Introduzione alle curve di Bezier e alle spline

Luciano Battaia

22 gennaio 2020

In queste pagine proponiamo una sintetica introduzione alle curve di Bezier e alle spline, con particolare riguardo alla loro implementazione in Geogebra. Lo scopo è quello di proporre una introduzione elementare alle curve NURBS, tipiche dei moderni sistemi CAD.

1 Equazioni parametriche di curve

Tra le varie tecniche utilizzate per scrivere le equazioni delle curve del piano[1] (o dello spazio), sono particolarmente importanti le equazioni parametriche. L'idea è la seguente: per ottenere un insieme di punti da rappresentare nel piano (e sarà analogo il problema nello spazio), cioè un insieme di coppie di numeri reali possiamo pensare all'ascissa e all'ordinata come funzioni di una variabile, abitualmente indicata con t, e detta parametro.

Per capire il problema consideriamo un esempio. Sia $x = t^2$ e $y = t^3$, con $-2 \leq t \leq 2$. Per ogni valore di t nell'intervallo considerato si ottengono un valore per x e un valore per y che possono essere considerati l'ascissa e l'ordinata di un punto del piano, come mostrato per alcuni valori nella seguente tabella.

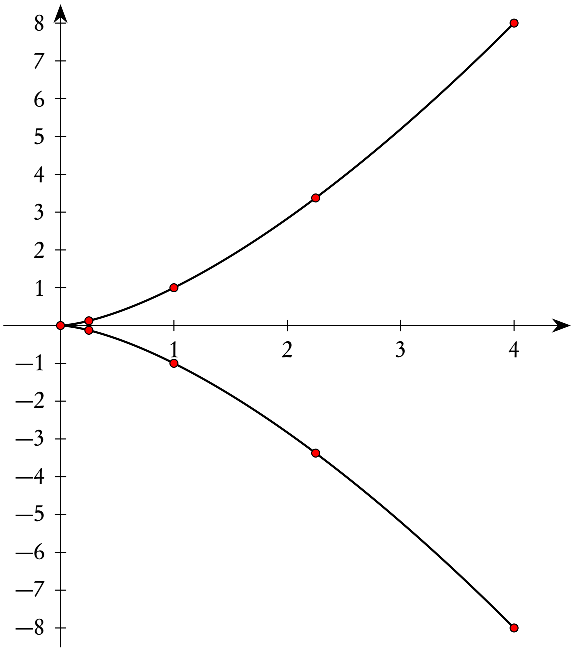
t	x	y
-2	4	-8
$-3/2$	$9/4$	$-27/8$
-1	1	-1
$-1/2$	$1/4$	$-1/8$
0	0	0
$1/2$	$1/4$	$1/8$
1	1	1
$3/2$	$9/4$	$27/8$
2	4	8

Si otterranno dunque i punti seguenti $(4, -8)$, $(9/4, -27/8)$, $(1, -1)$, $(1/4, -1/8)$, $(0, 0)$, $(1/4, 1/8)$, $(1, 1)$, $(9/4, 27/8)$, $(4, 8)$.

Se consideriamo tutti i valori del parametro t nell'intervallo considerato otterremo infiniti punti che, nei casi di nostro interesse (cioè per funzioni "non patologiche") si disporranno su una curva.

[1]Ricordiamo in particolare i grafici delle funzioni di una variabile e le curve soluzioni di equazioni in due incognite.

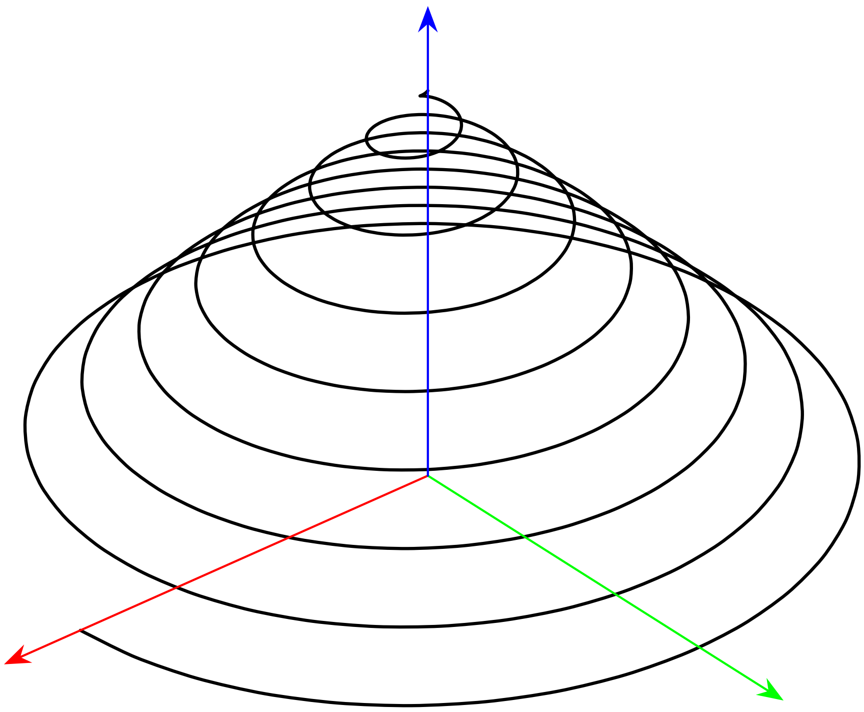
La figura che segue, dove sono state usate unità di misura diverse sui due assi, mostra il grafico della curva ottenibile dalle equazioni sopra considerate, con evidenziati i punti ottenuti nella precedente tabella.



Se invece di due coordinate x e y se ne considerano tre, x, y e z, si otterranno curve dello spazio.

La figura che segue illustra una spirale, a passo costante, tracciata su un cono rotondo e retto dove le tre coordinate del generico punto sono date da

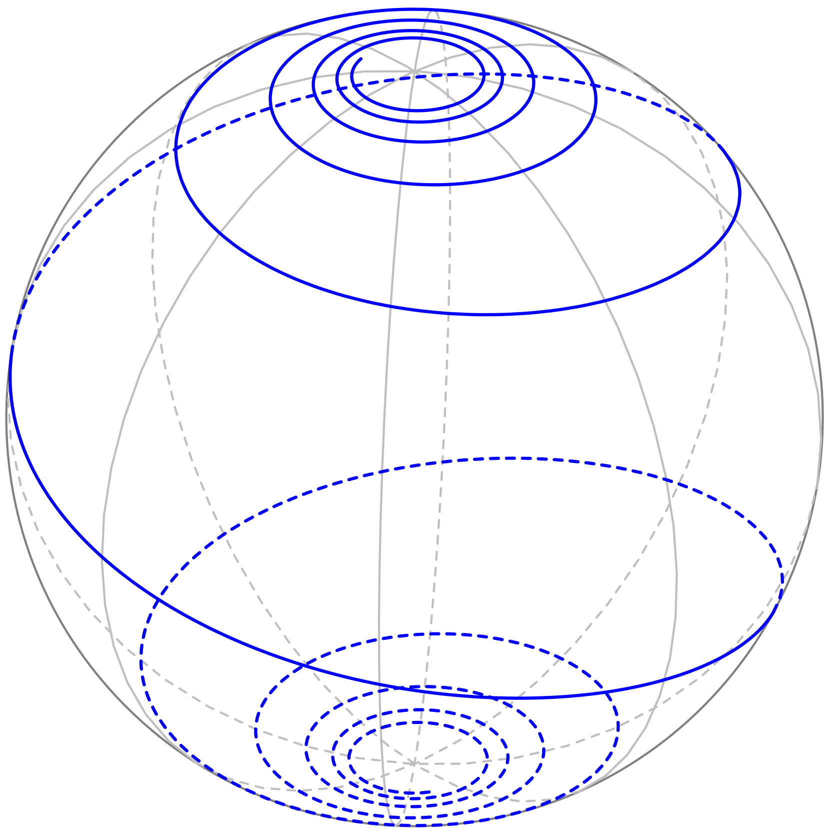
$$x = (5-t)\cos(10t) \quad , \quad y = (5-t)\sin(10t \quad , \quad z = t \quad , \text{con} \quad 0 \leq t \leq 5.$$



La figura che segue illustra invece una spirale decisamente più complessa, tracciata su un sfera di raggio r, con passo $p = 0.2$ e con infinite spire che si avvolgono intorno al polo Nord e al polo Sud. Le coordinate del generico punto sono date da

$$x = \frac{r\cos(t)}{\sqrt{1+p^2t^2}} \quad , \quad y = \frac{r\sin(t)}{\sqrt{1+p^2t^2}} \quad , \quad z = \frac{-prt}{\sqrt{1+p^2t^2}} \quad , \text{con} \quad -33 \le t \le 33.$$

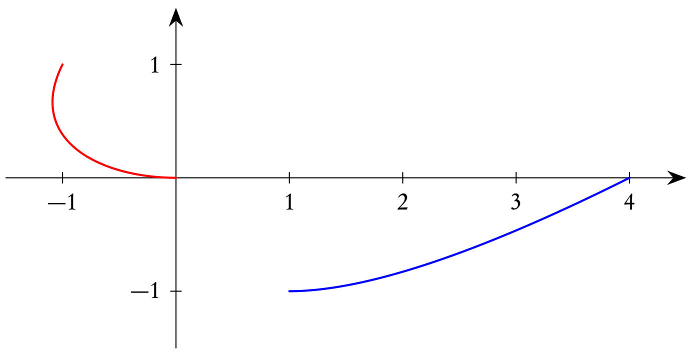
Per rendere più chiara la figura abbiamo anche rappresentato alcuni meridiani della sfera.



Le equazioni parametriche possono anche essere date "a pezzi", ovvero con diverse espressioni per diversi valori del parametro. Per esempio

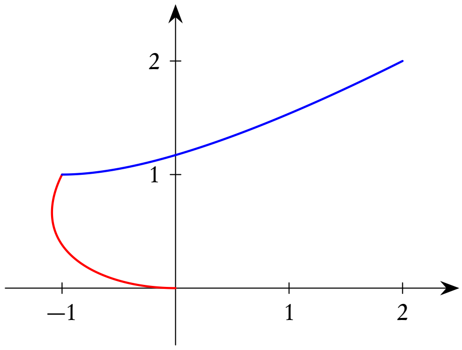
$$x = \begin{cases} t^3 - 2t, & \text{se } 0 \le t \le 1 \\ t^2, & \text{se } 1 < t \le 2 \end{cases} \qquad y = \begin{cases} t^2, & \text{se } 0 \le t \le 1 \\ t^2 - 2t, & \text{se } 1 < t \le 2 \end{cases}$$

che produce la seguente curva "a pezzi".



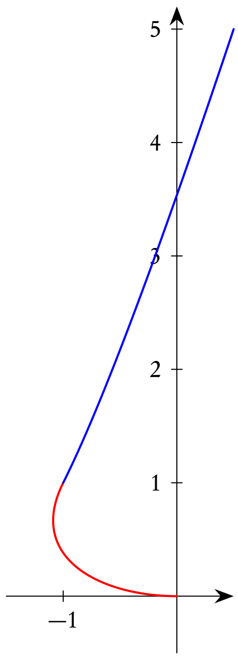
I due "rami" possono anche saldarsi in modo che la curva non presenti "strappi", come nell'esempio seguente, che modifica leggermente il precedente.

$$x = \begin{cases} t^3 - 2t, & \text{se } 0 \le t \le 1 \\ t^2 - 2, & \text{se } 1 < t \le 2 \end{cases} \qquad y = \begin{cases} t^2, & \text{se } 0 \le t \le 1 \\ t^2 - 2t + 2, & \text{se } 1 < t \le 2 \end{cases}$$



Una "saldatura" più efficiente si ha se i due rami hanno, nel punto di passaggio da un espressione all'altra, la stessa tangente. L'esempio che segue modifica ulteriormente l'esempio precedente in modo che il secondo "ramo" parta con la stessa tangente con cui termina il primo.

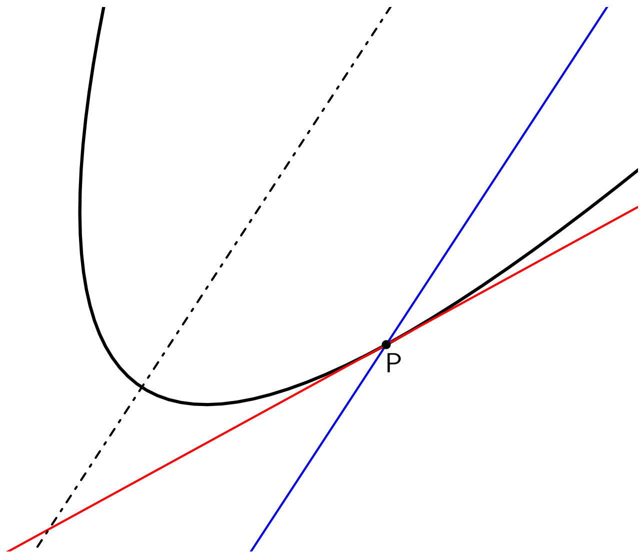
$$x = \begin{cases} t^3 - 2t, & \text{se } 0 \le t \le 1 \\ 0.5t^2 - 1.5, & \text{se } 1 < t \le 2 \end{cases} \qquad y = \begin{cases} t^2, & \text{se } 0 \le t \le 1 \\ 2t^2 - 2t + 1, & \text{se } 1 < t \le 2 \end{cases}$$



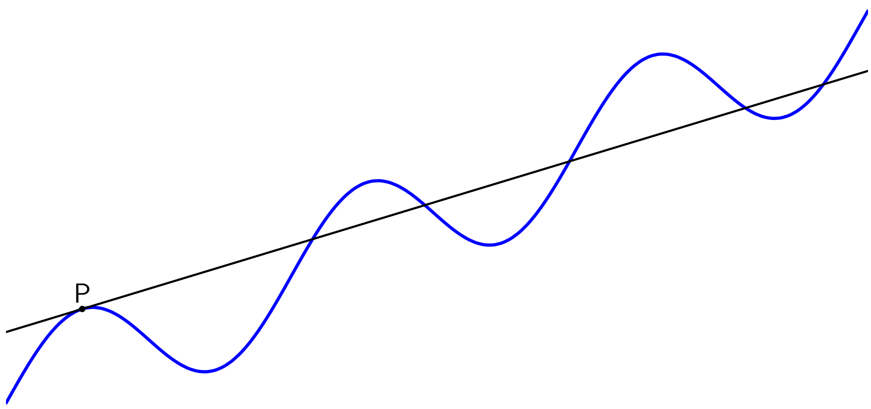
Si possono ottenere saldature ancora più "liscie" e ne riparleremo trattando la curvatura.

2 Richiami sulla tangente a una curva

È ben noto il concetto di tangente ad una circonferenza e anche la procedura grafica per tracciare la tangente a una circonferenza in un punto: basta tracciare la perpendicolare al raggio in quel punto. Tuttavia questo procedimento, fondamentale nel disegno geometrico, non è estendibile ad una curva qualunque, in quanto per una curva qualunque non esiste un centro e quindi un raggio con relativa perpendicolare. Si deve dunque cercare un diverso modo per estendere il concetto di retta tangente da una circonferenza ad una curva qualunque. Poiché in una circonferenza la retta tangente in un punto ha la caratteristica peculiare di avere un solo punto in comune con la circonferenza stessa (mentre tutte le altre rette per quel punto hanno due punti in comune con la circonferenza), verrebbe spontaneo dire che la tangente ad una curva in un punto è la retta che ha solo quel punto in comune con la curva. Si capisce facilmente che la cosa non è soddisfacente se si pensa ad una parabola: dato un suo punto esistono due rette che hanno un solo punto in comune con la parabola e solo una soddisfa a quello che, intuitivamente, è il concetto di tangente.

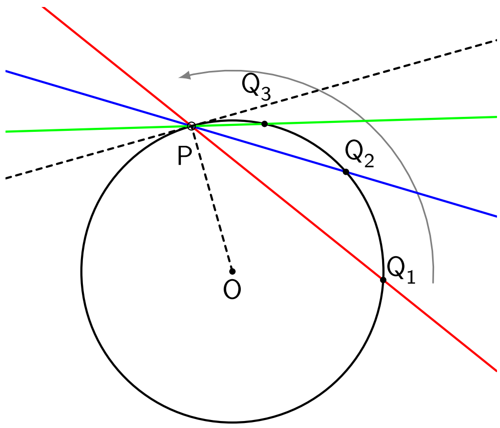


In altre situazioni la retta che, intuitivamente, gioca il ruolo di tangente può avere più punti in comune con un curva.

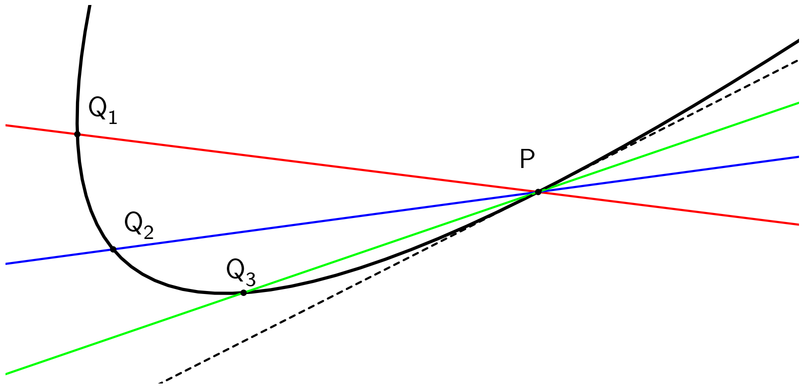


La costruzione della tangente ad una circonferenza che meglio si presta ad essere estesa a curve è la costruzione “dinamica” che ora illustriamo.

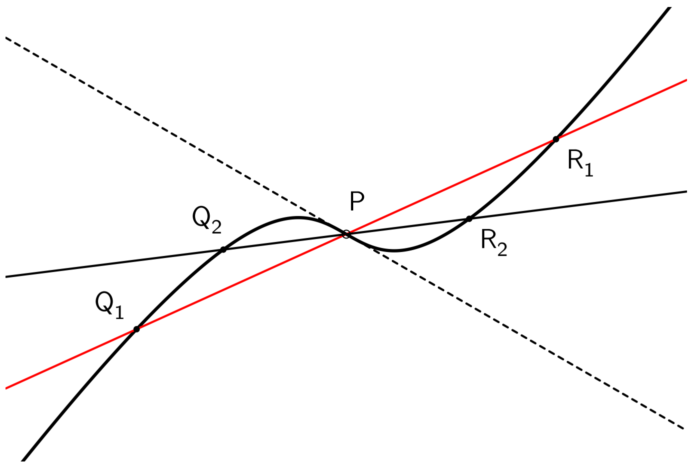
Fissato un punto P su una circonferenza, si consideri un altro punto Q e la retta PQ, che, come è ben noto, prende il nome di *secante* per P e Q. Supponiamo poi di muovere il punto Q verso P, fino a sovrapporlo a P stesso. Per ogni posizione di Q, distinta da P, si otterrà una diversa secante, mentre nel momento esatto in cui Q si sovrappone a P non si può più definire la retta PQ, in quanto per un unico punto passano infinite rette. Si può però considerare il fatto che, all’avvicinarsi di Q a P, la retta PQ tende ad assumere una posizione, detta *posizione limite*, ben definita, che coincide esattamente con la tangente costruita come perpendicolare al raggio per P. È proprio questa costruzione dinamica che può essere generalizzata ad una curva qualunque: si può dire, con un'espressione significativa a cui si potrebbe comunque dare anche un preciso significato geometrico, che la tangente è una secante passante per (*almeno*) due punti coincidenti, ovvero che una curva e la sua tangente in un punto, quando esiste, hanno, in quel punto, almeno due intersezioni coincidenti, o come dicono i geometri, un *contatto almeno bipunto*. Abbiamo utilizzato l'avverbio *almeno*, perché in generale si possono avere anche contatti di ordine più elevato, come vedremo su un esempio.



La figura che segue illustra questo processo applicato ad una parabola.



La figura che segue illustra un caso in cui la tangente ad una curva ha addirittura un contatto *tripunto* con la curva stessa



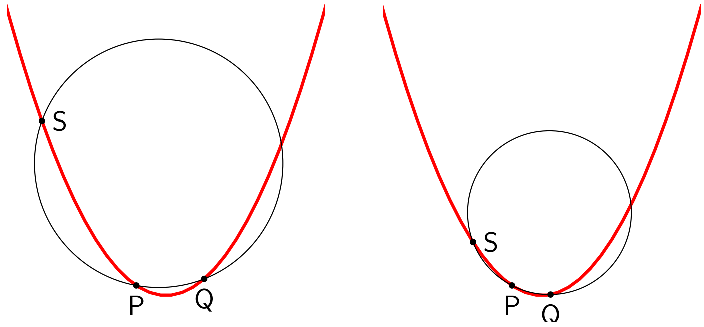
3 Il centro di curvatura di una curva

La prima differenza che viene in mente se si paragonano le rette con le circonferenze è il fatto che le rette non hanno alcuna curvatura, mentre le circonferenze sono linee curve e precisamente sono tanto più *incurvate* quanto più sono piccole, cioè quanto minore è il loro raggio. Questa considerazione intuitiva porta alla definizione seguente di *curvatura* per una circonferenza: si chiama *curvatura* di una circonferenza di raggio r il reciproco del raggio:

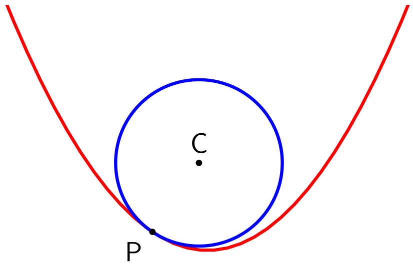
$$k = \frac{1}{r}.$$

Poiché una retta si può pensare come una circonferenza di raggio infinito, per la sua curvatura si avrà, come vuole l'intuizione, $k = 0$.

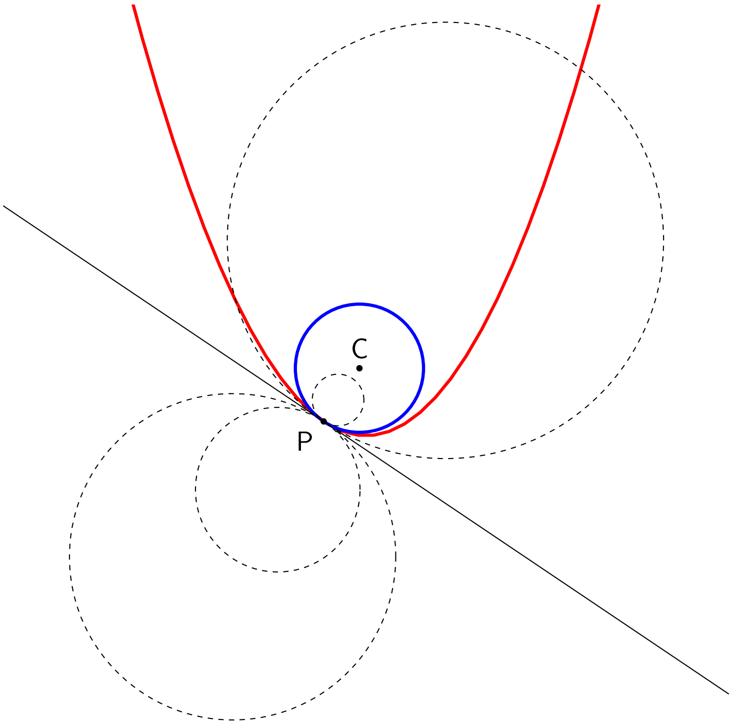
Anche per curve più complesse si può definire una curvatura, che potrà variare da punto a punto, con il seguente ragionamento. Sia P un punto di una curva $\mathcal{C}$, e siano Q e S altri due punti sulla stessa curva. Per questi tre punti (che supponiamo non allineati) passa una e una sola circonferenza, che varia al variare di Q e S, supponendo invece che P rimanga fisso. Si veda la figura seguente, dove abbiamo rappresentato due diverse situazioni corrispondenti allo stesso punto P, nel caso in cui la curva in questione sia una parabola.



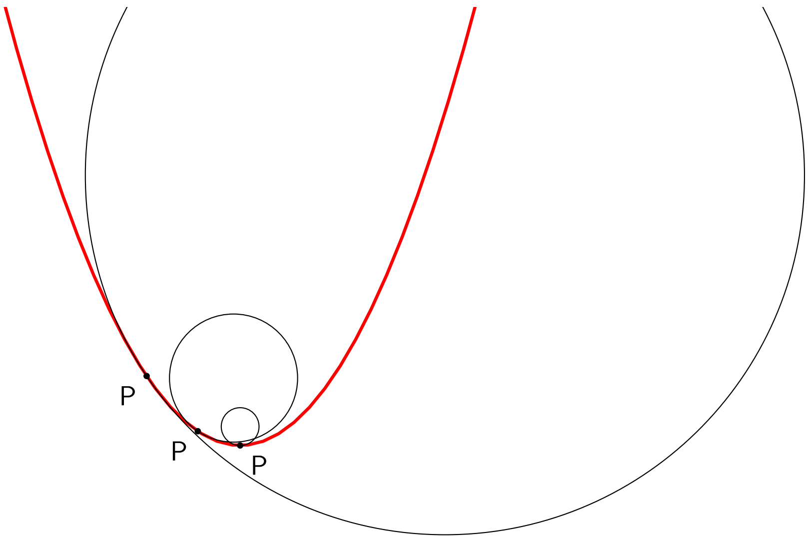
Se facciamo *avvicinare indefinitamente* Q e S a P, la circonferenza tende ad assumere, nei casi semplici di nostro interesse, un ben preciso centro e raggio: la corrispondente circonferenza, che è tangente alla curva data, è detta *circonferenza osculatrice* alla curva, il suo raggio è il *raggio di curvatura*, il reciproco del raggio è la *curvatura* e il suo centro il *centro di curvatura*, con riferimento al punto P scelto. Si veda la figura seguente dove è rappresentata questa speciale circonferenza per la parabola e il punto P già considerati nella figura precedente.



Come già osservato, questa circonferenza osculatrice è tangente alla curva in questione nel punto P; esistono però infinite circonferenze tangenti alla curva nello stesso punto, e precisamente tutte quelle tangenti alla retta tangente: la circonferenza osculatrice è, tra tutte le circonferenze tangenti, *la "più tangente di tutte"*, cioè quella che meglio approssima la curva nei pressi del punto. Per dirla con i geometri classici è quella che "ha il contatto più intimo con la curva". In termini un po' più formali, e senza entrare troppo nel dettaglio, si può dire che, mentre tutte le circonferenze tangenti alla curva in un punto hanno due punti coincidenti in comune con la curva, la circonferenza osculatrice ha (almeno) tre punti coincidenti in comune con la curva (si dice che ha un contatto almeno "tripunto" con la curva). Si veda la figura seguente dove, sempre con riferimento alla stessa situazione trattata nelle figure precedenti, sono visualizzate la retta tangente in P e alcune circonferenze tangenti sempre in P, oltre alla circonferenza osculatrice.



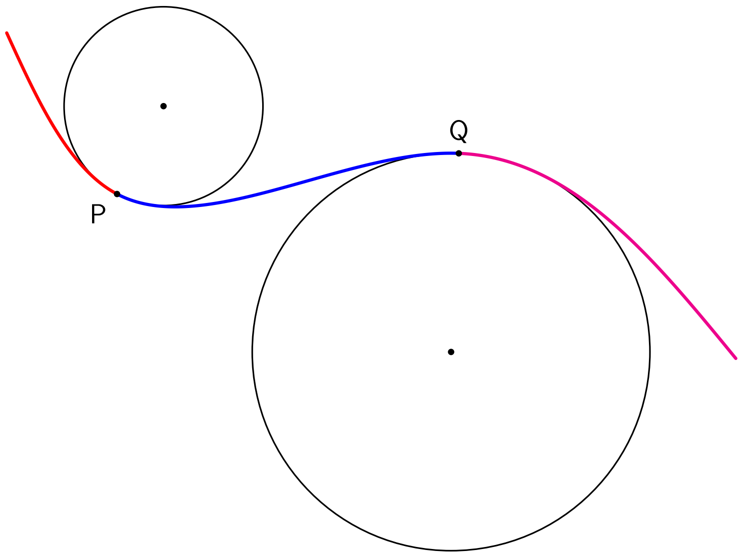
La circonferenza osculatrice cambia da punto a punto e di conseguenza cambia la curvatura della curva da punto a punto, come è intuitivamente evidente. Nella figura seguente abbiamo visualizzato le circonferenze osculatrici in due ulteriori punti, oltre a quello già precedentemente considerato. Si noti come la circonferenza osculatrice nel vertice sia molto piccola (e quindi la sua curvatura molto grande), mentre la circonferenza osculatrice in un punto distante dal vertice è molto grande (e quindi la sua curvatura molto piccola) in perfetto accordo con l'impressione "visiva".



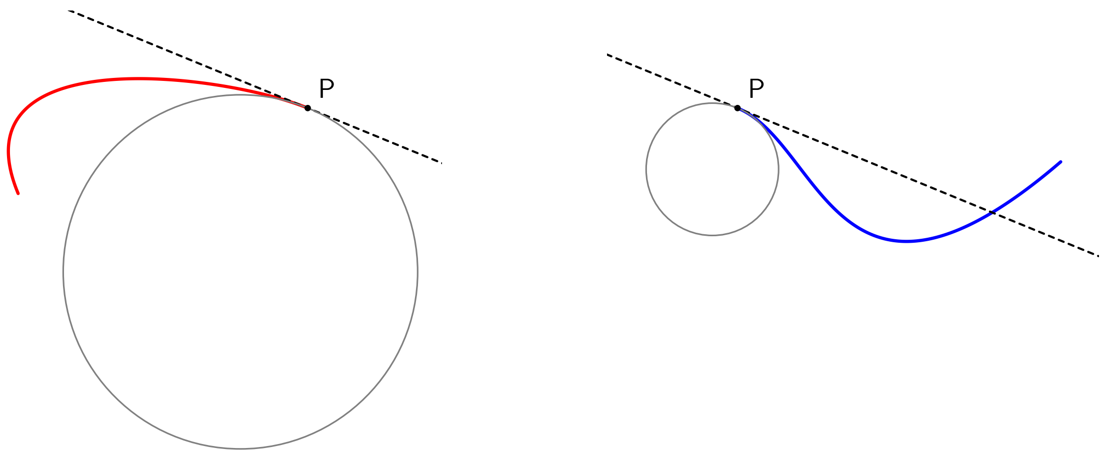
4 Curvatura e raccordi

Per diversi motivi le curve che interessano le applicazioni (anche nel design) sono costituite da archi di curve parametriche (cioè curve date "a pezzi") che si raccordano in maniera opportuna: abbiamo già visto precedentemente esempi di archi che non si raccordano affatto, oppure che si raccordano semplicemente in modo da non avere "strappi", o infine che si raccordano in modo da conservare la tangente (raccordi "lisci"). Nella applicazioni di solito servono raccordi ancora più lisci e, in particolare, che almeno conservino la curvatura.

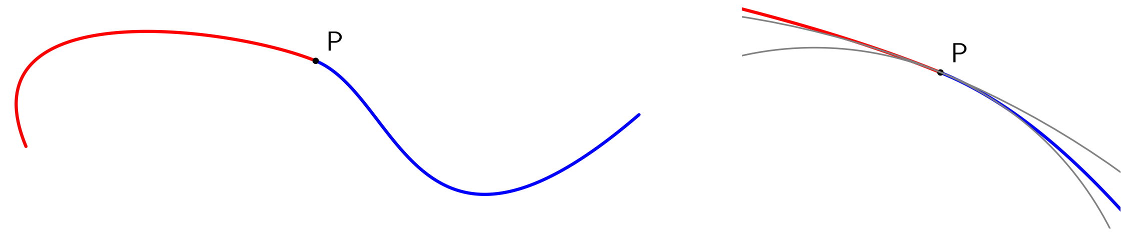
La figura seguente mostra una curva costruita con tre archi che si raccordano nei punti P e Q in modo da mantenere la curvatura: sono mostrate le circonferenze osculatrici nei punti di saldatura dei tre archi: queste circonferenze sono sia la circonferenza osculatrice finale di un arco che la circonferenza osculatrice iniziale dell'arco successivo.



La figura seguente illustra due archi di curva che si devono raccordare nel punto P: la tangente nel punto finale del primo arco e nel punto iniziale del secondo coincidono, mentre le due circonferenze osculatrici, e quindi le due curvature, sono diverse.



La figura seguente mostra i due archi saldati e uno zoom nei pressi del punto di saldatura, con evidenziate anche le due circonferenze oscultarici.

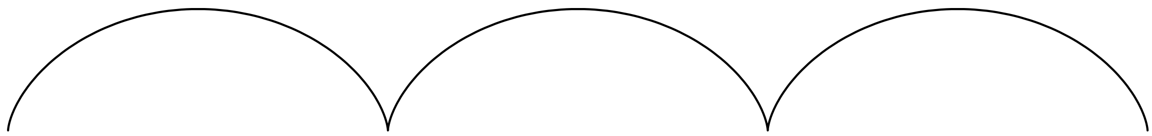


5 Qualche esempio di curva parametrica interessante

— La *Cicloide*, di equazioni parametriche

$$x = rt - r\sin t \quad , \quad y = r - r\cos t \quad , \quad 0 \le t \le 6\pi,$$

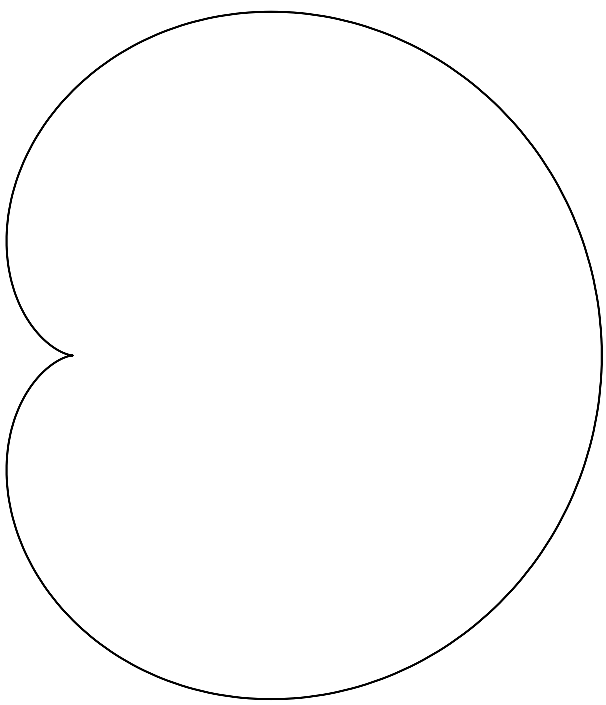
dove r è il raggio della circonferenza che rotola senza strisciare su una retta.



— La *Cardioide*, di equazioni parametriche

$$x = r(1 - 2\cos t + \cos 2t) \quad , \quad y = r(2\sin t - \sin 2t) \quad , \quad 0 \le t \le 2\pi,$$

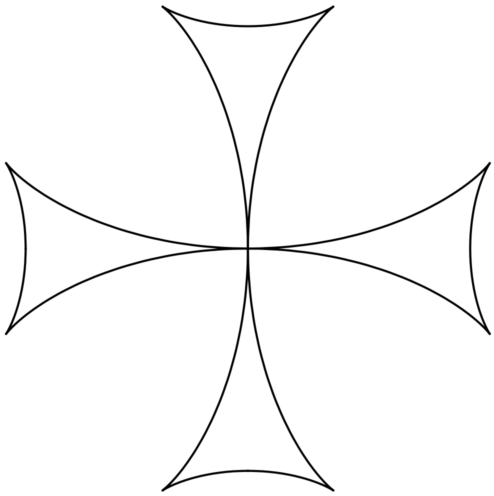
dove r è il raggio comune delle due circonferenze di cui una rotola sull'altra.



— La *Croce di Malta*, costituita da due rami, uno orizzontale e uno verticale, di equazioni parametriche

$$x = \begin{cases} a\cos t\left(\cos^2 t - 2\right), & \text{se } 0 \le t \le 2\pi \\ a\sin t\cos^2(t), & \text{se } 2\pi \le t \le 4\pi \end{cases} \quad , \quad y = \begin{cases} a\sin t\cos^2(t), & \text{se } 0 \le t \le 2\pi \\ a\cos t\left(\cos^2 t - 2\right), & \text{se } 2\pi \le t \le 4\pi \end{cases} ,$$

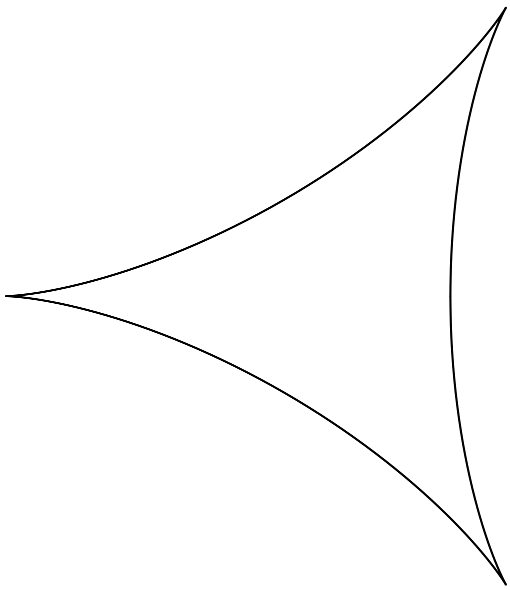
dove a è un parametro di dimensione (positivo). Si noti che i due rami sono ottenuti semplicemnte scambiando la x con la y.



— La *Deltoide*, di equazioni parametriche

$$x = a(2\cos t - \cos 2t) \quad , \quad y = a(2\sin t + \sin 2t) ,$$

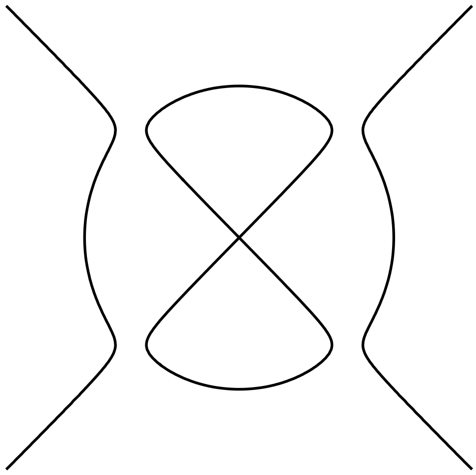
dove a è un parametro (positivo) di ampiezza.



— La *Curva del diavolo*, di equazioni parametriche

$$x = \cos t\sqrt{\frac{a^2\sin^2 t - b^2\cos^2 t}{\sin^2 t - \cos^2 t}} \quad , \quad y = \sin t\sqrt{\frac{a^2\sin^2 t - b^2\cos^2 t}{\sin^2 t - \cos^2 t}} \quad , \quad 0 \le t \le 2\pi,$$

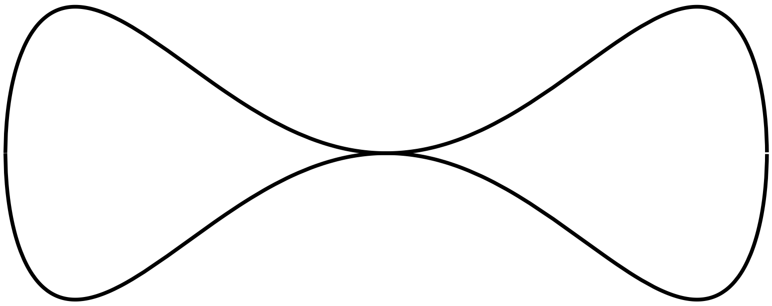
dove a e b sono due parametri di forma. Con $a = \sqrt{96}$ e $b = 10$ si ottiene la forma seguente, detta anche *motore elettrico*, per la sua forma simile al rocchetto rotante di un motore elettrico.



— La curva a *Manubrio*, o doppia goccia d'acqua, costituita da due rami, di equazioni parametriche

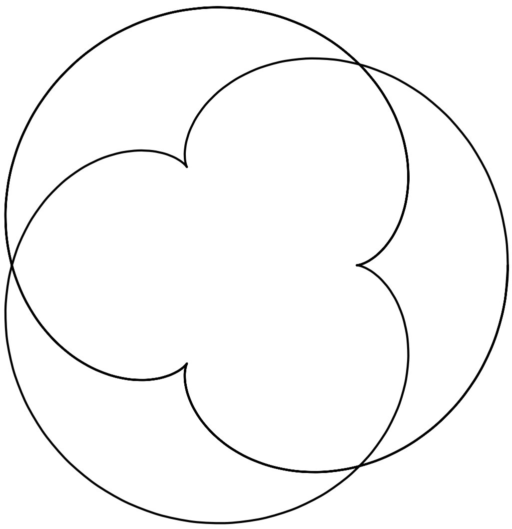
$$x = at \quad , \quad y = \pm a t^2 \sqrt{1-t^2} \quad , \quad -2\pi \le t \le 2\pi ,$$

dove a è un parametro di forma.



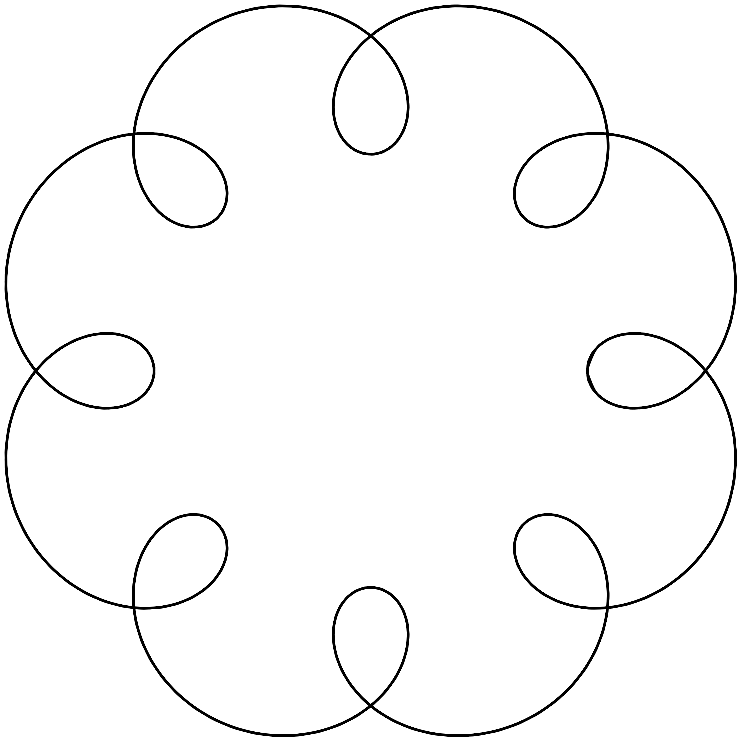
— Una epicicloide, precisamente l'*Epicicloide a tre punte*, di equazioni parametriche

$$x = 5\cos t - 2\cos(2.5t) \quad , \quad y = 5\sin t - 2\sin(2.5t) \quad , \quad 0 \le t \le 6\pi .$$



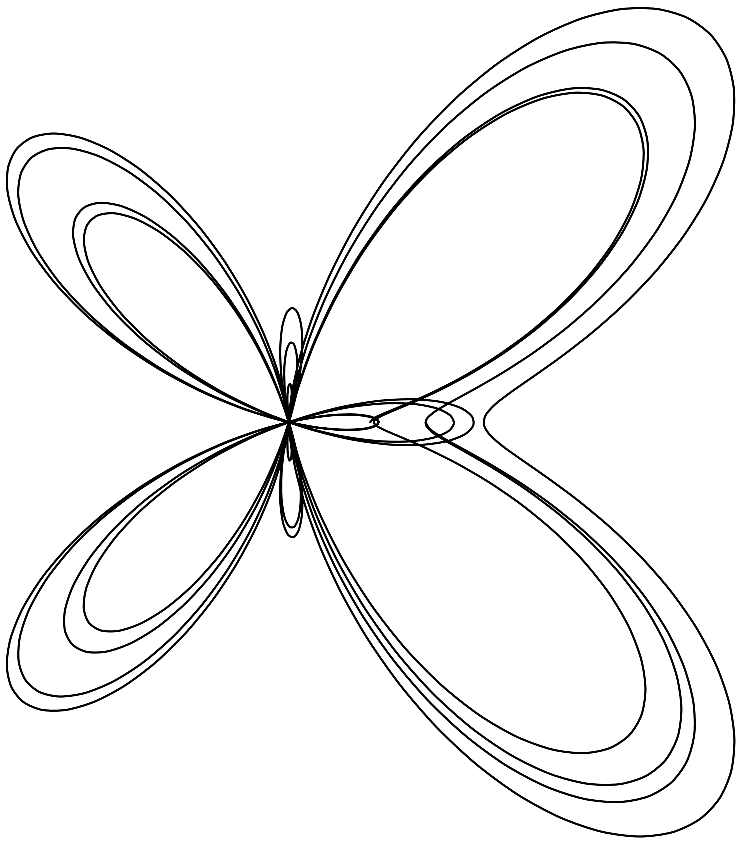
— Una epitrocoide, precisamente l'epitrocoide a otto lobi, di equazioni parametriche

$$x = 18\cos t - 5\cos(9t) \quad , \quad y = 18\sin t - 5\sin(9t) \quad , \quad 0 \le t \le 6\pi .$$



— Una curva a *farfalla* di equazioni parametriche

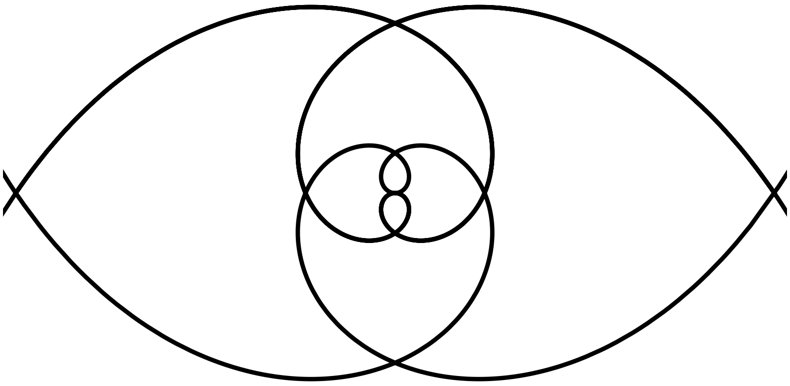
$$x = \cos t\left(e^{\cos t} - 2\cos 4t + \sin^5\frac{t}{12}\right) \quad , \quad y = \sin t\left(e^{\cos t} - 2\cos 4t + \sin^5\frac{t}{12}\right) \quad , \quad 0 \le t \le 8\pi.$$



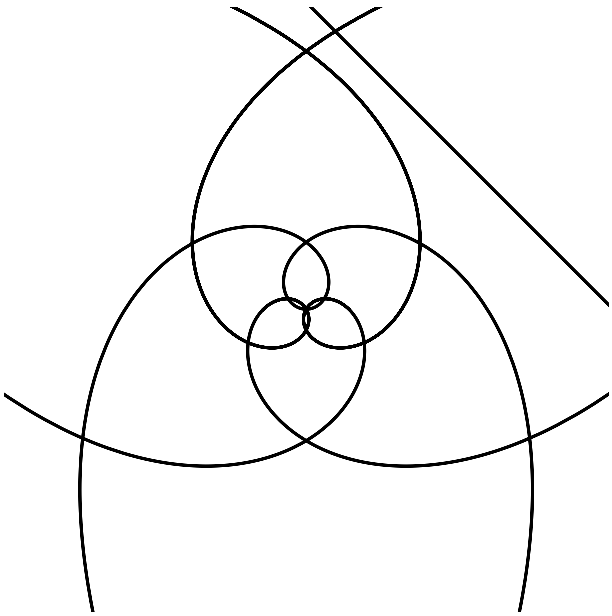
— Curve *Nodali*, di equazioni parametriche

$$x = 10\tan(nt)\cos t \quad , \quad y = 10\tan(nt)\sin t \quad , \quad -7\pi \leq t \leq 7\pi$$

dove n è un parametro reale. Con $n = 0.2$ si ottiene la figura seguente.

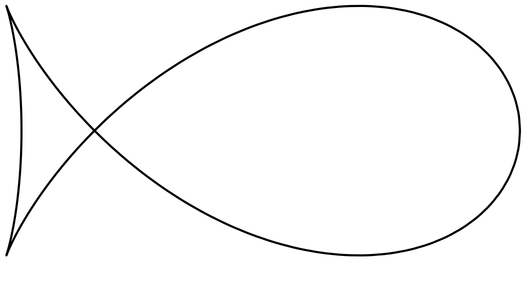


Con $n = 0.3$ si ottiene la figura seguente.



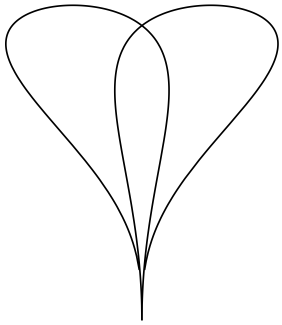
— Curva a *Pesce*, di equazioni parametriche

$$x = 3\left(\cos t - \frac{\sin^2 t}{\sqrt{2}}\right) \quad , \quad y = 3\sin t\cos t \quad , \quad 0 \leq t \leq 2\pi.$$



— Una curva a *Cuore*, di equazioni parametriche

$$x = \sin t \cos t \ln|t| \quad , \quad y = |t|^{3/10}\sqrt{\cos t} \quad , \quad -\pi/2 \leq t \leq \pi/2.$$



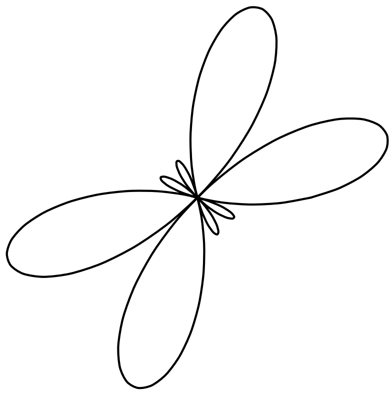
5.1 Curve in coordinate polari

Una curva in coordinate polari è della forma $r = f(\vartheta)$, dove r e il raggio e ϑ l'angolo. Si può facilmente passare dall'equazione in coordinate polari alle normali equazioni parametriche con le formule

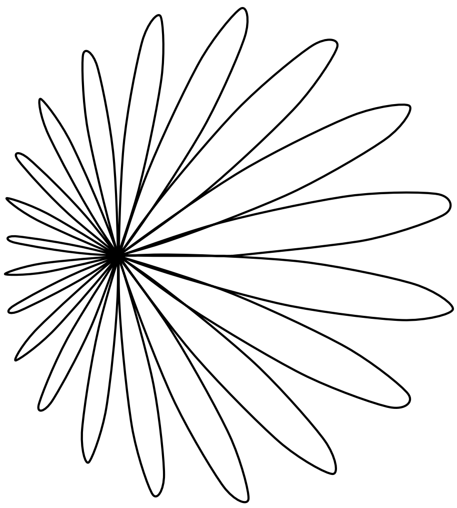
$$x = r\cos\vartheta = f(\vartheta)\cos\vartheta \quad , \quad x = r\sin\vartheta = f(\vartheta)\sin\vartheta.$$

Negli esempi che seguono useremo t al posto di ϑ per semplicità.

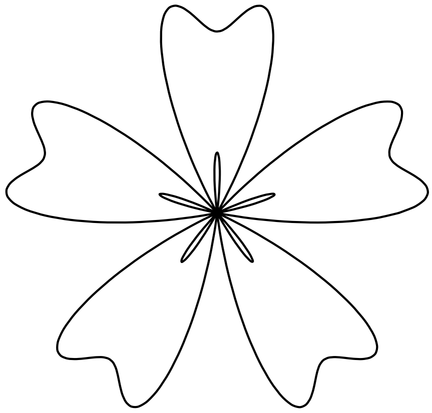
— $r = cos(2t) + sin(4t)$.



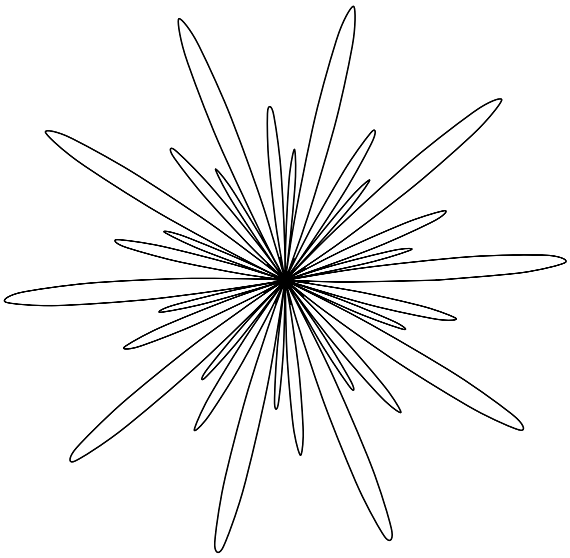
— $r = \cos(2t) + 2\sin(10t)$.



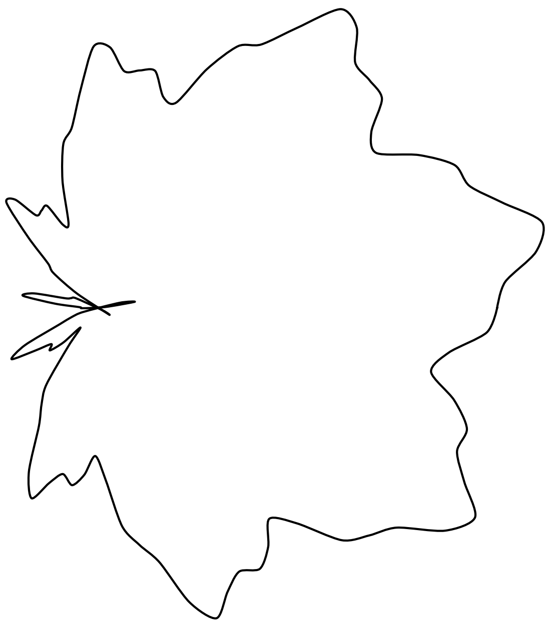
— $r = 2 + \cos(10t) + 2\sin(5t)$



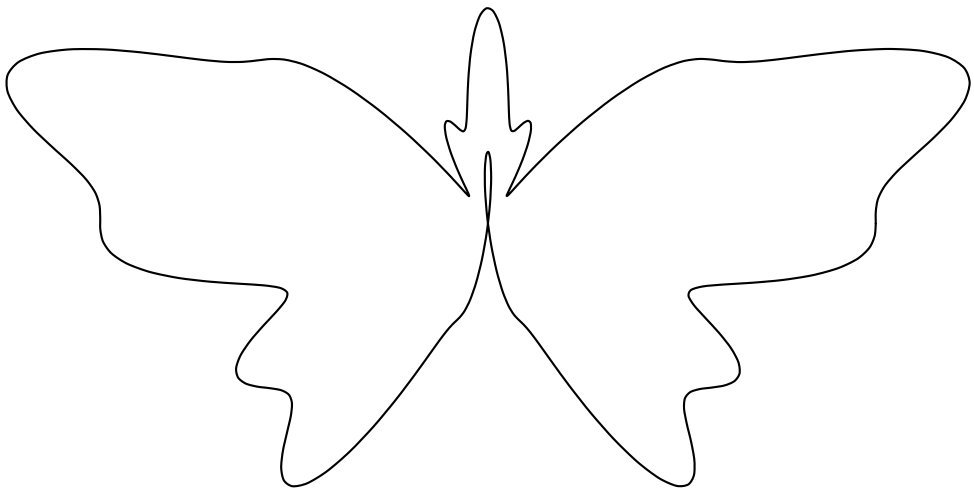
— $r = 1 + \cos(10t) + 2\sin(20t)$



— $r = 5 + 4.6\cos(t) + 1.4\sin^3(9t)$



— $r = 1.6 + 1.1\cos(2t) + \sin^3(5t)$



6 Codici Geogebra

Geogebra è in grado di rappresentare efficacemente curve in equazioni parametriche o polari, con i codici rispettivi seguenti.

1. Curve in equazione parametrica:

   ```
   Curva(<espr.x>,<espr.y>,<parametro>,<valore iniziale>,<valore finale>)
   ```

   dove `<espr.x>` e `<espr.y>` sono le formule per ottenere la coordinata x e la coordinata y rispettivamente, `<parametro>` è il nome che si usa per il parametro (di solito t), `<valor iniziale>` e `<valore finale>` i valori iniziali e finale del parametro. Per esempio per la curva

   $$x = a(2\cos t - \cos 2t) \quad , \quad y = a(2\sin t + \sin 2t)\,,$$

   si deve scrivere

   ```
   Curva(a(2 cos(t)-cos(2t)),a(2 sin(t)+sin(2t)),t,0,2pi)
   ```

   dopo aver introdotto il valor di a con un slider o come numero. Ricordarsi di racchiudere le variabili del cos e del sin tra parentesi, come per ogni altra funzione.

2. Curve in equazione polare:

   ```
   Curva((<espr.r>;<parametro>),<parametro>,<valore iniziale>,<valore finale>)
   ```

   dove `<espr.r>` è l'espressione per r, tutto il resto è come prima. Fare attenzione al punto e virgola in `(<espr.r>;<parametro>)`, infatti Geogebra usa il punto e virgola per identificare le coordinate polari. Per esempio per la curva

   $$r = 1.6 + 1.1\cos(2t) + \sin^3(5t)$$

   si deve scrivere

   ```
   Curva((1.6+1.1 cos(2t)+(sin(5t))^3;t),t,0,2pi)
   ```

3. Curve "a pezzi". Nel caso di curve "a pezzi" occorre usare il costrutto logico `Se( <Condizione>, <Allora>, <Altrimenti> )`, per le espressioni per la x, la y o la r rispettivamente. Se ci sono solo due "pezzi", basta questa semplice costruzione, se ce ne sono più di due occorre annidare più costrutti `Se`. Vediamo il tutto su esempi.

Per la curva

$$x = \begin{cases} t^3 - 2t, & \text{se } 0 \le t \le 1 \\ t^2, & \text{se } 1 < t \le 2 \end{cases} \qquad y = \begin{cases} t^2, & \text{se } 0 \le t \le 1 \\ t^2 - 2t, & \text{se } 1 < t \le 2 \end{cases}$$

il codice è

```
Curva(Se(t<=1,t^3-2t,t^2),Se(t<=1,t^2,t^2-2t),t,0,2)
```

Per la curva

$$x = \begin{cases} t^3 - 2t, & \text{se } 0 \le t \le 1 \\ t^2, & \text{se } 1 < t \le 2 \\ -t, & \text{se } 2 < t \le 3 \end{cases} \qquad y = \begin{cases} t^2, & \text{se } 0 \le t \le 1 \\ t^2 - 2t, & \text{se } 1 < t \le 2 \\ t^2 & \text{se } 2 < t \le 3 \end{cases}$$

il codice è

```
Curva(Se(t<=1,t^3-2t,Se(t<=2,t^2,-t)),Se(t<=1,t^2,Se(t<=2,t^2-2t,t^2)),t,0,3)
```

4. Per la curva

$$r = \begin{cases} \sin(t), & \text{se } 0 \le t < \pi/2 \\ \cos(t), & \text{se } \pi/2 \le t \le \pi \end{cases}$$

Il codice è

```
Curva((Se(t <=pi/2, sin(t), cos(t)); t), t, 0, pi)
```

Naturalmente in questi casi sui possono anche tracciare separatamente i vari pezzi come curve diverse, cosa che è indispensabile per esempio se si vogliono usare stile diversi per i vari pezzi (spessori diversi, colori diversi, ecc.). Logicamente però, in particolare quando i vari pezzi si devono saldare in modo "liscio", è meglio pensare ad una unica formula per definire la curva.

7 Curve per il CAD

Le curve utilizzate nel CAD devono avere una serie di proprietà tra le quali elenchiamo:

— sia semplice la valutazione numerica e la determinazione delle proprietà analitiche;
— la modifica possa essere immediata ed intuitiva ed abbia effetti locali;
— sia possibile avere un controllo locale della forma, specialmente nei punti di maggior interesse per il progettista;
— sia possibile risolvere problemi di approssimazione o interpolazione di nuvole di punti.

Le curve che interessano possono essere grossolanamente divise in due categorie:

1. curve *approssimanti*, ovvero che riproducano l'andamento di una serie di punti senza necessariamente passare per essi (o almeno per tuti);
2. curve *interpolanti*: curve che passano per un dato insieme di punti.

Le curve più comuni sono curve *polinomiali a tratti*, in quanto sono di abbastanza semplice valutazione e modifica, specie se di grado non elevato, e risolvono sia problemi di approssimazione che di interpolazione. In genere si richiede che le curve abbiano un grado di regolarità così definito:

1. C^0: due tratti successivi devono avere un estremo in comune;
2. C^1: due tratti successivi devono avere anche la tangente in comune nel punto di congiunzione;
3. C^2: due tratti successivi devono avere anche la curvatura in comune nel punto di congiunzione.

Si possono richiedere anche regolarità di grado più elevato, ma normalmente queste sono le più comuni.

Le implementazioni più semplici di curve con molte di queste caratteristiche, anche se non tutte, sono le curve di *Bézier* e le *Spline*, di cui ci occuperemo brevemente.

8 Curve di Bézier

L'idea che sta alla base delle curve di Bézier[2] è di cercare di definire una curva a partire da un numero finito di punti P_0, P_1, P_2, ..., P_n, in modo che segua l'andamento della poligonale che ha per vertici questi punti.

Le curve di Bézier non hanno il problema della regolarità nei punti di raccordo, in quanto sono costituite da un unico ramo: per contro hanno un grado che dipende dal numero dei punti, precisamente sono di grado n se si hanno $n+1$ punti.

Senza entrare nei dettagli della loro costruzione geometrica, riportiamo la formula generale dell'equazione parametrica di una curva di Bézier definita da $n+1$ punti $\mathsf{P}_0, \mathsf{P}_1, \mathsf{P}_2, \ldots, \mathsf{P}_n$:

$$B(t) = \sum_{i=0}^{n} \binom{n}{i} \mathsf{P}_i (1-t)^{n-i} t^i.$$

Siccome i punti P_i hanno due coordinate (se lavoriamo nel piano come noi supponiamo qui), la formula comprende sia l'espressione per la x che per la y, mettendo successivamente al posto di P_i prima la x e poi la y di P_i.

Per $n = 2$ si ottengono le curve di Bézier *lineari*:

$$B_2(t) = (1-t)\mathsf{P}_0 + t\mathsf{P}_1 \quad \text{ovvero} \quad \begin{cases} x = (1-t)x_0 + tx_1 \\ y = (1-t)y_0 + ty_1 \end{cases},$$

e questo è semplicemente il segmento per P_0 e P_1. Per il codice Geogebra non occorre separare la x dalla y, si può direttamente digitare quanto segue[3], dove abbiamo usato A e B per i nomi dei punti (che naturalmente devono essere preventivamente costruiti!), per semplicità:

```
Curva((1-t)A+t B,t,0,1)
```

Prestare attenzione allo spazio tra t e B, altrimenti Bt viene interpretato come un unico oggetto e il codice fornisce errore.

Per $n = 3$ si ottengono le curve di Bézier *quadratiche*:

$$B_3(t) = (1-t)^2\mathsf{P}_0 + 2t(1-t)P_1 + t^2\mathsf{P}_2.$$

Il relativo codice Geogebra è:

```
Curva((1-t)^2 A+2 t (1-t)B+t^2 C,t,0,1)
```

[2]Pierre Bézier (1910-1999), ingegnere e matematico francese, sviluppò la teoria delle curve che portano il suo nome durante il suo lungo periodo di lavoro alla Renault, in particolare studiate per disegnare le carrozzerie delle automobili. Sviluppò UNISURF, uno dei software pionieristici nel CAD/CAM, progettato per assistere i progettisti d'auto.

[3]Questo fatto è legato alle proprietà delle operazioni sui vettori, di cui noi non ci siamo occupati.

Numerosi font *TrueType* utilizzano curve di Bézier quadratiche.

Per $n = 4$ si ottengono le curve di Bézier cubiche, tra le più usate:

$$B_4(t) = (1-t)^3 \mathsf{P}_0 + 3t(1-t)^2 \mathsf{P}_1 + 3t^2(1-t)\mathsf{P}_2 + t^3 \mathsf{P}_3,$$

con codice Geogebra

```
Curva((1-t)^3A+3t (1-t)^2B+ 3t^2(1-t)C+t^3D,t,0,1)
```

Anche se è ormai chiaro come si procede con l'aumentare del numero di punti esplicitiamo in dettaglio il caso con $n = 5$ e il relativo codice Geogebra.

$$B_5(t) = (1-t)^4 P_0 + 4t(1-t)^3 P_1 + 6t^2(1-t)^2 P_2 + 4t^3(1-t)P_3 + t^4 P_4$$

```
Curva((1-t)^4 A+4 t (1-t)^3 B+6 t^2 (1-t)^2 C+4 t^3 (1-t) D+ t^4 E,t,0,1)
```

I polinomi

$$\binom{n}{i}(1-t)^{n-i}t^i$$

sono conosciuti come *polinomi di base di Bernstein*.

I punti P_0, P_1, P_2, $\dots$, P_n sono chiamati *punti di controllo*, mentre il poligono formato connettendo i punti di controllo si chiama *poligono di Bézier*: la curva di Bézier è contenuta nella chiusura convessa di questo poligono.

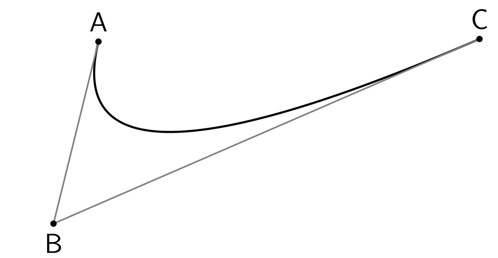
Alcune proprietà delle curve di Bézier sono elencate di seguito.

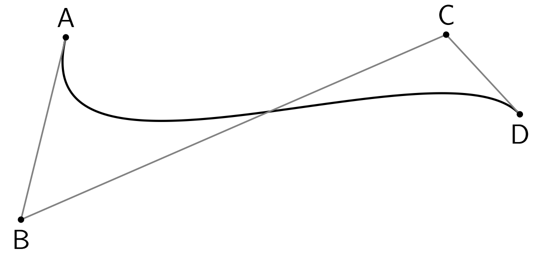
1. La curva di Bézier inizia sempre nel punto P_0 e termina nel punto P_n.
2. Le tangenti nel punto iniziale e, rispettivamente, nel punto finale coincidono con il primo e l'ultimo lato del poligono di Bézier.
3. Un cerchio o una sua parte non possono mai essere rappresentati con curve di Bézier.

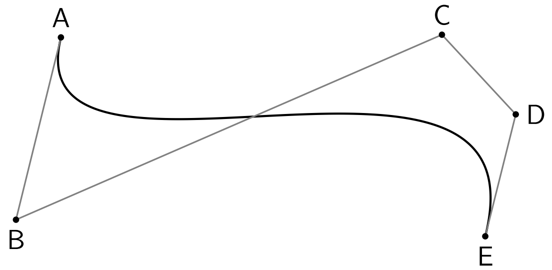
I problemi maggiori nell'uso delle curve di Bézier sono legati a due fatti

— il grado della curva aumenta all'aumentare del numero di punti;
— la variazione anche di un solo punto può comportare modifiche sostanziali sull'intera curva: si dice che queste curve non ammettono un controllo locale.

Proponiamo alcuni esempi di curve di Bézier con tre, quattro o cinque punti ci controllo.







Un miglioramento dell'efficienza delle curve di Bézier si ha utilizzando le *curve razionali di Bézier*. In sostanza si tratta di introdurre dei pesi da assegnare ai singoli punti, in modo da modificare il loro "grado di attrazione" per la curva. Ad ogni punto P_0, P_1, P_2, …, P_n viene assegnato un peso w_1, w_2, …, w_n e la forma generale dell'equazione della curva si modifica in

$$RB(t) = \frac{\sum_{i=0}^{n} w_i \binom{n}{i} \mathsf{P}_i (1-t)^{n-i} t^i}{\sum_{i=0}^{n} w_i \binom{n}{i} (1-t)^{n-i} t^i},$$

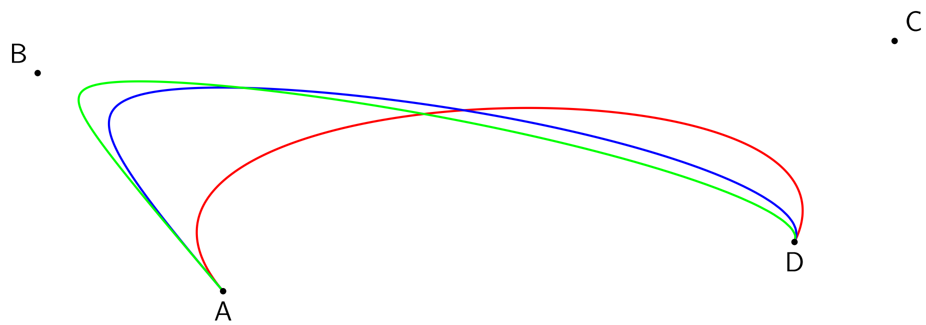
dove il denominatore è stato introdotto per "normalizzare" i pesi introdotti. Per esempio nel caso di quattro punti, con pesi $\{w_0, w_1, w_2, w_3\}$ si ottiene:

$$RB_4(t) = \frac{w_0(1-t)^3\mathsf{P}_0 + w_1 3t(1-t)^2\mathsf{P}_1 + w_2 3t^2(1-t)\mathsf{P}_2 + w_3 t^3\mathsf{P}_3}{w_0(1-t)^3 + w_1 3t(1-t)^2 + w_2 3t^2(1-t) + w_3 t^3},$$

con relativo codice Geogebra, dove abbiamo indicato i pesi con a, b, c, d e i punti con A, B, C, D per semplicità:

```
Curva((a(1-t)^3A+3b t (1-t)^2B+ 3c t^2(1-t)C+d t^3D)/
                        (a(1-t)^3+3b t (1-t)^2+ 3c t^2(1-t)+d t^3),t,0,1)
```

Per esempio nella figura seguente sono rappresentate le curve corrispondenti ai pesi $\{1,1,1,1\}$, $\{1,5,1,1\}$, $\{1,10,1,1\}$ con quattro punti di controllo: come si vede il valore di "attrazione" del punto B è via via crescente.



Il nome di curve *razionali* deriva dal fatto che le equazioni non sono semplici polinomi ma rapporti di polinomi, ovvero funzioni razionali.

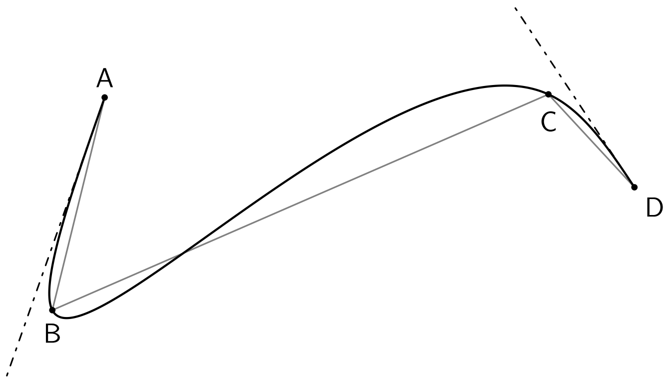
9 Spline

Spline è un termine inglese che, letteralmente, significa striscia di metallo o legno: infatti in origine una spline era uno strumento da disegno in materiale sufficientemente elastico per tracciare, su carta, curve di interpolazione per dati punti.

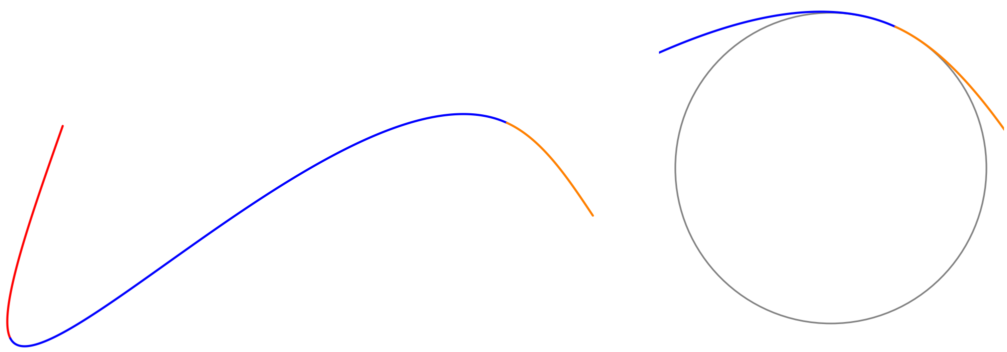
Dal punto di vista matematico sono curve polinomiali, definite a tratti, in modo che gli archi di curva che si succedono passino per per un dato insieme di punti, detti *nodi*, e che i raccordi tra un arco e il successivo siano opportunamente lisci, in generale mantenendo la tangente e la curvatura.

Le spline più comuni sono quelle cubiche, cioè quelle in cui ogni arco è un polinomio di grado 3. Le condizioni di interpolazione sui nodi e di regolarità sugli stessi non sono sufficienti a determinare in maniera univoca la spline interpolante diversi punti. Occorrono ulteriori condizioni: per esempio nell'implementazione dell'algoritmo per costruire le spline in Geogebra la scelta è quella di imporre che agli estremi iniziale e finale della curva la curvatura sia nulla, ma ci sono altre possibilità.

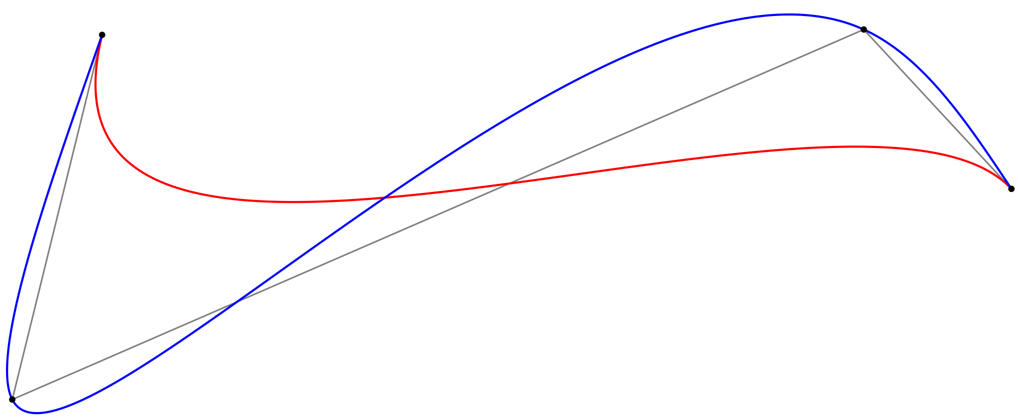
La figura seguente illustra una spline, ottenuta con Geogebra, per 4 nodi: si può notare, anche visivamente, che nei punti A e B la curva è praticamente coincidente con la tangente, cioè è praticamente rettilinea, e quindi con curvatura nulla.



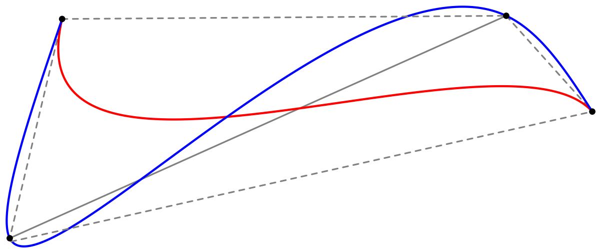
La spline è stata ottenuta unendo tre rami polinomiali cubici. La figura seguente mostra i tre rami con tre diversi colori ed uno zoom nei pressi del secondo punto di saldatura, con evidenziata la circonferenza osculatrice.



La figura seguente mostra sia la Bézier cubica che la spline cubica con gli stessi 4 punti.



Si noti che, come già osservato, la curva di Bézier è interna alla chiusura convessa del poligono di controllo, mentre la spline no.



Il codice Geogebra per ottenere queste figure è elementare, in quanto Geoegbra ha un apposito comando per ottenere questo tipo di spline:

```
Spline({A,B,C,D},3)
```

Se al posto di 3 si scrive 4 si ottiene una spline di grado 4, ancora più "liscia" nei raccordi, cosa che raramente si usa nelle applicazioni. Se si hanno più di quattro punti basta inserirli, nell'ordine giusto, all'interno delle parentesi graffe che, in Geogebra, identificano le liste di oggetti.

Purtroppo anche le spline costruite in questo modo hanno alcuni inconvenienti, tra cui quello che la modifica di un punto modifica l'intera curva, anche se con effetto meno drastico di quanto non avvenga con le curve di Bézier.

10 Verso le NURBS

La risoluzione dei problemi accennati per le curve di Bézier e le spline è demandata alle cosiddette curve NURBS (Non Uniform Rational Basis Splines). Non possiamo entrare nei dettagli matematici di queste curve, troppo complessi per il nostro corso. Segnaliamo solo che esse possono essere considerate una generalizzazione delle curve di Bézier e delle spline che abbiamo visto. La formula matematica generale non è molto diversa da quella che abbiamo visto per le curve di Bézier razionali, solo che al posto dei polinomi di base di Bernstein compaiono diverse funzioni di base la cui definizione avviene per ricorsione.

Solo a titolo di esempio proponiamo le equazioni, direttamente in formato Geogebra, che si ottengono per un quarto di cerchio di raggio 1, con 3 e rispettivamente 5 punti di controllo. Nel primo caso le equazioni sono rapporti di polinomi di grado 2, nel secondo caso rapporti di polinomi di grado 4.

```
Curva((sqrt(2)(1-u)^2+2u (1-u))/(sqrt(2)(1-u)^2+2u (1-u)+sqrt(2) u^2),
           (2u (1-u)+sqrt(2)u^2)/(sqrt(2)(1-u)^2+2u (1-u)+sqrt(2)u^2),u,0,1)

 Curva(((1-u)^4+sqrt(2)(1-u)^2 u^2)/((1-u)^4+sqrt(2)(1-u)^2 u^2+u^4),
           (sqrt(2)(1-u)^2 u^2+u^4)/((1-u)^4+sqrt(2)(1-u)^2 u^2+u^4),u,0,1)
```

Si confrontino queste equazioni con quelle, semplicissime dal punto di vista matematico, che si ottengono utilizzando le funzioni seno e coseno:

```
Curva(cos(t),sin(t),t,0,2pi)
```