

OPTIMIZACIÓN DE MALLAS NO EUCLÍDEAS

Nestor A. Calvo y Pablo J. Novara

*Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral, Ciudad Universitaria,
Santa Fe, Argentina*

Palabras Clave: Mallas, Superficies, Suavizado, Optimización.

Resumen. Se presenta el desarrollo de algunos algoritmos para optimización de forma y densidad de elementos en mallas sobre superficies curvas. Dichas mallas poseen condicionantes que deben respetarse: los vértices angulosos no deben moverse, los nodos pertenecientes a aristas angulosas o de frontera sólo pueden moverse por las curvas que los definen y, finalmente, los nodos de las zonas suaves de la superficie son libres de moverse por la superficie curva. La geometría curva original del diseño se desconoce y debe deducirse de la malla dada, para posicionar correctamente los nodos movidos o agregados. Se muestran algunos resultados obtenidos para el proceso de optimización de forma, refinamiento y desrefinamiento de este tipo de mallas.

1 INTRODUCCION

La optimización geométrica de mallas es un procedimiento con dos objetivos: Por un lado se pretende que la forma de los elementos tienda a la regularidad; esto es para minimizar los errores numéricos en las integrales e interpolaciones que se harán en el uso de la malla para cálculos. También se busca que la densidad de nodos o el tamaño de los elementos sea el adecuado para un debido muestreo del dominio, de acuerdo a la variabilidad de su propia geometría o al error de discretización o de cálculo aceptable en cada zona.

En una clasificación se suele hablar de optimización con y sin cambios topológicos, cambios que pueden consistir en: a) modificación de las conectividades, comúnmente llamadas operaciones de *swapping* o *flipping*; y/o b) el agregado o eliminación de nodos, operaciones que suelen denominarse refinamiento o des-refinamiento (*refinement / coarsening*). Las modificaciones sin cambios topológicos mueven los nodos sin cambiar sus conexiones y el proceso suele denominarse *smoothing* o suavizado y a veces: *mesh-moving*.

En este trabajo se utilizan ambos tipos de operaciones que realiza el programa *meshsuite* sobre mallas en superficies curvas, de las cuales se desconoce la geometría “exacta”, es decir las superficies del diseño original.

En los programas que permiten diseñar el dominio (CAD) y generar la malla con el mismo paquete de software, la geometría curva de la superficie suele estar definida mediante un conjunto de *trimmed-NURBS* o algún otro tipo de superficie analíticas, ya sean paramétricas o implícitas, diseñadas u obtenidas como frontera de un árbol de operaciones booleanas entre sólidos simples. La malla de frontera termina siendo la unión de varias mallas, cada una producida en una superficie individual (patch) y definida por una misma formulación analítica. En dichos programas la malla suele realizarse en dimensiones sucesivas, primero las líneas de interfaz entre las superficies, luego las superficies y finalmente la malla de volumen.

Cuando se recibe una malla sin la información geométrica subyacente hay que reconstruir dicha información para mover adecuadamente los nodos sin perder la forma de la superficie. Una técnica estándar consiste en considerar la superficie localmente suave, en tal caso hay muchas técnicas analíticas (*fitting*) y discretas (*subdivision*) para reconocer la superficie. Para las mallas que se utilizarán para el cálculo se prefiere el *fitting* que interpola los nodos a las aproximaciones que tienen más en cuenta a las caras que a los nodos; esto es debido a que, para el cálculo, toda la información de las variables y condiciones de borde se almacena en los nodos.

Como es habitual en los algoritmos geométricos, cuando los datos tienen suficiente regularidad muchas técnicas, aún las más simples, funcionan muy bien; mientras que en la realidad es mínima la frecuencia de aparición de datos regulares y cualquier técnica tiene un límite de robustez y se pueden encontrar casos para los cuales no funciona adecuadamente.

Este trabajo está orientado a dos procedimientos distintos pero relacionados: Un caso se puede observar en la [Figura 1](#), es la optimización de mallas más o menos regulares, pero que contienen elementos de mala calidad de forma o no respetan la densidad de nodos requerida. El otro procedimiento, mostrado en la [Figura 2](#), es el caso del remallado de mallas que definen muy bien la geometría pero tienen muy mala calidad de forma y una densidad de nodos que sólo depende de la curvatura, es el caso de las mal llamadas “mallas STL”; porque si bien STL (*STereoLithography*) es un formato de almacenamiento, el término se utiliza para las mallas que proveen algunos sistemas de CAD que graban, de ese modo, una malla de triángulos y/o cuadriláteros, más adecuada para procesos de renderizado o visualización que para el cálculo de elementos finitos.

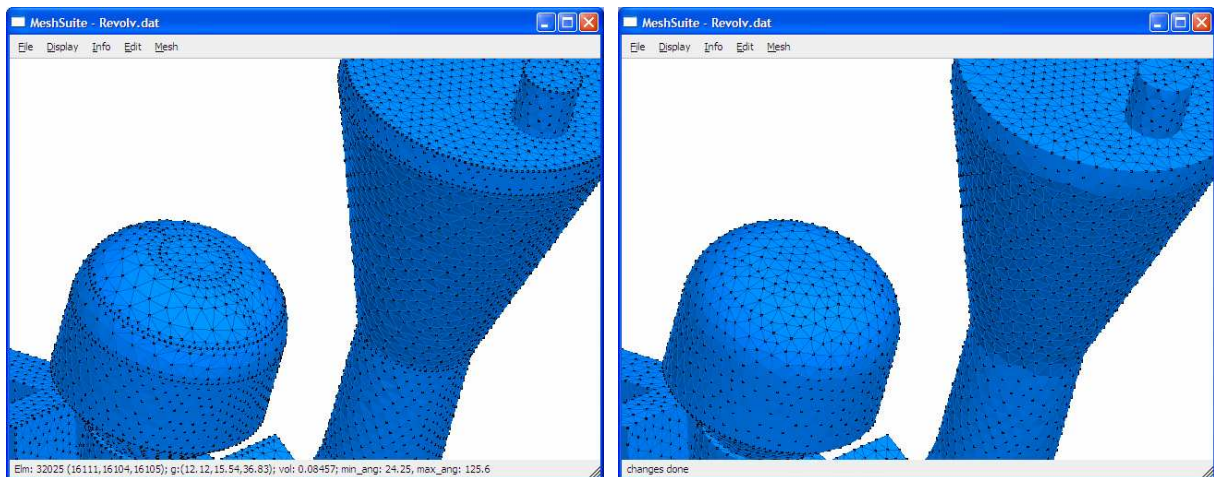


Figura 1: Optimización. Izq.: malla original. Der.: malla resultante

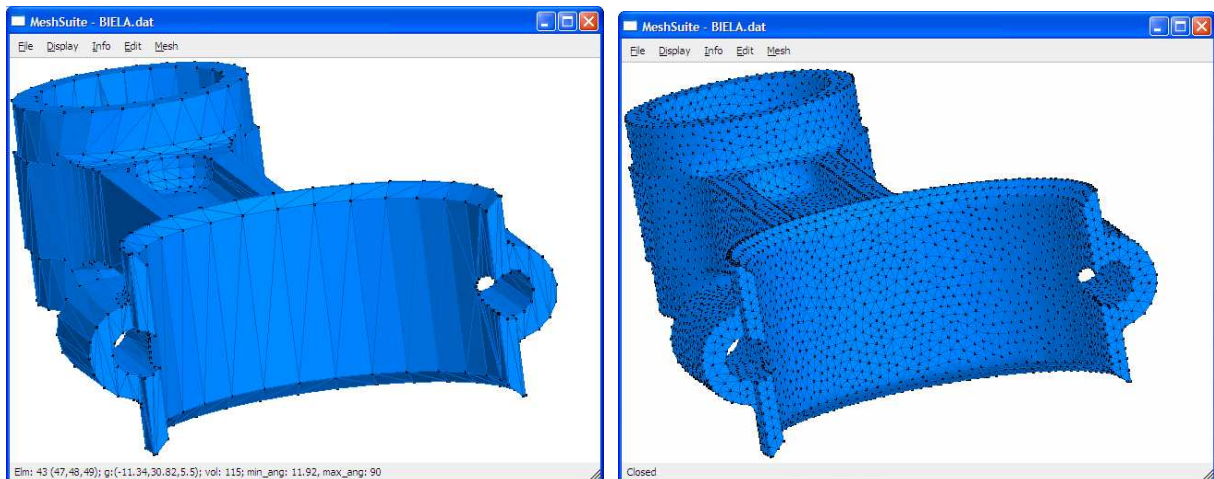


Figura 2: Remallado. Izq.: malla original. Der.: malla resultante

2 PREPROCESO

Antes del proceso de optimización se hace necesario adecuar la malla original mediante procedimientos que se describen a continuación.

2.1 Densidad de nodos

En los dos ejemplos mostrados más arriba, la densidad de nodos es constante, pero no necesariamente debe ser así. Para especificar la densidad en cada región, *meshsuite* asigna un valor real h a cada nodo; que indica, como es habitual, a que distancia se espera que haya otro nodo. Naturalmente el valor de h depende, por un lado, de las necesidades del cálculo respecto a un muestreo adecuado del dominio y las variables para garantizar una cota admisible del error; por otro lado, también depende de la potencia y tiempo de cálculo disponibles; datos que no pueden ser adivinados por el programa de mallado el cual solo dispone de la malla de entrada para una deducción razonable de dicho valor.

La asignación puede venir desde el archivo de datos o puede calcularse de diversas maneras, siendo la habitual un proceso que parte de medir la distancia menor de entre los nodos interconectados por aristas de elementos; para luego “suavizar” este valor,

promediándolo en múltiples pasos (5) con los valores de los vecinos. Otros dos procedimientos posibles son: a) Desde una malla background: El usuario define una malla muy gruesa de elementos volumétricos con h definido en sus (pocos) nodos, se averigua en que gran elemento cae cada nodo y se asigna el valor interpolado. b) Adaptividad: en procesos recursivos o de varios pasos de tiempo, la densidad se calcula en base a los gradientes de las soluciones y las máximas variaciones admisibles.

En cualquier caso, para *meshsuite*, son los nodos y no los elementos los que tienen la información de densidad de muestreo h , y esa información es estática: si el proceso de optimización mueve o agrega nodos, cada valor de h se interpola de los vértices del triángulo original en la posición del nodo actual.

2.2 Detección de aristas y vértices

Para *patches* contiguos de superficies analíticas las aristas son las uniones o intervalos sin continuidad G^1 , es decir: donde quiera que no haya continuidad del plano tangente. Para una malla de elementos planos nunca hay tal continuidad, por lo tanto hay que establecer un criterio de ángulo límite.

Se definen como aristas geométricas:

- a) Las aristas que separan dos elementos con un ángulo suficientemente alejado del llano. El cálculo se realiza mediante el producto escalar de las normales unitarias de los elementos vecinos, que se compara con un coseno límite, cuyo valor *default* para *meshsuite* es de 0.71 y corresponde a considerar un octógono como una circunferencia.
- b) Las aristas de frontera, es decir las aristas que pertenecen a un solo elemento.
- c) Las aristas *non-manifold*: aquellas que son compartidas por más de dos elementos.

En la [Figura 1](#) se puede observar que por efecto del ángulo límite utilizado, hay una transición suave entre el cono y los cilindros adyacentes. Sin intervención del usuario no hay forma de establecer un valor universal, que se adecue a todos los casos.

Los vértices geométricos o nodos fijos, son los nodos que:

- a) Unen dos aristas geométricas con un ángulo suficientemente alejado del llano, que se calcula con el mismo coseno límite antes citado.
- b) Unen más de dos aristas geométricas.
- c) Pertenecen a una sola arista geométrica.
- d) No tienen ninguna arista geométrica asociada, pero todos los elementos adyacentes forman la punta de un cono. Para detectar ésta situación se compara la “normal del nodo” con las normales de los elementos adyacentes. (Las normales por nodo se calculan según se explica en [Calvo 2009](#): minimizando el máximo ángulo con las caras).
- e) Son extremos de aristas *non-manifold*, o tienen elementos de distintas zonas, donde una zona de la superficie se intercepta o es tangente a otra zona. El caso se detecta porque no se puede definir un orden circular para los nodos y elementos adyacentes.

Como puede verse, se analizan condiciones métricas y topológicas para discriminar los nodos y las aristas especiales.

Una vez detectados los vértices inmóviles y las aristas angulosas se identifican las secuencias continuas de aristas sucesivas, ya sean cerradas o que terminan en uno o dos vértices geométricos. La finalidad es que los nodos vértice permanezcan invariables en la malla final, mientras que los bordes angulosos pueden cambiar la ubicación y cantidad de nodos que la definen, pero éstos deben estar siempre sobre las líneas curvas que se obtienen de ellas.

Los nodos admiten una clasificación de acuerdo a los diedros de los elementos que lo tienen por vértice:

- Nodo fijo, definido anteriormente.
- Nodo de arista: en el que convergen exactamente dos aristas geométricas y con un ángulo más llano que el límite impuesto, el nodo puede moverse sobre la secuencia de aristas y puede ser eliminado.
- Nodo suave: no es fijo ni está en ninguna arista, se mueve por la superficie y también puede ser eliminado.

2.3 Mesh-healing

Muchas veces la malla dato está malformada, dando lugar a la presencia de aristas que no deberían serlo. Estas aristas falsas pueden ser interiores, provenientes de elementos muy deformes; o bien de frontera, provenientes de *patches* mal unidos. Un ejemplo se muestra en la siguiente figura, donde se muestra un borde suave, conocido como *suitcase corner*, en el que aparecen errores del proceso de teselación con el que el CAD generó el archivo STL. Como consecuencia de esos errores, aparecen triángulos muy aplastados que se conectan con sus vecinos con ángulos diedros muy marcados. En la figura se han resaltado en rojo tres zonas con dicho error.

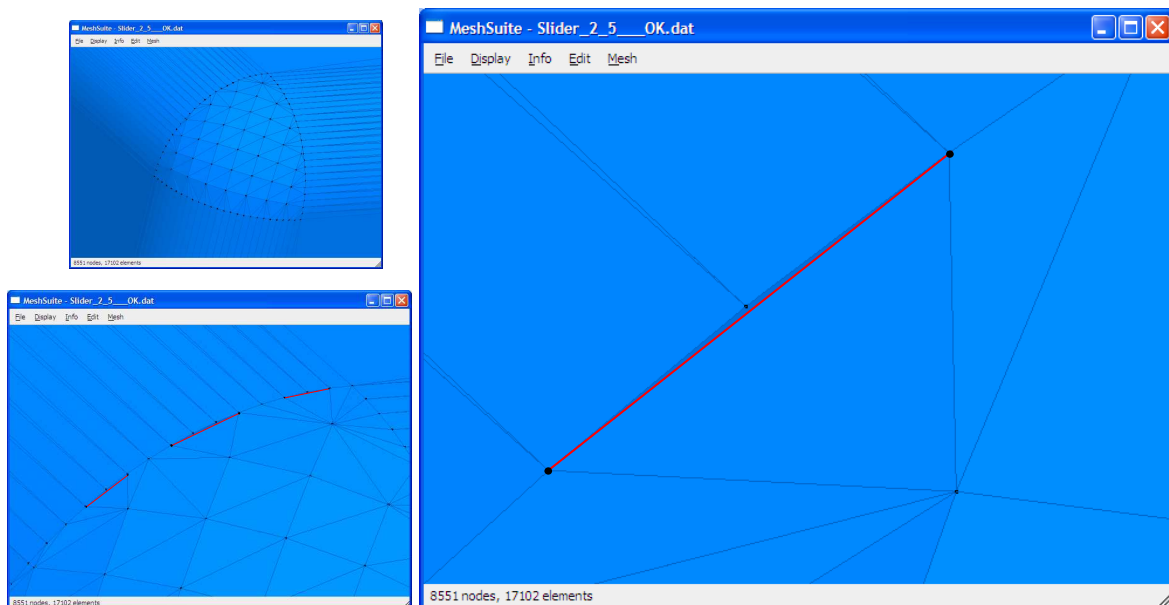


Figura 3: Detalle de aristas falsas causadas por defectos de la malla original.

La solución automática para estos y otros problemas del mismo tipo, es un proceso conocido como *mesh-healing*, que en este caso se divide en tres subprocesos:

- Unificación o *welding* de nodos cuya distancia sea menor que una tolerancia ($1/50 h$). Este proceso se realiza mediante consulta a un *octree* que devuelve el nodo más cercano a un punto dado.
- Cierre o *zipping* de los elementos separados, pero que tienen aristas prácticamente coincidentes. El elemento con la arista mayor se divide en dos y una de las partes se une al elemento adyacente.
- Swapping*: Cuando se identifica una arista geométrica, se verifica si el nodo opuesto de cada uno de los dos elementos que se unen en la arista, dista de la arista en menos que

una cierta tolerancia ($1/50 \mathbf{h}$). En tal caso, de ser posible, se intercambian las diagonales del cuadrilátero formado por el triángulo aplastado y el vecino por dicha arista. Ese es el caso de la figura de arriba.

Resulta evidente que cada vez que se utiliza una tolerancia se resuelven parcialmente los problemas, puesto que pueden quedar algunos casos en los que el error sea mayor que el valor tolerado. Lo que es menos evidente es que algunas veces, por más pequeña y “segura” que sea la tolerancia, la solución para una zona genera problemas en otra, por lo tanto estas son herramientas de *mesh healing* están disponibles para ser utilizadas, pero sin garantías de que sirvan al aplicarlas indiscriminadamente.

3 DETERMINACIÓN DE LA GEOMETRÍA SUBYACENTE

Para mover o agregar nodos en una malla de superficie hay dos opciones:

- En los triángulos de la malla dada.
- En una geometría curva deducida de la malla dada.

La primera consiste en mantener una copia de la malla original y ubicar los nodos en los planos de sus elementos. Si la geometría original está muy bien representada (STL) o bien las flechas (plano-superficie) no son muy grandes frente al \mathbf{h} requerido, éste método no es una mala elección. Por otro lado, si hay que refinar una malla de partida gruesa se hace necesario recrear la superficie, puesto que si los nodos se ubican en los planos originales se perderá volumen en las convexidades y se ganará en las concavidades.

El método elegido consiste en construir una superficie Bezier por cada elemento de la malla original, ya sean triángulos o cuadriláteros. Se eligió ese método considerando otras opciones implementadas (varios esquemas de *subdivisión surfaces*, *PN-triangles*) o analizadas (*Gregory patches* y *patches* de Bezier degenerados de 4° y 5° grado, con continuidad G^1) por un compromiso entre simplicidad y resultados efectivamente obtenidos. El método que se muestra en 3.2 presenta continuidad en los nodos pero no en las aristas, de todos modos la formulación adoptada para calcular los puntos de control genera interfaces suficientemente suaves.

3.1 Líneas Curvas

En primer lugar se analizan las curvas (1D) que se obtienen de las poligonales de aristas geométricas. Se realizan mediante splines compuestas de curvas individuales de Bezier de tercer grado unidas con continuidad G^1 , es decir con tangente suave. Para ello utilizamos una modificación del método de Overhauser (Buss 2003, Cap. 7) diseñada para adaptarse mejor a mallas para elementos finitos y no para *rendering* fotorealista. Cada arco de Bezier de 3° grado se define mediante cuatro puntos de control, los dos nodos extremos y dos puntos intermedios, que no son nodos de la malla, y se calculan a partir de la secuencia de nodos.

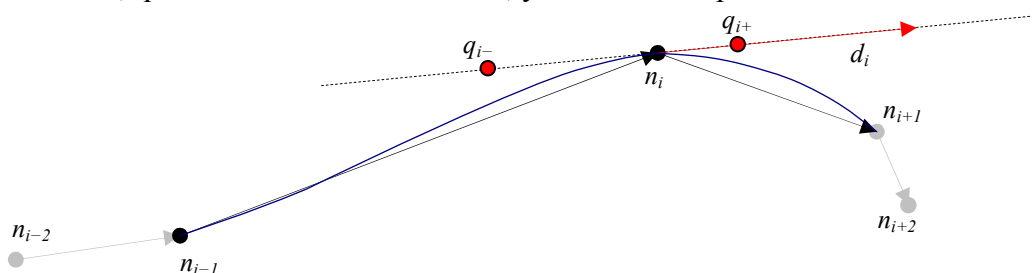


Figura 4: Puntos de control asociados a cada nodo en 1D

La Figura 4 muestra los elementos para el cálculo de los puntos de control q_{i-} y q_{i+} ,

correspondientes a los arcos de Bezier anterior y posterior al nodo n_i respectivamente. En primer lugar se obtiene la dirección d_i de la recta que pasa por el nodo, contiene a los dos puntos de control asociados y será la tangente común a los dos arcos que se unen en el nodo.

$$d_i^1 = [(n_i - n_{i-1}) \cdot |n_i - n_{i-1}| + (n_{i+1} - n_i) \cdot |n_{i+1} - n_i|]^1 \quad (1)$$

Donde $|v|$ representa el módulo del vector v y v^1 indica que el vector es unitario: $v^1 = v/|v|$.

Como puede verse en la ecuación 1, la dependencia con la longitud de cada segmento es cuadrática, otorgando mucha más importancia a la dirección del segmento más largo.

En la recta tangente, definida por el nodo n_i y la dirección d_i^1 , se definen los puntos de control extra del siguiente modo:

$$q_{i-} = n_i - d_i^1 |n_i - n_{i-1}|/3; \quad q_{i+} = n_i + d_i^1 |n_{i+1} - n_i|/3; \quad (2)$$

El factor $1/3$ no es del todo arbitrario, en la bibliografía está justificado por el hecho de que el parámetro variable se puede hacer global y varía una unidad entre cada par de nodos sucesivos (la derivada paramétrica es el triple de la longitud del segmento de control, de modo que es aproximadamente uno en toda la curva). Pero el parámetro es en sí arbitrario. De todos modos ese valor proporciona curvas bien formadas y por lo tanto así se utiliza.

En los nodos terminales se cuenta con un único nodo adyacente para definir el punto de control intermedio asociado al nodo terminal.

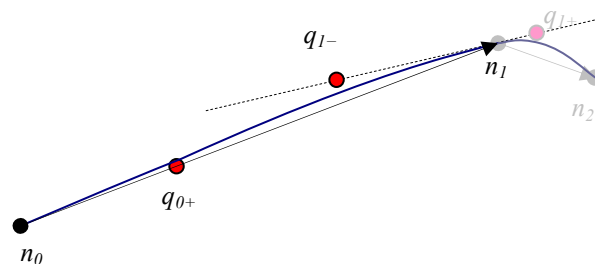


Figura 5: Puntos de control asociados a los nodos terminales.

El punto de control se ubica sobre el segmento que une los nodos, tal como muestra la figura, a un tercio del segmento. Suponiendo m segmentos:

$$q_{0+} = n_0 + (n_1 - n_0)/3 \quad q_{m-} = n_m - (n_m - n_{m-1})/3 \quad (3)$$

El principio rector de la elección de los parámetros utilizados es permitir que la curva represente una geometría razonable cuando hay mucha diferencia en la longitud de aristas adyacentes y el mismo criterio será utilizado para grandes relaciones de aspecto en mallas de superficie. Cuando la malla es regular cualquier esquema razonable produce resultados razonables. Cuando coexisten aristas largas con aristas pequeñas, se supone que es porque la geometría tiene muy poca flecha sobre las aristas largas.

3.2 Superficies

Así como en las aristas hay una sola dirección asociada a cada nodo, en las superficies hay una única normal asociada a cada nodo. En los nodos suaves la normal única determina el plano tangente común a las superficies elementales adyacentes.

En la figura siguiente se esquematizan los puntos de control que corresponden a un cuadrilátero y a un triángulo de Bezier, que son los elementos admitidos en este trabajo.

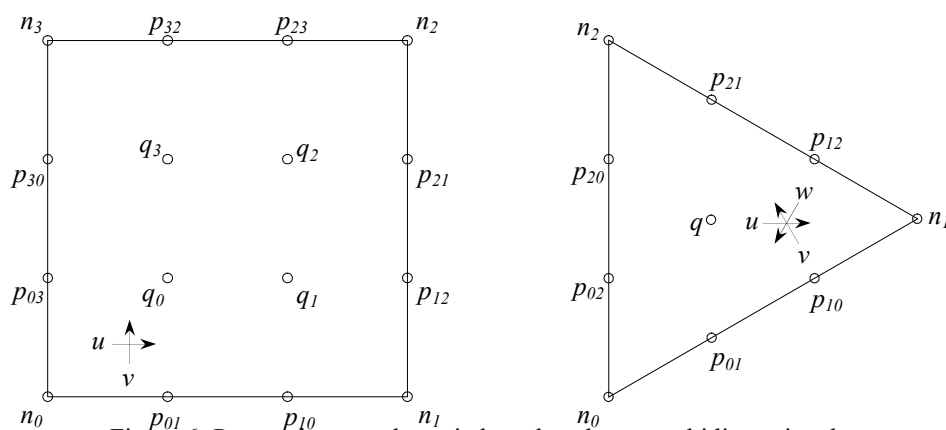


Figura 6: Puntos de control asociados a los elementos bidimensionales.

Para el caso de los cuadriláteros la situación es muy benigna puesto que todos los puntos de control asociados a un nodo se pueden poner en un mismo plano definido por la normal del nodo. Es decir que, por ejemplo, los puntos p_{03} , q_0 y p_{01} estarán en el plano que pasa por el nodo correspondiente n_0 ; al igual que todos los puntos equivalentes de los elementos que rodean al mismo nodo si es un nodo suave. En el caso del triángulo no es así, en el plano normal en n_0 sólo se encuentran los puntos p_{02} y p_{01} y, sin elevar el grado, no hay otra alternativa que desarrollar una estrategia cuidadosa para ubicar el punto de control central q .

La normal d_i en el nodo n_i se calcula según se explica en Calvo, 2009, de modo de asegurar que la normal “vea bien” a todos los elementos del nodo; no sería así si se utilizara un esquema parecido al de las líneas y se calculase por áreas o longitudes.

Para todos los puntos p , de triángulos o cuadriláteros, vale la misma ecuación, equivalente del caso 1D, que consiste en proyectar la arista sobre el plano normal y en esa dirección alejar el punto a $1/3$ de la longitud de la arista:

$$p_{ij} = n_i + [(n_j - n_i) - d_i \cdot ((n_j - n_i) \cdot d_i)] |n_j - n_i|/3 \quad (4)$$

Los puntos q de los cuadrángulos se calculan según:

$$q_i = p_{ij} + p_{ik} - n_i \quad (5)$$

Donde j y k son los índices de los nodos vecinos del nodo i en el cuadrángulo.

Para el triángulo, el punto central es un promedio ponderado de tres puntos intermedios calculados del mismo modo que para el cuadrilátero (ec. 5). El promedio se pondera para lograr que la superficie curva se aproxime al plano en los triángulos muy agudos. Cada punto q_i , calculado como se dijo, se pondera con la longitud de la arista opuesta al nodo n_i .

$$q = \sum q_i |n_j - n_k| / \sum |n_j - n_k| \quad (6)$$

En los nodos que contienen aristas angulosas no hay un único plano. Se utiliza el cálculo equivalente al caso 1D (ecs. 1 a 3). Sobre las aristas geométricas se utiliza el punto 1D calculado para la línea (ecs. 1 y 2) y en el resto de las aristas de los elementos del nodo se utiliza la ecuación 3, que es el mismo criterio usado para los nodos terminales de las líneas que se colocan sobre las aristas.

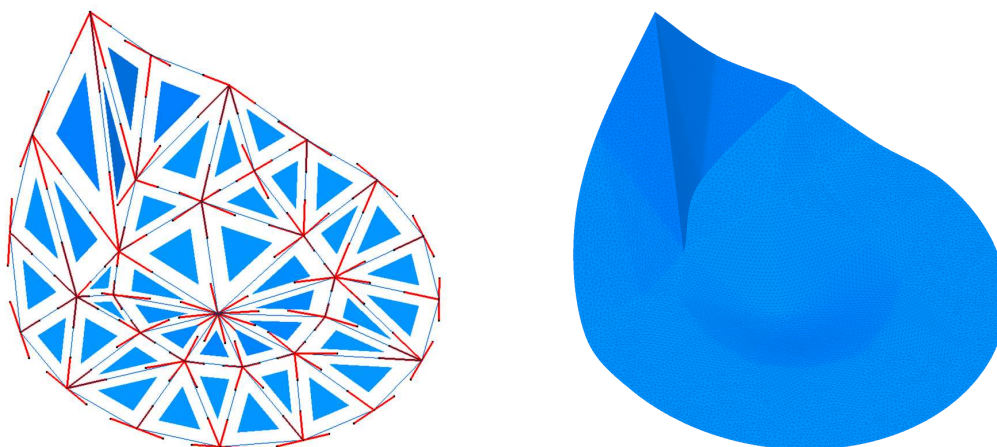


Figura 7: Ejemplo de definición de los puntos de control de aristas y resultado de un fuerte refinamiento.

A modo de ejemplo, la figura anterior muestra una malla de prueba sobre un disco donde se han movido algunos nodos. Se puede apreciar el grado de suavidad obtenida al densificar la malla. Los segmentos rojos de la izquierda son los que unen cada nodo con sus puntos tipo p , los puntos de control interiores no se muestran (no se almacenan).

3.3 Evaluación

Para evaluar la superficie curva en un punto, se deben conocer los parámetros u y v del punto en la superficie de Bezier elemental. Se debe identificar primero el triángulo que contiene al punto y sus coordenadas baricéntricas en el mismo, para luego utilizarlas como parámetros en la superficie. Para calcular las coordenadas baricéntricas, todos los elementos cuadriláteros se dividen en dos triángulos y, sólo si el cuadrilátero contiene al punto, se invierten las funciones de forma bilineales del cuadrilátero según (Buss 2003, Cap. 4).

El punto que se quiere evaluar es un punto nuevo o movido que no está, en general, apoyado en el plano ni en la superficie. Para ubicarlo se utilizan varios criterios que, comenzando por un punto cercano conocido, van aproximando la ubicación del punto.

El criterio principal consiste en particionar el espacio alrededor de la malla de acuerdo a las normales por nodo, como se explica en (Novara 2010): se calculan primero las coordenadas baricéntricas en un triángulo paralelo que pasa por el punto y se extiende según las normales de cada nodo; de acuerdo a los valores se avanza por vecindades. Si el criterio principal falla (inversión por lejanía) se ensayan otras aproximaciones utilizando el *octree* y las conexiones que provee la malla. Existen muchas fuentes de posibles fallas, en general en puntos muy alejado de la superficie con mucha curvatura y/o aristas angulosas en la zona.

Finalmente, las funciones de forma calculadas para el punto en el plano permiten evaluar su posición sobre la superficie de Bezier correspondiente.

4 OPTIMIZACIÓN

El proceso de optimización en sí consta de varias fases recursivas y concurrentes de técnicas estándar:

- a) Eliminar nodos cercanos.
- b) Agregar nodos donde hagan falta según el h local.
- c) Swap de diagonales.
- d) Suavizado laplaciano.

Si se pudiese mantener una malla Delaunay en una superficie curva sería mucho mejor

utilizar el centro de las circunferencias para agregar nodos. Dado que es imposible, sobretodo en superficies con aristas y vértices angulosos, se utilizan las longitudes de aristas para analizar si la distancia es pobre o excesiva frente al h requerido.

En caso de que se detecte que dos nodos están muy cerca se unifican, determinando si sus posiciones (y valores asociados) se pueden promediar o no, dependiendo de si los nodos son vértice, arista o suaves. Se realizan dos tipos de detección de cercanía, una es la longitud de arista de dos nodos conectados, la otra se utiliza cuando se agrega un nodo nuevo, se consulta al *octree* si hay algún nodo muy cercano, en cuyo caso el nodo no se agrega. Esto es así porque el punto central de una arista puede estar muy cerca de un nodo preexistente que no esté conectado con los extremos de la arista larga.

Las longitudes tienen un problema grave en mallas curvas: miden la cuerda y no el arco de superficie, por lo tanto, cuando es posible, se evalúa el punto medio sobre la superficie y se suman longitudes hasta que la flecha sea aceptable en relación al h local.

El *swap* de diagonales es el proceso por el cual se elige la diagonal más favorable del cuadrilátero formado por dos triángulos adyacentes. En este trabajo el criterio utilizado consiste en maximizar la menor altura de los triángulos. Para realizar el *swap* antes se realizan varios análisis, pues debe mantenerse la condición de ser o no arista geométrica para las aristas involucradas y debe evitarse dejar nodos no-frontera con menos de cuatro elementos asociados. Si las alturas son razonablemente parecidas, se utiliza la diagonal mejor “apoyada” sobre la superficie, evaluando el punto medio.

El suavizado laplaciano es otra técnica estándar utilizada. Consiste en mover un nodo al centroide de sus vecinos. Dicho centroide puede estar muy alejado de la superficie cuando ésta tiene mucha curvatura, por lo tanto siempre se analiza si puede encontrarse el punto correspondiente y en caso contrario el nodo no se mueve.

5 RESULTADOS

Además de los mostrados en la introducción, aquí se agregan algunos ejemplos más de los resultados que se logran con el procedimiento explicado.

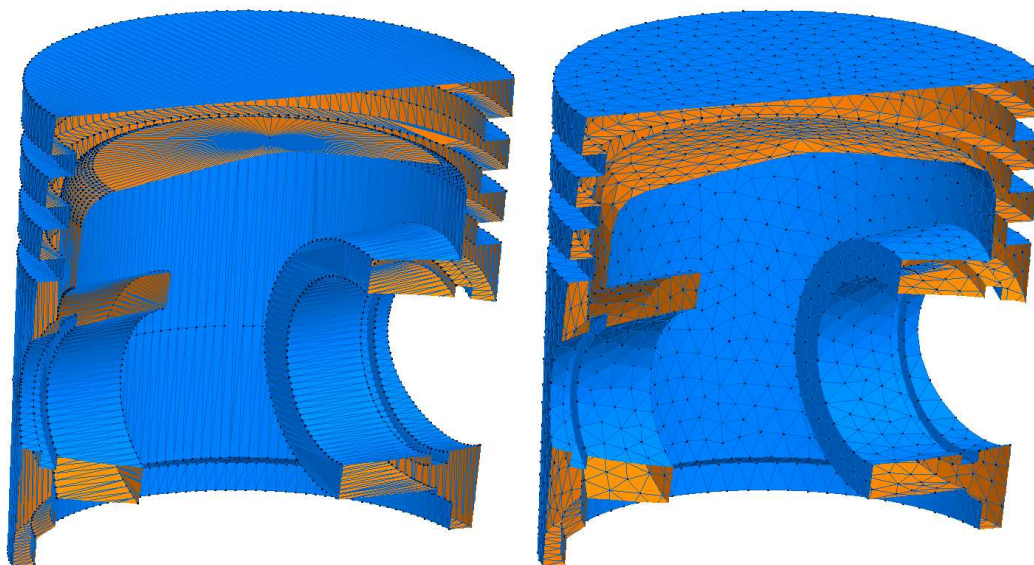


Figura 8: Remallado de la superficie de un pistón (corte).

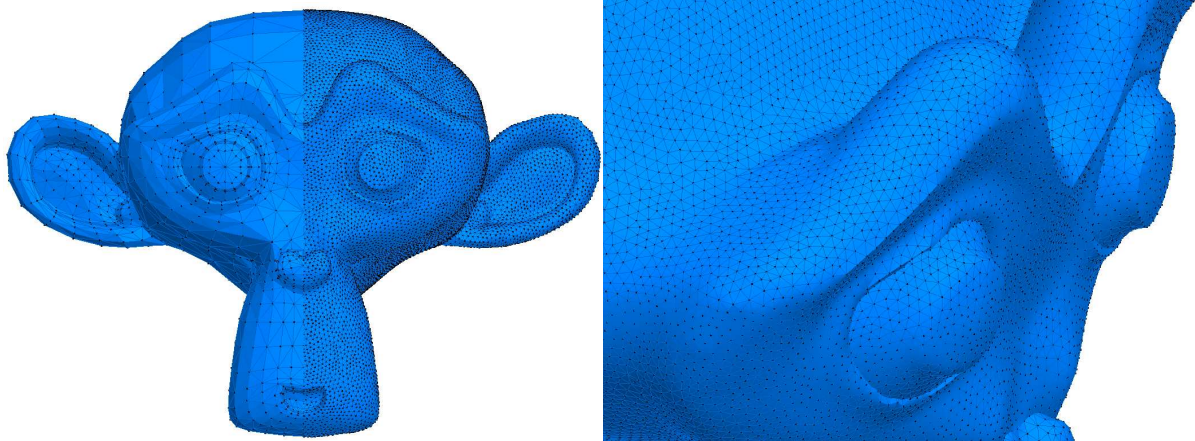


Figura 9: *Monkey*.

Izquierda, comparación original/resultado. Derecha: muestra de la superficie evaluada.

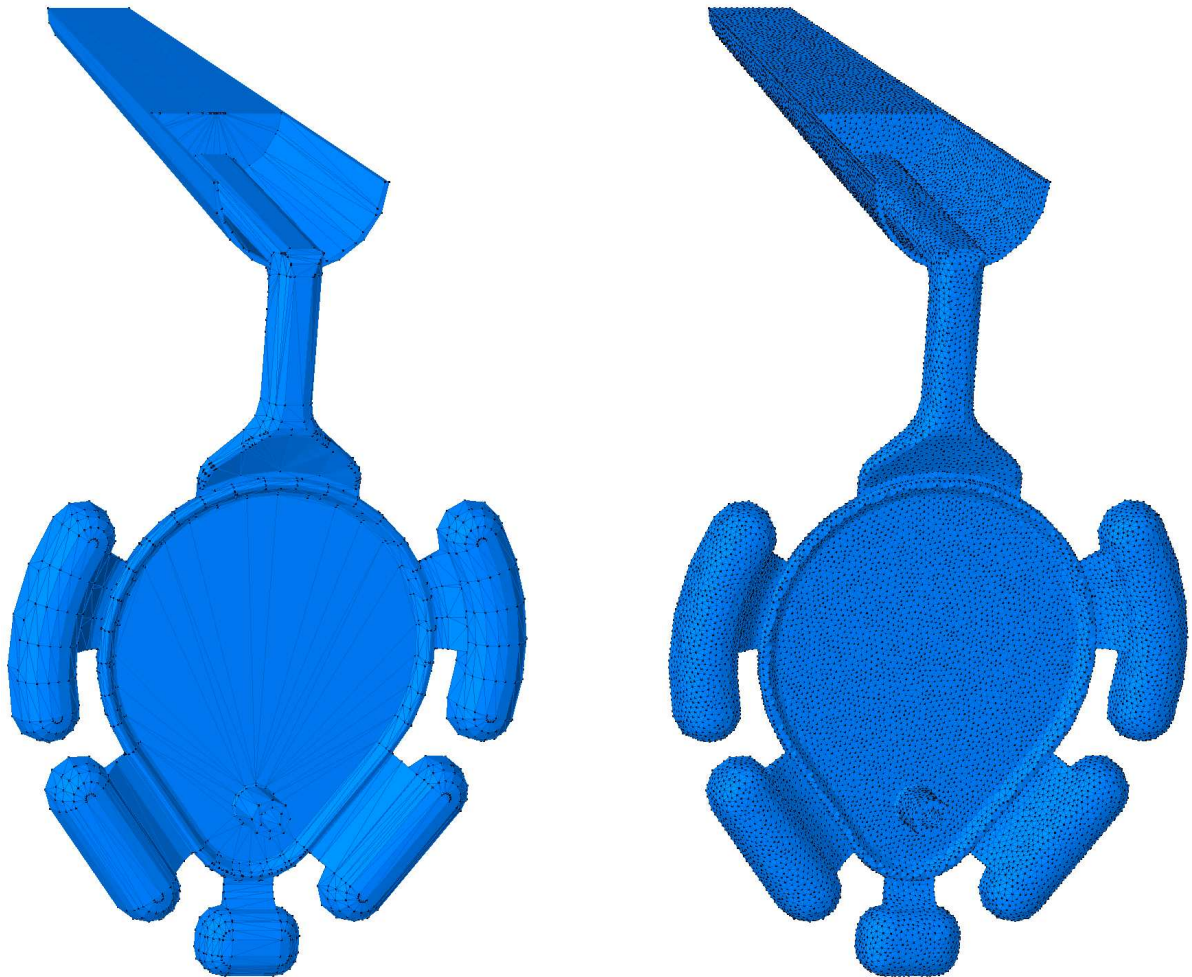


Figura 10: Molde de pieza con canales y mazarotas para fundición metálica.

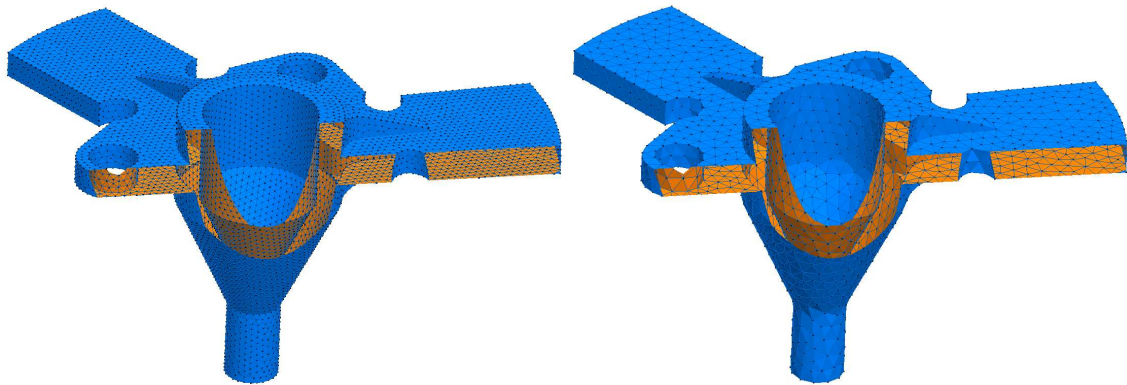


Figura 11: Desrefinamiento de la malla de una pieza mecánica.

6 CONCLUSIONES

Se ha desarrollado un procedimiento que permite optimizar mallas de superficie respetando una geometría subyacente que se identifica en forma automática. El proceso sirve para la optimización de forma de los elementos y densidad de los nodos en mallas de partida muy irregulares, como es el caso de las mallas STL.

Las técnicas originales desarrolladas para deducir la superficie subyacente han dado resultados muy satisfactorios en cuanto a interpretar la forma de una superficie apta para el método de elementos finitos.

Los procedimientos implementados para *mesh-healing* requieren mucho trabajo futuro para aumentar su robustez y poder aplicarlos a un mayor número de mallas erróneas.

REFERENCIAS

- Calvo, N., y Novara, P., Aplicaciones del kernel de un polígono a problemas de edición de mallas. *Mecánica Computacional*, 28,15:1167-1179, 2009.
- Buss, S., *3D Computer Graphics. A mathematical introduction with OpenGL*. Cambridge University Press, 2003.
- Novara, P., y Calvo, N., Corte de mallas de elementos finitos 2.5-dimensionales. *Mecánica Computacional*, 29: 6091-6101, 2010.