UNIVERSITAT POLITÈCNICA DE CATALUNYA

APORTACIÓN AL ESTUDIO DEL SOFTWARE NECESARIO PARA LA INFORMATIZACIÓN DE LOS MÉTODOS DE APRENDIZAJE DE LAS TÉCNICAS DE EXPRESIÓN GRÁFICA, Y SU POSTERIOR IMPLEMENTACIÓN

Autor: Miquel Castillo i Ballarda
Director: Jordi Mestres i Sardà

1988

Introducción.

En el capitulo anterior hemos estudiado el caso de las cónicas y las secciones planas de las cuádricas como curvas de las que podemos hallar su ecuación; veremos en este capitulo como se tratarán las curvas cuando no tienen una ecuación preestablecida. Al respecto puede ser interesante reproducir, aunque sólo sea parcialmente, la introducción al tratamiento de curvas planas que realiza un clásico del Computer Graphics, el texto de David Rogers (1):

> "Hoy en dia existen multitud de técnicas para el dibujo manual y el diseño de curvas. Los instrumentos a usar los encontramos entre una amplia variedad de plumas, lápices, pinceles, etc.(.....) pueden ser usados para la construcción de curvas. En Computer Graphics, también existen gran cantidad de técnicas diferentes, asi como herramientas tanto para la generación como para la presentación gráfica (display) de las curvas.(......)"

> "(......). Una curva puede venir representada por una colección de puntos, de forma que están tanto más juntos cuanto más pequeño es el radio de curvatura que los une. De

todas formas, existen varias razones por las que una representación matemática de una curva sustituye con ventaja a la representación de los mismos como una malla de puntos. Varias de estas razones son las siguientes:"

"1.- Una representación matemática es precisa, y las propiedades de la curva tales como la pendiente y el radio de curvatura son facilmente calculables."

"2.- Una representación matemática de una curva puede ser compactamente almacenada en un ordenador."

"3.- Los dibujos de un objeto representado matemáticamente en un ordenador pueden ser facilmente producidos."

"4.- Cuando tenemos curvas analiticamente definidas sobre una región de puntos en estudio, no es necesario realizar operaciones de interpolación para determinar (to compute en el original) puntos intermedios."

"5.- El uso de puntos para representar curvas se ha demostrado inoperante cuando queremos alterar con-

> tinuamente la forma de dichas curvas según un cierto criterio de diseño. Una representación matemática de las curvas se ha demostrado mucho más operativo para este tipo de curvas."

El problema de ajustar una curva a un conjunto de puntos dados tiene, desde hace tiempo, muchas soluciones. Es necesario, no obstante, distinguir entre dos grandes tipos de formas de acercarse a dicho estudio:

- Si pretendemos encontrar una curva que pase por el conjunto de puntos dados, se trata de un problema de interpolación.

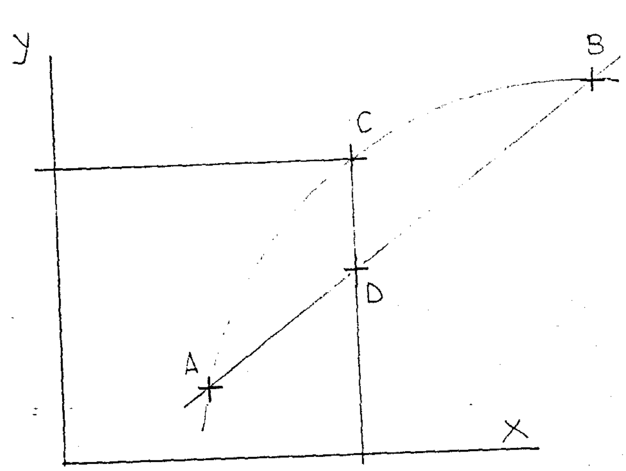
- Si pretendemos encontrar una curva que pase **cerca** del conjunto de puntos dados, el problema es de aproximación.

Paradójicamente, este segundo caso será el que se usará con más profusión; hablando un poco tópicamente de los problemas de diseño de la carroceria de un coche, de la ala de un avión o de la quilla de un barco, desde un punto de vista matemático será sencillo encontrar las curvas que interpolen los puntos obtenidos mediante experimentación, pero la imagen que nos ofrecen es demasiado realista, sobre todo si la obtención de los puntos, como es probable, está sujeta a una serie más o menos grande de errores. Normalmente, se adopta un compromiso que consiste en escoger un conjunto de puntos directores, que pueden ser interactivamente modificados por el usuario, de forma que la curva pase lo más cerca posible de ellos. Entre las curvas que permiten este tipo de aproximación se encuentran los Splines y las curvas de Bezier.

A pesar de ello, y sin pretender profundizar excesivamente en el tema considero que el orden lógico de estudio de esta problemática del tratamiento de curvas debe ser enfocado empezando por los polinómios de Lagrange (2).

Interpolación polinómica.

El concepto más elemental de interpolación es la interpolación lineal, sea la curva de la figura:



Estudio de curvas. Segunda parte.

Tenemos dos puntos A, B sobre una curva, una interpolación lineal tendría como resultado trabajar con el punto D en vez de con C, veámos por que.

La curva está referida a unos ejes X e Y coordenados, el hecho de asociarla con una linea recta quiere decir que consideramos que la función implicita $y=f(x)$ durante el intervalo estudiado es proporcional a x. Ya se ve que la aproximación es grosera pues aunque la linea interpoladora pasa por A y B, como pretendemos, la diferencia entre C y D puede ser inadmisible en muchos casos; lo que habitualmente se hace es escoger, despues de un estudio del problema a representar, n puntos en vez de dos, como en el ejemplo, y hacer pasar por ellos un polinomio de grado n-1. Será este polinomio el que nos permitirá saber el valor que toma la función en cada punto que nos interese, veámos como:

Sean X_0, X_1, X_2, X_3,......,X_n los puntos por los que queremos obligar que pase el polinomio interpolador, de dichos puntos conocemos, lógicamente, sus ordenadas Y_i, se deberá cumplir $Y_i=P(X_i)$ en cada punto. La expresión del polinomio es:

$$
\begin{aligned}
P(X)=&C_0.(X-X_1).(X-X_2).(X-X_3)....(X-X_{n-1}).(X-X_n)+\\
&C_1.(X-X_0).(X-X_2).(X-X_3).........(X-X_{n-1}).(X-X_n)+\\
&C_2.(X-X_0).(X-X_1).(X-X_3).........(X-X_{n-1}).(X-X_n)+\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&Cn.(X-X_0).(X-X_1).(X-X_2).........(X-X_{n-2}).(X-X_{n-1})
\end{aligned}
$$

El conocimiento de los valores de los coeficientes C_i nos dará la expresión del polinomio que estamos buscando. La forma de hallarlos es imponer que se cumpla, como he indicado antes, que: $Y_0=P(X_0)$; $Y_1=P(X_1)$;......;$Y_n=P(X_n)$, lo que no reviste ninguna dificultad especial, veámos de obtener por ejemplo C_5:

imponiendo en la expresiòn anterior que X=X5 tendremos:

$$C_5= Y_5/(X_5-X_0).(X_5-X_1).......(X_5-X_n)$$

Se trataria ahora de hallar la ecuaciòn general, de hecho, la ley de recurrencia, que nos permita hallar el polinomio de Lagrange en cualquier caso y de cualquier grado.

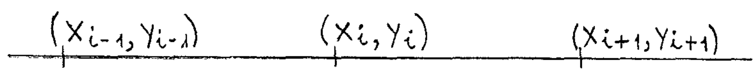
Otros autores parten directamente de una ecuaciòn implicita del tipo:

$$A.X^2+2.B.XY+CY^2+2.CY+F=0$$

de la cual una serie de condiciones de contorno, o la tècnica de los minimos cuadrados, nos permitiràn encontrar los valores de las constantes que la definen.

Finalmente, y para acabar el tema de la Interpolaciòn, refiràmonos a los Splines Cùbicos. El nombre de Spline es familiar a los diseñadores anglo-sajones como la plantilla de madera que les permitia dibujar una curva pasando por una serie de puntos dados con la caracteristica de que dicha curva fuera lo màs suave posible. La energia de flexiòn de la curva que proporcionaba el spline era la minima posible. De ahi el nombre de spline para los instrumentos matemáticos que vamos a comentar a continuación, veremos que vuelven a aparecer cuando hablemos de interpolación, de hecho cuando se habla de splines se considera que el problema de que se trata se refiere a interpolación, de ahi que se añada la palabra cúbico para diferenciarlos claramente.

El polinomio que usaremos será un cúbica, pero a diferencia de los casos vistos anteriormente, esta cúbica no será única, sino que trabajaremos con una distinta en cada intervalo. Se planteará, como es lógico, el problema de unir las cúbicas de intervalos contiguos. La solución que nos dará la continuidad deseada será que las pendientes y las curvaturas en los puntos comunes sean las mismas.



Sean los dos intervalor contiguos de la figura, la expresión de la cúbica es:

$$Y=A_i.(X-X_i)^3+B_i.(X-X_i)^2+C_i.(X-X_i)+D_i$$

llamando IN_i a la diferencia $X_{i+1}-X_i$, veàmos los valores que toma la cúbica en los extremos de dicho intervalo:

$$Y_i=A_i.(X_i-X_i)^3+B_i.(X_i-X_i)^2+C_i.(X_i-X_i)+D_i=D_i$$
$$Y_{i+1}=A_i.IN_i{}^3+B_i.IN_i{}^2+C_i.IN_i+D_i \quad (a)$$

vamos a imponer la igualdad de pendiente (PE) y curvatura (CU) en los extremos de los intervalos, sabiendo que son iguales, respectivamente a la primera y segunda derivadas de la ecuación de la cúbica.

$$PE=Y'=3.A_i.(X-X_i)^2+2.B_i.(X-X_i)+C_i$$
$$CU=Y''=6.A_i.(X-X_i)+2.B_i$$

en los extremos del intervalo, tomarán los valores:

$$CU_i=6.A_i.(X_i-X_i)+2.B_i=2.B_i$$
$$CU_i+1=6.A_i.IN_i+2.B_i$$

de donde se infiere que:

$$B_i=CU_i/2$$
$$\text{y } A_i=(CU_{i+1}-CU_i)/6.IN_i$$

para hallar el último coeficiente que nos falta sustituamos los que ya hemos obtenido en (a):

$$Y_{i+1}=((CU_{i+1}-CU_i)/6.IN_i).IN_i{}^3+$$
$$+(CU_i/2).IN_i{}^2+C_i.IN_i+Y_i$$

de donde:

$$C_i=((Y_{i+1}-Y_i)/IN_i)-$$
$$-(2.IN_i.CU_i-IN_i.CU_{i+1})/6$$

será con estos valores, y con las leyes de recurrencia que de ellos se infieran como se implementará el cálculo de los Splines Cúbicos.

Habitualmente se añaden algunas condiciones más, siempre en función del tipo de problema sobre el que se está diseñando, como son:

- Las curvaturas en el primer y en el último punto iguales entre si e iguales a cero. Lo que quiere decir que en sus extremos el Spline se acerca a una recta. Estas curvas reciben el nombre de Splines Naturales.

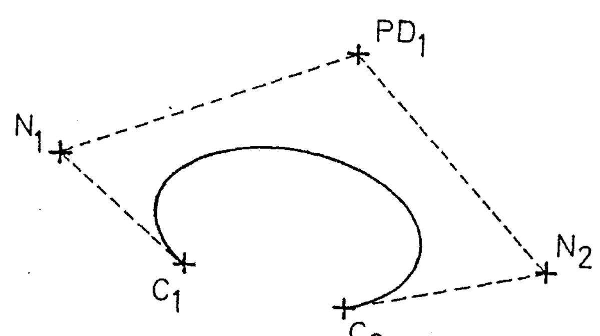
- Haciendo CU1=CU2 y CUn=CUn-1, imponemos que el Spline empiece y acabe de una forma parabólica.

Aproximación polinómica.

Dado que la relación existente entre las aproximaciones de Bezier y de los B-Splines guarda cierta similitud con las interpolaciones de Lagrange y los Splines Cúbicos, debido al carácter global de los primeros enfrente al no global de los segundos, el comentario de ambos lo haré en un orden discutiblemente anacrónico, pues historicamente los Splines vieron la luz primero, ya que los primeros estudios sobre ellos datan de 1946, aunque la puesta en funcionamiento en el ambiente del CAD fue alrededor de 1973, lo cual considerando que la factoria Renault usa las curvas de Bezier, oficialmente, desde 1972, da lugar a la polèmica. Para terciar en ella, me apoyo, como he dicho anteriormente, en la similitud al proceso seguido por la interpolaciòn.

La interpolaciòn, como hemos dicho màs arriba, no pretende dar una serie de puntos y hallar una curva que se ajuste a ellos, sino dar una serie de puntos de control o directores y conseguir una curva que pase cerca de ellos, cumpliendo una serie de condiciones de suavidad y continuidad. El problema radicarà en la forma de obtener dichos puntos directores, pero eso sale fuera del àmbito de esta tesis, para entrar en el diseño propiamente dicho, nosotros cogemos el problema una vez que se encuentra en el estadio de tener los puntos y unirlos convenientemente. De todas formas, el que suscribe tiene cierta experiencia en el tema, ya que he realizado un programa en el que se pretende diseñar un canal semicircular al que se le imponen unas ciertas condiciones de pendiente y curvatura, presentando como resultado el eje de dicho canal en Axonomètrico, y con la posibilidad de efectuar otras proyecciones de dicho eje, como planta, alzado, perfil, etc. El càlculo de los puntos directores se realiza en un papel de croquizar en donde se calcula, en planta y alzado, la forma que debe tener el eje del canal para que se cumplan las condiciones fijadas. Una vez introducidos los valores en el ordenador, èste dibuja mediante Splines la curva pedida calculando los valores del radio de curvatura y pendiente. La respuesta gráfica viene complementada por una respuesta escrita en otra pantalla en donde se dan los valores límites de dichos conceptos, así como su posición sobre la curva. Lógicamente, el programa ofrece la posibilidad de modificar los puntos directores hasta que obtengamos el resultado apetecido (3).

Las curvas de Bezier trabajan con los puntos directores o vértices de un polígono que nos define la forma de la curva, sólo los puntos extremos del polígono pertenecerán a la curva a diseñar, los otros vértices nos sirven para definir las derivadas, orden y forma de la curva. Dado que la forma de la curva tiende a seguir la forma del polígono, el cambio de los vértices de dicho polígono da al usuario una idea **intuitiva** de las relaciones input/output (4).



Es suficiente añadir un nuevo vértice interior para aumentar el orden de la curva, cualquier cambio de posición de uno de los vértices del polígono afecta a la forma de la curva entera.

Estudio de curvas. Segunda parte.

Las curvas de Bezier se apoyan para su definición en los polinómios de Bernstein de expresióm:

$$B_{i,n}(t)= C_n{}^i.t^i.(1-t)^{n-i}$$

con $C_n{}^i= n!/i!.(n-i)!$, que son las combinaciones de n objetos tomados de i en i.

Los puntos de la curva de Bezier nos vienen dados por:

$P(t)= \Sigma P_i.B_{i,n}(t)$, donde Σ representa, en este caso: sumatorio para i valiendo desde 0 hasta n. $0<t<1$ (b)

Dado que las curvas de Bezier son combinaciones de los polinomios de Bernstein, veámos las caracteristicas de estos:

1.- El número de vértices del poligono determina el orden del polinomio definido. Por ejemplo, si tenemos un poligono de 6 vértices, dado que, en principio consideraremos los poligonos abiertos, el número de lados es de 5, consecuentemente el grado de la curva será también de 5. Para variar dicho grado no hay más remedio que alterar el número de vértices.

2.- Los polinómios de Bernstein son siempre distintos de cero cualquiera que sea el valor de t, consecuentemente como todos los puntos de la curva son el resultado de los valores de los vértices de los poligonos y en estos nunca pueden valer cero, cualquier variación de la posición de uno de ellos provoca un efecto en toda la curva.

Implementación.

Ya hemos visto que las curvas de Bezier se apoyan sobre los polinómios de Bernstein, veámos de que forma damos valores a P_i. La expresiòn (b) es vectorial, y nosotros lo que tendremos son parejas de puntos a las que hay que aproximar, luego lo que hemos de hallar son otras parejas de puntos funciòn de los que ya tenemos. P_i seràn los valores de las coordenadas correspondientes, asi:

$$X(t)=\Sigma X_i.(n!/i!.(n-i)!).t_i.(1-t)^{n-i}$$
$$Y(t)=\Sigma Y_i.(n!/i!.(n-i)!).t_i.(1-t)^{n-i},$$

Σ tiene el significado comentado en (b).

Propiedades de las curvas de Bezier.

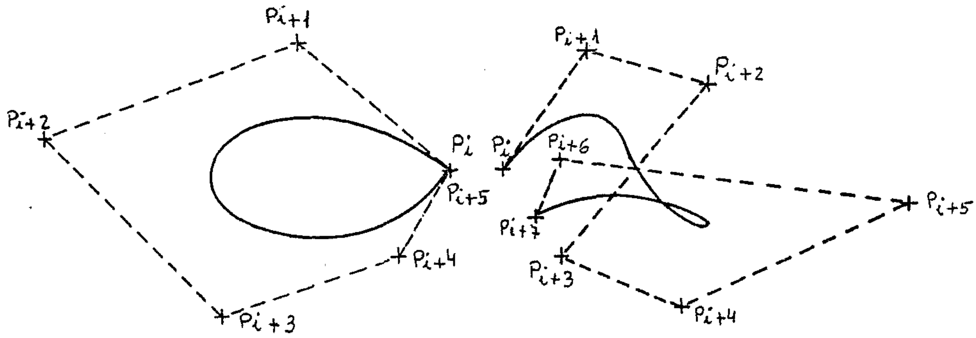
- Una curva de Bezier pasa siempre por el primer y el ùltimo punto director, mientras que no tiene porque pasar por ningùn otro.

- La curva de Bezier es tangente a los lados del polìgono en sus puntos extremos.

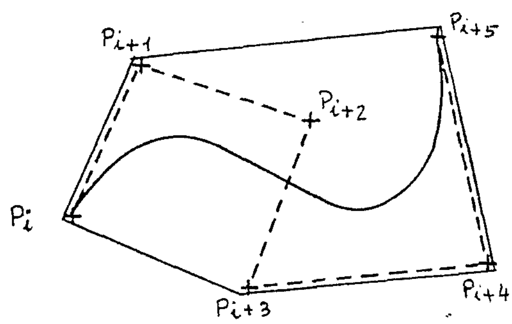
- Cada punto director ejerce una atracciòn de la parte de curva que se encuentra cerca de èl, proporcional al coeficiente de Bernstein correspondiente. (ver nota (4)).

Estudio de curvas. Segunda parte.

- Dado el carácter paramétrico de las curvas, éstas pueden cerrarse e incluso cortarse.



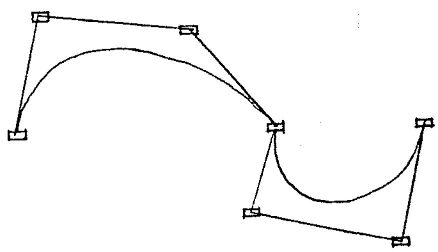
- Una curva de Bezier se encontrará siempre contenida dentro del poligono convexo formado por sus puntos directores.



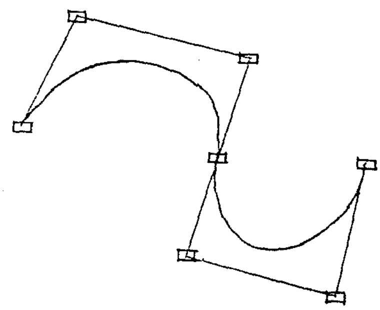
- El control de una curva de Bezier es global, como ya hemos dicho anteriormente, la variación de un punto hace que varie la forma de toda la curva.

- Existe la posibilidad de unir varias curvas de Bezier, para ello podemos imponer ciertas normas de continuidad:

- Continuidad de orden cero. Para ello sòlo es necesario que el ùltimo punto de una de las curva coincida con el primero de la siguiente.



- Continuidad de primer orden. Para ello es necesario que los **lados** de los dos poligonos estèn alineados. Esto añade una condiciòn al caso anterior, además de coincidir el primer punto de una curva con el último de la otra, es necesario que los "penultimos" de cada curva estèn alineados con èste.



Por último, si tenemos un interès especial en que la curva se acerque o incluso pase por un punto, o varios, dado, basta con aumentar la multiplicidad de este punto introducièndolo varias veces. De esta forma, la fuerza con que atrae el punto a la curva aumenta, siguiendo este razonamiento si hacemos que la multiplicidad de todos los puntos de un poligono director tiendan al infinito, la curva de Bezier asociada tenderá al poligono creado.

La aproximaciòn por B-Splines es más general que la de Bezier, hasta el punto de que èsta se puede considerar un caso particular de aquella. Las curvas de Bezier se basan en la base de los polinòmios de Bernstein, los B-Splines son ellos mismos una base, que a diferencia de la de Bernstein no son globales, quiere esto decir que cada vèrtice o punto director està asociado a una funciòn de base ùnica, consecuentemente la modificaciòn de la posiciòn de un punto director tendrà una influencia muy localizada en su propio entorno.

La base de los B-Splines es:

$$S_{i,j}(t)=((t-X_i).S_{i,j-1}(t))/(X_{i+j-1}-X_i)+$$
$$+((X_{i+j}-t).S_{i+1,j-1}(t))/(X_{i+j}-X_{i+1})$$

cumplièndose que: $S_{i,1}(t)=1$ si $X_i<=t=<X_{i+1}$ y $S_{i,1}(t)=0$ en el resto de casos.

Los polinòmios se expresan:

$P(t)=\Sigma P_i.S_{i,j}(t)$ (c), Σ anàlogo comentario que en (b).

A efectos de programaciòn serà necesario implementar que en el caso de obtener fracciones del tipo 0/0, se les asigne como resultado el valor 0. j es el orden de la curva, i es el nùmero de puntos directores, los valores Xi son los elementos de un vector llamado vector de nodos que se calculan mediante las expresiones:

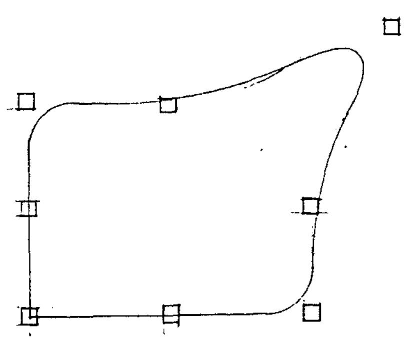
$Xi=0$ si $i<j$
$Xi=i-j+1$ si $j<=i<=n$
$Xi=n-j+2$ si $i>n$

Este mètodo es màs flexible que el de Bezier, que hemos visto anteriormente, dado que nos permite construir curvas con tramos de naturaleza diferente; especificando el grado de los polinomios que nos definiràn el B-Spline, y el nùmero de intervalos que deseamos, y en ellos podemos determinar los condiciones de continuidad que nos sean precisas. Podemos, de esta forma, especificar en un determinado intervalo la continuidad de la curva, o la continuidad ademàs de la pendiente, o, incluso, la continuidad de la curva, la pendiente y la curvatura, etc (5).

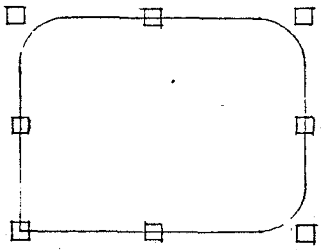
Propiedades de las curvas de B-Splines.

Ademàs de las propiedades de las curvas de Bezier vistas anteriormente las curvas B-Spline tienen algunas propiedades caracterìsticas:

- El control de la curva es local. En la figura se aprecia el efecto de desplazar un sòlo punto.



- Es posible cambiar el orden de la curva sin tener que añadir o eliminar puntos directores.



- Si el orden de la curva es el mismo que el número de vertices del polígono, nos encontramos en el caso, visto anteriormente, de curvas de Bezier.

- Curvas con puntos singulares, entendiendo por puntos singulares aquellos en los que encontramos dos tangentes.

- Las curvas de B-Spline de orden 3 son tangentes a los lados del poligono formado por los puntos directores.

Finalmente, es importante destacar que todo lo dicho es aplicable tanto a dos como a tres dimensiones entrando el número conveniente de coordenadas y posteriormente sometiéndolas a las proyecciones necesarias. Es interesante, a tal efecto, leer e incluso probar las rutinas en BASIC que se presentan en el articulo de Steve Enns, "Free-Form Curves On Your Micro", publicado en Byte en Diciembre de 1986, páginas 225 a 230.

Notas y Referencias.

(1) pags. 89 y 90 de "Mathematical Elements For Computer Graphics" por David Rogers. McGraw-Hill. 1976

(2) Uno de los textos más claros, a nivel divulgativo, sobre estos temas es "Calcul des parties cachées. Aproximation par la metode de BEZIER (sic) et dès B-Splines" de Robert Dony publicado por la editorial Masson en 1986. Paradójicamente, realiza el estudio justo a la inversa de lo que he expuesto, supongo que la justificación se debe al carácter de gloria nacional "viva" de Bezier, dado que Lagrange también es francés.

(3) En el capitulo de Conclusiones, se puede encontrar un ejemplo del funcionamiento de este programa paso a paso, como ilustración de como se establece el diàlogo ordenador-usuario.

(4) Este caràcter intuitivo màs que ser una virtud es, en cierta forma, un inconveniente, como se ha visto con la experiencia realizada a travès de cuatro años con distintas variantes del programa comentado anteriormente. El usuario eventual, puede tener dificultades para establecer una proporcionalidad correcta sobre el efecto de la posición de un punto director sobre la forma de la curva; el razonamiento humano es lineal y puede conducir a errores. Otro caso serà cuando estemos tratando de un diseñador que trabaja con un programa cierto tiempo, insconcientemente, se va adaptando a la relación input/output y llega a establecer una satisfactoria relación, no es el caso de un usuario que le dedique al programa una hora, como fue el caso experimentado con nuestro programa del canal.

(5) Para màs información, autorizada, sobre el tema es muy interesante consultar las publicaciones del "Deapartament d'Informatica" de la Escola Técnica Superior d'Enginyers Industrials de Barcelona, donde bajo la dirección del Dr. Pere Brunet se està realizando una gran labor al respecto. En el caso concreto de los B-Splines, se recomienda la consulta de las pàginas 104 a 116 de los apuntes editados con motivo del "Curs sobre TECNIQUES DE REPRESENTACIO GRAFICA I DISSENY UTILITZANT PETITS

ORDINADORS", a cargo, además del nombrado Pere Brunet, de Lluis Sanz, Isabel Navazo, Xavier Pueyo y Dolors Ayala.