

Universitat Politècnica de Catalunya
Dep. Teoria del Senyal i Comunicacions
E.T.S. d'Enginyeria de Telecomunicació

**APLICACION DEL MODELADO GEOMETRICO A
PROBLEMAS DE DIFRACCION ELECTROMAGNETICA**

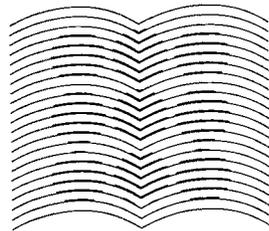
Mercè Vall-Ilossera Ferran

Director: Dr. Juan Manuel Rius Casals

Tesis doctoral presentada a la
Universitat Politècnica de Catalunya
para la obtención del título de
Doctor Ingeniero de Telecomunicación

Barcelona, Marzo 1994

UNIVERSITAT
POLITÈCNICA
DE CATALUNYA



BIBLIOTECA
EX - LIBRIS

Universitat Politècnica de Catalunya
Dep. Teoria del Senyal i Comunicacions
E.T.S. d'Enginyeria de Telecomunicació

**APLICACION DEL MODELADO GEOMETRICO A
PROBLEMAS DE DIFRACCION ELECTROMAGNETICA**

Mercè Vall-llossera Ferran

Director: Dr. Juan Manuel Rius Casals

UNIVERSITAT POLITÈCNICA DE CATALUNYA
ADMINISTRACIÓ D'ASSUMPTE S ACADÈMI S

Aquesta Tesi ha estat onregistrada
a la pàgina 61 amb el número 566

Barcelona,

L'ENCARREGAT DEL REGISTRE,

Tesis doctoral presentada a la
Universitat Politècnica de Catalunya
para la obtención del título de
Doctor Ingeniero de Telecomunicación

Barcelona, Marzo 1994

A Juan Manuel

El trabajo realizado en esta tesis ha sido subvencionado por la CICYT en el marco de:

La Acción Especial:

- *"Programa de investigación aplicada para el desarrollo y validación de métodos de cálculo numérico para la predicción y análisis de las características de los ecos radar (RCS) y su reducción"*

TIC 88-288E

Gestor: Dirección General de Telecomunicaciones

Coordinador: Empresa C.A.S.A

Los proyectos:

- *"Antenas para comunicaciones móviles y VSAT"*

TIC 91-1034

- *"Análisis numérico mediante ordenadores vectoriales y paralelos de sistemas radiantes en entornos reales"*

TIC 93-0518

AGRADECIMIENTOS

Son muchas las personas que han contribuido a realizar esta tesis y a todas ellas quiero expresarles mi más sincero agradecimiento. Sin embargo, quisiera hacer especial mención de la inestimable ayuda aportada por el director de esta tesis, Juan Manuel Rius en todo momento. Además, quisiera agradecer a Angel Cardama sus valiosos consejos, a Nuria Duffo su colaboración en el manejo del programa I-DEAS y, en general, a todos los compañeros del Grupo A.M.R. el soporte profesional y humano que me han proporcionado.

Finalmente, quiero agradecer a toda mi familia y muy especialmente a mis padres el afecto y soporte incondicional que me han dado a lo largo de los años.

1 INTRODUCCION	1-1
2 TECNICAS DE MODELADO GEOMETRICO TRIDIMENSIONAL	2-1
2.1 MODELADO GEOMETRICO DE SOLIDOS	2-2
2.2 METODOS DE MODELADO DE SOLIDOS	2-3
2.2.1 Objetos definidos mediante barrido	2-4
a) Traslación o extrusión:	2-4
b) Rotación o revolución:	2-4
c) Barrido generalizado:	2-5
2.2.2 Objetos definidos mediante parámetros	2-7
2.2.3 Modelado por fronteras	2-8
a) Modelado por aristas	2-9
b) Modelado por hilos	2-10
c) Modelado por placas planas (facetas)	2-11
d) Modelado mediante parches curvos	2-13
2.2.4 Representación mediante particiones del espacio	2-17
a) Descomposición mediante celdas (o elementos finitos)	2-18
b) Enumeración de ocupación del espacio	2-19
c) Arboles octales (<i>Octrees</i>)	2-21
d) Arboles de partición binaria del espacio	2-23
2.2.5 Geometría constructiva de sólidos	2-25
2.3 COMPARACION DE METODOS Y CONCLUSIONES	2-26
2A CURVAS Y SUPERFICIES PARAMETRICAS	2A-1
2A.1 MODELADO DE SUPERFICIES ARBITRARIAS	2A-1
a) Mallas de polígonos	2A-1
b) Superficies paramétricas	2A-2
c) Superficies algebraicas implícitas	2A-2
d) Superficies algebraicas explícitas	2A-2
2A.2 TIPOS DE CURVAS Y SUPERFICIES PARAMETRICAS	2A-3
2A.3 FORMULACION DE LAS CURVAS PARAMETRICAS	2A-6
2A.3.1 Curvas de Hermite	2A-10
2A.3.2 Curvas de Bézier	2A-13
2A.3.3 Curvas <i>spline</i>	2A-17
a) <i>B-splines</i> uniformes	2A-18
b) <i>B-splines</i> no uniformes	2A-23
c) <i>B-splines</i> racionales	2A-30
d) Otras formas <i>spline</i>	2A-34
2A.4 SUPERFICIES PARAMETRICAS <i>B-SPLINE</i>	2A-36
a) Superficies paramétricas	2A-36
b) Superficies <i>B-spline</i>	2A-37
c) Curvas definidas sobre superficies <i>B-spline</i>	2A-39

3	VISUALIZACION GRAFICA TRIDIMENSIONAL	3-1
3.1	PARAMETROS Y DATOS DE ENTRADA	3-3
a)	Información básica	3-3
b)	Información opcional para mejora del realismo	3-4
c)	Librerías gráficas	3-6
3.2	TRANSFORMACIONES GEOMETRICAS	3-7
3.3	MODELOS DE ILUMINACION	3-10
3.3.1	Modelos de iluminación local	3-11
a)	Fuentes de luz	3-11
b)	Reflexión de la luz en las superficies	3-12
3.3.2	Modelos de iluminación global	3-16
a)	Ray tracing	3-17
b)	Radiosity	3-19
c)	Conclusiones	3-22
3.3.3	Suavizado de la imagen	3-22
a)	Sombreado constante	3-22
b)	Sombreado con interpolación	3-23
3.3.4	Aplicación de los métodos de iluminación al análisis electromagnético	3-24
3.4	GENERACION DE PIXELS	3-25
3.4.1	Memoria de pantalla y tabla de colores	3-25
3.4.2	Generación de pixels	3-26
3.5	ELIMINACION DE PARTES OCULTAS	3-30
a)	Algoritmo del z-buffer	3-31
b)	Algoritmos de listas de prioridades	3-32
3.6	HARDWARE DEL ACELERADOR GRAFICO	3-33
3.6.1	Subsistema <i>front-end</i>	3-35
3.6.2	Subsistema <i>back-end</i>	3-38
a)	Generación de pixels (<i>scan conversion</i>)	3-38
b)	Memoria de pantalla (<i>Raster Engine</i>)	3-39
4	MODELADO DE OBJETOS PARA ANALISIS NUMERICO	4-1
4.1	METODO DE LOS ELEMENTOS DE CONTORNO	4-2
4.1.1	Formulación	4-3
a)	Equivalencia superficial	4-3
b)	Análisis de cuerpos con regiones homogéneas	4-4
c)	Formulación de las ecuaciones integrales	4-7
4.1.2	Discretización	4-8
a)	Método de los momentos	4-9
b)	Método de los elementos de contorno	4-11
4.1.3	Modelado	4-14
4.2	METODO DE LOS ELEMENTOS FINITOS	4-19
4.2.1	Formulación	4-19
a)	Formulación diferencial	4-19

b) Formulación variacional	4-20
4.2.2 Discretización	4-21
4.2.3 Modelado	4-25
4.3 METODO DEL GRADIENTE CONJUGADO+FFT	4-27
4.3.1 Formulación	4-29
a) Equivalencia volumétrica	4-29
b) Formulación de las ecuaciones integrales	4-31
4.3.2 Discretización	4-31
a) Gradiente conjugado	4-31
b) Funciones base "tejado" (<i>roof top</i>)	4-34
4.3.3 Modelado	4-34

PARTE II

APLICACION DE LOS METODOS DE MODELADO Y VISUALIZACION A LA TECNICA DE PROCESADO GRAFICO

5 PREDICCIÓN DE RCS	5-1
5.1 TECNICAS CLASICAS DE PREDICCIÓN DE RCS	5-1
5.2 METODO DE PROCESADO GRAFICO	5-3
5.3 ANALISIS GEOMETRICO	5-5
5.3.1 Identificación de superficies ocultas	5-7
5.3.2 Cálculo de las normales a la superficie	5-7
5.3.3 Detección de aristas	5-11
5.3.4 Cálculo de los radios de curvatura	5-12
5.4 ANALISIS ELECTROMAGNÉTICO	5-13
5.4.1 Aproximación de óptica física	5-13
5.4.2 Condición de contorno de impedancia	5-15
5.4.3 Método de las corrientes equivalentes	5-17
5.4.4 Reflexiones múltiples por trazado de rayos	5-22
6 ANALISIS DE ANTENAS DE REFLECTOR	6-1
6.1 ANTENAS DE REFLECTOR	6-2
6.1.1 Reflectores planos	6-3
6.1.2 Reflectores diédricos	6-3
6.1.3 Reflectores cilíndricos	6-4
6.1.4 Reflectores con simetría de revolución	6-5
6.1.5 Reflectores <i>offset</i>	6-6
6.1.6 Reflectores conformados	6-7
6.2 MODELADO DE LAS ANTENAS DE REFLECTOR	6-7
6.3 METODOS DE ANALISIS DE ANTENAS DE REFLECTOR	6-10
6.3.1 Método de la distribución de campo en una apertura	6-11
6.3.2 Método de la distribución de corrientes	6-13

6.3.3 Bloqueo	6-13
6.3.4 Difracción en bordes	6-17
6.4 PROCESADO GRAFICO Y OPTICA FISICA	6-18
6.4.1 Obtención de las coordenadas del modelo	6-19
6.4.2 Obtención de los vectores normales a la superficie	6-19
6.4.3 Análisis del bloqueo	6-19
a) Sombras móviles	6-20
b) Sombras fijas	6-20
6.4.4 Cálculo de la integral de óptica física	6-22
6.5 VALIDACION Y RESULTADOS	6-24
6.5.1 Reflector circular con alimentador dipolo corto	6-24
6.5.2 Reflector parabólico simétrico	6-29
6.5.3 Reflector esférico simétrico	6-32
6.5.4 Reflector parabólico <i>offset</i>	6-32
7 OPTIMIZACION DE FORMAS	7-1
7.1 OPTIMIZACION DE RCS POR PROCESADO GRAFICO	7-2
7.1.1 Métodos de reducción de RCS	7-2
7.1.2 Optimización interactiva	7-4
7.1.3 Optimización automática	7-5
7.1.4 Planteamiento del problema	7-5
a) Restricciones	7-8
b) Criterios	7-8
c) Métodos	7-8
7.2 METODOS DE OPTIMIZACION NO LINEAL	7-8
7.2.1 Método del gradiente descendiente	7-10
7.2.2 Métodos de las direcciones conjugadas	7-12
a) Método de Powell	7-14
b) Método del gradiente conjugado	7-16
7.3 OPTIMIZACION DE FORMAS SIMPLES	7-19
7.3.1 Optimización de diedros	7-22
7.3.2 Optimización de un elipsoide	7-29
7.3.3 Optimización de un conjunto de dos elipsoides	7-30
7.3.4 Conclusiones	7-35
7.4 OPTIMIZACION DE OBJETOS COMPLEJOS	7-35
7.4.1 Planteamiento del problema	7-35
7.4.2 Medidas geométricas para el cálculo de las restricciones	7-37
a) Medidas sobre <i>NURBS</i>	7-37
b) Medidas sobre facetas	7-40
7.4.3 Implementación de los métodos de optimización	7-43
a) Estimación del gradiente	7-44
b) Selección de los puntos a modificar	7-44
7.4.4 Resultados	7-45

CAPITULO 1

INTRODUCCION

Desde la aparición de los primeros ordenadores en la década de los 50 se ha realizado una gran actividad en el campo de análisis numérico y asintótico de radiación, difracción, guiado y compatibilidad electromagnéticos. Sin embargo, hasta hace pocos años la escasa capacidad de memoria y velocidad de proceso de los ordenadores disponibles limitaba el conjunto de objetos analizables a tamaños eléctricos pequeños o geometrías relativamente simples.

El reciente desarrollo de la tecnología informática permite en la actualidad analizar objetos muy complejos y de gran tamaño eléctrico, como por ejemplo circuitos de microondas, aviones y helicópteros. Esta posibilidad abre un nuevo campo de investigación: ¿Cómo introducir en el ordenador toda la información de una geometría tan compleja, de la forma más general y a la vez más concisa, y con el menor coste de ingenieros x hora?.

La solución se encuentra en las técnicas de modelado geométrico de sólidos, que se han introducido recientemente para aplicaciones de diseño industrial y control numérico. En la actualidad se dispone de paquetes de CAD/CAM muy potentes capaces de modelar objetos enormemente complejos, visualizar los modelos en la pantalla de una estación de trabajo y realizar mallados de elementos finitos para análisis de estructuras y dinámica de fluidos. Estas herramientas software se utilizan en multitud de campos de la ciencia, la tecnología y las artes gráficas. Sin embargo, su utilización en el campo del análisis electromagnético está muy poco extendida.

Hasta hace relativamente poco tiempo, los métodos de análisis electromagnético se basaban en técnicas de modelado muy particulares, dependientes de la geometría del objeto a analizar y difícilmente generalizables. Habitualmente el modelado era un problema particular para cada objeto y de solución muy laboriosa, debido a no disponer de un método de modelado automático de objetos totalmente arbitrarios. Las soluciones más relevantes al problema del modelado generalizado son relativamente recientes:

modelado por hilos [Burke'81], por facetas triangulares planas [Rao'82], por facetas poligonales planas [Youssef'89] y por elementos finitos/elementos de contorno [Jin'88].

En esta Tesis se estudia la aplicación de las técnicas y del software de modelado de sólidos a los problemas de análisis electromagnético, tanto los relacionados con métodos numéricos como los basados en aproximaciones asintóticas de alta frecuencia. Por ello, el capítulo 2 se dedica a repasar las técnicas de modelado geométrico más extendidas y su posible aplicación al problema de análisis electromagnético.

Un tema muy relacionado con el modelado geométrico y, según estudios recientes, con el análisis electromagnético es la visualización realista tridimensional. En los últimos años se ha dedicado un esfuerzo muy importante a la creación de imágenes realistas para aplicaciones de simulación, diseño, publicidad, educación y, particularmente, visualización de los modelos geométricos construidos con los paquetes CAD. Con objeto de obtener estas imágenes en tiempo real, también se ha desarrollado paralelamente el hardware gráfico de las estaciones de trabajo.

La relación de estas técnicas de visualización con el campo del análisis electromagnético es más reciente: el método de procesado gráfico de una imagen realista del objeto simplifica enormemente la aplicación de las técnicas de alta frecuencia para analizar la difracción y reflexión electromagnéticas [Rius'91,93b,93d]. Por este motivo, en el capítulo 3 se estudian las técnicas de visualización realista y las configuraciones de hardware gráfico más habituales.

Una vez presentadas las herramientas de modelado y visualización en los dos primeros capítulos, en el siguiente se estudia su aplicación a los métodos de análisis numérico. Para ello se estudian los tres métodos que permiten analizar objetos de geometría arbitraria con la máxima generalidad posible: el método de los elementos de contorno (BEM), el método de los elementos finitos (FEM) y el método del gradiente conjugado con transformada rápida de Fourier.

Los tres últimos capítulos se dedican a las nuevas técnicas de análisis en alta frecuencia por procesado gráfico. Este nuevo método es, posiblemente, la herramienta más general para análisis de objetos de geometría arbitraria con aproximaciones de alta frecuencia y constituye un excelente ejemplo de aplicación eficiente de las técnicas de modelado geométrico y visualización realista al análisis electromagnético.

Las técnicas de alta frecuencia, desarrolladas entre las décadas de los 50 y los 80 para el estudio de objetos eléctricamente grandes, hasta hace muy pocos años únicamente eran capaces de procesar geometrías muy simples. Con el advenimiento de las técnicas de modelado por facetas poligonales planas [Youssef'89] la aplicación de los métodos de alta frecuencia se extendió a objetos complejos. En contrapartida, los modelos de

facetas constituían una representación muy poco concisa de la realidad y requerían un pre-procesado geométrico muy costoso previo al análisis electromagnético. Estos problemas se solucionan con el nuevo método de procesado gráfico, el cual permite utilizar un modelo geométrico más conciso y realizar el pre-procesado geométrico mediante técnicas gráficas de visualización realista.

Actualmente, la posibilidad de estudiar en alta frecuencia objetos de geometría arbitraria permite, por ejemplo, la predicción de sección recta (RCS) de blancos radar complejos y el análisis de antenas de reflector conformado. Además, al procesarse modelos geométricos más concisos en un tiempo mucho más reducido, puede plantearse el problema de optimización de formas para cumplir unas especificaciones de RCS o de diagrama de radiación. Este problema, cuya solución era muy difícil con las técnicas clásicas, puede ser abordado si se combinan adecuadamente las técnicas de modelado geométrico, visualización realista y análisis electromagnético por procesado gráfico.

En el capítulo 5 se presenta muy brevemente el método de procesado gráfico y su aplicación al cálculo de sección recta radar. A las nuevas aplicaciones del procesado gráfico se dedican los dos capítulos siguientes: en el sexto se estudia el análisis las antenas de reflector y el séptimo se dedica a la optimización de formas en alta frecuencia.

CAPITULO 2

TECNICAS DE MODELADO GEOMETRICO TRIDIMENSIONAL

Cuando se genera un modelo de un objeto o un comportamiento real, se pretende que sea lo más exacto posible. Es decir, que se aproxime al máximo a la realidad que representa. Como la representación exacta de la realidad en la mayoría de los casos resulta imposible lo que se hace generalmente es imponer unas restricciones e hipótesis bajo las cuales el modelo se comporta satisfactoriamente. En el problema electromagnético que nos ocupa, el modelo es satisfactorio cuando introduce un error aceptable en los resultados del cálculo numérico.

La rápida evolución de la tecnología informática ha posibilitado simulaciones por ordenador de problemas enormemente complejos, gracias a las mejoras en velocidad de proceso y memoria de almacenamiento. Sin embargo, la resolución de estos problemas requiere técnicas de modelado de sólidos cada vez más sofisticadas y precisas. También hay que tener en cuenta que en cualquier simulación por ordenador además de los errores debidos a la aproximación del objeto real por un modelo aparecen errores debidos a aproximaciones en la formulación matemática junto con los errores numéricos. Estos tipos de errores no están relacionados con las técnicas de modelado y, por tanto, no se consideran en este trabajo.

El análisis electromagnético precisa de técnicas de modelado geométrico tridimensional que permitan modelar cualquier geometría. Las primeras técnicas de modelado de que se disponía sólo permitían modelar objetos muy simples. Sin embargo, en la actualidad se han desarrollado potentes herramientas CAD capaces de generar formas mucho más sofisticadas y por tanto, modelos que se ajustan más a la realidad. Estos modelos más rigurosos permiten modelar cualquier objeto con gran precisión y, lo que es más importante, posibilitan el análisis electromagnético de una estructura cualquiera.

En este capítulo se van a exponer las técnicas de modelado geométrico tridimensional que se utilizan en el diseño de objetos para análisis electromagnético. En la práctica no existe una técnica de modelado idónea para todos los problemas electromagnéticos, sino que según la naturaleza del problema, el método de análisis aplicado y las características del objeto a analizar será más adecuada una u otra técnica.

2.1 MODELADO GEOMETRICO DE SOLIDOS

El objetivo del modelado geométrico tridimensional es proporcionar unas herramientas potentes, eficientes y flexibles para representar objetos tridimensionales, que faciliten, además, una posterior manipulación de los mismos. Manipular objetos es crear objetos nuevos aplicando transformaciones a objetos ya existentes u operaciones booleanas entre ellos.

En el desarrollo de una cierta técnica de modelado debe tenerse en cuenta que el modelo geométrico de un cierto objeto real debe cumplir las siguientes propiedades [Brunet'89]:

- **No ambigüedad:** El modelo debe corresponder a un solo objeto real. A partir del modelo debe poder ser deducido el objeto real al que se corresponde. Una representación no ambigua también se denomina completa.
- **Unicidad:** A cada objeto real le debe corresponder un solo modelo. Si se asegura la unicidad, resultan más sencillas operaciones tales como la determinación de si 2 modelos son iguales.
- **Dominio:** Debe permitir representar una variedad de objetos reales lo más amplia posible. Se busca que el método de modelado permita el diseño de objetos complejos con características muy distintas entre sí, evitando un coste computacional demasiado elevado.
- **Validez:** Sólo debe permitir representar objetos reales.
- **Precisión:** La representación debe ser lo más rigurosa posible. Muchas de las representaciones sólo permiten el trazado de líneas rectas, superficies planas o poliedros, lo cual obliga a aproximar objetos en los que se encuentren superficies curvadas
- **Invarianza:** Se dice que un método es invariante cuando al aplicar rotaciones, traslaciones y/u otras operaciones sobre objetos modelados por este método se obtienen nuevos objetos también representables por el mismo método.

Además, en cualquier método es conveniente optimizar los siguientes aspectos:

- **Concisión:** La cantidad de información necesaria para la descripción del modelo debe ser mínima.
- **Facilidad de diseño de objetos y formas arbitrarios:** Las técnicas de modelado deben ser suficientemente flexibles para que no se incrementen de forma exagerada los tiempos de modelado de objetos complejos.
- **Control local de los modelos:** Las técnicas que poseen esta propiedad permiten modificar unas zonas concretas del objeto sin variar el resto del modelo. De esta forma pueden conseguirse diseños muy precisos de objetos complejos.
- **Facilidad de interrogación para validar el modelo:** El proceso de modelado de un objeto complejo suele ser iterativo. Se realiza un modelo inicial, el cual se somete a interrogación (se compara con el objeto real) para evaluar su precisión. Si no es satisfactorio, se modifica adecuándolo más a la realidad y se interroga el nuevo diseño. Este proceso se repite hasta llegar a un modelo satisfactorio.
- **Facilidad de imposición de propiedades físicas:** En muchas ocasiones el modelado geométrico tridimensional no es un objetivo, sino un medio para la simulación del comportamiento del objeto en una situación física real. Este es el motivo por el que en general es necesario imponer ciertas propiedades físicas sobre el modelo. Ejemplos típicos en electromagnetismo son las propiedades de los materiales (constante dieléctrica, conductividad, etc).
- **Facilidad de visualización:** En muchas ocasiones los objetos modelados han de formar parte de una imagen fotorrealista. Por ello, es necesario poderlos incorporar a algoritmos de visualización de imágenes, imponiendo sobre los mismos características de color, reflexión de las superficies, etc. (ver capítulo 3). Por ejemplo, algunas técnicas de análisis electromagnético se basan en la visualización del objeto, como el cálculo de RCS en alta frecuencia por procesado gráfico (*GRECO*) (ver capítulo 5).

2.2 METODOS DE MODELADO DE SOLIDOS

Los diferentes métodos de modelado de sólidos pueden clasificarse de una forma muy general según el tipo de elementos constitutivos de que están formados. En la figura 2-1 se muestra una posible clasificación de este tipo [Brunet'89]. La casilla que se ha dejado en negro en la figura significa que no es posible representar objetos de dos dimensiones mediante elementos tridimensionales.

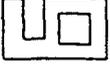
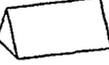
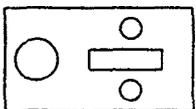
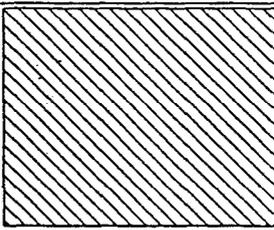
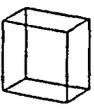
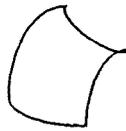
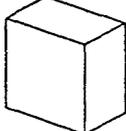
Dimensión de los elementos básicos			
	1 D 	2 D 	3 D 
modelos 2 D	2D Drafting 	Trazado de paneles 	
modelos 3 D	modelado alámbrico 	superficies 	sólidos 

Figura 2-1: Clasificación de los modelos geométricos según el tipo de elementos constitutivos

En la presente tesis se va a considerar únicamente el modelado de objetos tridimensionales, pues el caso bidimensional es más sencillo y puede considerarse incluido en el anterior. Por tanto, los métodos que se estudian a continuación se encuentran dentro de los tipos de la segunda fila de la figura. Los métodos incluidos en las dos primeras columnas de la segunda fila, modelado alámbrico y superficies, se conocen como métodos de "modelado por fronteras". La tercera columna corresponde a los métodos que utilizan sólidos como elementos básicos del modelo.

2.2.1 Objetos definidos mediante barrido

Esta técnica consiste en trasladar o rotar una estructura básica a través de un camino, de forma que todo punto del espacio atravesado por dicha estructura en su movimiento forma parte del modelo.

a) Traslación o extrusión:

El método más sencillo de barrido consiste en trasladar una estructura bidimensional a lo largo de una recta perpendicular al plano donde está definida, generándose así un objeto tridimensional. El *barrido traslacional o extrusión* es muy útil para modelar todo tipo de cilindros, lo cual lo hace muy apropiado para la resolución numérica de problemas electromagnéticos que incluyen cuerpos con simetría cilíndrica. En este caso, se plantea un problema bidimensional sobre un corte transversal del objeto, el cual es igual a la estructura bidimensional que se ha utilizado para generar el cilindro. En la

figura 2-2 se muestra un objeto modelado por traslación junto a la estructura bidimensional utilizada.

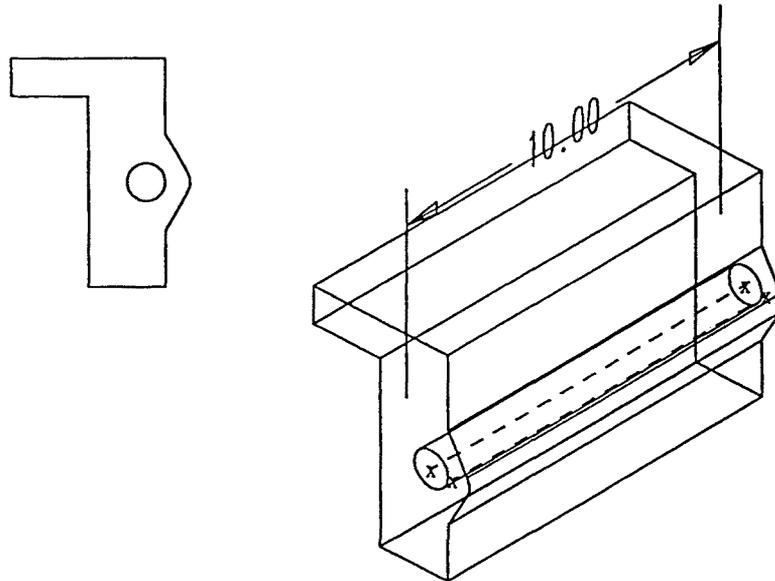


Figura 2-2: objeto modelado por extrusión de una estructura 2D

b) Rotación o revolución:

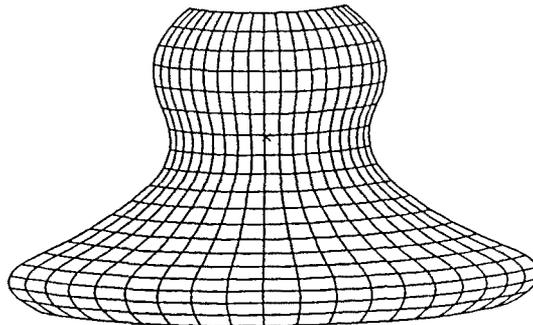


Figura 2-3: Objeto de revolución

El *barrido rotacional* genera objetos 3D mediante la rotación de un cuerpo bidimensional alrededor de un eje, de forma que los objetos modelados poseen simetría de revolución. Análogamente al caso anterior, la aplicación electromagnética se basa en reducir el problema tridimensional al análisis de la estructura básica 2D. En la figura 2-3 se muestra un ejemplo de barrido rotacional.

c) Barrido generalizado:

En esta generalización de la técnica de barrido el área o sólido básico que se traslada puede cambiar de tamaño, forma y orientación a lo largo de la trayectoria. Los modelos obtenidos mediante barrido generalizado de perfiles 2D se conocen como *cilindros generalizados*. Los perfiles 2D y la trayectoria se representan matemáticamente como curvas paramétricas (ver apéndice 2A), y el modelo tridimensional definido se almacena como un objeto modelado mediante fronteras (ver sección 2.2.3), representadas matemáticamente como superficies paramétricas. La figura 2-4 muestra el modelo del fuselaje de un avión Boeing 727, obtenido mediante la técnica de barrido generalizado y el software I-DEAS [I-DEAS].

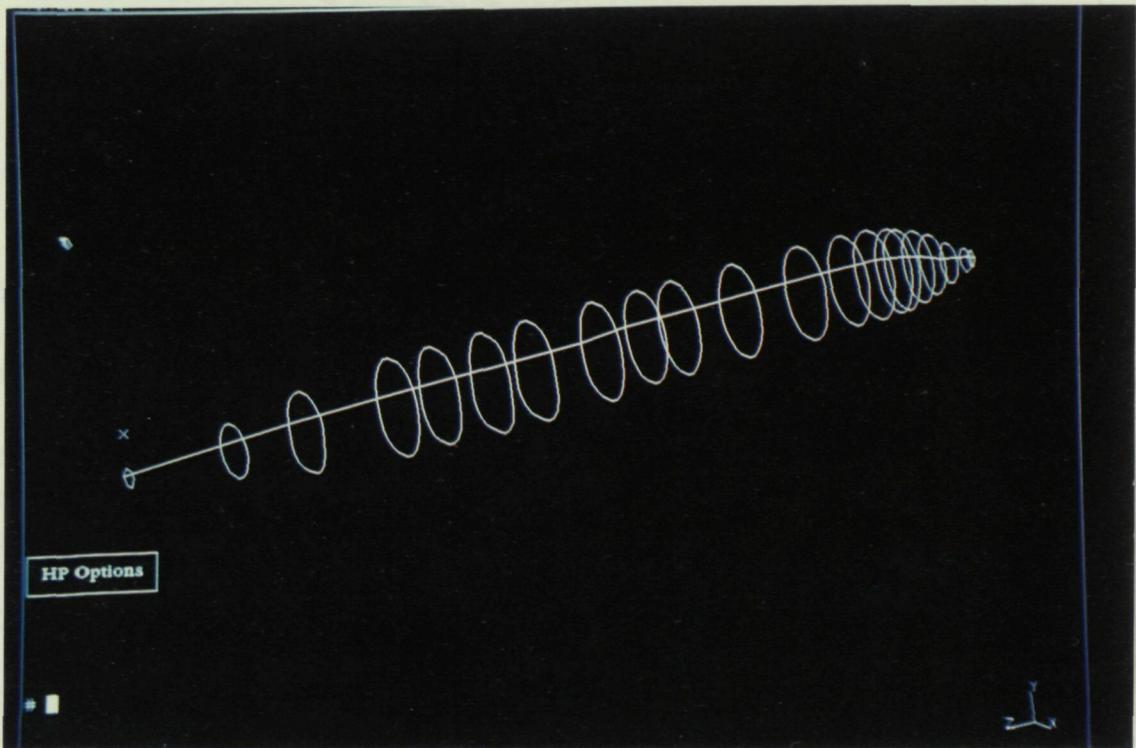


Figura 2-4: Modelo del fuselaje de un avión Boeing 727

Los inconvenientes más importantes de los métodos de barrido son:

- Es difícil aplicar operaciones booleanas (suma, intersección, diferencia) sobre objetos representados mediante esta técnica sin antes convertirlos a otro tipo de representación.

- No es un método invariante puesto que, tal como se demuestra en la figura 2-5, operaciones booleanas entre objetos modelados mediante barrido no obtienen siempre un objeto representable por esta técnica.

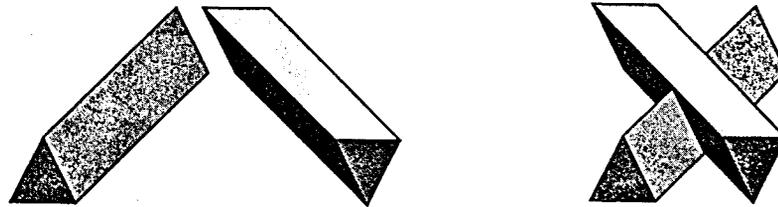


Figura 2-5: El objeto que se obtiene mediante la unión de dos objetos modelados por barrido habitualmente no se puede representar por barrido.

A pesar de estos inconvenientes, resulta ser un método muy intuitivo de modelado apto para una variedad bastante amplia de objetos. Esta es la razón por la que muchas herramientas de CAD permiten al usuario construir objetos mediante barrido, aunque después los modelos se almacenen a través de otras técnicas como modelado por fronteras.

2.2.2 Objetos definidos mediante parámetros

Este sistema de modelado se basa en un conjunto de formas, a las cuales se denomina primitivas, definidas mediante unos pocos parámetros que las caracterizan totalmente. Una primitiva parametrizada puede interpretarse como una familia de objetos cuyos miembros se diferencian en los valores de algunos parámetros. En general, una elección adecuada de las magnitudes a parametrizar proporciona una definición compacta de los modelos, por lo que se trata de un método muy conciso.

En este método no se prevé el uso de operaciones booleanas en la generación de un nuevo objeto a partir de objetos ya definidos anteriormente. Por tanto, siempre que se pretenda generar un nuevo objeto hay que definirlo mediante el conjunto de parámetros que lo caractericen totalmente.

El inconveniente principal es la imposibilidad de modelar geometrías arbitrarias: el dominio del método se reduce a la familia de primitivas utilizada. Sin embargo, es muy adecuado para modelar geometrías sencillas, por lo que es el método utilizado con más frecuencia en electromagnetismo, particularmente para el análisis de difracción en alta

frecuencia. La figura 2-6 muestra una ojiva bidimensional [Jina'90] definida mediante dos arcos de circunferencia. Los parámetros son dos: el radio de las circunferencias y la distancia entre los centros.

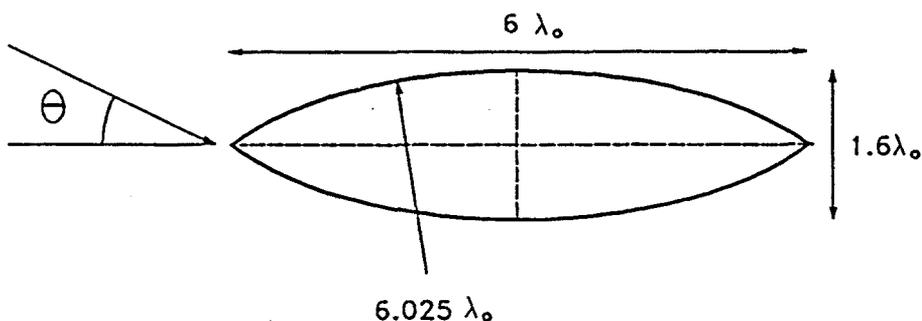


Figura 2-6: Modelo paramétrico de una ojiva bidimensional [Jina'90]

2.2.3 Modelado por fronteras

En los métodos de modelado por fronteras el tipo de información almacenada para describir al sólido corresponde a los elementos geométricos que definen su frontera exterior. Como se muestra en la figura 2-7, éstos últimos pueden ser adimensionales (vértices), lineales (aristas) o superficiales (parches). Para modelar objetos tridimensionales, los elementos frontera más utilizados en la práctica son las aristas y las superficies.

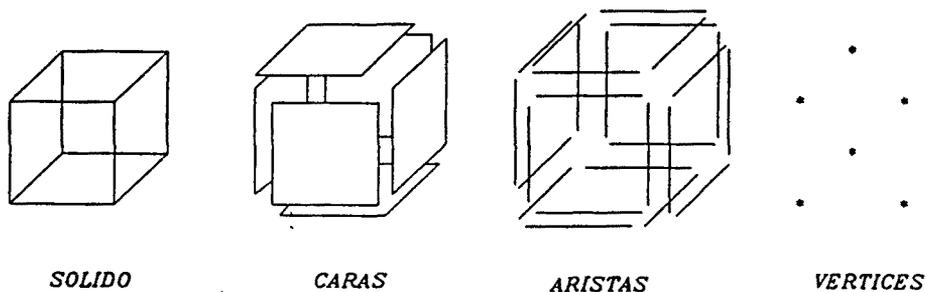


Figura 2-7: Distintas posibilidades de modelado por fronteras

El modelado alámbrico ofrece dos métodos de diseño claramente diferenciados. Uno de ellos representa los objetos mediante las aristas que existen sobre la superficie del mismo, mientras que el otro aproxima las superficies curvadas del objeto por una malla de hilos (aristas artificiales, existentes en el modelo pero no en la realidad).

El modelado mediante parches superficiales ofrece dos grupos de técnicas diferentes: el primero consiste en modelos de parches planos (facetas), mientras que el segundo engloba todos los métodos que utilizan parches curvos (superficies paramétricas). Estas técnicas son muy flexibles y permiten el diseño de objetos con geometrías arbitrarias.

Los métodos de modelado por fronteras presentan ciertas ventajas sobre otros métodos de modelado de sólidos:

- Facilidad en la realización de traslaciones y rotaciones del modelo.
- Facilidad en la realización de cálculos geométricos.
- Rapidez en la visualización realista tridimensional.

Los principales inconvenientes son:

- Representación poco concisa: se precisa una gran cantidad de información para definir el objeto.
- Dificultad en la realización de operaciones de intersección entre dos sólidos.

a) Modelado por aristas

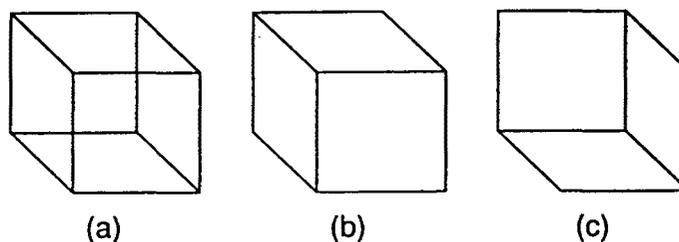


Figura 2-8: a) Modelado alámbrico de un cubo. b) y c) representan dos orientaciones posibles del cubo

La representación mediante aristas de objetos tridimensionales es ambigua: un mismo modelo puede corresponder a varios objetos reales que tengan las mismas aristas pero distintas superficies, como se demuestra en las figuras 2-8 y 2-9. En la figura 2-8 se muestra el modelado alámbrico de un cubo (a) y dos posibles interpretaciones del mismo en (b) y (c). La figura 2-9 presenta el ejemplo de la escalera de Schröder: a partir de esta imagen uno no es capaz de decidir si corresponde a una escalera vista desde arriba

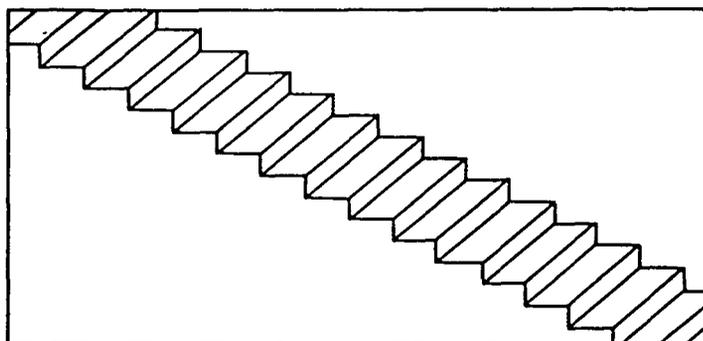


Figura 2-9: Escalera de Schröder. Obsérvese que la posición del punto de vista es ambigua.

o vista desde abajo.

Por el mismo motivo, el método no es capaz de discernir entre un objeto real y un objeto imposible caracterizado con las mismas aristas. Por todo ello, esta técnica se usa muy raramente en aplicaciones prácticas.

b) Modelado por hilos

Es la técnica de discretización de superficies arbitrarias más antigua utilizada para análisis electromagnético. Su origen data de los años 60, y aproxima la superficie del objeto por una malla de hilos o aristas artificiales. El helicóptero de la figura 2-10 es un ejemplo clásico de modelado por hilos.

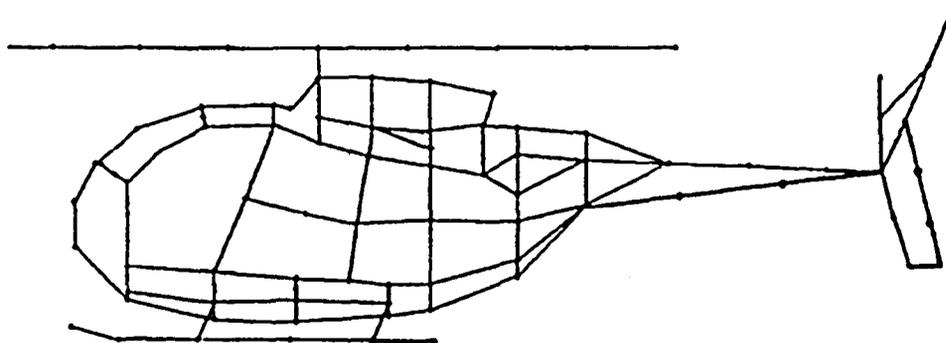


Figura 2-10: Modelado de superficies mediante hilos de corriente

Desde el punto de vista electromagnético, el modelado por hilos equivale a aproximar la radiación de las corrientes superficiales por la radiación de hilos de corriente situados sobre la superficie. Este método permite obtener buenos resultados cuando se calculan parámetros de campo lejano, como diagrama de radiación de una antena o sección recta de un blanco radar. Sin embargo, no es adecuado para el cálculo de parámetros de campo próximo, como impedancias de entrada o corrientes inducidas sobre las superficies. El programa NEC (Numerical Electromagnetic Code) [Burke'81] es el ejemplo más conocido de modelado por hilos para análisis mediante el Método de los Momentos. En España también se han desarrollado trabajos en este área, por ejemplo, en la Universidad Politécnica de Madrid [Heras'90a] [Heras'90b].

c) Modelado por placas planas (facetas)

Este método consiste en aproximar las superficies frontera mediante mallas de polígonos, generando un sólido en forma de poliedro. Las caras del poliedro son facetas poligonales planas, cuyo número de lados es en general variable. La base de datos que define los objetos contiene la información de los vértices, las aristas y las facetas que forman las superficies. Para que el modelo realmente represente un sólido y no un objeto imposible, debe cumplir una serie de propiedades, como por ejemplo: cada arista debe conectar 2 vértices y debe estar compartida por dos caras, al menos 3 aristas deben encontrarse en un vértice y las caras no deben interceptarse entre sí.

Este método permite la representación de cualquier geometría. Sin embargo, mientras que la representación de objetos cuyas superficies son caras planas es idéntica al objeto real, la representación de objetos en las que hay alguna superficie curvada es siempre una aproximación. Esta aproximación será tanto más exacta cuanto menor sea el tamaño de las facetas que se utilicen. En el avión que se presenta en la figura 2-11 se demuestra que el número de facetas aumenta enormemente al aumentar la complejidad del objeto. Esta técnica de modelado presenta las siguientes ventajas:

- **Dominio muy amplio:** No existen limitaciones en la geometría del objeto.
- **Control local:** Proporciona una gran flexibilidad para las modificaciones de los modelos. Permite modificar una zona concreta del objeto dejando el resto intacto, lo cual facilita la obtención de modelos precisos.

Los inconvenientes principales son:

- **Poco conciso:** La aproximación de superficies curvadas requiere una gran cantidad de facetas. Cuando los objetos son grandes y complejos esta enorme cantidad de información puede penalizar apreciablemente los tiempos de entrada/salida y de cálculo.

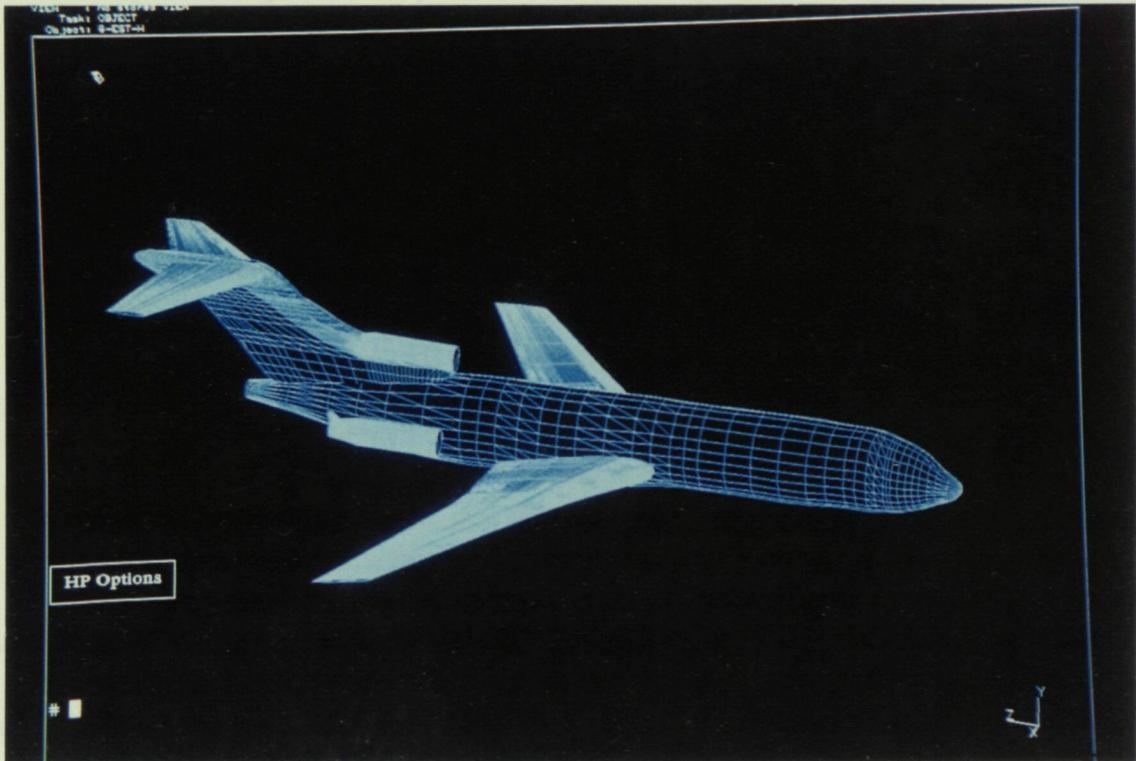


Figura 2-11: Modelado de un avión mediante facetas

- **Precisión limitada:** Al aproximar una superficie curvada mediante facetas, no puede evitarse la aparición de aristas no reales en la superficie del objeto, correspondientes a las uniones entre facetas (ver figura 2-12).

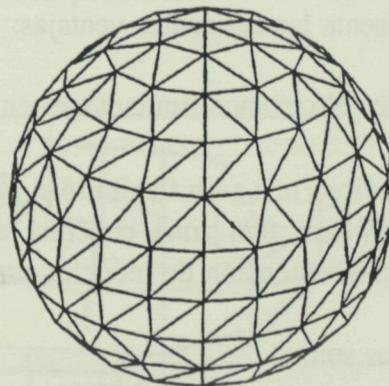


Figura 2-12: Modelado de una esfera mediante facetas triangulares

Esta técnica se utiliza en muchas herramientas CAD de modelado de sólidos por la flexibilidad que ofrece para el diseño de cualquier tipo de objetos, sea cual sea su

geometría. Es además, una de las técnicas más utilizadas en la actualidad en el diseño de blancos complejos para el análisis electromagnético en alta frecuencia. El programa RECOTA desarrollado en Boeing Aerospace [Youssef'89] utiliza modelado mediante facetas y aristas para un posterior estudio mediante técnicas de alta frecuencia. También se ha utilizado este tipo de modelado en España, como por ejemplo en el programa TOTAL desarrollado en la Universidad de Cantabria [Domingo'90], [Abad'90], [Domingo'92].

Los modelos de facetas planas también se utilizan para el análisis electromagnético en baja frecuencia. Por ejemplo, los triángulos de Rao, Glisson y Wilton [Rao'81] son facetas triangulares de tamaño eléctricamente pequeño que aproximan la superficie del objeto. En la figura 2-12 se muestra la discretización de una esfera utilizando estos parches triangulares.

d) Modelado mediante parches curvos

La técnica de facetas sólo puede modelar exactamente objetos poliédricos. Para aproximar superficies curvadas con una precisión aceptable requiere un número muy elevado de facetas y, además, no puede evitarse la aparición de aristas artificiales. Por ello, en la actualidad se han desarrollado nuevas técnicas que utilizan parches curvos (habitualmente superficies paramétricas) que se ajustan mucho mejor a las superficies curvadas y, además, la cantidad de información que precisan es en general mucho menor. En la figura 2-13 se muestra una superficie curva modelada mediante 100 facetas y en la figura 2-14 la misma superficie modelada mediante un único parche curvo.

Las dos formas matemáticas no lineales más comunes en el modelado geométrico para representación de curvas y superficies son las formas polinomiales paramétricas y las implícitas o algebraicas. Estas últimas son de desarrollo muy reciente y poseen la ventaja de que los círculos, cónicas y cuádricas pueden ser representados exactamente y de manera muy concisa. Sin embargo, las curvas y superficies de forma arbitraria -tan importantes en modelado geométrico- son representadas más fácilmente mediante polinomios paramétricos a trozos. Además, el hardware de las estaciones gráficas disponibles en la actualidad sólo permite representar en tiempo real objetos modelados por superficies paramétricas de tipo B-Spline racional no uniforme (NURB) (ver apéndice 2A), por lo que las superficies algebraicas no se utilizan todavía en la práctica.

Las ventajas más importantes que presenta el modelado geométrico con parches curvos son:

- **Concisión:** No precisa almacenar gran cantidad de datos en los modelos de superficies curvadas. Al ser una representación más concisa que la de facetas se reduce en varios órdenes de magnitud la cantidad de información a procesar, con

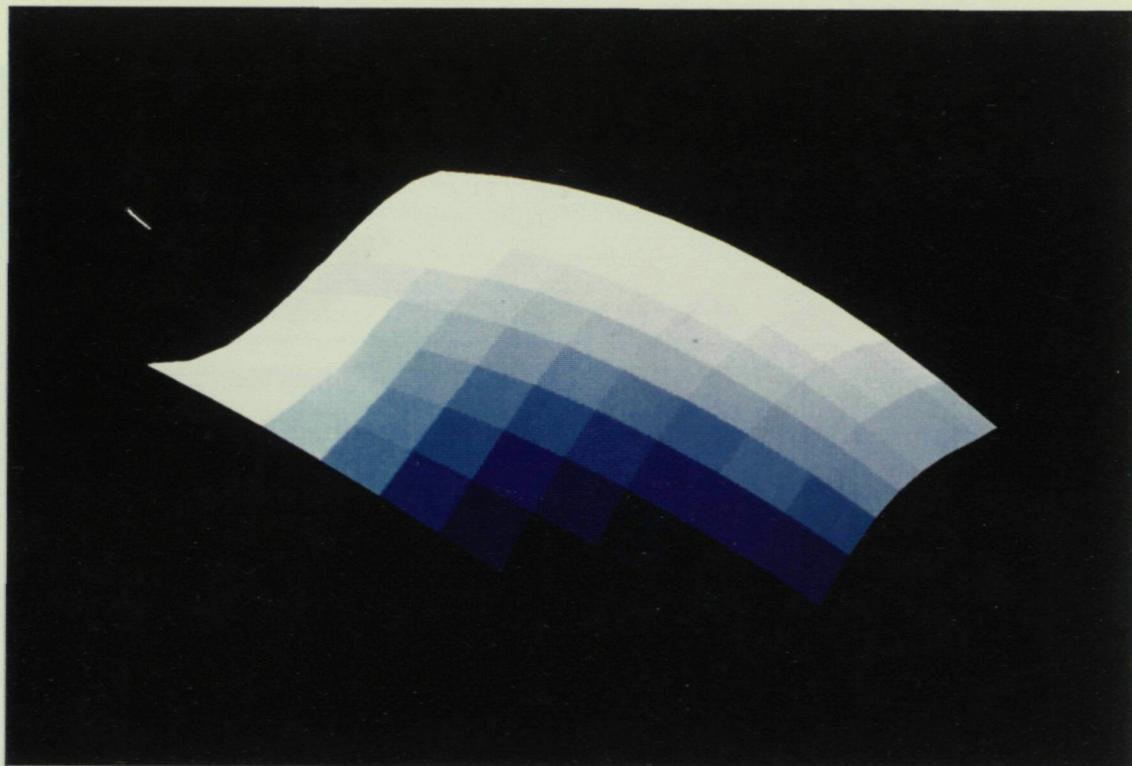


Figura 2-13: Superficie modelada mediante facetas

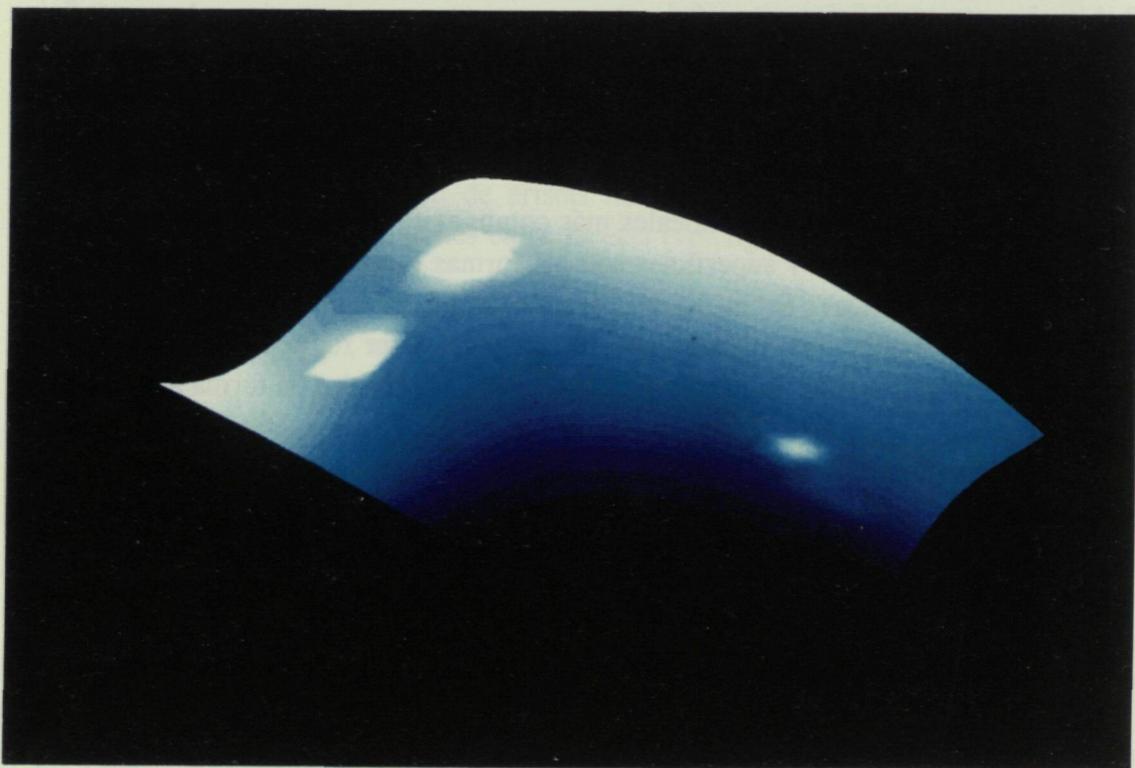


Figura 2-14: Parche superficial curvo

todas las ventajas que ello conlleva.

- **Precisión:** Permite aproximar las superficies reales con gran exactitud.
- **Facilidad de visualización:** El hardware gráfico de las estaciones de trabajo más potentes permite la visualización realista tridimensional de superficies paramétricas NURBS en tiempo real.

Los programas de diseño asistido por ordenador para el modelado de objetos tridimensionales incluyen el diseño mediante superficies paramétricas. Dada la importancia actual de esta técnica en el apéndice 2A se desarrolla la formulación matemática de las curvas y superficies paramétricas.

El diseño de objetos mediante superficies paramétricas se suele llevar a cabo de una de las siguientes formas:

- Se definen directamente las superficies paramétricas frontera del objeto, situándolas después con la orientación adecuada en el lugar preciso.
- **Barrido generalizado:** Se diseñan los perfiles bidimensionales, en general distintos, los cuales se sitúan con una cierta orientación en unas posiciones determinadas a lo largo de una trayectoria concreta. Esta estructura forma el esqueleto del objeto, de manera que la superficie frontera lo recubre ajustándose al mismo como si se tratara de una piel. Por ello, a este proceso de modelado se le denomina en terminología inglesa *skin group*. El objeto que se obtiene mediante barrido generalizado se almacena como un conjunto de superficies paramétricas. En las figuras siguientes se detallan los distintos pasos que intervienen en este proceso de modelado: 2-15 diseño de perfiles, 2-16 situación de los perfiles a lo largo de una trayectoria y 2-17 generación de las superficies frontera.

Esta técnica de modelado se está utilizando cada vez más en electromagnetismo, habiéndose realizado en España los primeros trabajos importantes:

- El programa *GRECO* de cálculo de sección recta radar por alta frecuencia desarrollado en la Universidad Politécnica de Cataluña [Rius'93b] utiliza superficies paramétricas NURB para el modelado de blancos radar reales (ver capítulo 5). La técnica de procesado gráfico de una imagen del blanco evita la formulación directa de las aproximaciones de alta frecuencia sobre las superficies paramétricas, lo cual permite una enorme velocidad de cálculo.
- En la Universidad de Vigo se ha desarrollado un programa de análisis de cavidades mediante óptica geométrica y trazado de rayos [Hermida'90a,90b,91a,91b,93] basado en modelos de superficies paramétricas.

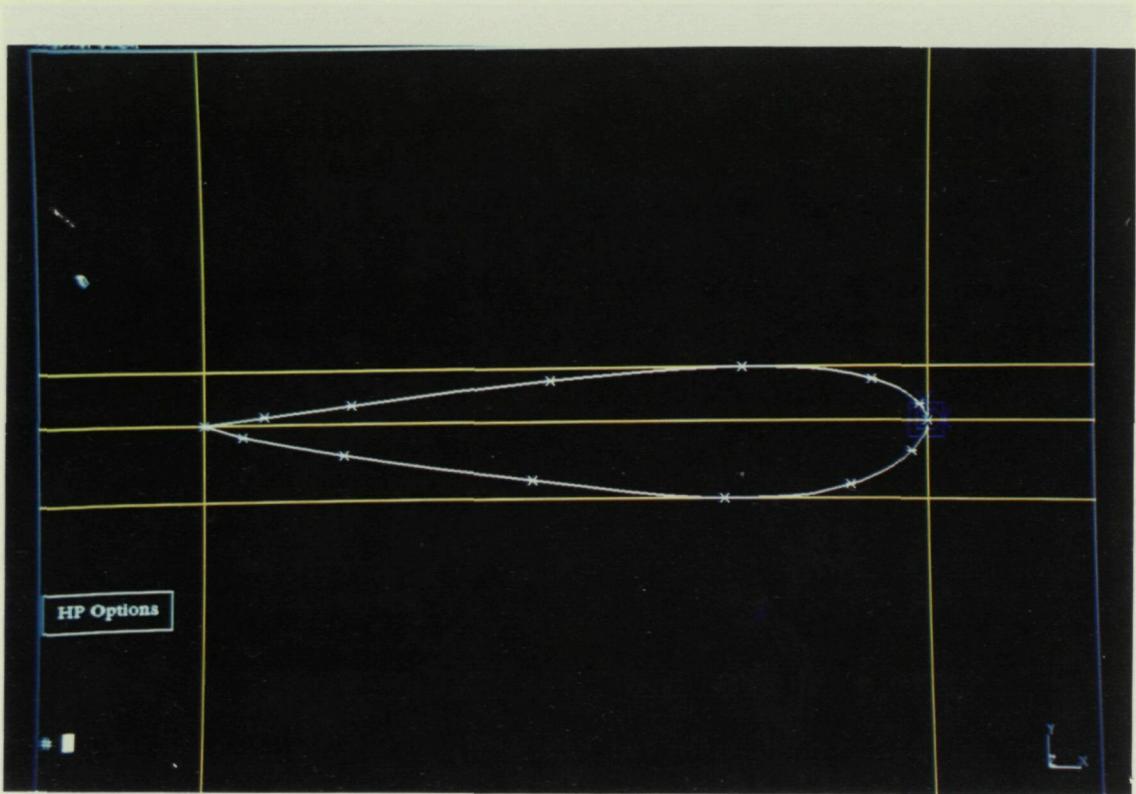


Figura 2-15: Diseño de un perfil del ala de un avión

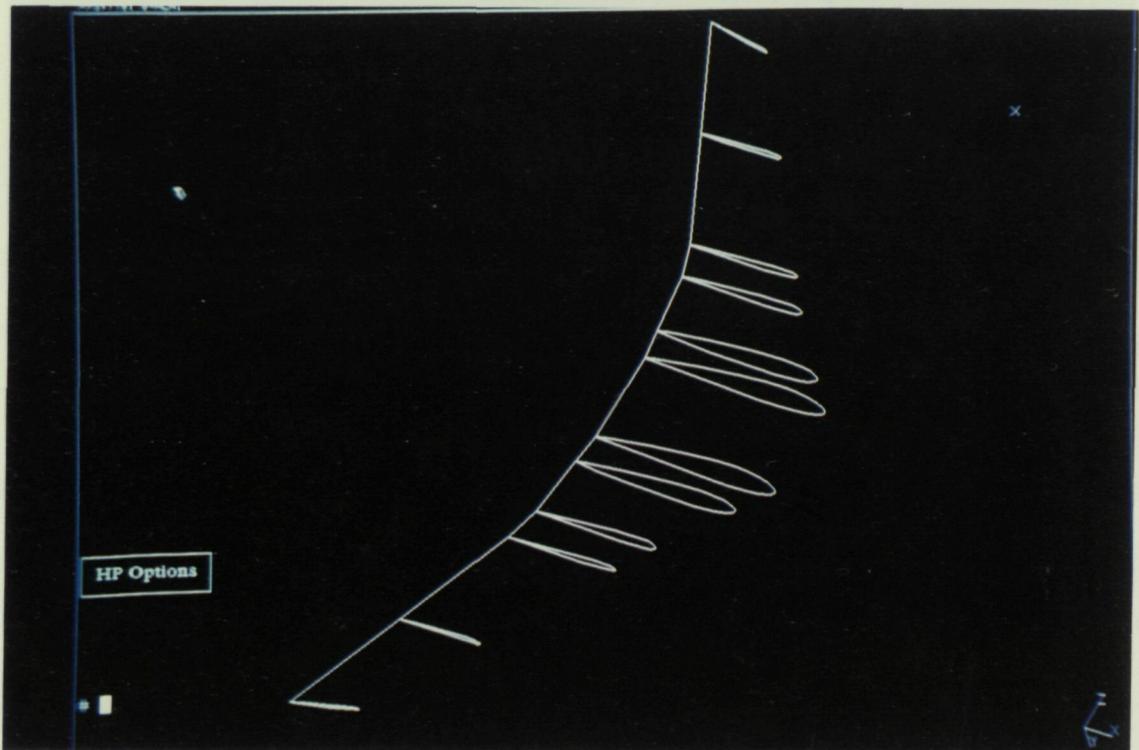


Figura 2-16: Situación de los perfiles a lo largo de una trayectoria

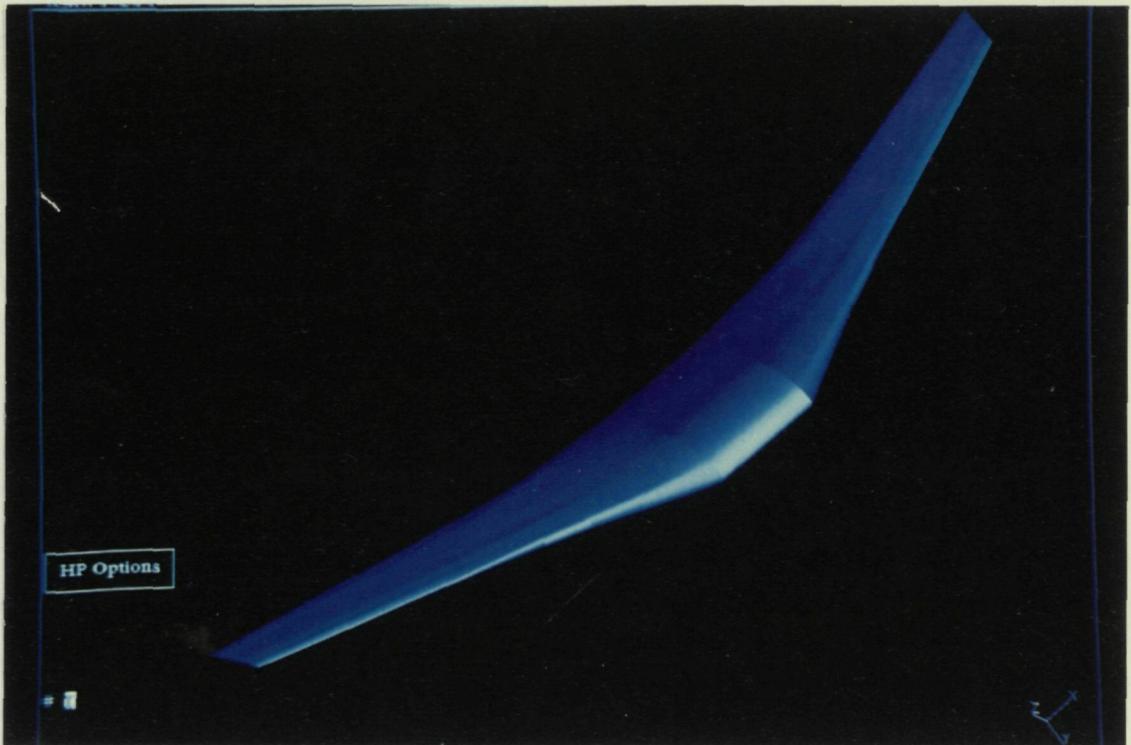


Figura 2-17: Modelo del ala del avión por superficies paramétricas

- En la Universidad de Cantabria se ha realizado la formulación e implementación de las aproximaciones de alta frecuencia sobre superficies paramétricas NURB [Perez'92] [Cátedra'93].
- La aplicación de técnicas de baja frecuencia a modelos de superficies paramétricas también se ha llevado a cabo en la Universidad de Cantabria, mediante el método de los momentos [Valle'92].

2.2.4 Representación mediante particiones del espacio

Estos métodos se basan en la descomposición del objeto en una colección de sólidos simples no intersectados y adjuntos entre sí. Estos sólidos más sencillos se denominan primitivas o elementos. Las primitivas que se utilizan pueden ser de distintos tipos, tamaños, posición, parametrización y orientación; lo cual da lugar a diferentes técnicas de descomposición con sus características particulares.

Algunos de estos métodos, como el de los elementos finitos o el mallado regular, son muy utilizados en aplicaciones de cálculo numérico, tanto para el análisis mecánico

como electromagnético. Estos métodos se presentan a continuación, junto con otros no tan utilizados.

a) Descomposición mediante celdas (o elementos finitos)

Una de las formas más generales de partición del espacio es la descomposición mediante celdas o elementos. Éstas se sitúan una junto a otra, sin que se intersecten, llenando todo el espacio ocupado por el objeto. Según el tipo de celdas que se utilicen pueden obtenerse distintos modelos de un mismo objeto. Estos métodos permiten, en general, el modelado de cualquier objeto arbitrariamente complejo.

La formulación matemática de las celdas o elementos básicos suele realizarse mediante polinomios de interpolación. Las celdas o elementos se pueden clasificar según el grado de los polinomios que se utilicen, siendo los más comunes los lineales y cuadráticos. Estos últimos aproximan bien las superficies de curvatura suave sin presentar una complejidad excesiva. Otra clasificación habitual se basa en el número de caras de estas celdas, siendo los más comunes los sólidos no regulares de 4 ó 6 caras. Un ejemplo de discretización mediante celdas de 6 caras y polinomios de orden uno (lineales) se muestra en la figura 2-18.

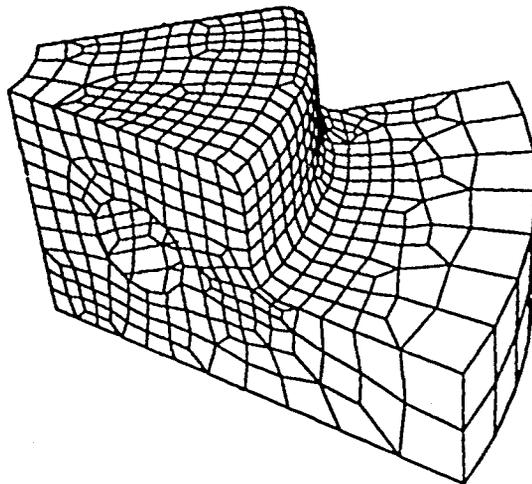


Figura 2-18: Ejemplo de discretización mediante celdas de 6 caras y orden 1

Esta técnica de modelado corresponde a lo que se conoce en ingeniería como mallado por elementos finitos. Estos mallados constituyen, posiblemente, la herramienta más potente y flexible para discretizar con precisión superficies arbitrarias muy complejas. Su utilización está muy extendida en problemas de ingeniería mecánica y civil (análisis

de deformaciones de piezas sometidas a esfuerzos, presas, puentes, etc.). Por este motivo, existe un gran número de programas de diseño asistido por ordenador que incorporan modelado geométrico, mallado en elementos finitos, análisis numérico (de tipo mecánico) y postprocesado de resultados. La figura 2-19 muestra un mallado bidimensional en elementos finitos cuadrangulares de orden 2 (cuadráticos) realizado con el software I-DEAS.

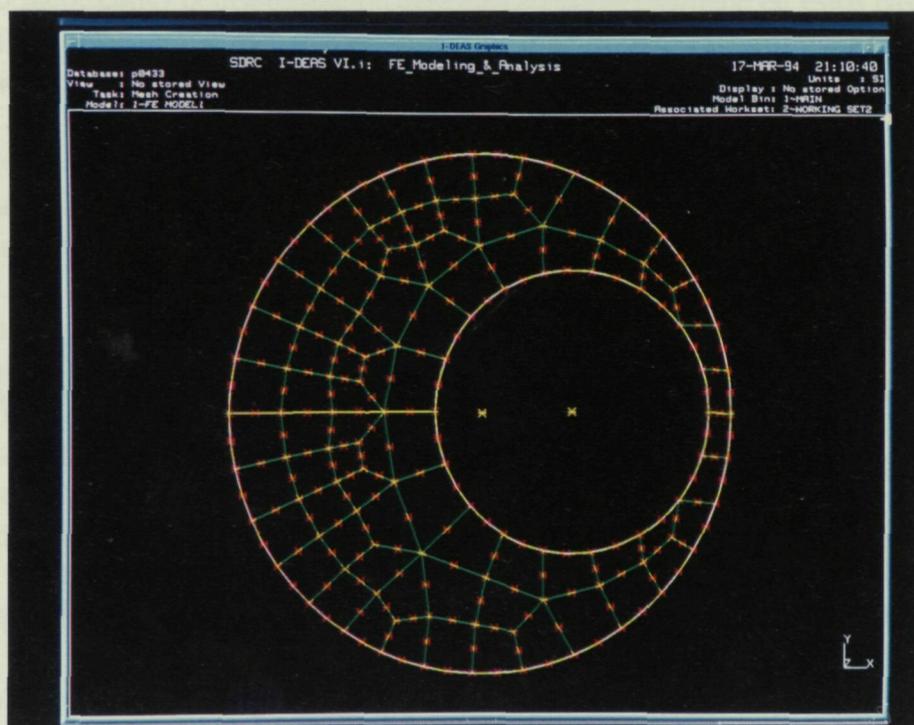


Figura 2-19: Modelo de elementos finitos cuadrangulares de orden 2.

En el campo electromagnético también se han realizado en España numerosas implementaciones de éste método, entre las que pueden citarse las de las Universidades Politécnicas de Catalunya [Cruellas'89,93] [Duffo'90,93a,93b] [Prieto'92] [Abad'92,93] [Alvarez'90] [Vallmitjana'92], Valencia [Nuño'93] y Madrid [Salazar'89].

b) Enumeración de ocupación del espacio

Es un caso especial de descomposición en celdas en el cual los elementos son todos idénticos y se sitúan formando una malla regular y fija. La celda de este tipo más común es el cubo. El modelo consiste únicamente en indicar la presencia o ausencia de objeto, o el tipo de material del mismo, en una posición determinada de la malla.

En caso de utilizar el cubo como celda base sólo es posible la representación exacta de objetos cuyas caras externas sean paralelas a las caras del cubo y cuyos vértices

coincidan con los de la malla. Cualquier otro tipo de objeto sólo puede modelarse de forma aproximada. La precisión del modelo es tanto mejor cuanto menor sea el tamaño de la celda utilizada. En la figura 2-20 se muestra la aproximación de un toro mediante una malla regular de cubos.

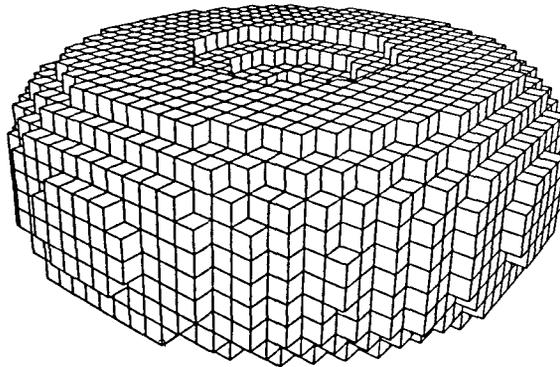


Figura 2-20: Discretización de un toro por la técnica de enumeración de ocupación del espacio

Este método presenta las siguientes ventajas:

- **Dominio:** Es muy adecuado para modelar objetos extraordinariamente complejos.
- **Obtención de las propiedades físicas:** El cálculo de volúmenes y áreas es trivial. Basta conocer las dimensiones de la celda básica y el número de celdas en el interior y la superficie del objeto.
- **Invarianza frente a operaciones booleanas,** las cuales son triviales.

Sin embargo, también presenta algunos inconvenientes, la mayoría de ellos derivados de la gran cantidad de celdas a procesar una por una:

- **Poca exactitud:** debido a la aproximación por poliedros regulares (cubos).
- **Poca concisión:** por ejemplo, para modelar objetos sencillos como una esfera, precisa de mucha información a almacenar.
- **Lentitud en las transformaciones geométricas** como rotaciones o traslaciones.
- **Lentitud en la visualización realista tridimensional.**

La posibilidad de representar objetos muy complejos hace este método muy apropiado para aplicaciones biomédicas. En la Universidad Politécnica de Catalunya se han realizado trabajos de tomografía de microondas [Rius'87,92a] modelando por este método tanto los resultados de las reconstrucciones como las maquetas para simulación numérica [Llorens'93].

Por otro lado, como la discretización volumétrica que se obtiene en estos casos forma una malla regular, la matriz que define el sistema de ecuaciones a resolver en el análisis por el método de los momentos es de Toeplitz o de Toeplitz a bloques, lo cual reduce el coste computacional del almacenamiento e inversión de la matriz de impedancias. Por otra parte, para este tipo de matrices las operaciones que aparecen al aplicar el método del gradiente conjugado tienen la forma de una convolución discreta y pueden implementarse eficientemente mediante la transformada rápida de Fourier (FFT).

Otro método que aprovecha las ventajas de la malla regular para obtener una formulación discreta muy simple es el de diferencias finitas en el dominio del tiempo (FDTD).

En España se han realizado últimamente importantes trabajos con esta técnica de modelado y análisis por el método del gradiente conjugado más transformada rápida de Fourier [Cátedra'89,90a] [Vaquer'93] o análisis por diferencias finitas en el dominio del tiempo [Gonzalez'90a,90b].

c) Árboles octales (*Octrees*)

Ésta es una variante del método de la enumeración de ocupación del espacio, que optimiza la cantidad de memoria requerida para almacenar el objeto. El *Octree* deriva del *quadtree*, representación bidimensional desarrollada originalmente para codificación de imágenes [Foley'90].

La idea de ambos métodos es la de dividir el espacio a representar en forma de árbol de decisión sobre la ocupación de las celdas. En el *quadtree* se divide el espacio bidimensional en dos direcciones perpendiculares para formar cuatro cuadrantes. Cada cuadrante puede estar lleno, parcialmente lleno o vacío (también se denominan cuadrante negro, gris o blanco respectivamente). Todo cuadrante parcialmente lleno se subdivide recursivamente en cuatro subcuadrantes. Las subdivisiones se suceden hasta que todos los cuadrantes sean homogéneos (totalmente llenos o totalmente vacíos) o hasta que el tamaño de los mismos sea menor que un valor de precisión prefijado. Las sucesivas subdivisiones se representan mediante un árbol con cuadrantes parcialmente llenos en los nodos y cuadrantes homogéneos en las hojas.

En la figura 2-21 se muestra el modelado de un objeto bidimensional mediante *quadrees*. Las celdas blancas corresponden a zonas externas del objeto, mientras que las grises corresponden a zonas parcialmente ocupadas. Puede observarse que solamente se subdividen las zonas del objeto que presentan una geometría más compleja, con lo que se consigue una subdivisión adaptativa. La figura 2-22 presenta el árbol correspondiente a la subdivisión mostrada en la figura 2-21. De cada nodo gris del árbol se bifurcan cuatro ramas, cada una de las cuales corresponde a un subcuadrante obtenido en la subdivisión.

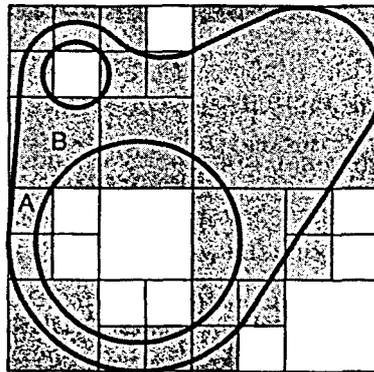


Figura 2-21: Modelado de una estructura bidimensional mediante *quadrees*

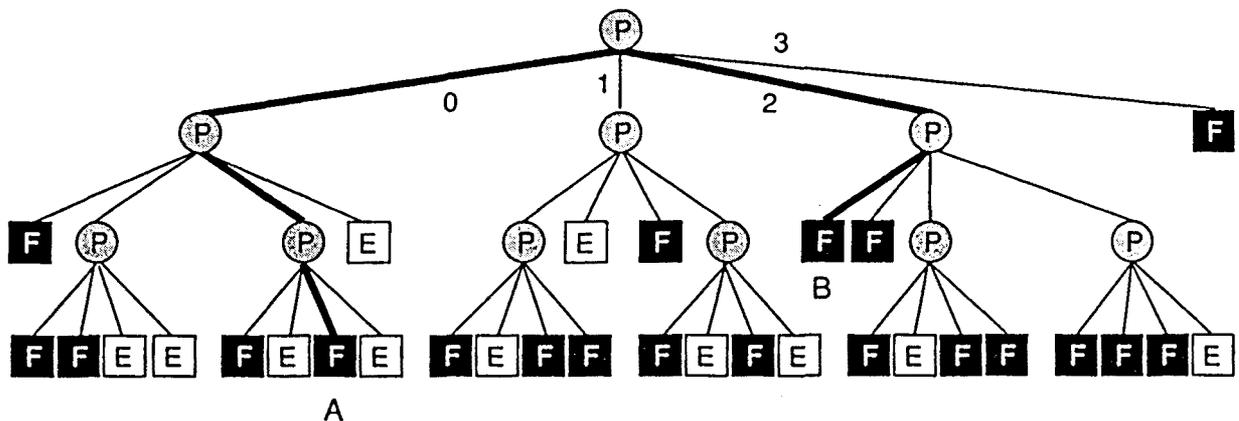


Figura 2-22: Árbol de la representación en *quadrees* para el objeto de la figura 2-21. "F" indica totalmente lleno, "E" indica totalmente vacío y "P" parcialmente lleno

El *octree* se basa en el mismo concepto que el *quadtree*, pero en este caso las tres dimensiones del espacio se dividen recursivamente en octantes. En la figura 2-23 se presenta la subdivisión espacial en octantes para el caso tridimensional.

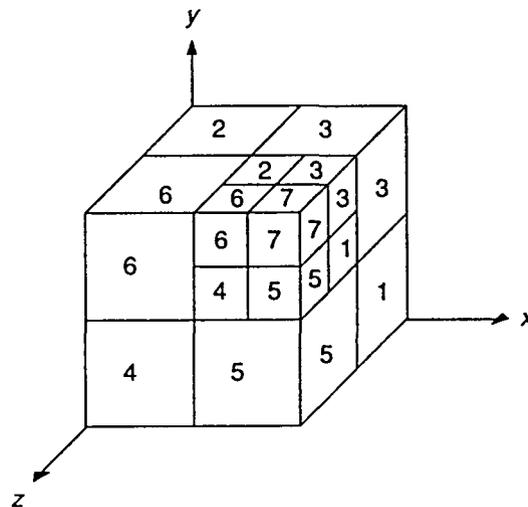


Figura 2-23: Subdivisión del espacio tridimensional en octantes

Este método presenta ventajas e inconvenientes similares a los del método de enumeración de ocupación del espacio, por tratarse de una variante del mismo. Sin embargo, es un método adaptativo en el sentido de que las zonas de formas más complicadas quedan más divididas y las de variación suave menos, con lo que se optimiza la cantidad de información necesaria para modelar con precisión un objeto muy complejo. Se trata, por tanto, de un método más conciso que el anterior. Un ejemplo típico de aplicación de esta técnica es el modelado del cerebro en biomedicina.

No se conocen aplicaciones electromagnéticas de este método, aunque su campo de aplicación potencial se encuentra en el análisis por FDTD o CGM + FFT de modelos formados por regiones homogéneas separadas por superficies extremadamente complejas. Ello permitiría reducir el espacio de almacenamiento a costa de aumentar el tiempo de proceso.

d) Árboles de partición binaria del espacio

Los árboles de partición binaria del espacio (BSP) dividen recursivamente el espacio en pares de subespacios, separados por un plano de posición y orientación arbitrarias. La estructura de datos del árbol binario fue originariamente utilizada para determinar si las superficies eran visibles o no en los gráficos (ver capítulo 3), sin embargo Naylor y Thibault posteriormente lo adaptaron como técnica de representación de poliedros no regulares [Foley'90].

El modelo geométrico está formado por un árbol tal que cada nodo corresponde a un plano de subdivisión del espacio el cual contiene una de las caras del poliedro. De cada nodo parten dos hijos que contienen la geometría que queda en cada semiespacio. Si se asume que las normales siempre apuntan hacia fuera del objeto se puede decidir si la región definida por cada hijo (delimitada por la superficie que la genera y por las superficies ya estudiadas con anterioridad) está totalmente dentro del objeto, totalmente fuera o en una situación mixta. En los dos primeros casos esta rama del árbol termina en una hoja. En el tercer caso se coloca un nuevo nodo correspondiente a una nueva subdivisión del espacio a partir de otra cara del poliedro. El proceso finaliza cuando queda perfectamente delimitado el volumen que ocupa el poliedro.

La figura 2-24 muestra un polígono bidimensional modelado mediante esta técnica, junto con el árbol binario que representa el modelo.

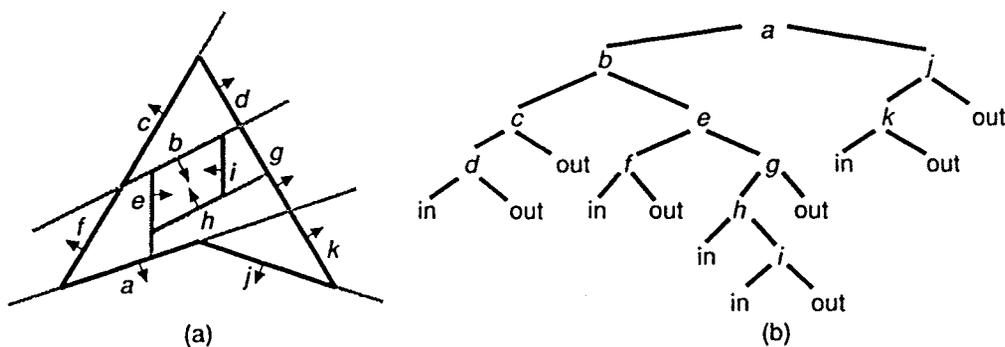


Figura 2-24: a) Modelado en árbol BSP de un polígono bidimensional, b) El árbol BSP

Aunque este método sea una forma sencilla y elegante de modelar poliedros, la representación mediante un árbol impide que muchas operaciones geométricas y booleanas sean compactas y eficientes. La principal ventaja del árbol BSP consiste en que se basa en una definición independiente del número de dimensiones del espacio y, además, es muy apropiado para determinar si un punto está dentro o fuera del objeto, así como para detectar intersecciones entre poliedros.

La aplicación de esta técnica a los problemas electromagnéticos está en sus inicios. La principal aplicación potencial reside en la determinación de facetas visibles y oscuras de un modelo de facetas (2.2.3 c) para los métodos de alta frecuencia. Sin embargo, existen otras técnicas más adecuadas para esta tarea (ver capítulo 5) por lo que la utilización de los árboles BSP no se ha extendido.

2.2.5 Geometría constructiva de sólidos

El método de geometría constructiva de sólidos (CSG) se basa en la combinación de objetos sencillos, como esferas, cilindros o cubos, para obtener un objeto complejo mediante operaciones booleanas. Estos sólidos primitivos son parametrizables: pueden especificarse sus dimensiones, posición y orientación.

El objeto complejo modelado se almacena como un árbol formado por nodos internos que contienen operaciones booleanas. Las hojas del árbol contienen los sólidos primitivos junto con las operaciones de traslación, rotación y escalado que se han realizado sobre los mismos. En la construcción del árbol hay que tener en cuenta que la mayoría de operaciones booleanas no son conmutativas, por lo que es muy importante mantener el orden adecuado en la definición del mismo. En la figura 2-25 se presenta un ejemplo de modelado por CSG.

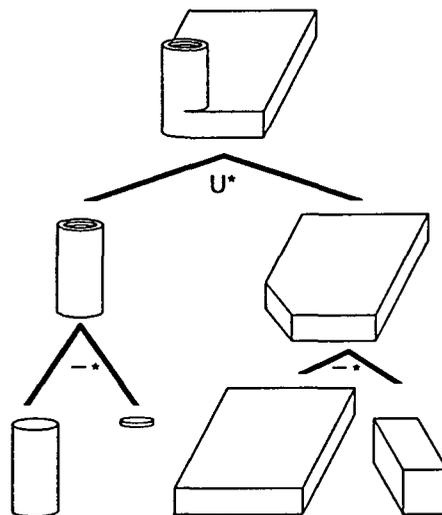


Figura 2-25: Diseño de una pieza mediante el método de geometría constructiva de sólidos

Las principales ventajas de la geometría constructiva de sólidos son:

- **Concisión:** basta muy poca información para definir el objeto.
- **Operaciones booleanas:** gran sencillez en la realización de operaciones booleanas entre objetos.

- **Transformaciones geométricas:** facilidad en la realización de transformaciones geométricas.

Los mayores inconvenientes residen en:

- **Dominio:** dificultad en el modelado de formas complejas.
- **Obtención de propiedades físicas:** los cálculos geométricos sobre el objeto son muy costosos.
- **Facilidad de visualización:** el dibujo bidimensional y la visualización realista tridimensional no son sencillas.

Este método se utiliza principalmente en software de diseño mecánico y no se conocen aplicaciones electromagnéticas.

2.3 COMPARACION DE METODOS Y CONCLUSIONES

En este capítulo se han presentado los métodos de modelado geométrico tridimensional más comunes. Un estudio comparativo de las ventajas e inconvenientes de los distintos métodos, según las bases expuestas en la sección 2.1, podría ser el siguiente:

- **Precisión:** Los métodos de partición del espacio (excepto cuando se utilizan celdas de caras curvadas) y el método modelado por fronteras mediante placas planas no pueden modelar con suficiente precisión objetos muy complejos. Esto no es un inconveniente para muchas aplicaciones. Sin embargo, en aquellos casos en los que se requiera una precisión muy elevada, la gran cantidad de información que se necesita almacenar para definir objetos complejos por estas técnicas puede hacerlas inviables.

Los métodos de geometría constructiva de sólidos y los métodos de fronteras que utilizan parches curvos obtienen aproximaciones muy buenas, para objetos que pertenecen al dominio del método, con una buena concisión en la definición de los mismos.

La técnica de modelado de objetos mediante parámetros consigue modelos de gran precisión, dentro de su dominio, pero no es un método invariante frente a operaciones booleanas.

Dominio: El dominio de los objetos que se pueden representar mediante los métodos de barrido y parámetros es limitado. En cambio, los métodos de partición del espacio y modelado por fronteras (excepto aristas) pueden representar cualquier sólido, aunque en muchas ocasiones no se pueda obtener la suficiente precisión.

Unicidad: Solamente el método de los árboles octales y el de la enumeración de ocupación del espacio garantizan la unicidad de una representación: existe una única representación de un objeto con un tamaño y posición concretos. Esto no así para los modelos de fronteras.

Los modelos de parámetros no presentan unicidad porque un mismo objeto puede definirse mediante conjuntos de primitivas distintos, por ejemplo una esfera puede representarse como una primitiva esférica o como una primitiva elíptica. Sin embargo, sí puede existir unicidad dentro de un sólo conjunto de primitivas escogido adecuadamente.

Validez: Los objetos modelados por fronteras son los más difíciles de validar, pues no hay un algoritmo eficiente que pueda determinar si estos modelos se corresponden a un sólido válido o no. Por otra parte, los modelos mediante árboles de partición binaria del espacio siempre consiguen volúmenes válidos, aunque no necesariamente un sólido cerrado.

Para validar los modelos obtenidos mediante geometría constructiva de sólidos (CSG) y árboles octales sólo es necesaria una comprobación sintáctica local del árbol que define el modelo. Además, en el caso de CSG si las primitivas empleadas son objetos cerrados se garantiza que los nuevos objetos también lo son.

El método de enumeración de ocupación del espacio garantiza siempre la obtención de sólidos reales.

Concisión y eficiencia: Las principales ventajas del método CSG son que genera modelos muy compactos y es capaz de realizar y almacenar operaciones booleanas y transformaciones geométricas con mucha facilidad. Esto es debido a que estas manipulaciones sólo implican añadir nuevos nodos al árbol. Por esta misma razón, tanto el método de árboles octales como el de árboles binarios (BSP) también son muy adecuados para la realización de operaciones booleanas entre modelos.

Por otro lado, tanto los métodos de modelado mediante fronteras como el de enumeración de ocupación del espacio requieren de procesos mucho más

complejos para la realización de operaciones booleanas. Sin embargo, no es así para otras operaciones tales como determinar si un punto del espacio está dentro o fuera del objeto, que resultan mucho más simples con este tipo de modelos que con los basados en árboles.

En general, los modelos de fronteras con parches curvos resultan mucho más compactos que los de facetas cuando los cuerpos a modelar presentan superficies curvadas. Esto los hace muy atractivos para el modelado de objetos de gran complejidad y por ello este método se ha implementado en la gran mayoría de las herramientas CAD comerciales.

A partir de estas características de los distintos métodos de modelado se deduce que no hay ningún método que sea claramente superior a los demás, sino que según sea la aplicación para la que se vayan a utilizar los modelos unas u otras técnicas resultan más apropiadas. En el problema de análisis electromagnético que nos ocupa cada método de cálculo utiliza una técnica particular de modelado de sólidos y de discretización. Esto ha repercutido en que muchas de las herramientas CAD disponibles permiten el diseño de objetos por distintas técnicas, facilitando al usuario la elección de las mismas.

Finalmente, es necesario remarcar que aunque actualmente existen algoritmos y sistemas muy potentes como herramientas de modelado de sólidos todavía quedan muchos problemas sin resolver. Uno de los más importantes es la robustez de los métodos, pues muchos de ellos están afectados por las inestabilidades numéricas de los ordenadores: algunos algoritmos requieren más precisión en la representación en coma flotante que la que puede proporcionar el hardware de los ordenadores actuales.

APENDICE 2A

CURVAS Y SUPERFICIES PARAMETRICAS

2A.1 MODELADO DE SUPERFICIES ARBITRARIAS

Para llevar a cabo el modelado de sólidos mediante fronteras son necesarios modelos matemáticos de curvas y superficies. En este apéndice se va a introducir el área del modelado por superficies. Este área es muy amplia, siendo los métodos de modelado más comunes: superficies mediante mallas de polígonos, superficies paramétricas, representación implícita o algebraica de superficies y la representación mediante funciones explícitas. El modelado mediante mallas poligonales es una aproximación lineal, mientras que los demás tipos de representaciones consiguen aproximaciones de orden superior.

a) Mallas de polígonos

Las mallas poligonales son un conjunto de vértices, aristas y polígonos conectados de tal forma que una arista es compartida por dos polígonos. Se puede hacer un modelado muy intuitivo mediante estas placas planas de cualquier superficie. Sin embargo, tanto los polígonos como las polilíneas para curvas se representan por funciones lineales, por lo que sólo pueden modelar con exactitud formas representables por funciones de primer orden. Por ello, a no ser que las curvas y superficies a representar sean también lineales es necesario generar y almacenar gran cantidad de información para conseguir aproximaciones de una precisión aceptable. Esto provoca que cualquier manipulación interactiva de los datos para aproximar una geometría resulte costosa porque hay que posicionar muchos puntos con una precisión muy elevada.

b) Superficies paramétricas

Las curvas y superficies paramétricas ($x = x(t)$, $y = y(t)$, $z = z(t)$) consiguen una representación más compacta y manipulable de curvas y superficies de variación suave. Estos métodos utilizan habitualmente polinomios a trozos de grado mayor a 1. Estas funciones se ajustan mejor a las superficies del modelo original que las mallas poligonales logrando una definición más compacta de superficies arbitrarias, con lo que la manipulación interactiva resulta más sencilla que en los modelos mediante funciones lineales. Al presentar estas ventajas estos son los métodos que más se utilizan en la actualidad como herramienta de modelado de sólidos de alta calidad. En esta sección se va a hacer un estudio exhaustivo de las representaciones mediante curvas y superficies paramétricas desarrolladas hasta la actualidad con más repercusión práctica, tanto en el modelado de curvas como en el modelado de superficies. Se hace un estudio de las curvas porque por un lado son útiles por si mismas en el modelado tridimensional de intersecciones de superficies en los objetos y porque, además, las superficies paramétricas son una generalización de las curvas, las cuales resultan más manejables, tanto para el estudio de la definición para cada una de las técnicas, como de las propiedades de las geometrías obtenidas.

c) Superficies algebraicas implícitas

La representación de superficies mediante funciones implícitas o algebraicas ($f(x,y,z)=0$) son de desarrollo muy reciente. Presentan la ventaja de que los círculos, esferas, cónicas y cuádricas pueden ser representados exactamente de forma muy concisa. Sin embargo, las curvas y superficies de formas arbitrarias, tan importantes en el modelado geométrico, son representadas más fácilmente mediante polinomios paramétricos a trozos. En muchas ocasiones a la ecuación implícita de modelado del objeto habrá que añadir otras ecuaciones o inecuaciones que modelen restricciones sobre las mismas (un ejemplo es el modelado de una semiesfera, en el cual se requiere de la ecuación implícita de la esfera más la inecuación que indique que hemisferio se está modelando).

d) Superficies algebraicas explícitas

Finalmente, también cabe la posibilidad de representar objetos mediante funciones explícitas de y y z respecto a x ($y = f(x)$, $z = g(x)$). Las principales desventajas que presentan este tipo de representaciones son: (1) Es imposible diseñar curvas con múltiples valores de y o z para un mismo valor de x , por tanto curvas tales como círculos y elipses sólo pueden representarse mediante la unión de múltiples segmentos. (2) Una definición de este tipo no es invariante respecto a rotaciones (la descripción de una curva o superficie rotada requiere en muchas ocasiones de la subdivisión en varios segmentos de la curva original). (3) La representación de curvas y superficies con

alguna tangente infinita resulta difícil de definir. Estas desventajas conducen a que estas funciones no se utilicen como técnica de modelado.

2A.2 TIPOS DE CURVAS Y SUPERFICIES PARAMETRICAS

En la introducción que se ha expuesto al principio de este apéndice se han presentado las distintas formas de modelado mediante superficies señalando algunas de las propiedades de las mismas, a partir de ellas se deduce que dentro de las aproximaciones de órdenes superiores a 1 la formulación paramétrica es la más versátil; lo que ha producido que, en la actualidad, sea la herramienta más utilizada para el modelado de curvas y superficies arbitrarias. Esta es la razón por la que en esta tesis se dedica un apéndice a la formulación de las mismas. En realidad, dentro del término *curvas y superficies paramétricas* quedan englobadas gran cantidad funciones, pero sólo algunas resultan realmente útiles como técnicas de modelado. A continuación, se introducen los criterios de discriminación entre unas y otras funciones y en los demás apartados se exponen los modelos matemáticos de las formulaciones más utilizadas.

Las propiedades básicas que se exigen a cualquier formulación matemática de curvas y superficies son:

- **Flexibilidad:** Debe adaptarse con facilidad a curvas arbitrarias muy complejas.
- **Intuitividad:** A partir de los parámetros que definen una curva, debe poderse intuir su forma.

El problema más común en el diseño de curvas es la interpolación de unos puntos que constituyen los datos del problema. A cada punto de los datos P_i se le asocia un valor del parámetro t_i propio de la representación paramétrica. Las primeras técnicas que resuelven este problema encuentran una polinomial de interpolación de estos puntos. Entre otros algoritmos que se han desarrollado se puede nombrar el algoritmo de Aitken y los polinomios interpoladores de Lagrange [Farin'88].

El algoritmo de Aitken es una técnica recursiva que obtiene una polinomial interpolante a través de una secuencia de interpolaciones lineales, sin embargo, no obtiene una única representación por objeto y es difícil encontrar una expresión cerrada de las curvas diseñadas.

Los polinomios de Lagrange, por otro lado, presentan las ventajas de obtener una interpolación simple y única y de conseguir una buena interpolación geométrica. Por tanto, parecería lógico que se usara frecuentemente en el diseño de curvas que interpolen un conjunto de puntos; sin embargo, no es así. La razón principal es que los polinomios

interpolantes generan curvas con oscilaciones extrañas, no intrínsecas a los puntos de datos, al utilizar polinomios de grado elevado. En la figura 2A-1 se muestra una curva que al interpolar unos puntos mediante el polinomio correspondiente de Lagrange, la curva obtenida presenta unos recovecos que son totalmente imprevisibles a partir de los puntos de datos a interpolar.

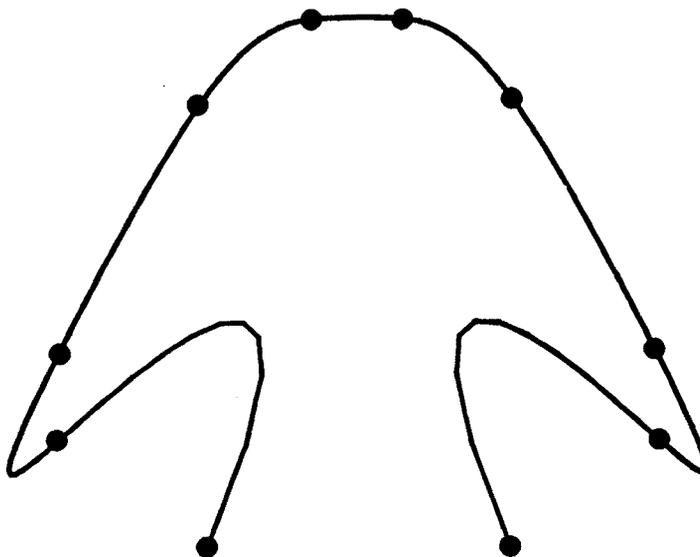


Figura 2A-1: Interpolación de Lagrange. Mientras los puntos de datos sugieren una curva interpolante convexa, la interpolación de Lagrange presenta unas oscilaciones extrañas y no inherentes a los datos.

Este fenómeno no es debido a efectos numéricos; está, en realidad, inherente al proceso de interpolación polinómica de Lagrange. De ahí que, aún cuando lo lógico sería esperar que en el diseño de una curva al ir incrementando el número de puntos intermedios a interpolar, se generaran curvas interpolantes de orden cada vez mayor, las cuales convergieran a la curva a diseñar, en general, no es cierto al utilizar los polinomios de Lagrange. Ello también repercute en tiempos de cálculo muy elevados, tanto en el proceso de diseño de estas curvas como en la evaluación de las mismas. Por tanto, es una técnica muy poco flexible y no resulta adecuada para el diseño de curvas y superficies.

Las técnicas de interpolación no están restringidas a la interpolación de puntos de datos; sino que también se pueden interpolar otras informaciones, tales como datos de las derivadas en puntos de la curva. Esto conduce a esquemas de interpolación más útiles que la interpolación de Lagrange. El método más conocido que interpola información de las derivadas en los puntos extremos de la curva son los *polinomios de Hermite*, a los que a continuación se les va a dedicar una sección.

La principal desventaja de los polinomios de Hermite es que no son invariantes bajo transformaciones en el espacio afín. Esta es la razón principal por la que, cuando es posible, se utilizan otras técnicas de interpolación como pueden ser los polinomios de Bézier.

Las curvas de Bézier proporcionan una potente herramienta de diseño de curvas, pero tienen algunas limitaciones. Entre ellas cabe destacar la siguiente: si la curva a modelar presenta una geometría muy compleja, entonces su representación mediante curvas de Bézier requiere un grado muy elevado de la función polinómica, el cual puede llegar a ser prohibitivo (las curvas de Bézier de grado superior a 10 se convierten en intratables). Estas curvas tan complejas pueden, sin embargo, ser modeladas utilizando curvas de Bézier compuestas. Uniendo polinomios de grado reducido puede conseguirse una polinomial a trozos muy flexible y adaptable a formas arbitrariamente complejas, sin necesidad de elevar el grado de los polinomios. Este tipo de curvas se denominan *B-splines*.

En las curvas *B-splines* la intuitividad se consigue expresando los coeficientes de los mismos en función de las coordenadas espaciales de unos "puntos de control", que son los vértices de una poligonal cuya forma es similar a la curva descrita por los polinomios.

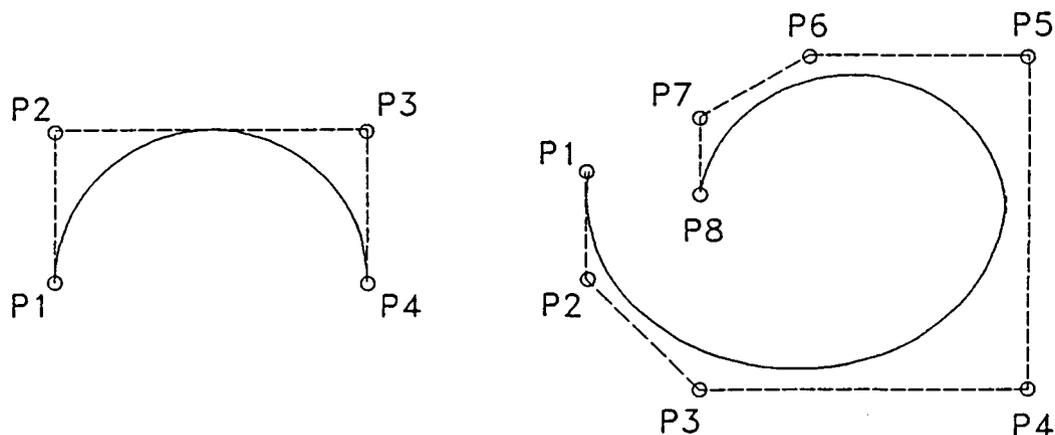


Figura 2A-2: Ejemplo de curvas paramétricas B-spline

En la figura 2A-2 se representan 2 ejemplos de curva polinómica a trozos, "*B-spline*", con las correspondientes poligonales definidas por los puntos de control. Obsérvese que la curva "*B-spline*" queda siempre dentro de la poligonal.

Recientemente se han desarrollado varios métodos, además de los *B-splines*, para generar polinomios paramétricos con gran intuitividad. Entre ellos destacan: los γ -*splines*, V^2 -*splines*, β -*splines*, v -*splines*, etc. [Brunet'90]. Sin embargo, se ha decidido utilizar los "*B-splines*" para modelar los objetos complejos en esta tesis por ser la herramienta de modelado clásica, sólidamente establecida, que se encuentra implementada en muchos sistemas CAD/CAM, en las librerías gráficas software y en el hardware de los aceleradores gráficos de las estaciones de trabajo actuales. Solamente se ha optado por los modelos de facetas cuando la superficie de los objetos está constituida por caras planas, caso en el que éste método es el más conciso.

2A.3 FORMULACION DE LAS CURVAS PARAMETRICAS

La representación mediante curvas paramétricas aproxima la curva por secciones polinomiales en vez de utilizar trozos de recta. Cada segmento de curva Q de la curva se define mediante tres funciones x , y y z , las cuales son polinomios de grado superior a 1 en el parámetro t . La figura 2A-3 muestra dos segmentos de curva paramétricos y sus polinomiales, para el caso de dos dimensiones con coordenadas $x(t)$, $y(t)$. Además se demuestra la posibilidad de representar curvas con múltiples valores de y , para un mismo valor de x . Las gráficas $x(t)$ y $y(t)$ para el segundo segmento han tenido que ser trasladadas y tienen el inicio en $t = 1$, para demostrar la continuidad de la curva entre segmentos en el punto de unión.

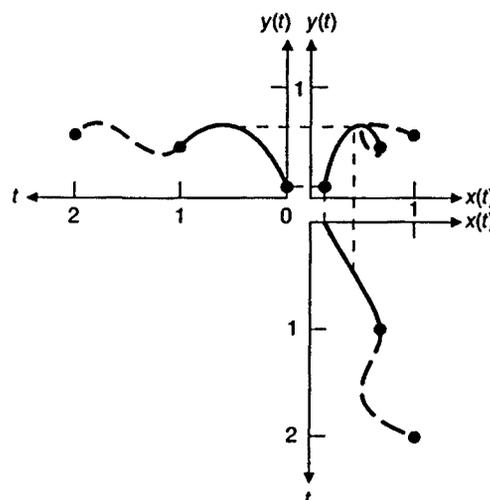


Figura 2A-3: Representación de 2 segmentos de curva y las polinomiales que las definen en el espacio 2D.

Las curvas polinomiales cúbicas son las más utilizadas por las siguientes razones:

- Los polinomios de grado menor ofrecen poca flexibilidad, mientras que los polinomios de mayor grado pueden introducir variaciones indeseadas en la curva y además requieren mayor carga computacional.
- En la definición de un polinomio de tercer grado sólo hay que imponer cuatro restricciones, mientras que en polinomios de grado mayor es necesario imponer más condiciones para determinar los coeficientes que los definen, lo que los hace más difíciles de generar.
- Representaciones de grado menor que las cúbicas no permiten obligar a una curva a pasar por dos puntos específicos con una derivada determinada en cada uno de ellos.
- Los polinomios cúbicos son los primeros que permiten definir curvas no planas en un espacio de tres dimensiones.

Esto justifica el que solamente se usen polinomios de grado mayor al cúbico en aplicaciones concretas, tales como el diseño en aeronáutica y automoción, en las que se requiere garantizar la continuidad de las derivadas de orden alto para diseñar superficies aerodinámicamente eficientes. Por este motivo, los paquetes de software CAD usualmente permiten la definición de curvas de hasta grado 6.

Debido a las propiedades expresadas arriba, al desarrollar las expresiones matemáticas de los polinomios de las curvas paramétricas se trabajará con polinomios cúbicos, para obtener de forma intuitiva y elegante la formulación de los mismos. Al final de cada sección se definirá la fórmula general para un polinomio de grado arbitrario "n".

Los polinomios cúbicos que definen un segmento de curva $Q(t) = [x(t) \ y(t) \ z(t)]$ se definen por el sistema de ecuaciones siguiente:

$$\begin{aligned} x(t) &= a_x t^3 + b_x t^2 + c_x t + d_x \\ y(t) &= a_y t^3 + b_y t^2 + c_y t + d_y \\ z(t) &= a_z t^3 + b_z t^2 + c_z t + d_z \end{aligned} \quad 0 \leq t \leq 1 \quad (2A.1)$$

Si se define $T = [t^3 \ t^2 \ t \ 1]$ y la matriz de los coeficientes de los tres polinomios como

$$C = \begin{bmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \\ d_x & d_y & d_z \end{bmatrix} \quad (2A.2)$$

se puede escribir la ecuación (2A.1) de forma más compacta como:

$$Q(t) = [x(t) \quad y(t) \quad z(t)] = T \cdot C \quad (2A.3)$$

La derivada de $Q(t)$ es el vector paramétrico tangente a la curva. Aplicando esta definición a la ecuación (2A.3) se obtiene:

$$\begin{aligned} \frac{d}{dt}Q(t) &= Q'(t) = \left[\frac{d}{dt}x(t) \quad \frac{d}{dt}y(t) \quad \frac{d}{dt}z(t) \right] = \frac{d}{dt}T \cdot C \\ \frac{d}{dt}Q(t) &= [3t^2 \quad 2t \quad 1 \quad 0] \cdot C \\ \frac{d}{dt}Q(t) &= [3a_x t^2 + 2b_x t + c_x \quad 3a_y t^2 + 2b_y t + c_y \quad 3a_z t^2 + 2b_z t + c_z] \end{aligned} \quad (2A.4)$$

La derivada de las curvas se calcula con objeto de poder garantizar la continuidad de la derivada en los puntos de unión entre dos segmentos de curva. Si dos segmentos de curva se unen en un punto de la curva generada se tiene *continuidad geométrica* G^0 . Si las direcciones (pero no necesariamente las magnitudes) de los dos vectores tangentes al punto de unión en cada segmento son iguales, la curva posee *continuidad geométrica* G^1 . En el diseño de objetos por ordenador muy a menudo se exige la *continuidad geométrica* G^1 de las curvas. Cuando los vectores tangentes a cada segmento en el punto de unión son idénticos (la dirección y la magnitud son iguales), la curva posee *continuidad de primer grado en el parámetro t* , o *continuidad paramétrica de orden 1*, se dice que posee *continuidad* C^1 . Cuando la dirección y la magnitud de $d^n [Q(t)]/dt^n$ en el punto de unión, se dice que la curva es C^n *continua*.

En la figura 2A-4 se muestra una curva formada por un segmento S unido en un punto a otros tres segmentos de curva C_0 , C_1 y C_2 con *continuidad paramétrica de orden 0*, *1* y *2*, respectivamente. La diferencia visual entre las curvas C_1 y C_2 es tan pequeña que prácticamente no se aprecia, pero se hace obvia en puntos más alejados.

Obsérvese que cada segmento de curva paramétrica mantiene la continuidad en cualquier punto; la continuidad sólo debe ser controlada en los puntos de unión entre segmentos. De las definiciones de *continuidad visual* y de *continuidad paramétrica* se deduce que, en general, tener *continuidad paramétrica* C^1 implica tener *continuidad visual* G^1 , sin embargo el caso contrario no es generalmente cierto.

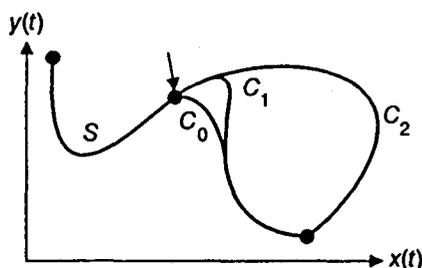


Figura 2A-4: El segmento de curva S se une a los segmentos C_0 , C_1 y C_2 , con continuidad paramétrica de orden 0, 1 y 2, respectivamente.

Un segmento de curva $Q(t)$ se define por restricciones en los puntos extremos, los vectores tangentes y la continuidad entre segmentos de curva. Cada polinomio cúbico tiene cuatro coeficientes, de ahí que sea necesario imponer 4 restricciones, que se utilizan para hallar los coeficientes de las cuatro ecuaciones que permiten encontrar los demás puntos de la curva. A continuación se van a desarrollar los tipos de curvas más utilizados:

- **Curvas de Hermite:** Definidas, en el caso cúbico, por los dos puntos extremos del segmento y las tangentes a ambos puntos.
- **Curvas de Bézier:** Definidas (caso cúbico) por las coordenadas de los dos puntos extremos y por otros dos puntos que controlan la derivada en los puntos extremos.
- **Curvas splines:** Cada una definida por cuatro puntos de control. Los splines de orden cúbico garantizan la continuidad de C^1 y C^2 en los puntos de unión y quedan cerca de la polinomial definida por los puntos de control, pero generalmente no pasan por los puntos de control. Los tipos de splines que se van a comentar son: los *B-splines no racionales uniformes*, *B-splines no racionales no uniformes* y los *B-splines racionales*.

Para ver como los coeficientes de la ecuación (2A.1) pueden depender de las cuatro restricciones impuestas, se considera que la curva paramétrica cúbica general se representa de forma compacta por $Q(t) = T \cdot C$, y se reescriben los coeficientes de la matriz como $C = M \cdot G$, donde M es una *matriz base* 4×4 , y G es una matriz 4×3 cuyas filas son 4 vectores que contienen la información de las cuatro restricciones. A la matriz G se la conoce como *vector geométrico*. Se usan G_x para referir al vector columna de las componentes x del vector geométrico, G_y y G_z tienen definiciones paralelas para las componentes y y z , respectivamente. M o G , o ambos, son distintos para cada tipo de curva. Para cada una de ellas los elementos de M y G son constantes,

por tanto el producto $T \cdot M \cdot G$ es un polinomio cúbico en t . Expandiendo el producto $Q(t) = T \cdot M \cdot G$ se llega a la ecuación

$$Q(t) = [x(t) \quad y(t) \quad z(t)] = [t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} g_{1x} & g_{1y} & g_{1z} \\ g_{2x} & g_{2y} & g_{2z} \\ g_{3x} & g_{3y} & g_{3z} \\ g_{4x} & g_{4y} & g_{4z} \end{bmatrix} \quad (2A.5)$$

Obteniendo la curva de la componente $x(t)$

$$x(t) = (t^3 m_{11} + t^2 m_{21} + t m_{31} + m_{41})g_{1x} + (t^3 m_{12} + t^2 m_{22} + t m_{32} + m_{42})g_{2x} + (t^3 m_{13} + t^2 m_{23} + t m_{33} + m_{43})g_{3x} + (t^3 m_{14} + t^2 m_{24} + t m_{34} + m_{44})g_{4x} \quad (2A.6)$$

La ecuación (2A.6) demuestra que la curva es una suma ponderada de los elementos de la matriz geométrica. Cada uno de los pesos son polinomios cúbicos en el parámetro t , y se conocen como *funciones blending*. Las funciones blending B vienen dadas por $B = T \cdot M$. El cálculo de la matriz base M se va a desarrollar para cada tipo de curvas paramétricas que se estudian a continuación.

2A.3.1 Curvas de Hermite

Las curvas cúbicas de Hermite se diseñan imponiendo la posición de los puntos extremos de la misma P_1 y P_4 y los vectores tangentes a dichos puntos extremos R_1 y R_4 (se utilizan los subíndices 1 y 4 en vez de 1 y 2 para obtener una notación más acorde con la utilizada para los demás tipos de curvas).

Para encontrar los coeficientes de la *matriz base de Hermite* M_H , la cual relaciona el *vector geométrico de hermite* G_H con los coeficientes del polinomio, se escriben cuatro ecuaciones, una para cada una de las cuatro restricciones, siendo las incógnitas los coeficientes de las ecuaciones; una vez hallados se pueden encontrar los demás puntos de la curva.

Se define G_{Hx} , la componente x de la matriz geométrica de Hermite como

$$G_{Hx} = \begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix}_x \quad (2A.7)$$

y reescribiendo $x(t)$ a partir de las ecuaciones (2A.1) y (2A.5) se llega a

$$\begin{aligned} x(t) &= a_x t^3 + b_x t^2 + c_x t + d_x = T \cdot C_x \\ x(t) &= T \cdot M_H \cdot G_{Hx} = [t^3 \quad t^2 \quad t \quad 1] \cdot M_H \cdot G_{Hx} \end{aligned} \quad (2A.8)$$

Las condiciones a cumplir en $x(0)$ y $x(1)$ (puntos extremos del segmento porque se considera $0 \leq t \leq 1$) se encuentran directamente de la sustitución en la ecuación (2A.8):

$$x(0) = P_{1x} = [0 \quad 0 \quad 0 \quad 1] M_x \cdot G_{Hx} \quad (2A.9)$$

$$x(1) = P_{4x} = [1 \quad 1 \quad 1 \quad 1] M_x \cdot G_{Hx} \quad (2A.10)$$

Derivando la ecuación (2A.8)

$$x'(t) = [3t^2 \quad 2t \quad 1 \quad 0] M_H \cdot G_{Hx} \quad (2A.11)$$

se pueden aplicar las dos restricciones de los vectores tangentes en los puntos extremos del segmento:

$$x'(0) = R_{1x} = [0 \quad 0 \quad 1 \quad 0] M_H \cdot G_{Hx} \quad (2A.12)$$

$$x'(1) = R_{4x} = [3 \quad 2 \quad 1 \quad 0] M_H \cdot G_{Hx} \quad (2A.13)$$

De forma que las cuatro restricciones se pueden incluir en una única ecuación

$$\begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix}_x = G_{Hx} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} M_H \cdot G_{Hx} \quad (2A.14)$$

Por tanto, a partir de esta ecuación (y lo mismo ocurriría para las coordenadas y y z) se obtiene que M_H debe ser la matriz inversa de la que aparece en la ecuación anterior

$$M_H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (2A.15)$$

Esta matriz es única y puede utilizarse en $x(t) = T \cdot M_H \cdot G_H$ para encontrar $x(t)$ con el vector geométrico G_{Hx} (para las coordenadas y y z se obtienen expresiones paralelas). Si se expande el producto $T \cdot M_H$ de $Q(t) = T \cdot M_H \cdot G_H$ se obtienen las

funciones blending de Hermite como polinomios de ponderación de cada uno de los elementos del vector geométrico:

$$\begin{aligned} Q(t) &= T \cdot M_H \cdot G_H = B_H \cdot G_H \\ &= (2t^3 - 3t^2 + 1)P_1 + (-2t^3 + 3t^2)P_4 + (t^3 - 2t^2 + t)R_1 + (t^3 - t^2)R_4 \end{aligned} \quad (2A.16)$$

La figura 2A-5 muestra las 4 funciones Blending de Hermite. Notar que para $t = 0$, solamente la función etiquetada por P_1 es distinta de cero, por tanto en $t = 0$ sólo afecta el punto P_1 . Para valores de t mayores de 0, R_1 , P_4 y R_4 empiezan a tener influencia.

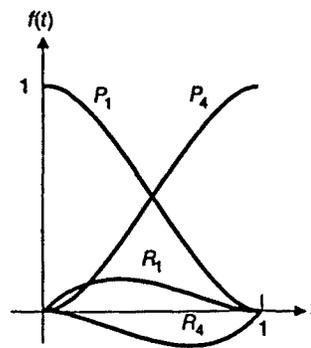


Figura 2A-5: Funciones Blending de Hermite, etiquetadas por los elementos del vector geométrico a los que ponderan

El diseño de curvas de Hermite de grado quinto imponen como restricciones de diseño la segunda derivada de los puntos extremos, además de la posición de los mismos y primera derivada en dichos puntos. Por tanto, la generación de curvas de grados mayores van introduciendo restricciones sobre la derivada de grados superiores.

Como estas curvas son combinaciones lineales (sumas ponderadas) de los elementos del vector geométrico, se pueden hacer transformaciones de la curva aplicando la transformación sobre el vector geométrico y utilizar este nuevo vector para generar la curva transformada, lo cual es equivalente a decir que las curvas obtenidas son equivalentes bajo rotación, escalado y traslación, pero no bajo proyecciones de perspectiva. Esta estrategia resulta mucho más eficiente que la de aproximar la curva por una serie de segmentos de recta y después hacer la transformación sobre cada línea individual.

La principal desventaja de los polinomios de Hermite es que no son invariantes bajo transformaciones en el espacio afín. Ello conduce a la utilización generalizada de otras técnicas de modelado como las que a continuación se comentan [Farin'88].

2A.3.2 Curvas de Bézier

Las curvas de Bézier imponen como restricciones el valor de los puntos extremos y determinan de forma indirecta el valor de las tangentes a dichos puntos especificando las coordenadas de dos puntos intermedios adicionales que no pertenecen a la curva. En la figura 2A-6 se muestran 2 ejemplos de curvas de Bézier con los correspondientes puntos de control. Notar que la curva queda definida en el interior de la poligonal que se forma a partir de la unión por rectas de los cuatro puntos. Los vectores tangentes a los puntos extremos se determinan por los vectores P_1P_2 y P_3P_4 y se relacionan con R_1 y R_4 por

$$\begin{aligned} R_1 &= Q'(0) = 3(P_2 - P_1) \\ R_4 &= Q'(1) = 3(P_4 - P_3) \end{aligned} \tag{2A.17}$$

por tanto, las curvas de Bézier interpolan 2 puntos de control y aproximan los otros 2 (caso cúbico).

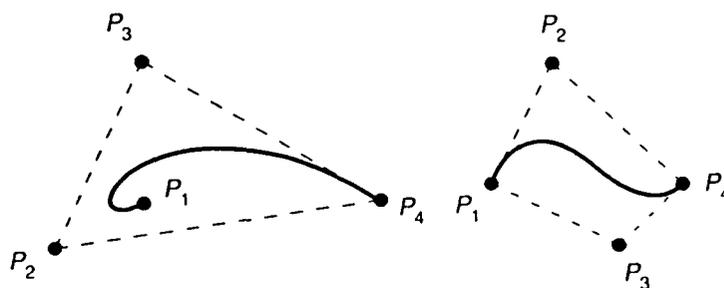


Figura 2A-6: Dos curvas de Bézier con los correspondientes puntos de control

El vector geométrico de Bézier G_B , tiene por elementos a los cuatro puntos

$$G_B = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} \tag{2A.18}$$

Entonces la matriz M_{HB} que define la relación $G_H = M_{HB} \cdot G_B$ entre el vector geométrico de Hermite G_H y el vector geométrico de Bézier G_B es la matriz 4x4 que representa la ecuación (4.17) en forma matricial:

$$G_H = \begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -3 & 3 & 0 & 0 \\ 0 & 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = M_{HB} \cdot G_B \quad (2A.19)$$

Para encontrar la matriz base de Bézier M_B , puede usarse la ecuación $Q(t) = T \cdot M_H \cdot G_H$ para la forma de Hermite con las siguientes sustituciones

$$\begin{aligned} G_H &= M_{HB} \cdot G_B \\ M_B &= M_H \cdot M_{HB} \end{aligned} \quad (2A.20)$$

quedando

$$\begin{aligned} Q(t) &= T \cdot M_H \cdot G_H = T \cdot M_H \cdot (M_{HB} \cdot G_B) \\ &= T \cdot (M_H \cdot M_{HB}) \cdot G_B = T \cdot M_B \cdot G_B \end{aligned} \quad (2A.21)$$

Llevando a cabo el producto $M_B = M_H \cdot M_{HB}$ se obtiene

$$M_B = M_H \cdot M_{HB} = \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (2A.22)$$

y el producto $Q(t) = T \cdot M_B \cdot G_B$ es

$$Q(t) = (1-t)^3 P_1 + 3t(1-t)^2 P_2 + 3t^2(1-t) P_3 + t^3 P_4 \quad (2A.23)$$

Las cuatro polinomiales $B_B = T \cdot M_B$, las cuales son los pesos de la ecuación (2A.23) se conocen como *polinomios de Bernstein* y son los que se presentan en la figura 2A-7.

La figura 2A-8 muestra dos segmentos de curva de Bézier con un punto final en común. Obsérvese en la figura que la continuidad G^1 se obtiene en el punto final cuando $P_3 - P_4 = k(P_4 - P_5)$, $k > 0$, es decir los tres puntos P_3 , P_4 y P_5 deben ser distintos y colineales. Si se cumple además para $K = 1$, se consigue también continuidad C^1 .

Si consideramos los dos segmentos de curva en el punto de unión como x^l (segmento izquierdo) y x^r (segmento derecho) las condiciones de continuidad C^0 y C^1 se consiguen en el punto de unión cuando:

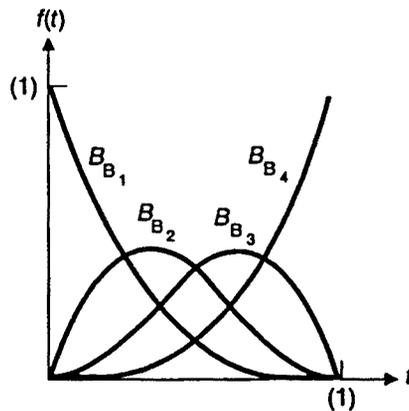


Figura 2A-7: Polinomios de Bernstein, que representan las funciones peso de las curvas de Bézier

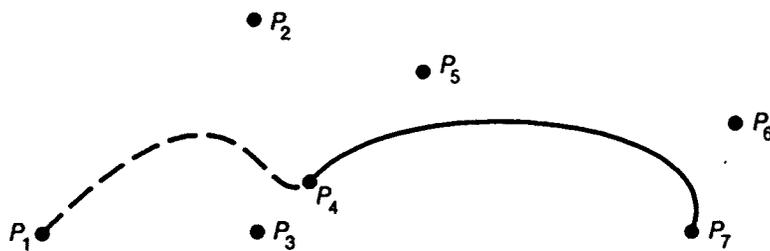


Figura 2A-8: Dos curvas de Bézier unidas en el punto P_4 . Los puntos P_3 , P_4 y P_5 son colineales

$$\begin{aligned} x'(1) &= x'(0) \\ \frac{d}{dt} x'(1) &= \frac{d}{dt} x'(0) \end{aligned} \tag{2A.24}$$

Lógicamente, estas condiciones de continuidad se consiguen para $P_4 - P_3 = P_5 - P_4$.

En la definición de las polinomiales de Bézier de grado "n" se utilizan los polinomios de Bernstein de grado "n" como

$$\begin{aligned} B_i^n(t) &= \binom{n}{i} t^i (1-t)^{n-i} \\ \binom{n}{i} &= \begin{cases} \frac{n!}{i!(n-i)!} & \text{si } 0 \leq i \leq n \\ 0 & \text{en otro caso} \end{cases} \end{aligned} \tag{2A.25}$$

quedando la curva de Bézier de grado "n" expresada de la siguiente forma

$$Q^n(t) = \sum_{i=0}^n P_i B_i^n(t) \quad (2A.26)$$

donde los P_i son los $n+1$ puntos a interpolar y $B_i^n(t)$ son los polinomios de Bernstein de grado " n ".

Al examinar los 4 polinomios de Bernstein utilizados en la definición de las curvas de Bézier cúbicas se aprecia que su suma para cualquier valor del parámetro t es la unidad y, además, que todas las polinomiales toman valores comprendidos en el margen $0 \leq t \leq 1$. Por tanto, $Q(t)$ resulta ser la media ponderada de los cuatro puntos de control (caso cúbico).

En general, cualquier curva de Bézier, sea cual sea el grado de la misma, se define por la suma ponderada de los puntos de control que la diseñan. Esto significa que cada segmento de curva queda definido dentro del espacio convexo de los puntos de control. En el diseño de curvas 2D el espacio convexo es el polígono convexo formado a partir de la unión mediante rectas de los puntos de control. Para curvas 3D el espacio convexo es el poliedro generado por los puntos de control. El conocimiento de este espacio convexo, en el interior del cual queda definida la curva, facilita el estudio de intersecciones entre objetos y la visualización de los mismos por pantalla.

Las curvas de Bézier resultan ser las más estables de entre todas las bases polinomiales corrientemente usadas en sistemas CAD. Sin embargo, si la geometría del objeto es muy compleja, se requiere una polinomial de Bézier de un grado tan elevado que la hace intratable (grados de la curva superiores a diez ya no son aconsejables). La solución es diseñar la curva a trozos, mediante la composición de varias curvas de Bézier de grado menor, habitualmente cúbicas.

2A.3.3 Curvas *spline*

El término *spline* proviene de las cuerdas flexibles de metal usadas por los diseñadores para dibujar las superficies curvadas de coches, barcos, aviones, etc. Los splines metálicos, a no ser que se deformen severamente, poseen la derivada de segundo orden continua.

El equivalente matemático de estas cuerdas son las curvas *spline*. El *spline cúbico natural* es una polinomial cúbica con continuidad C^0 , C^1 , C^2 que interpola (pasa por) los puntos de control. Estos poseen un grado de continuidad mayor que las curvas del mismo orden de Hermite y Bézier. Además, las curvas *splines* generan formas más suaves. Los coeficientes de la polinomial para los *splines naturales*, sin embargo,

dependen de los " n " puntos de control y sus cálculos implican invertir una matriz $(n+1) \times (n+1)$, lo cual repercute en dos inconvenientes:

- mover un punto de control cualquiera afecta a la curva entera (no hay control local)
- los tiempos de cálculo necesarios para invertir la matriz pueden impedir un modelado interactivo de la curva que resulte sencillo y rápido.

Las curvas *B-spline* son una alternativa a los *splines naturales*. Están formadas por segmentos de curva cuyos coeficientes polinomiales dependen de unos pocos puntos de control. Esto garantiza el control local de la curva, porque mover un punto de control afecta solamente a una zona reducida de la misma. Además, poseen el atractivo de que los tiempos de cálculo de los coeficientes se reducen muchísimo. Asimismo, los *B-splines* mantienen el mismo grado de continuidad que los *splines naturales*, pero no interpolan los puntos de control que los definen, aunque la curva pasa por puntos cercanos a los mismos.

En los desarrollos matemáticos que a continuación se detallan se cambia la notación que se ha utilizado hasta aquí, puesto que se trabaja con una curva entera formada por varios segmentos de curva, en vez de trabajar con segmentos individuales. Un segmento de curva no tiene que pasar necesariamente por los puntos de control que lo definen y las condiciones de continuidad en un segmento se obtienen de los segmentos adyacentes, ya que los puntos de control son compartidos por segmentos contiguos. Por ello, es mejor describir el proceso de diseño con todos los segmentos al mismo tiempo.

a) *B-splines* uniformes

Los *B-splines cúbicos* uniformes aproximan un conjunto de $m+1$ puntos de control $P_0, P_1, P_2, \dots, P_m$, con $m \geq 3$, mediante una curva formada por $m-2$ segmentos polinomiales cúbicos Q_3, Q_4, \dots, Q_m . Aunque cada una de estas curvas cúbicas podría definirse en su propio dominio $0 \leq t < 1$, se puede ajustar el parámetro (sustituyendo t por $t+k$) para que el dominio del parámetro para varios segmentos de curva sea secuencial. Por tanto, se concluye que el rango del parámetro para Q_i está comprendido entre $t_i \leq t \leq t_{i+1}$, para $3 \leq i \leq m$. En particular para el caso $m=3$, que es un segmento de curva simple Q_3 , el parámetro está definido en el intervalo $t_3 \leq t \leq t_4$, que corresponde a los cuatro puntos de control P_0 a P_3 .

Para cada $i \geq 4$, hay un punto de unión o *nodo* entre Q_{i-1} y Q_i en el valor del parámetro t_i , denominado "*valor de nodo*". Los puntos inicial y final situados respectivamente en t_3 y t_{m+1} también corresponden a nodos de la curva, con lo que en total se utilizan $m+1$ nodos en la definición de una curva. Para generar una curva *B-*

spline cerrada los puntos de control P_0, P_1, P_2 se repiten al final de la secuencia $P_0, P_1, P_2, \dots, P_m, P_0, P_1, P_2$.

El término *uniforme* significa que los nodos están espaciados en intervalos iguales del parámetro t . Sin pérdida de generalidad, se puede asumir que $t_3 = 0$ y el intervalo $t_{i+1} - t_i = 1$. Los *B-splines no uniformes* permiten espaciados desiguales entre nodos y se discuten en la sección siguiente. Estas curvas son *no racionales*, a diferencia de las *curvas polinomiales cúbicas racionales* que se presentarán en la sección 2A.3.3.c, donde $x(t)$, $y(t)$ y $z(t)$ se definen como cocientes de polinomios cúbicos. La "B" se refiere a base, ya que los *B-splines* pueden representarse como sumas ponderadas de funciones base polinomiales, en contraposición de los *splines naturales*, los cuales no permiten este tipo de descripción.

Cada uno de los $m-2$ segmentos de curva de una curva *B-spline* está definida por cuatro de los $m+1$ puntos de control. En particular, el segmento de curva Q_i se define por los puntos $P_{i-3}, P_{i-2}, P_{i-1}, P_i$. Además, el *vector geométrico del B-spline* G_{bsi} para el segmento Q_i es

$$G_{Bsi} = \begin{bmatrix} P_{i-3} \\ P_{i-2} \\ P_{i-1} \\ P_i \end{bmatrix}, \quad 3 \leq i \leq m \quad (2A.27)$$

El primer segmento de curva, Q_3 , se define por los puntos del P_0 al P_3 a través del rango de valores del parámetro $t_3 = 0$ a $t_4 = 1$, Q_4 se define por los puntos del P_1 al P_4 a través del rango de valores del parámetro $t_4 = 1$ a $t_5 = 2$, y el último segmento, Q_m , se define por los puntos $P_{m-3}, P_{m-2}, P_{m-1}$ y P_m a través del rango de valores del parámetro $t_m = m-3$ a $t_{m+1} = m-2$. Generalizando, un segmento Q_i empieza en algún lugar cerca del punto P_{i-2} y acaba cerca del punto P_{i-1} .

Así, cada segmento de curva se define por cuatro puntos de control y cada punto de control (excepto los del principio y final de la secuencia $P_0, P_1, P_2, \dots, P_m$) influye sobre cuatro segmentos de curva. Al mover un punto de control en una dirección dada se mueven los cuatro segmentos de curva en los que este punto interviene en la definición, mientras que el resto de segmentos se ven inalterados. En la figura 2A-9 se muestra un ejemplo en el que se varía solamente la posición del punto de control P_4 : puede observarse como hay zonas de la curva que permanecen inalteradas. Por tanto, se puede concluir que las curvas *B-spline* tienen control local, propiedad que las hace atractivas en el diseño interactivo de curvas arbitrarias.

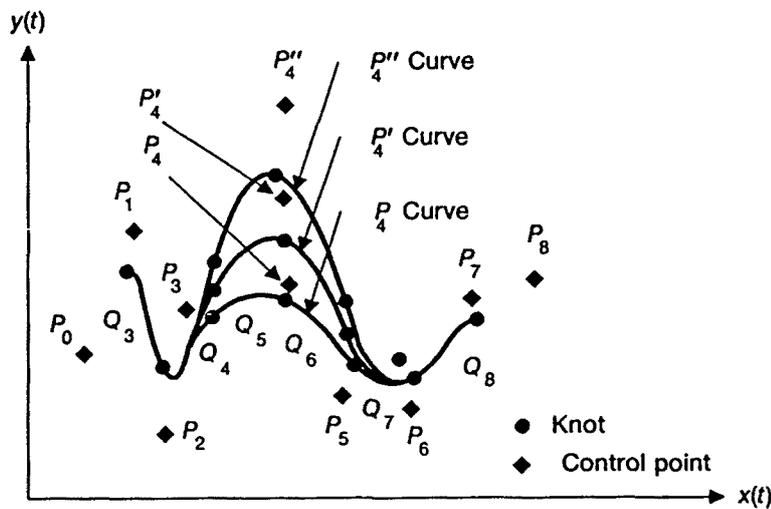


Figura 2A-9: Efecto de la variación del punto P_4 en una curva B-spline.

Si se define T_i como el siguiente vector fila $[(t - t_i)^3 \ (t - t_i)^2 \ (t - t_i) \ 1]$, entonces la formulación utilizada para el segmento de curva i es

$$Q_i(t) = T_i \cdot M_{Bs} \cdot G_{Bsi} \quad \text{con } t_i \leq t < t_{i+1} \quad (2A.28)$$

La curva entera se genera al aplicar la ecuación anterior para $3 \leq i \leq m$. La *matriz de base B-spline*, M_{Bs} , relaciona las restricciones geométricas G_{Bs} con las *funciones blending* y los coeficientes polinomiales:

$$M_{Bs} = \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \quad (2A.29)$$

El desarrollo matemático necesario para llegar a esta matriz se encuentra en [Bart'87].

Las *funciones blending* para los B-spline B_{Bs} vienen dadas por el producto $T_i \cdot M_{Bs}$, de forma análoga a lo que se ha hecho en las secciones previas para las formulaciones de Bézier y de Hermite. Nótese que las *funciones blending* para cada segmento de curva son exactamente iguales, porque en cada segmento i el rango de valores de $t - t_i$ va desde 0 en $t = t_i$ a 1 en $t = t_{i+1}$. Si se reemplaza $t - t_i$ por t , y se sustituye el intervalo $[t_i, t_{i+1}]$ por $[0, 1]$, se obtienen las *funciones blending*:

$$\begin{aligned}
 B_{bs} &= T \cdot M_{Bs} = [B_{Bs-3} \ B_{Bs-2} \ B_{Bs-1} \ B_{Bs0}] \\
 &= \frac{1}{6} [-t^3+3t^2-3t+1 \quad 3t^3-6t^2+4 \quad -3t^3+3t^2+3t+1 \quad t^3] \quad (2A.30) \\
 &= \frac{1}{6} [(1-t)^3 \quad 3t^3-6t^2+4 \quad -3t^3+3t^2+3t+1 \quad t^3], \quad 0 \leq t < 1
 \end{aligned}$$

La figura 2A-10 muestra las funciones *blending* de los *B-spline*, B_{Bs} . Como la suma de las cuatro funciones es 1 y todas ellas son no negativas, la propiedad de mantenerse en el espacio convexo definido por el polígono creado al unir los puntos de control se cumple para cada segmento de curva.

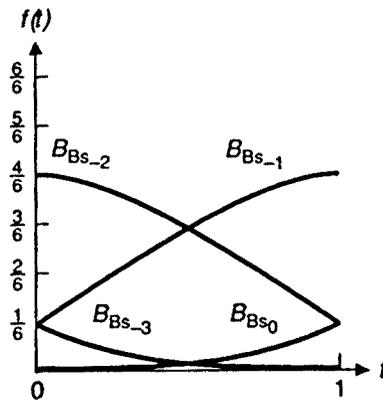


Figura 2A-10: Las cuatro funciones Blending que definen las curvas *B-splines*. En $t=0$ y $t=1$, tres de las funciones son no nulas.

Si se expande la ecuación (4.28), reemplazando $t - t_i$ por t en el segundo igual, tenemos

$$\begin{aligned}
 Q_i(t - t_i) &= T_i \cdot M_{Bs} \cdot G_{Bsi} = T \cdot M_{Bs} \cdot G_{Bsi} \\
 &= B_{Bs} \cdot G_{Bsi} = B_{Bs-3} \cdot P_{i-3} + B_{Bs-2} \cdot P_{i-2} + B_{Bs-1} \cdot P_{i-1} + B_{Bs0} \cdot P_i \quad (2A.31) \\
 &= \frac{(1-t)^3}{6} P_{i-3} + \frac{3t^3-6t^2+4}{6} P_{i-2} + \frac{-3t^3+3t^2+3t+1}{6} P_{i-1} + \frac{t^3}{6} P_i \quad \text{para } 0 \leq t < 1
 \end{aligned}$$

Se puede demostrar fácilmente que Q_i y Q_{i+1} tienen continuidad C^0 , C^1 y C^2 en los puntos de unión para el caso analizado, en el que se utilizan segmentos de curva cúbicos. Si se estudia la componente x de los dos segmentos adyacentes, los cuales son $x_i(t - t_i)$ y $x_{i+1}(t - t_{i+1})$ (el planteamiento es análogo para las componentes y y z) sólo hay que demostrar que en el nodo t_{i+1} donde se produce la unión se cumple,

$$\begin{aligned}
 x_i(t_{i+1}) &= x_{i+1}(t_{i+1}) \\
 \frac{d}{dt}x_i(t_{i+1}) &= \frac{d}{dt}x_{i+1}(t_{i+1}) \\
 \frac{d^2}{dt^2}x_i(t_{i+1}) &= \frac{d^2}{dt^2}x_{i+1}(t_{i+1})
 \end{aligned}
 \tag{2A.32}$$

Sustituyendo t por $t - t_i$ se encuentra la siguiente ecuación equivalente

$$\begin{aligned}
 x_i|_{t-t_i=1} &= x_{i+1}|_{t-t_{i+1}=0} \\
 \frac{d}{dt}x_i|_{t-t_i=1} &= \frac{d}{dt}x_{i+1}|_{t-t_{i+1}=0} \\
 \frac{d^2}{dt^2}x_i|_{t-t_i=1} &= \frac{d^2}{dt^2}x_{i+1}|_{t-t_{i+1}=0}
 \end{aligned}
 \tag{2A.33}$$

y si se opera con la ecuación 2A.31 y sus dos primeras derivadas se llega a

$$\begin{aligned}
 x_i|_{t-t_i=1} &= x_{i+1}|_{t-t_{i+1}=0} = \frac{P_{i-2_x} + 4P_{i-1_x} + P_{i_x}}{6} \\
 \frac{d}{dt}x_i|_{t-t_i=1} &= \frac{d}{dt}x_{i+1}|_{t-t_{i+1}=0} = \frac{-P_{i-2_x} + P_{i_x}}{2} \\
 \frac{d^2}{dt^2}x_i|_{t-t_i=1} &= \frac{d^2}{dt^2}x_{i+1}|_{t-t_{i+1}=0} = P_{i-2_x} - 2P_{i-1_x} + P_{i_x}
 \end{aligned}
 \tag{2A.34}$$

La continuidad de segundo orden que ofrecen los *B-splines* es atractiva, pero ello repercute en un control menor de los puntos por los que pasa la curva. Si se quiere también se puede forzar a la curva a pasar por los puntos de control, lo cual se consigue repitiendo dichos puntos de control. Esto tanto es válido tanto para los puntos extremos como para los puntos intermedios de la curva. Por ejemplo, si $P_{i-2} = P_{i-1}$, la curva pasa por una posición más cercana a este punto porque el segmento Q_i se define por tres puntos diferentes, y el punto $P_{i-2} = P_{i-1}$ se pondera dos veces en la ecuación (2A.31) - una vez con B_{Bs-2} y otra por B_{Bs-1} .

Al utilizar tres veces un mismo punto de control -por ejemplo $P_{i-2} = P_{i-1} = P_i$ - la ecuación (2A.31) del caso cúbico se transforma en

$$Q_i(t) = B_{Bs-3} \cdot P_{i-3} + (B_{Bs-2} + B_{Bs-1} + B_{Bs0}) \cdot P_i
 \tag{2A.35}$$

de la que se deduce que el segmento Q_i se transforma en una línea recta. Además, el punto P_{i-2} queda interpolado en la línea para $t = 1$, donde los tres pesos aplicados sobre P_i suman 1. En cambio P_{i-3} , en general, no queda interpolado en $t = 0$. Otra

forma de ver lo mismo es pensar que el espacio convexo Q_i se genera solamente con dos puntos distintos, lo que lleva a que Q_i sea una recta.

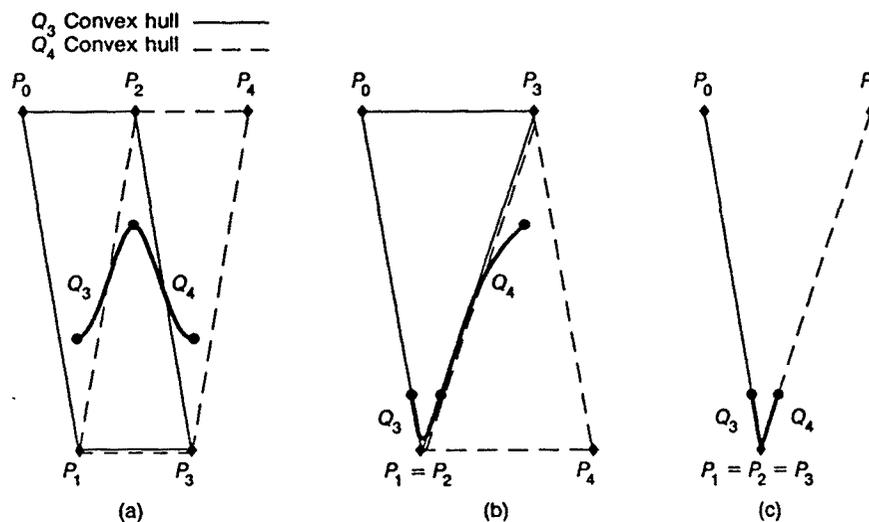


Figura 2A-11: Efecto de los puntos de control múltiples.

En la figura 2A-11 se muestra el efecto que produce el tener puntos de control múltiples en el interior de un *B-spline*: (a) no hay puntos de control múltiples y los espacios convexos de los dos segmentos que forman la curva se solapan; (b) hay un punto de control doble y, por tanto, los dos espacios convexos comparten la arista P_2P_3 ; (c) Hay un punto de control triple y los espacios convexos son dos rectas que comparten este punto de control triple: al ser los espacios convexos líneas rectas, también lo son los segmentos de curva.

El precio de interpolar los puntos de control es la pérdida de continuidad de orden G^l , aún cuando a partir de la ecuación (2A.34) se deduzca que la continuidad C^l se mantiene (pero se mantiene con valor cero de la derivada). Este es el único caso en que se tiene continuidad C^l y no G^l . En la curva de la figura 2A-11 (c) se tiene continuidad paramétrica de orden C^2 , mientras que sólo se tiene continuidad visual de orden G^0 . Los *B-splines no-uniformes*, que se comentan a continuación, interpolan los puntos de control de forma más natural y sencilla, permitiendo un mejor control de la continuidad.

b) *B-splines* no uniformes

Los *B-splines* no-uniformes difieren de los *B-splines* uniformes presentados en la sección anterior en que los intervalos del parámetro entre valores de nodos sucesivos no están uniformemente distribuidos. Una secuencia de valores de nodos no-uniforme significa que las funciones *blending* ya no son idénticas para todos los intervalos, y varían de un segmento de curva a otro.

Estas curvas tienen algunas ventajas respecto a los *B-splines* uniformes:

- La continuidad en algunos puntos seleccionados puede reducirse, para polinomios cúbicos, desde C^2 a C^1 , C^0 o ninguna.
- Si la continuidad se reduce a C^0 , entonces la curva interpola un punto de control, pero sin el efecto indeseable de los *B-splines* uniformes, en donde los segmentos de curva que se unen en el punto de control interpolado están obligados a ser rectas.
- Los puntos de control inicio y final se pueden interpolar exactamente de forma muy sencilla en la curva, sin provocar la aparición de segmentos de curva lineales.
- Es posible añadir nodos y puntos de control adicionales a un *B-spline* no uniforme, con lo que la curva resultante puede deformarse localmente de forma muy simple.

El aumento de generalidad de los *B-splines* no-uniformes requiere una notación algo distinta de la utilizada para los *B-splines* uniformes. Como en el caso anterior, el *spline* es una curva continua a trozos generada a partir de curvas polinomiales cúbicas, que aproximan los puntos de control P_0 a P_m . La secuencia de valores de los nodos es no decreciente: t_0 a t_{m+4} , es decir, hay 4 nodos más que puntos de control. Como el mínimo número de puntos de control es 4, la secuencia menor de puntos de control tiene 8 valores de los nodos y la curva se define en el intervalo del parámetro desde t_3 a t_4 .

La única restricción a considerar en la secuencia del vector de nodos es la de no ser decreciente, lo cual permite que varios valores consecutivos del vector de nodos sean iguales. En este caso se dice que se tiene un *nodo múltiple* y el número de veces en se repite el valor del nodo se denomina *multiplicidad del nodo*.

El segmento de curva Q_i se define por los puntos de control P_{i-3} , P_{i-2} , P_{i-1} , P_i y por las funciones *blending* $B_{i-3,4}(t)$, $B_{i-2,4}(t)$, $B_{i-1,4}(t)$, $B_{i,4}(t)$, como la suma ponderada siguiente:

$$Q_i(t) = P_{i-3} \cdot B_{i-3,4}(t) + P_{i-2} \cdot B_{i-2,4}(t) + P_{i-1} \cdot B_{i-1,4}(t) + P_i \cdot B_{i,4}(t) \quad (2A.36)$$

$$3 \leq i \leq m, \quad t_i \leq t < t_{i+1}$$

La curva no está definida fuera del intervalo t_3 a t_{m+1} . Cuando $t_i = t_{i+1}$ (un nodo múltiple), el segmento de curva Q_i es un único punto. Esta idea de reducir un segmento de curva a un punto es la que proporciona una mayor flexibilidad a los *B-splines* no-uniformes respecto a los uniformes.

No existe un único conjunto de funciones *blending*, como en otros tipos de curvas *splines*. Las funciones *blending* dependen de los intervalos entre valores del vector de nodos y se definen recursivamente en términos de las funciones *blending* de órdenes menores. Sea $B_{i,j}(t)$ la función *blending* de orden j para ponderar el punto de control P_i . Como se está trabajando con *B-splines* de cuarto orden (grado 3 ó cúbico), la definición recursiva acaba con $B_{i,4}(t)$. La recurrencia para *B-splines* cúbicos es

$$\begin{aligned}
 B_{i,1}(t) &= \begin{cases} 1, & t_i \leq t < t_{i+1} \\ 0, & \text{otro caso} \end{cases} \\
 B_{i,2}(t) &= \frac{t - t_i}{t_{i+1} - t_i} B_{i,1}(t) + \frac{t_{i+2} - t}{t_{i+2} - t_{i+1}} B_{i+1,1}(t) \\
 B_{i,3}(t) &= \frac{t - t_i}{t_{i+2} - t_i} B_{i,2}(t) + \frac{t_{i+3} - t}{t_{i+3} - t_{i+1}} B_{i+1,2}(t) \\
 B_{i,4}(t) &= \frac{t - t_i}{t_{i+3} - t_i} B_{i,3}(t) + \frac{t_{i+4} - t}{t_{i+4} - t_{i+1}} B_{i+1,3}(t)
 \end{aligned} \tag{2A.37}$$

La figura 2A-12 muestra como la ecuación de recurrencia (2A.37) puede usarse para encontrar las funciones *blending* de un *B-spline* no-uniforme cúbico, usando el vector de nodos (0,0,0,0,1,1,1,1) como ejemplo. Dicha figura también pone de manifiesto que es necesario utilizar 8 valores del vector de nodos para generar cuatro funciones *blending*. $B_{3,1}(t)$ es la unidad en el intervalo $0 \leq t < 1$. Los restantes $B_{i,1}(t)$ son cero. $B_{2,2}(t)$ y $B_{3,2}(t)$ son rampas lineales, y son las funciones *blending* para interpolación lineal entre dos puntos. Análogamente, $B_{1,3}(t)$, $B_{2,3}(t)$ y $B_{3,3}(t)$ son cuadráticas y forman el conjunto de funciones *blending* para interpolación cuadrática.

Para este vector de nodos en particular, los $B_{i,4}(t)$ son polinomios de Bernstein, es decir, las funciones *blending* de Bézier para el caso cúbico (comparar con las funciones presentadas en la figura 2A-7). Para este vector de nodos, la curva interpola los puntos de control P_0 y P_3 , y es de hecho una curva de Bézier con el vector tangente a los puntos extremos determinado por los vectores P_0P_1 y P_2P_3 . De todo ello se deduce que las curvas de Bézier son un caso particular de los *B-splines* no uniformes en los que se interpolan los puntos de control extremos del segmento.

Calcular las funciones *blending* requiere un coste computacional demasiado elevado para una visualización interactiva de estas curvas. Por ello, se restringen las secuencias de los nodos de los *B-splines* a tener intervalos que sean 0 ó 1, de forma que es posible almacenar un pequeño número de matrices correspondientes a la ecuación (2A.37), la cual cubre todas las configuraciones de nodos posibles. Así se elimina la necesidad de recalcular la ecuación (2A.37) para cada segmento de curva.

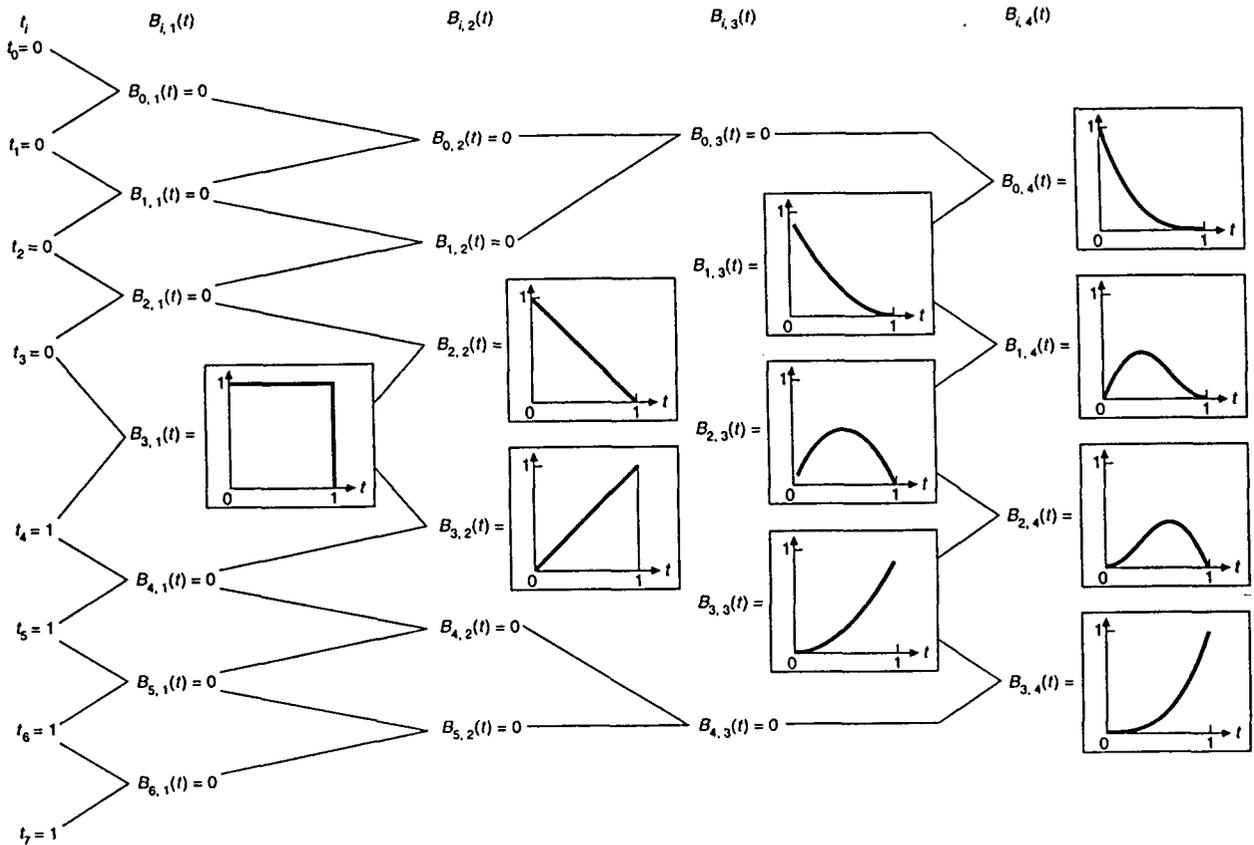


Figura 2A-12: Relaciones definidas en la ecuación (2A.37) entre el vector de nodos $(0,0,0,0,1,1,1,1)$ y las funciones *blending*.

Se puede demostrar que las funciones *blending* son no negativas y que, además, suman 1, con lo que los segmentos de curva de los *B-spline* no uniformes quedan dentro de los espacios convexos generados por los cuatro puntos de control que los definen. Para nodos de multiplicidad mayor a 1, los denominadores de la ecuación (2A.37) pueden ser cero porque valores de nodos sucesivos pueden ser iguales: en este caso se impone que las divisiones por cero son siempre cero.

El aumento de multiplicidad de los nodos tiene dos efectos:

- Cada valor del nodo t_i queda automáticamente dentro del espacio convexo de los puntos P_{i-3} , P_{i-2} y P_{i-1} . Si t_i y t_{i+1} son iguales, ambos deberán estar en el espacio convexo de P_{i-3} , P_{i-2} y P_{i-1} y en el espacio convexo P_{i-2} , P_{i-1} y P_i . Esto significa que se sitúan en el segmento de línea entre P_{i-2} y P_{i-1} . Análogamente,

si $t_i = t_{i+1} = t_{i+2}$ entonces este nodo debe situarse en P_{i-1} . Si $t_i = t_{i+1} = t_{i+2} = t_{i+3}$, entonces el nodo debe situarse al mismo tiempo en P_{i-1} y en P_i y la curva presenta un salto.

- Los nodos múltiples reducen la continuidad paramétrica: desde C^2 a C^1 cuando el nodo posee multiplicidad doble; a C^0 en nodos de multiplicidad triple; y a discontinuidad en la curva cuando la multiplicidad del nodo es cuádruple.

El efecto de la multiplicidad de los nodos en la continuidad de la curva puede observarse en la figura 2A-13:

2A-13 (a): Todos los nodos tienen multiplicidad simple. Cada segmento de curva se define por cuatro puntos de control y cuatro funciones *blending*, y los segmentos de curva adyacentes comparten tres puntos de control. Por ejemplo el segmento de curva Q_3 se define por los puntos de control P_0, P_1, P_2 y P_3 ; el segmento de curva Q_4 se define por los puntos de control P_1, P_2, P_3 y P_4 ; el segmento de curva Q_5 se define por los puntos de control P_2, P_3, P_4 y P_5 ; y así sucesivamente.

2A-13 (b): Hay un nodo con multiplicidad doble, $t_4 = t_5$, por lo que la curva Q_4 tiene longitud nula. Los segmentos Q_3 y Q_4 pasan a ser adyacentes, pero sólo comparten 2 puntos de control, P_2 y P_3 ; estos dos segmentos de curva tienen menos en común que el resto de segmentos de curva adyacentes lo que provoca la pérdida de un grado de continuidad en el punto de unión.

2A-13 (c) Hay un nodo triple. Esto provoca que sólo haya un punto de unión entre los dos segmentos adyacentes en P_3 ; el cual queda interpolado por los dos segmentos de curva. Al compartir un único punto de control, sólo habrá una restricción a cumplir y, por tanto, sólo se tendrá continuidad paramétrica C^0 , a ser satisfecha en este punto.

2A-13 (d) Hay un nodo de multiplicidad 4, que provoca una discontinuidad en la curva.

En consecuencia, varios *splines* distintos pueden representarse por una única secuencia de nodos y de puntos de control, si la multiplicidad de los nodos es distinta. La tabla que se adjunta a continuación resume, para el caso cúbico, los efectos de tener puntos de control múltiples y la de tener nodos múltiples:

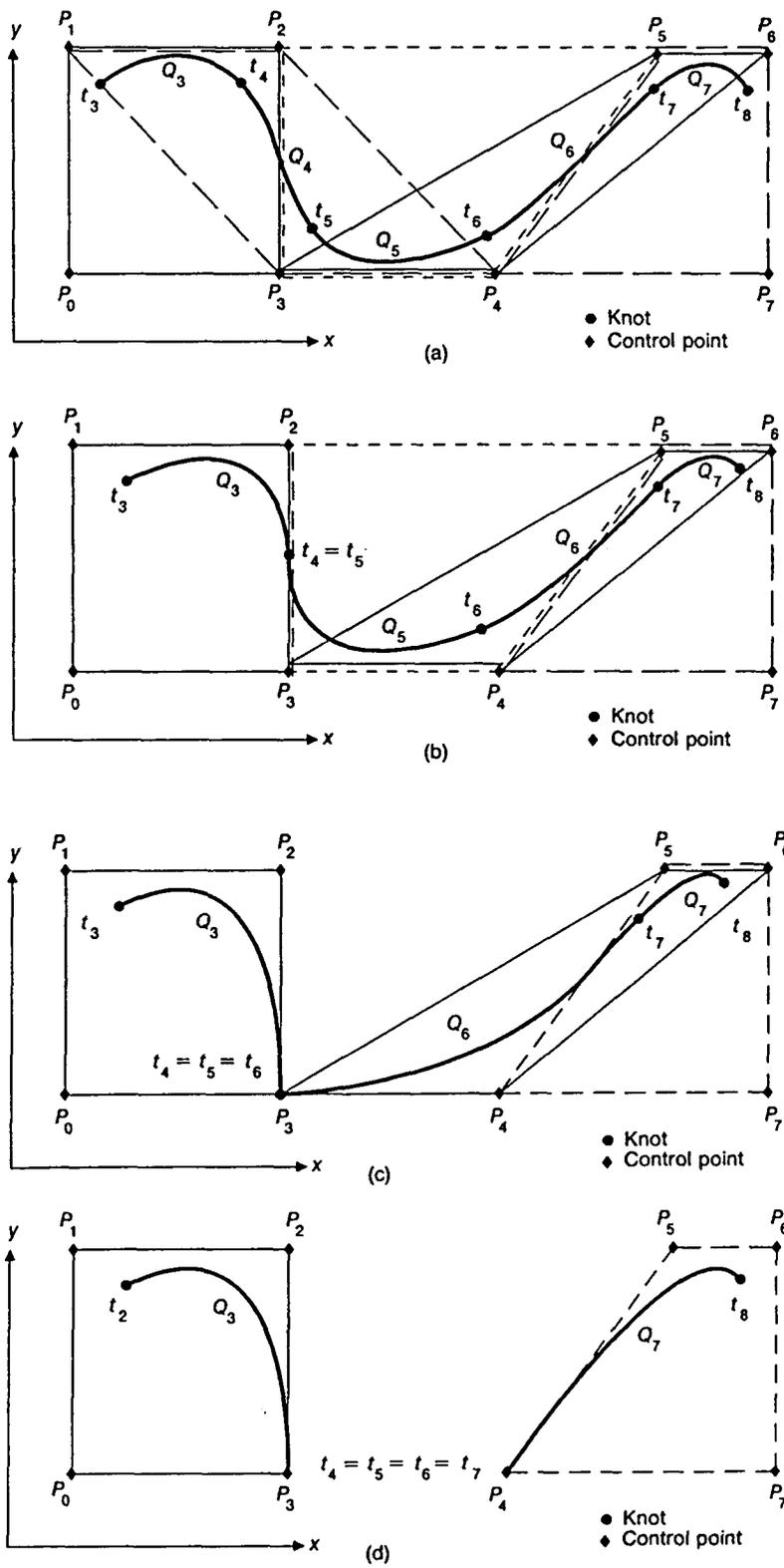


Figura 2A-13: Efecto de la multiplicidad de nodos en un B-spline no-uniforme .

Multiplicidad	Puntos de control múltiples	Nodos múltiples
1	$C^2 G^2$	$C^2 G^2$
2	$C^2 G^1$	$C^1 G^1$
3	$C^2 G^0$	$C^0 G^0$
4	$C^2 G^0$	discontinuidad

Tabla 2A-1: Comparación de los efectos en la continuidad de la curva de puntos de control múltiples y nodos múltiples, para los B-spline cúbicos no uniformes.

c) B-splines racionales

Los segmentos de curva racionales son cocientes de polinomios:

$$x(t) = \frac{X(t)}{H(t)}, \quad y(t) = \frac{Y(t)}{H(t)}, \quad z(t) = \frac{Z(t)}{H(t)} \quad (2A-38)$$

donde $X(t)$, $Y(t)$, $Z(t)$ y $H(t)$ son curvas polinomiales cuyos puntos de control están definidos en coordenadas homogéneas (4 componentes para cada punto). La ecuación (2A.38) equivale a la proyección de la curva

$$Q(t) = [X(t) \ Y(t) \ Z(t) \ H(t)] \quad (2A-39)$$

definida en un espacio homogéneo de 4 dimensiones al espacio de 3 dimensiones, lo que implica dividir por $W(t)$ todas las componentes. De la misma forma, cualquier curva no racional puede transformarse en una curva racional añadiendo $W(t)=1$ como cuarta coordenada.

En general, las polinomiales en una curva racional pueden ser de Bézier, de Hermite o de cualquier otro tipo. Cuando son B-spline, se denominan curvas B-spline racionales no uniformes (NURBS). Los NURBS son las curvas más comúnmente utilizadas en la actualidad para el diseño de curvas y superficies arbitrarias.

En conclusión, las curvas spline racionales se definen en un sistema de coordenadas homogéneo: si $P = (x,y,z)$ es un punto en el espacio euclídeo de tres dimensiones, se define el punto correspondiente a aquél en el espacio homogéneo de 4 dimensiones como $P^h = (hx,hy,hz,h)$, donde $h > 0$. A h se le conoce como coordenada homogénea o peso. Al proyectar el espacio homogéneo de 4 dimensiones en el euclídeo de 3, $P^h = (hx, hy, hz, h)$ se transforma en $P = (x, y, z)$.

La definición de una curva *B-spline* polinomial en el espacio homogéneo se hace de forma paralela al caso de los *B-splines* no racionales en el espacio euclídeo

$$Q^h(t) = \sum_{i=1}^n B_{i,k}(t) P_i^h \quad (2A-40)$$

donde $B_{i,k}(t)$ son las mismas funciones base polinomiales de orden k que en el caso de los *B-splines* no racionales, y P_i^h son los puntos de control en el espacio homogéneo. También queda implícito en la definición un vector de nodos $\{t_j\}_{j=1}^{n+k}$ como en el caso de los *B-splines* no racionales.

La coordenada h (o peso) de cada punto de control define cuánto afecta éste a la forma de la curva. En los *B-splines* no racionales todos los puntos de control influyen por igual sobre la forma de la curva, salvo los puntos de control múltiples que "atraen" hacia sí la curva. En los *B-splines* racionales el parámetro " h " se conoce como parámetro de forma, ya que al aumentar su valor en un punto se obtiene una nueva curva que pasa más cerca del punto afectado a por dicho peso.

La figura 2A-14 demuestra que el efecto producido sobre la curva al modificar el peso " h " de un punto de control es distinto del que se obtiene cuando en una curva *B-spline* no racional se modifica la posición del mismo.

Para $a \leq t \leq b$, la curva $Q^h(t)$ es un conjunto de puntos en el espacio homogéneo 4-D. La proyección de $Q^h(t)$ en el espacio euclídeo 3-D es la curva *B-spline* racional $Q(t)$, que se obtiene por división de cada una de las tres primeras coordenadas de cada punto de $Q^h(t)$ por la coordenada homogénea:

$$Q(t) = \frac{\sum_{i=1}^n B_{i,k}(t) h_i P_i}{\sum_{i=1}^n B_{i,k}(t) h_i} \quad (2A-41)$$

La ecuación (2A.41) es una curva polinomial a trozos con nodos en t_j . La curva polinomial tiene la máxima flexibilidad cuando el vector de nodos es no uniforme. En este caso la forma matemática recibe el nombre de *B-spline* racional no uniforme (NURB).

Las ventajas adicionales que presentan las curvas racionales sobre las no racionales son:

- Las curvas racionales son invariantes bajo las transformaciones de rotación, escalado, traslación y cambio de perspectiva de los puntos de control, mientras que las curvas no racionales sólo lo son bajo las transformaciones de rotación,

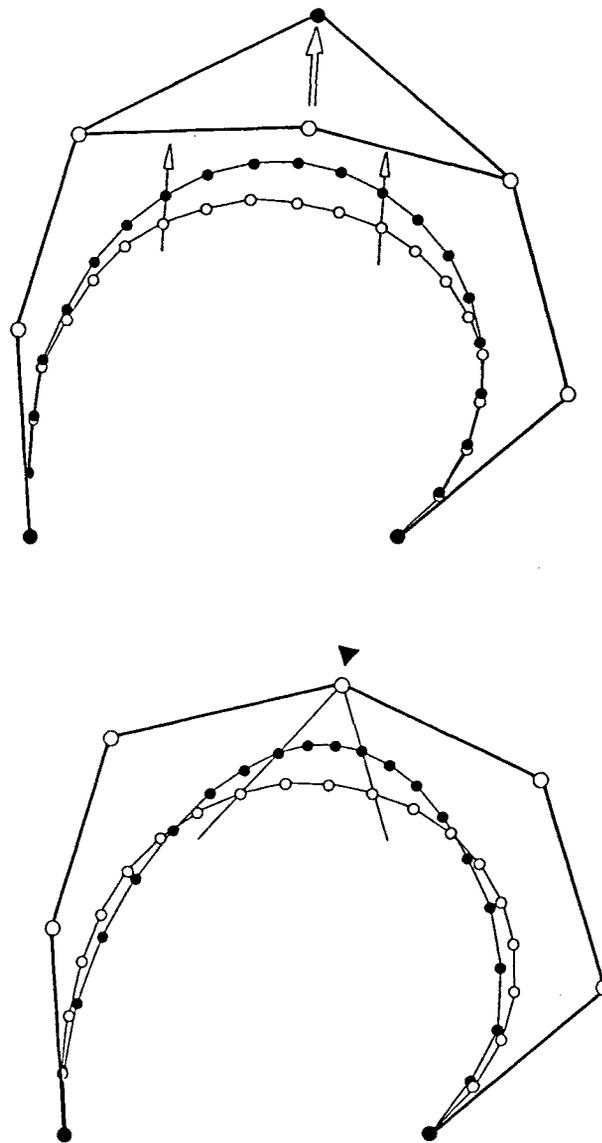


Figura 2A-14: Efecto de los pesos. Arriba, modificación de la posición de un punto de control en una curva no racional. Abajo, modificación del peso de un punto de control en una curva racional.

escalado y traslación. Ello conduce a que los cambios de perspectiva sólo se apliquen sobre los puntos de control que definen la curva, generándose la nueva curva con mayor rapidez que si se tuviera que aplicar estas transformaciones a todos los puntos de la curva. Para poder aprovechar esta característica importantísima una curva no racional puede convertirse en una racional antes de aplicar la transformación de perspectiva, aplicar este cambio a los puntos de control que definen la curva expresada como una polinomial racional, para finalmente volverla a escribir en el espacio de tres dimensiones. El resultado de este proceso es mucho más rápido que el de aplicar la transformación directamente sobre los puntos que definen la curva *B-spline* no racional.

- Estos *B-splines*, a diferencia de los no racionales, pueden definir con precisión una sección cónica cualquiera. Los splines no racionales sólo pueden obtener aproximaciones más o menos buenas de las mismas según la elección de los puntos de control. Esta segunda propiedad es muy útil en aplicaciones CAD, en las que es necesario representar tanto curvas y superficies de geometrías totalmente arbitrarias como curvas cónicas. Gracias a esta propiedad de los *B-splines racionales*, se evita el tener que implementar un método mixto que combine las formas matemáticas implícitas con las polinomiales.

La definición de cónicas sólo requiere polinomios cuadráticos. Las funciones *blending* de la ecuación de recurrencia (2A.37) se usan en la definición de la curva de la siguiente forma:

$$Q_i(t) = P_{i-2}B_{i-2,3}(t) + P_{i-1}B_{i-1,3}(t) + P_iB_{i,3}(t) \tag{2A-42}$$

$$2 \leq i \leq m, \quad t_i \leq t < t_{i+1}$$

En la figura 2A-15 se muestran dos posibles maneras de representar un círculo unidad centrado en el origen de coordenadas. Las coordenadas homogéneas del espacio bidimensional son (x,y,h) . El vector de nodos es $(0, 0, 0, 1, 1, 2, 2, 3, 3, 4, 4)$, con el primer y último punto de control iguales. La elección de este primer punto de control es arbitraria. Nótese que con los B-splines cuadráticos un nodo doble provoca que un punto de control quede interpolado en la curva y los nodos triples se utilizan para marcar el inicio y final de la curva en los puntos de control.

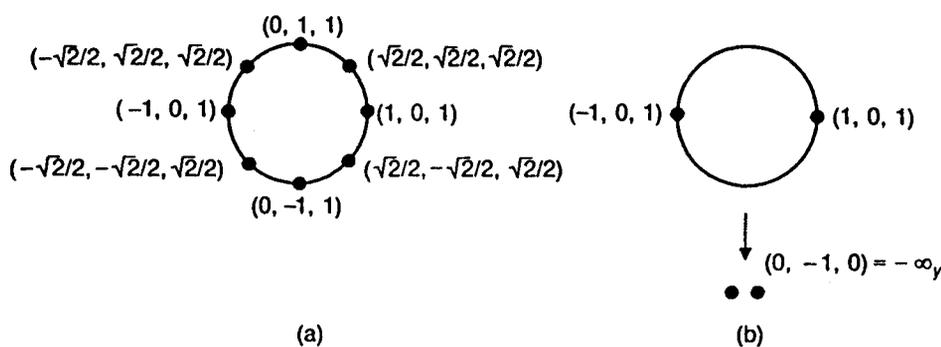


Figura 2A-15: Dos conjuntos de puntos de control para definir una circunferencia mediante un B-spline racional.

Los *B-splines* racionales no uniformes están adquiriendo una importancia creciente y ya son muchas las estaciones gráficas con hardware dedicado que son capaces de visualizarlos en tiempo real (HEWLETT_PACKARD, SILICON GRAPHICS, etc.), así como los paquetes software de modelado geométrico que incorporan los NURBS: I-

DEAS, CATIA, ALIAS, etc. (los dos primeros muy extendidos en la industria aeronáutica). El paquete I-DEAS es el adquirido por el departamento para el modelado de objetos para el análisis electromagnético.

d) Otras formas *spline*

Splines interpoladores

En el modelado de curvas muy a menudo el problema consiste en el diseño de una curva suave que interpole un conjunto de puntos determinado. La familia de splines interpoladores de Catmull-Rom, también conocidos como *splines de Overhauser*, son muy útiles en estos casos. Un miembro de esta familia de *splines* es capaz de interpolar los puntos P_1 al P_{m-1} de una secuencia de puntos P_1 al P_m . Además, el vector tangente al punto P_i es paralelo a la línea que conecta los puntos P_{i-1} y P_{i+1} , como se muestra en la figura 2A-16. Sin embargo, estos *splines* no poseen la propiedad del espacio convexo. Los *splines naturales* también interpolan puntos, pero no poseen el control local que consiguen los *splines de Catmull-Rom*.

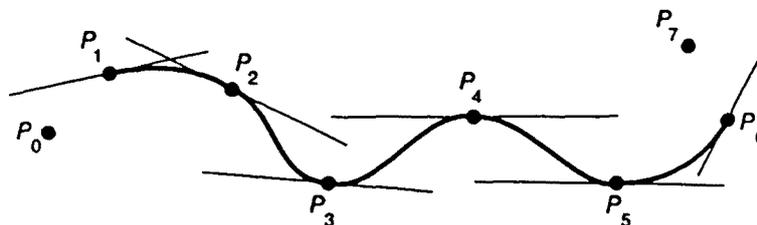


Figura 2A-16: *Spline de Catmull-Rom*

La representación de los *splines de Catmull-Rom* con una matriz base y utilizando el mismo vector de geometría en el caso cúbico es la siguiente (una forma rápida de obtener las curvas Catmull-Rom se encuentra en [Barr'88b]):

$$Q^i(t) = T \cdot M_{CR} \cdot G_{Bsi} = \frac{1}{2} \cdot T \cdot \begin{bmatrix} -1 & 3 & -3 & 1 \\ 2 & -5 & 4 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_{i-3} \\ P_{i-2} \\ P_{i-1} \\ P_i \end{bmatrix} \quad (2A-43)$$

β -splines

Otro tipo de *spline* es el *uniformly shaped β -spline* que tiene dos parámetros adicionales, β_1 y β_2 , que proporcionan un mayor control sobre la forma de la curva. Utiliza el

mismo vector de geometría G_{Bsi} que los B -splines, y mantiene la propiedad de ocupar un espacio convexo. La *matriz base de los β -splines* M_β es

$$M_\beta = \frac{1}{\delta} \begin{bmatrix} -2\beta_1^3 & 2(\beta_2 + \beta_1^3 + \beta_1^2 + \beta_1) & -2(\beta_2 + \beta_1^2 + \beta_1 + 1) & 2 \\ 6\beta_1^3 & -3(\beta_2 + 2\beta_1^3 + 2\beta_1^2) & 3(\beta_2 + 2\beta_1^2) & 0 \\ -6\beta_1^3 & 6(\beta_1^3 - \beta_1) & 6\beta_1 & 0 \\ 2\beta_1^3 & \beta_2 + 4(\beta_1^2 + \beta_1) & 2 & 0 \end{bmatrix} \quad (2A-44)$$

con

$$\delta = \beta_2 + 2\beta_1^3 + 4\beta_1^2 + 4\beta_1 + 2 \quad (2A-45)$$

El primer parámetro, β_1 , se conoce como *bias*; el segundo, β_2 , se conoce como *tensión*. Si $\beta_1 = 1$ y $\beta_2 = 0$, M_β se reduce a M_{Bs} para B -splines. Al ir incrementando β_1 desde 1, el *spline* queda cada vez más influido por el vector tangente en el lado del parámetro que se incrementa. Para valores menores que 1, el *bias* es hacia la otra dirección. Al ir incrementando el parámetro de tensión β_2 el *spline* se va acercando a las líneas que conectan los puntos de control.

A estos β -splines se denominan *uniformly shaped* porque β_1 y β_2 se aplican uniformemente sobre todos los segmentos de la curva. Este efecto global de β_1 y β_2 viola la propiedad del control local de la curva. *Continuously shaped β -splines* y *discretely shaped β -splines* asocian distintos valores de β_1 y β_2 , en cada punto de control, proporcionando un efecto local y no global.

Aunque los β -splines diseñan una variedad mayor de formas que los B -splines, dichas curvas presentan continuidad de orden G^2 pero solamente continuidad C^0 en los puntos de unión. Esto no es una desventaja en el modelado geométrico, pero introduce discontinuidades de velocidad en los *animation motion path*. En el caso especial en que $\beta_1 = 1$, se obtienen curvas con continuidad C^1 .

2A.4 SUPERFICIES PARAMETRICAS B-SPLINE

a) Superficies paramétricas

Una superficie B -spline se define de forma análoga a las curvas B -spline. Para ello se utilizan dos familias de curvas B -spline que recorren la superficie en direcciones perpendiculares. El diseño de superficies requiere dos vectores de nodos, uno para cada familia de curvas, y una matriz rectangular de puntos de control. Los dos vectores de

nodos son necesarios para recorrer las dos dimensiones perpendiculares de la matriz rectangular de puntos de control.

La forma matemática de las superficies es un producto tensorial de dos parámetros s y t . La forma general de una curva paramétrica, como se ha señalado en la sección 2A.3, es $Q(t) = T \cdot M \cdot G$, donde G , el vector geométrico, es constante.

Una forma de expresar la superficies paramétricas es dejar que cada uno de los elementos del vector G varíe en el espacio tridimensional a lo largo de una camino también parametrizado (parámetro " s "). A partir de aquí se llega a la siguiente expresión general:

$$S(s,t) = T \cdot M \cdot G(s) = T \cdot M \cdot \begin{bmatrix} G_1(s) \\ G_2(s) \\ G_3(s) \\ G_4(s) \end{bmatrix} \quad (2A-46)$$

Para un valor fijo de s , $s=s_1$, se tiene que $S(s_1,t)$ corresponde a la ecuación de una curva porque $G(s_1)$ es constante. Al asignar a s otro valor, por ejemplo s_2 , donde s_2-s_1 es pequeño, se obtiene otra curva $S(s_2,t)$ ligeramente distinta de la anterior. Repitiendo el proceso para todos los valores de s entre 0 y 1 se habrá definido una familia de curvas. El conjunto de estas curvas definen una superficie.

Para la generación de superficies bicúbicas, las $G_i(s)$ son curvas cúbicas que se pueden expresar como $G_i(s) = S \cdot M \cdot G_i$, donde $G_i = [g_{i1} \ g_{i2} \ g_{i3} \ g_{i4}]^T$, siendo g_{ij} el elemento j -ésimo del vector geométrico de la curva $G_i(s)$. La matriz transpuesta de $G_i(s)$ es

$$G_i(s)^T = G_i^T \cdot M^T \cdot S^T = [g_{i1} \ g_{i2} \ g_{i3} \ g_{i4}] M^T \cdot S^T \quad (2A-47)$$

con lo que la ecuación 2A.46 permite representar una superficie en función de una malla de puntos de control de la siguiente forma:

$$S(s,t) = T \cdot M \cdot \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{bmatrix} \cdot M^T \cdot S^T \quad (2A-48)$$

ó bien

$$S(s,t) = T \cdot M \cdot G \cdot M^T \cdot S^T, \quad 0 \leq s,t \leq 1 \quad (2A-49)$$

La expresión de cada una de las componentes x , y y z de una superficie paramétrica es:

$$\begin{aligned} x(s,t) &= T \cdot M \cdot G_x \cdot M^T \cdot S^T \\ y(s,t) &= T \cdot M \cdot G_y \cdot M^T \cdot S^T \\ z(s,t) &= T \cdot M \cdot G_z \cdot M^T \cdot S^T \end{aligned} \quad (2A-50)$$

El proceso a seguir para llegar a las expresiones de cada tipo de superficie sería paralelo al que se ha realizado para la definición de las curvas en las secciones anteriores. Se pueden distinguir, entre otras, las superficies de Hermite, de Bézier y las superficies B-splines. Estas últimas son a las que se va a dedicar especial atención.

b) Superficies *B-spline*

Para el diseño de superficies paramétricas se utilizan dos familias de curvas que recorren la superficie en direcciones perpendiculares entre sí. Cuando estas curvas son del tipo *B-spline*, se requieren además dos vectores de nodos, uno para cada familia de curvas, y una matriz rectangular de puntos de control. Los dos vectores de nodos son necesarios para recorrer las dos dimensiones perpendiculares de la matriz rectangular de puntos de control.

En la figura 2A-17 se representa una superficie *B-spline* con las dos familias de curvas que la definen y la poligonal bidimensional correspondiente a la matriz de puntos de control.

La forma matemática de las superficies *B-spline* es un producto tensorial de los dos parámetros s y t . Para el caso no racional se obtiene:

$$S(s,t) = \sum_{i=1}^n \sum_{j=1}^m B_{i,k}(s) B_{j,l}(t) P_{i,j} \quad (2A-51)$$

donde

- $P_{i,j}$ es una matriz ($n \times m$) de puntos de control.
- $B_{i,k}(s)$ es el polinomio de Bernstein de orden k definido por el vector de nodos $\{s_p\}_{p=1}^{n+k}$
- $B_{j,l}(t)$ es el polinomio de Bernstein de orden l definido por el vector de nodos $\{t_q\}_{q=1}^{n+l}$

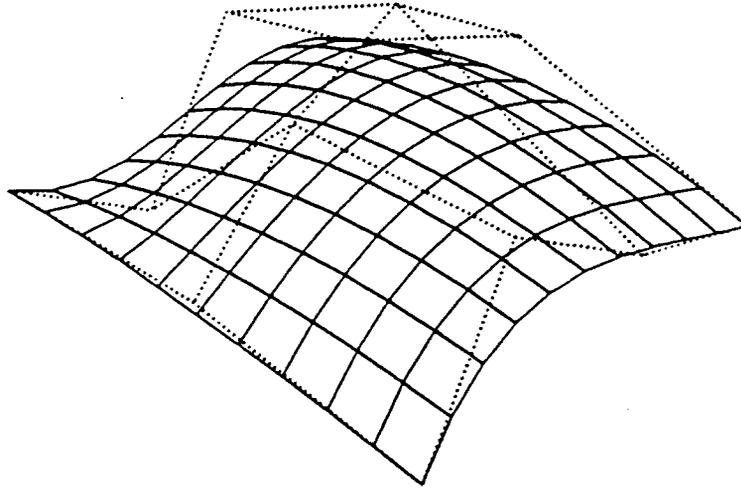


Figura 2A-17: Superficie B-spline

Para el caso racional se obtiene una expresión, paralela a la anterior, en coordenadas homogéneas:

$$S^h(s,t) = \sum_{i=1}^n \sum_{j=1}^m B_{i,k}(s) B_{j,l}(t) P^{h_{i,j}} \quad (2A-52)$$

que al proyectarla en el espacio euclídeo tridimensional se obtiene la forma polinómica racional equivalente:

$$S(s,t) = \frac{\sum_{i=1}^n \sum_{j=1}^m B_{i,k}(s) B_{j,l}(t) h_{i,j} P_{i,j}}{\sum_{i=1}^n \sum_{j=1}^m B_{i,k}(s) B_{j,l}(t) h_{i,j}} \quad (2A-53)$$

c) Curvas definidas sobre superficies B-spline: *Trimming curves*

Para modelar agujeros en las superficies *B-spline*, los límites externos de las mismas, o bien líneas de intersección entre superficies es necesario definir curvas sobre la propia superficie. A estas curvas se las denomina *trimming curves*.

Para asegurar que todos los puntos de las *trimming curves* pertenezcan realmente a la superficie sobre la que éstas están definidas, se representan en coordenadas paramétricas de la superficie. Para ello, basta expresar los puntos de control de las *trimming curves* en coordenadas paramétricas de la superficie *B-spline*, iguales al valor de los dos parámetros de la superficie en estos puntos.

En la figura 2A-18 se muestra un ejemplo de definición de una *trimming curve* sobre una superficie *B-spline* para recortar un agujero sobre la misma. Obsérvese que la superficie resultante puede ser tanto la exterior como la interior a la *trimming curve*.

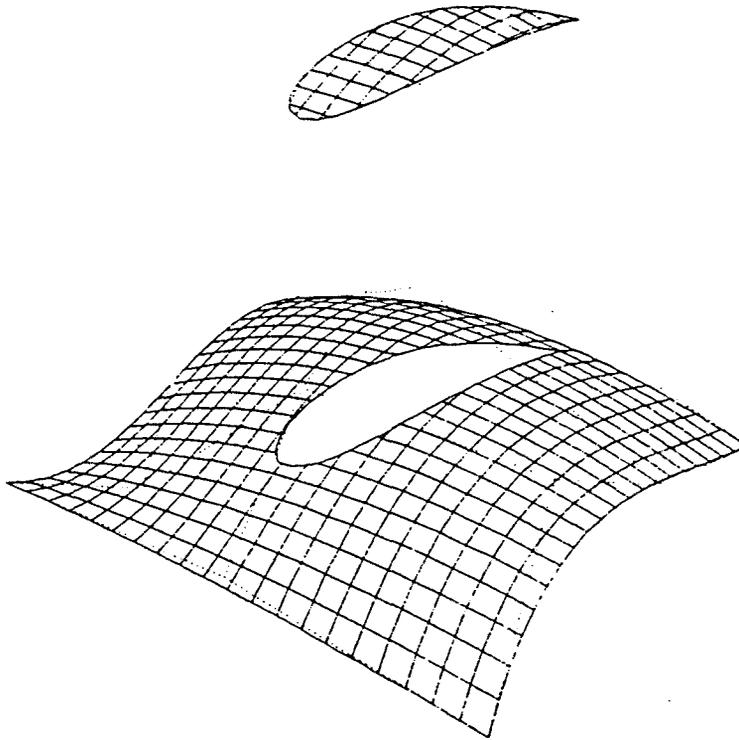


Figura 2A-18: *Trimming curves* definidas sobre una superficie *B-spline*

