000) und in Istron 6 (2000).

Sicher wird man fordern können, dass $X(0) = A$ und $X(1) = B$ ist. Aber welcher Parameter soll zu C gehören? Das Aussehen der Kurve hängt sehr davon ab, wofür man sich entscheidet.

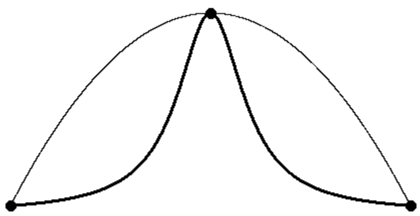
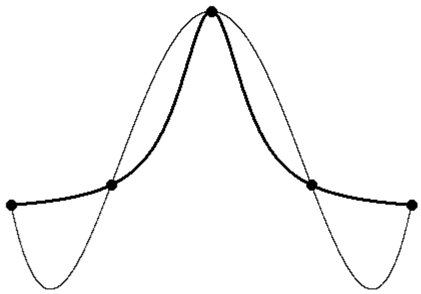




Auch wenn man zusätzlich zwei Tangenten vorschreibt, bekommt man keine eindeutige Kurve heraus. Im mittleren Bild hat man gar keine Parabel, sondern eine Ellipse, und im rechten Bild ist die Kurve eine Hyperbel.

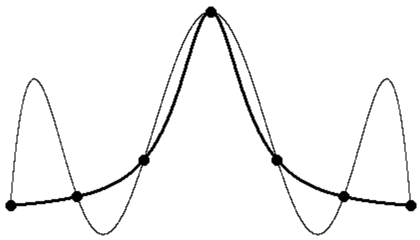
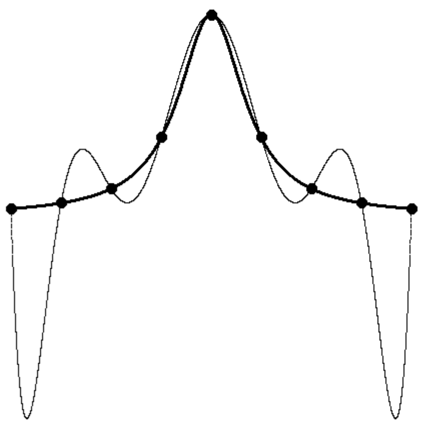
Nun gut; so geht es offenbar nicht. Aber warum wollen wir die Designer–Kurve auch mit nur 3 Punkten beschreiben? Nehmen wir an, der Designer habe sich folgende Gestalt ausgedacht:

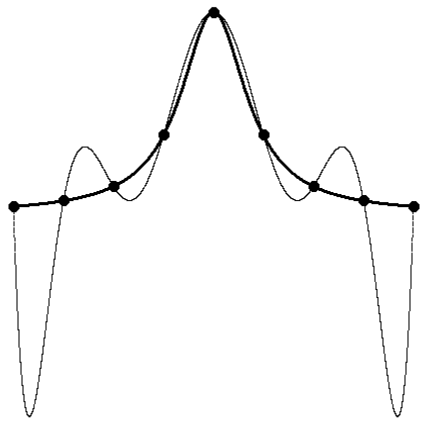
Auf die Kurve legen wir Punkte und ermitteln die sich aus diesen Punkten ergebende Kurve (Lagrange–Interpolation; es reicht, den Lernenden die folgenden Graphiken zu zeigen). Als Parameter wählen wir die jeweiligen x–Werte.

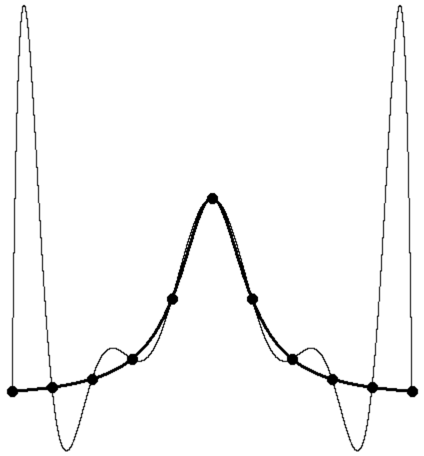
Bei 3 Punkten bekommt man einen Parabelbogen, der der gewünschten Kurve noch nicht ähnlich sieht. Das Aussehen lässt sich mit 2 weiteren Punkten verbessern.

Vermehrt man die Anzahl der Punkte, so wird das Ergebnis nicht unbedingt besser:

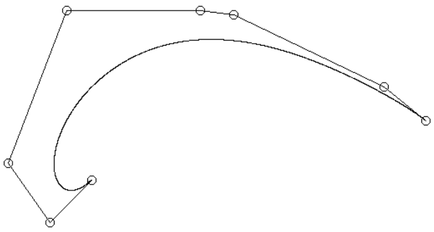



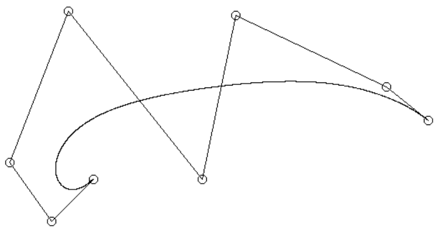




Solche Schwingungsphänomene können auftreten, müssen aber nicht. Allerdings gibt es keine praktikable Methode, vorherzusagen, ob ein solches Phänomen auftritt oder nicht.
Damit gibt es auch keine praktikable Methode, um aus Kurvenpunkten eindeutig und sicher die Kurve zu reproduzieren.

Dies war die Situation im Automobilbau um 1960 herum. Dann kamen Paul de Faget de Casteljau (von der Firma Citroën) und Pierre Bézier (von der Firma Renault) unabhängig voneinander auf die Idee, Kurven ganz anders zu erzeugen und zu beschreiben, nämlich ausgehend von wenigen Punkten, die gar nicht alle auf der Kurve zu liegen brauchen. Der Designer bekommt eine interaktive Software, zieht an den Punkten und sieht, wie sich die Kurve verändert:

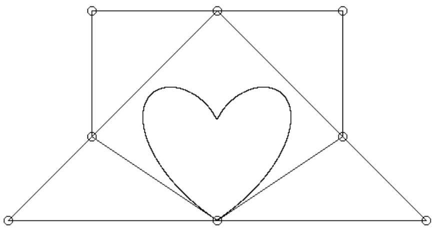




Da der Gebrauch dieser Software recht intuitiv ist, bekommt der Designer recht schnell ein Gefühl dafür, wie er die Punkte verändern muss, um zu der von ihm gewünschten Form zu gelangen. Zur Beschreibung der Kurve sind dann nur noch die Punkte notwendig. Man kann auch beweisen, dass die oben erwähnten Schwingungsphänomene bei dieser Art der Kurvenbeschreibung nicht auftreten können.

Die entstehenden Kurven sind nach Bézier benannt (obwohl de Casteljau etwas früher darauf gekommen war, aber von Citroën zur Geheimhaltung verpflichtet war). Die dahinter steckende Mathematik ist Gegenstand der nächsten Kapitel. Das Grundprinzip lässt sich auch anwenden, um (Karosserie–)Flächen zu beschreiben.

Aber auch außerhalb des Autobaus spielen Bézierkurven eine wachsende Rolle: So werden skalierbare Buchstaben mit Bézierkurven beschrieben, und Freihandkurven in Corel Draw oder auch in Word 97 werden im Béziermodus erzeugt.



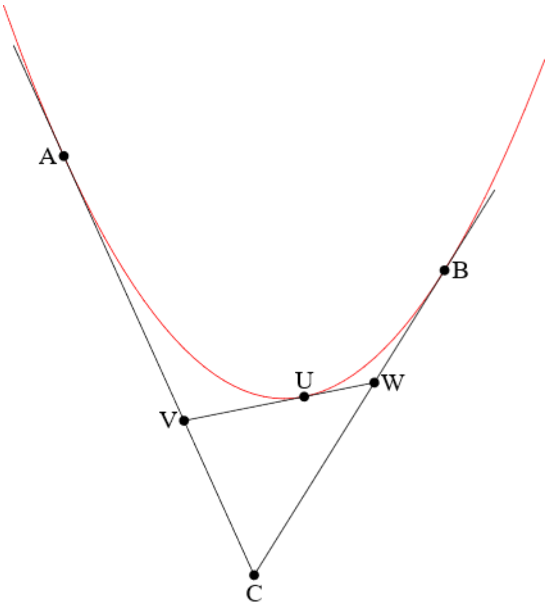
Vielleicht erinnerte sich de Casteljau an den folgenden Sachverhalt:

Die Tangenten zu $A=\begin{pmatrix} a \\ a^2 \end{pmatrix}$ und zu $B=\begin{pmatrix} b \\ b^2 \end{pmatrix}$ schneiden sich in $C=\begin{pmatrix} (a+b)/2 \\ a\cdot b \end{pmatrix}$.

Es sei $V=t\cdot A+s\cdot C,\ W=t\cdot C+s\cdot B$ mit $t+s=1$.
Dann ist

$$t\cdot V+s\cdot W=t^2\cdot\begin{pmatrix} a \\ a^2 \end{pmatrix}+2\cdot t\cdot s\cdot\begin{pmatrix} (a+b)/2 \\ a\cdot b \end{pmatrix}+s^2\cdot\begin{pmatrix} b \\ b^2 \end{pmatrix}$$
$$=\begin{pmatrix} t\cdot a+s\cdot b \\ (t\cdot a+s\cdot b)^2 \end{pmatrix}=:U$$

wieder ein Parabelpunkt.



VW hat wegen $W-V=t\cdot(b-a)\cdot\begin{pmatrix} 1/2 \\ a \end{pmatrix}+s\cdot(b-a)\begin{pmatrix} 1/2 \\ b \end{pmatrix}$ die Steigung $m=2\cdot(t\cdot a+s\cdot b)$; VW ist also Tangente in U.

Bei Bézierkurven geht man aus von A, C, B, konstruiert V und W und anschließend U wie oben. Variiert man t, beschreibt U eine Parabel. Es ist dann

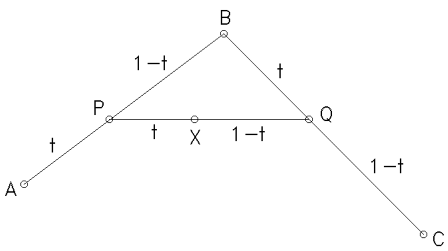
$$\begin{aligned} U&=t\cdot V+s\cdot W \\ &=t^2\cdot A+2\cdot t\cdot s\cdot C+s^2\cdot B \\ &=C+t^2\cdot(A-C)+s^2\cdot(B-C) \\ &=B+2\cdot t\cdot(C-B)+t^2\cdot(A-C+B-C) \end{aligned}$$

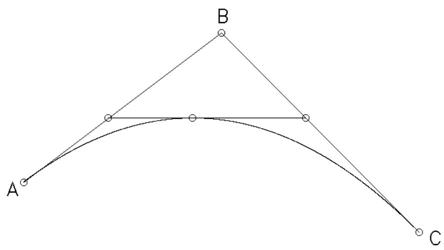
Da t und s nicht gleichzeitig verschwinden können, liegt C nicht auf der Parabel. Es ist $U(0)=B$ und $U(1)=A$.

1. Der einfachste Fall: Quadratische ganze Bézierkurven

Erzeugung

Gegeben seien 3 Punkte A, B und C, die nicht kollinear sind. P teile AB im Verhältnis $\frac{AP}{PB}=\frac{t}{1-t}$.
Q teile BC im selben Verhältnis. R teile PQ wiederum im selben Verhältnis.





Wird t variiert, so bewegt sich X auf einer Parabel.

Parabeln sind <u>quadratische</u> Kurven, da ein beliebiger Kurvenpunkt die Form $X(t)=P+t\cdot Q+t^2\cdot R$ hat.
In den folgenden Abschnitten werden auch kubische Kurven thematisiert werden.

Parabeln sind ferner <u>ganze</u> Kurven, da der Term für $X(t)$ keinen t enthaltenden Nenner hat. Ein Kreis ist auch eine quadratische Kurve, aber keine ganze, da ein allgemeiner Kreispunkt etwa die Form $X(t) = \frac{1}{t^2+1} \cdot \begin{pmatrix} 2\cdot t \\ t^2-1 \end{pmatrix}$ hat.

Jede quadratische ganze Kurve mit dem allgemeinen Punkt $X(t) = P + t\cdot Q + t^2 \cdot R$ lässt sich als Bézierkurve mit dem allgemeinen Punkt $X(t) = (1-t)^2 \cdot A + 2\cdot(1-t)\cdot t\cdot B + t^2 \cdot C$ und $A = P$, $B = P + \frac{Q}{2}$, $C = P + Q + R$.) darstellen.

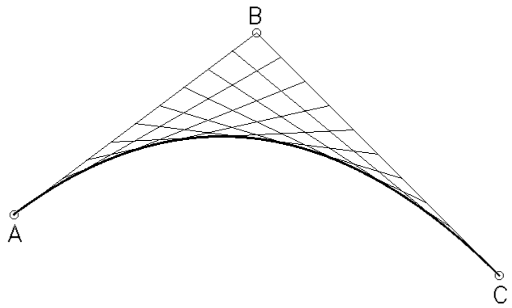
Tangenten

Man überlegt sich leicht, dass eine Kurve mit dem allgemeinen Punkt $X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ an der Stelle t den Tangentenrichtungsvektor $\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$ hat, der kurz mit $X'(t)$ abgekürzt sei. Die Tangente an der Stelle t hat somit den allgemeinen Punkt $X(t) + \lambda \cdot X'(t)$.

Die Gerade durch A und B ist Tangente an A; entsprechendes gilt für C. Die Gerade durch P und Q ist Tangente an X.

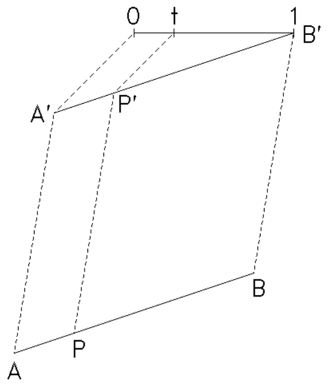
Die quadratische Bézierkurve als Hüllkurve

Da die Gerade durch P und Q Tangente an X ist, kann man eine quadratische Bézierkurve als Hüllkurve ihrer Tangenten auffassen.



2. Zur Konstruktion ganzer Bézierkurven mit Dynamischer Geometrie–Software

Bézierkurven kann man mit Hilfe eines Funktionsplotters erzeugen, aber auch mit Dynamischer Geometrie–Software. Der zweite Weg ist natürlich besonders instruktiv.
Nebenstehend die Konstruktion des Teilpunktes P.



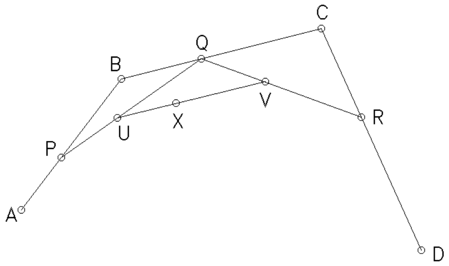
Oben befindet sich die Referenzstrecke 01. Die zu teilende Strecke AB wird parallel verschoben und Punkt P‘ nach dem Strahlensatz konstruiert. Die Konstruktion von P ist dann klar.
Nun schreibe man ein Makro, das zu 0, 1, t, A, B den Teilungspunkt P konstruiert. Bei quadratischen ganzen Bézierkurven kann man dies Makro dann anwenden auf AB, BC und PQ.
Dann schreibe man ein Makro, das zu 0, 1, t, A, B, C den Bézierpunkt X konstruiert.

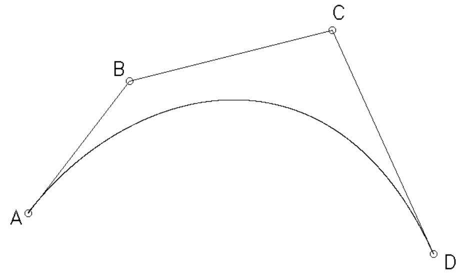
Die Konstruktion versagt, wenn die Referenzstrecke zu AB oder zu BC parallel ist. Das ist nicht weiter schlimm, da man die Referenzstrecke beliebig drehen kann.

3. Kubische ganze Bézierkurven

Erzeugung

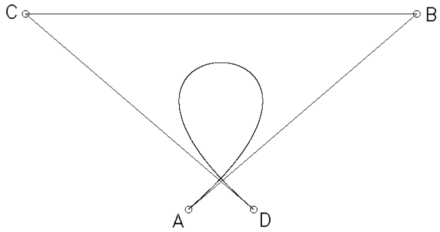
Man kann die Vorgehensweise bei quadratischen Bézierkurven auch mit 4 Punkten durchführen.

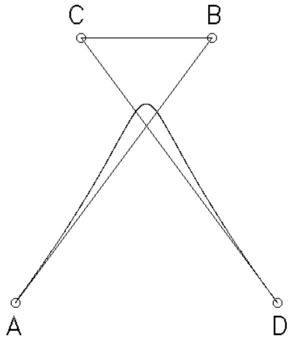




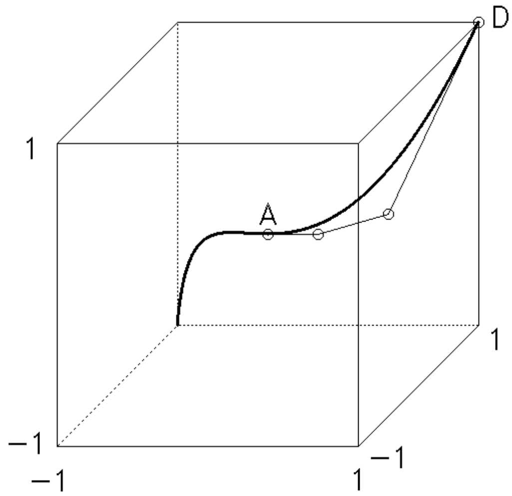
Man bekommt dann eine <u>kubische Kurve</u>. Weiter unten wird der allgemeine Punkt angegeben werden.

Die Kurve braucht allerdings kein Funktionsgraph zu sein. Die Kurve muss aber nicht unbedingt eine Schleife haben, wenn sich das Polygon ABCD überschlägt.





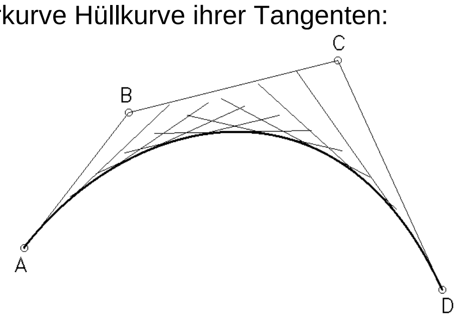
Im Gegensatz zu quadratischen Kurven kann eine kubische ganze Bézierkurve auch eine <u>Raumkurve</u> sein. So führt $A = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 2/3 \\ 1/3 \\ 0 \end{pmatrix}$ und $D = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ auf die räumliche Parabel mit dem allgemeinen Punkt $\begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}$.



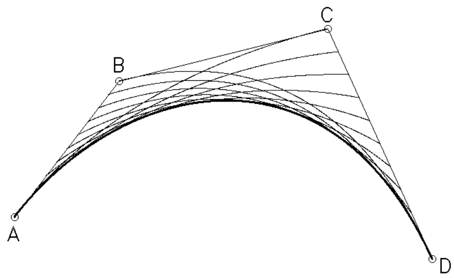
Wir werden im folgenden nur ebene Kurven betrachten.

Die kubische Bézierkurve als Hüllkurve

Wie im quadratischen Fall ist die kubische Bézierkurve Hüllkurve ihrer Tangenten:

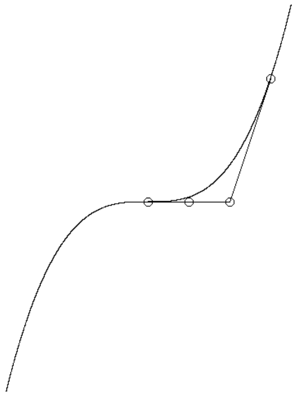


Sie ist aber auch Hüllkurve ihrer berührenden Parabeln:



4. Beispiele kubischer ganzer Bézierkurven

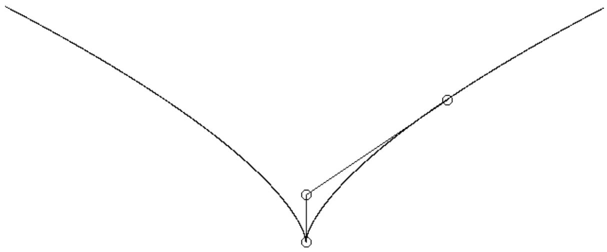
Kubische Parabel



Es ist $y = x^3$; für jeden Kurvenpunkt gilt mithin $X(t) = \begin{pmatrix} t \\ t^3 \end{pmatrix}$.

Also ist $A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 1/3 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix}$ und $D = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

<u>Neilsche Parabel</u>

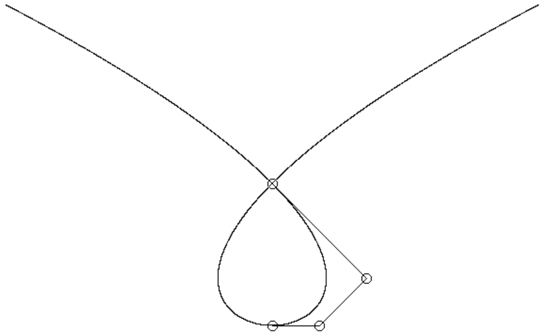


Es ist $y^3 = x^2$; für jeden Kurvenpunkt gilt $X(t) = \begin{pmatrix} t^3 \\ t^2 \end{pmatrix}$ und damit $A = B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 \\ 1/3 \end{pmatrix}$ und $D = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Die Kurve hat im Punkt A eine <u>Spitze</u>.

Tschirnhaus–Kubik

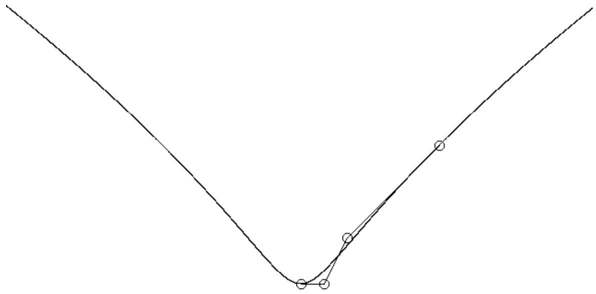
Es ist $x^2 = y\cdot(y-1)^2$; für jeden Kurvenpunkt gilt $X(t)=\begin{pmatrix} t-t^3 \\ t^2 \end{pmatrix}$, also $A=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $B=\begin{pmatrix} 1/3 \\ 0 \end{pmatrix}$, $C=\begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$ und $D=\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Die Kurve hat in D einen Doppelpunkt.



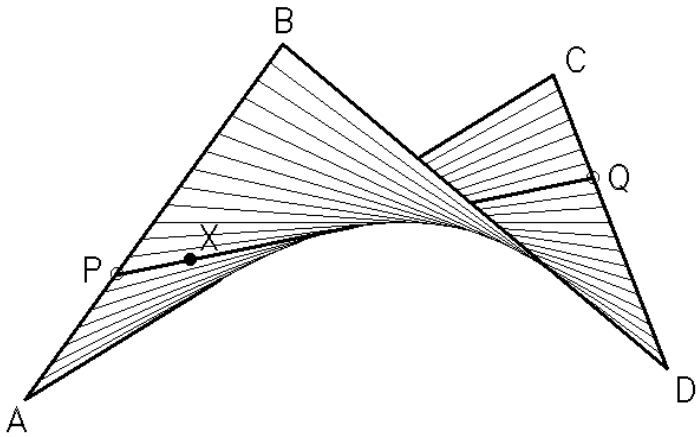
Kubische Duplikatrix

Es ist $x^2 = y\cdot(y+1)^2$; für jeden Kurvenpunkt gilt $X(t)=\begin{pmatrix} t+t^3 \\ t^2 \end{pmatrix}$ und damit $A=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $B=\begin{pmatrix} 1/3 \\ 0 \end{pmatrix}$, $C=\begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$ und $D=\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Das Bild ist in waagerechter Richtung gestaucht.
Die kubische Duplikatrix hat zwei Wendepunkte.



5. Ein Ausblick auf Flächen

Bézierkurven werden durch Punkte definiert, und in einem ersten Iterationsschritt werden aus den Punkten Geraden gewonnen. Auch Bézierflächen werden durch Punkte definiert. Aus den Raumpunkten werden dann aber keine Ebenen gewonnen – dies erweist sich aus programmiertechnischen Gründen nicht als günstig –, sondern aus je 4 Punkten (im unteren Bild A, B, C und D, die nicht in einer Ebene liegen sollen) wird eine gekrümmte Fläche erzeugt:



P durchläuft die Strecke AB, und Q durchläuft die Strecke CD, wobei jeweils $\frac{AP}{PB}=\frac{CQ}{QD}=\frac{t}{1-t}$ gelten soll. Somit gilt $P(t)=(1-t)\cdot A+t\cdot B$ und $Q(t)=(1-t)\cdot C+t\cdot D$.

Nun teile der Punkt X die Strecke PQ im Verhältnis $\frac{PX}{XQ}=\frac{s}{1-s}$, es sei also

$$\begin{aligned} X(t,\ s) &= (1-s)\cdot P+s\cdot Q \\ &= (1-s)\cdot\big((1-t)\cdot A+t\cdot B\big)+s\cdot\big((1-t)\cdot C+t\cdot D\big) \\ &= A+s\cdot(C-A)+t\cdot(B-A)+s\cdot t\cdot(A+D-B-C). \end{aligned}$$

Zu A, B, C, D und den Parametern t und s gehört mithin der allgemeine Punkt $X(t,\ s)$.

Das sich ergebende Gebilde heißt hyperbolisches Paraboloid, es ist eine Fläche mit einem Sattel.

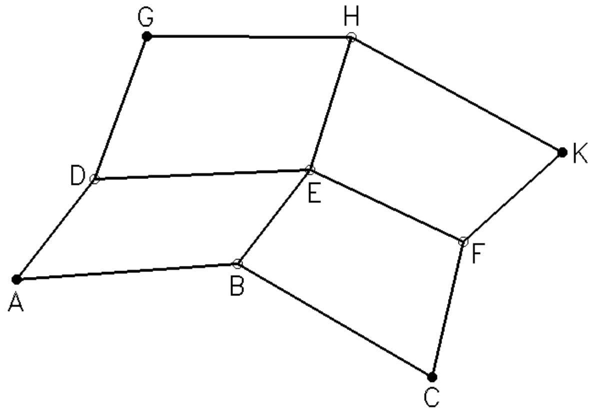
Eine solche Sattelfläche ist Grundbaustein von Bézierflächen.

Hier wird der zweite Iterationsschritt erläutert. Wir gehen aus von 9 Punkten A, ..., K, die nicht alle in einer Ebene liegen sollen.
Zu A, B, E, D und den Parametern t und s wird der allgemeine Punkt U(t, s) gebildet.

Zu B, C, F, E und den Parametern t und s wird der allgemeine Punkt V(t, s) gebildet.

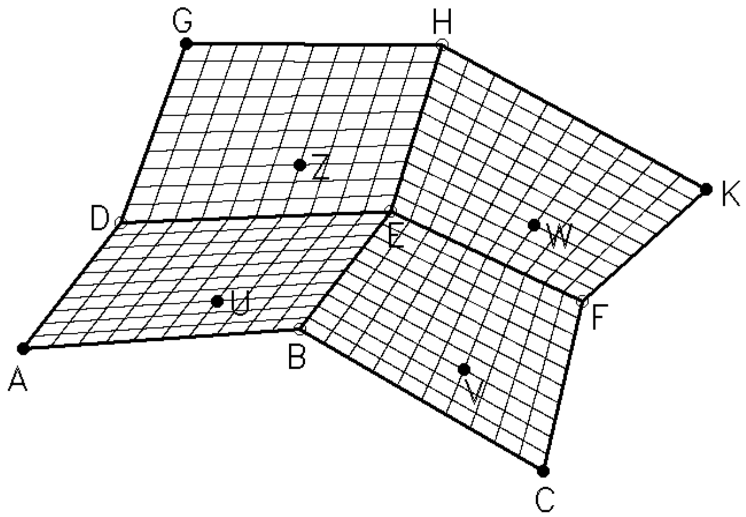
Zu E, F, K, H und den Parametern t und s wird der allgemeine Punkt W(t, s) gebildet.

Zu D, E, H, G und den Parametern t und s wird der allgemeine Punkt Z(t, s) gebildet.

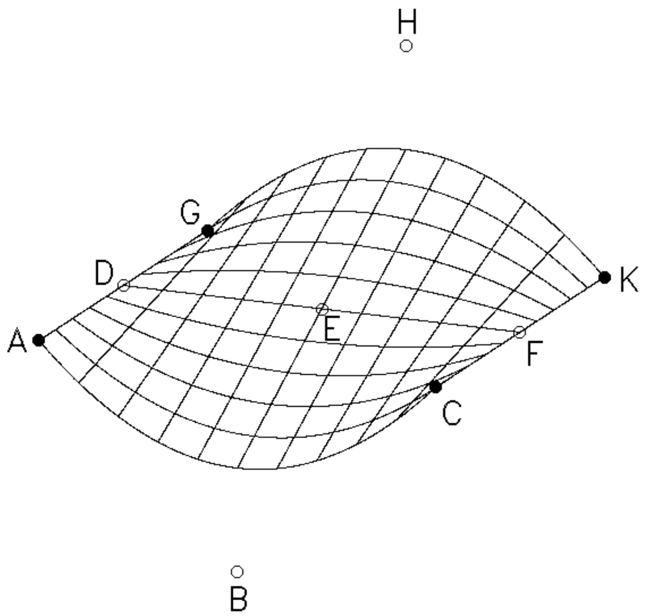


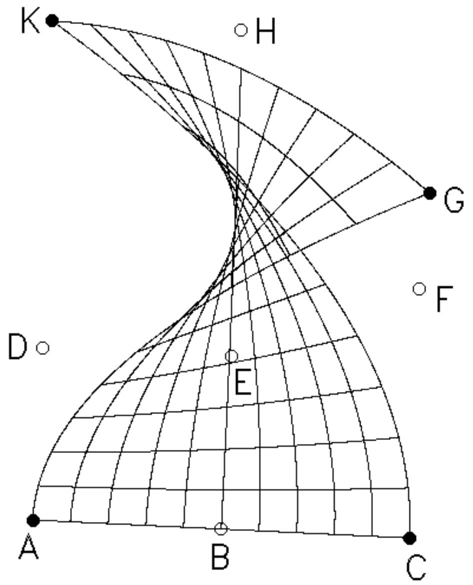
U, V, W und Z liegen auf den entsprechenden Sattelflächen.

Nun wird zu U(t, s), V(t, s), W(t, s), Z(t, s) und den Parametern t und s der allgemeine Punkt X(t, s) gebildet. X(t, s) ist der allgemeine Punkt der entstehenden Fläche.



Durch geeignete Wahlen von A, ..., K bekommt man interessante Flächen:





Literatur

Farin, G.: Kurven und Flächen im Computer Aided Geometric Design. Vieweg 1994.

Meyer, J.: Die Sattelfläche im Grundkurs. Praxis der Mathematik 1995, S. 250 – 255.
online: http://Mathematik-Meyer.de/Materialien/Sattelflaeche_Grundkurs.pdf

Meyer, J.: Von der Normalparabel zu kubischen Kurven. Mathematica didactica 1998, S. 84 – 108.
online: http://Mathematik-Meyer.de/Materialien/KubKurv.pdf