Proof.
[$\mathbf{p}_a \in G \Rightarrow \mathbf{p}_a \in$ set of minimal paths] If the point $\mathbf{p}_a$ is in G then $d_{\mathcal{S}}(\mathbf{p}_a, S) + d_{\mathcal{S}}(\mathbf{p}_a, D) = g_m$. Therefore, there exist a path from S to $\mathbf{p}_a$ and from $\mathbf{p}_a$ to D which together form a minimal length path that passes through $\mathbf{p}_a$.

[$\mathbf{p}_a \notin G \Rightarrow \mathbf{p}_a \notin$ any minimal path] If $\mathbf{p}_a \notin G$ then $d_{\mathcal{S}}(\mathbf{p}_a, S) + d_{\mathcal{S}}(\mathbf{p}_a, D) > g_m$. Recalling the d_z definition, all possible paths from S to D which pass through the point $\mathbf{p}_a$ are longer than g_m and, therefore, not minimal. ■

Now, we can prove the following result connecting geodesics to the trace of tangential points of the two equal distance contours α_S and α_D, propagating from the source and destination.

Lemma 6.4 *The tangential points of $\alpha_S(u,t)$ and $\alpha_D(\tilde{u},\tilde{t})$ for $\tilde{t} + t = g_m$, generate the minimal paths from point P_1 to point P_2; i.e lies on a constant parameter $u = u_0$ ($\tilde{u} = \tilde{u}_0$) of the propagating curve $\alpha_S(u,t)$ ($\alpha_D(\tilde{u},\tilde{t})$).*

Proof. The shortest path built by

$$\frac{d\alpha_S}{dt} = N \times \mathcal{T}^{\alpha_S}$$

and tangent at $t = a$ at $u = u_0$ to the path built by

$$\frac{d\alpha_D}{dt} = N \times \mathcal{T}^{\alpha_D}$$

at $\tilde{t} = b$ at $\tilde{u} = \tilde{u}_0$, is of length $a + b$ (see Lemma 6.3).

Assume that there is only one minimal path between P_1 and P_2, and assume that the trace of that minimal path passes through two different parameterization points u_0 and u_1. See Figure 6.3.

Let the length of the shortest distance from P_1 to $\alpha_S(a, u_0)$ be a, and the distance from P_2 to $\alpha_D(\tilde{u}_0, b)$ be b. The shortest path is of length $a + b$, and it is given by the path $P_1 \rightsquigarrow \alpha_S(u_0, a) = \alpha_D(\tilde{u}_0, b) \rightsquigarrow P_2$.

Assume there is a parameterization point $u_1 \neq u_0$ through which the minimal path passes at $t = a + \epsilon$ ($\tilde{t} = b - \epsilon$). According to that assumption there is a minimal path $P_1 \rightsquigarrow \alpha_S(u_1, a) \rightsquigarrow \alpha_S(u_1, a + \epsilon) \rightsquigarrow P_2$ of length $a + \epsilon + (b - \epsilon)$. However, part of this path, $P_1 \rightsquigarrow \alpha_S(u_1, a)$, is not equal to the original sub–path of the minimal path $P_1 \rightsquigarrow \alpha_S(u_0, a)$, and this contradicts the assumption that the minimal path should pass through $\alpha(u_0, a)$, and concludes the proof. ■

We also have the following result.

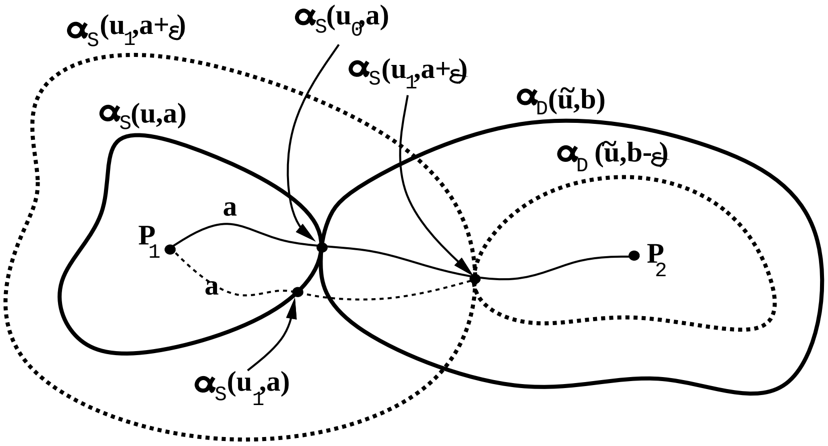

Figure 6.3: When considering only one minimal geodesic between two points, the path is the trace of a constant parameter along the evolution, see text.

Corollary 6.1 *All minimal paths between S and D which defined by G (equation (6.3)), are minimal geodesics.*

In the next section a numerical scheme based on the level set representation of the evolving planar curve is presented. Note that the shortest paths are the minimal value level set of the function $\mathcal{M}_S + \mathcal{M}_D$. This observation will later be used on to find the minimal paths.

6.3 The Numerical Approximation

Using the normal component of the velocity, in the Eulerian formulation as presented in Chapter 2, we get in our case

$$\phi_t = \sqrt{a(x,y)\phi_x^2 + b(x,y)\phi_y^2 - c(x,y)\phi_x\phi_y} \tag{6.4}$$

This equation describes the propagation rule for the surface ϕ.

6.3.1 Finite difference approximation

In our implementations, we use the following finite difference approximation [111, 119, 110]. The minmod definition, see Chapter 2, is used to approximate $\phi_x\phi_y$ by

$$\phi_x\phi_y|_{x=i\Delta x, y=j\Delta y} \approx minmod(D_+^x\phi_{i,j}, D_-^x\phi_{i,j})\; minmod(D_+^y\phi_{i,j}, D_-^y\phi_{i,j}).$$

A different approximation, see Chapter 2, is used for the squared partial derivatives,

$$\begin{aligned}\phi_x^2|_{x=i\Delta x, y=j\Delta y} &\approx (\max(D_+^x\phi_{i,j}, -D_-^x\phi_{i,j}, 0))^2\\ \phi_y^2|_{x=i\Delta x, y=j\Delta y} &\approx (\max(D_+^y\phi_{i,j}, -D_-^y\phi_{i,j}, 0))^2\end{aligned}$$

These finite difference approximations yield a first order numerical scheme for the equal distance contours evolution. Using a forward difference approximation in time gives the following numerical scheme for the propagation of the function $\Phi^n_{i,j} \approx \phi(i\Delta x, j\Delta y, n\Delta t)$ on the (x, y) rectangular grid

$$\begin{aligned} \Phi^{n+1}_{i,j} &= \Phi^n_{i,j} + \Delta t[a_{i,j}(\max(D^x_+\Phi_{i,j}, -D^x_-\Phi_{i,j}, 0))^2 + b_{i,j}(\max(D^y_+\Phi_{i,j}, -D^y_-\Phi_{i,j}, 0))^2 \\ &\quad - c_{i,j}\ minmod(D^x_+\Phi_{i,j}, D^x_-\Phi_{i,j})\ minmod(D^y_+\Phi_{i,j}, D^y_-\Phi_{i,j})]^{\frac{1}{2}} \qquad (6.5) \end{aligned}$$

which is the finite difference approximation of equation (6.4). This numerical scheme is stable and inherently overcomes topological changes in the evolving contour.

6.3.2 Initialization

The function $\phi(x, y, 0)$ is actually an implicit representation of $\partial\tilde{S}$ - the projection of the boundary of the source area ∂S to the (x, y)–plane. The first demand for $\phi(x, y, 0)$ is to follow

$$X(0) \equiv \{(x, y)|\phi(x, y, 0) = 0\} \ = \ \{(x, y)|(x, y, z(x, y)) \in \partial S\} \equiv \partial\tilde{S}.$$

Furthermore, $\phi(x, y, 0)$ should admit smoothness, continuity and be negative in the interior of $\partial\tilde{S}$, and positive in the exterior of $\partial\tilde{S}$. The $(2D)$ planar zero distance contour $\partial\tilde{S}$ is the projection of ∂S on the plane.

There are many ways to initialize $\phi(x, y, 0)$, for examples see [68, 66, 102]. It is possible, for example, to truncate the values of $\phi(x, y, 0)$ using the observation that we are interested in the function behavior only near the relevant contour, (the zero level set). Note that every ϕ function which obeys the demands described earlier is sufficient.

6.3.3 Distance assignment

After the initialization is completed, the ϕ function is propagated according to equation (6.5). While propagating the function, our goal is to find the distance of each grid point. A simple way of achieving (first order) accurate results is by interpolating the zero crossings. At every iteration step, for each grid point, check

If $(\Phi^n_{i,j} \cdot \Phi^{n-1}_{i,j} < 0)$ **then** $Distance_{i,j} = \Delta t(n - \frac{\Phi^n_{i,j}}{\Phi^n_{i,j} - \Phi^{n-1}_{i,j}})$.

Using this procedure, each grid point gets its distance at the "time" when the ϕ functions'–zero level passes through it.

6.3.4 Finding minimal geodesics

Having $\mathcal{M}_S$ and $\mathcal{M}_D$ on the grid, the minimal geodesic may be found in a simple way. Recalling that $g_m = \min(\mathcal{M}_S + \mathcal{M}_D)$, the projection of the minimal geodesic, G, on to the (x, y)–plane is $\tilde{G} = \{(x, y)|(\mathcal{M}_S(x, y) + \mathcal{M}_D(x, y)) = g_m\}$.

The desired minimal geodesics are achieved by applying a contour finder on $\mathcal{M}_S + \mathcal{M}_D$ to find the level set $g_m + \epsilon$, for some very small ϵ, and then applying a simple thinning algorithm that operates on the interior of the minimal level set. It is also possible to apply level sets based refinement methods [77, 78, 27], see Chapter 7.

In [76] we also show how to use level sets of combinations of the geodesic distance maps to solve the 'three point Steiner Problem', and how to compute Voronoi diagrams on $3D$ surfaces.

6.4 Examples and Results

We demonstrate the performance of the algorithm by applying it to a synthetic surface and finding the paths of minimal length given on a mesh of 256×256 points. The source and destination areas are points located at $(x, y) = (64, 64)$ and $(x, y) = (192, 192)$. The (black) minimal geodesics connecting the source and the destination on the egg–box surface is presented in Figure 6.4a. Figure 6.4b present the projection of the equal distance contours from the source point, on which the minimal (plus epsilon) level set is displayed as a smooth black curve and the two minimal geodesics as a chain of pixels. The sum of the distance maps from the source and destination is presented in Figure 6.4c and the geodesic distance from the destination point in Figure 6.4d. The minimal level set of that sum surface (Figure 6.4c) is the minimal geodesics.

Ways of achieving more accurate results are by increasing the grid resolution, and by decreasing the time step ($\Delta t = 0.21$ in our examples).

6.5 Chapter Summary

We have described a numerical method for calculating a geodesic distance map from a given area on a surface, so that topological problems in the propagated equal distance contours are inherently avoided. An algorithm for finding the minimal geodesics between two areas on the surface based on the distance mapping was constructed. The algorithm works on a grid, therefore it is easy to implement the algorithm in parallel using each mesh point as a small calculating device which communicates with its four close neighbors. In each iteration we need to calculate the values of $\phi(x, y, t)$ in those grid points close to the current contour and the rest of the grid points serve as sign holders. This can be exploited to reduce calculation effort. When not considering any possible redundancy, the calculation effort is of order $\mathcal{O}(\frac{L}{\Delta t} m \cdot n)$, where L is the length of the shortest geodesic path and $m \cdot n$ is the number of grid points.

It was shown that wavefront propagation methods in fluid dynamics also provide a nice approach to the problem of finding the minimal geodesics.

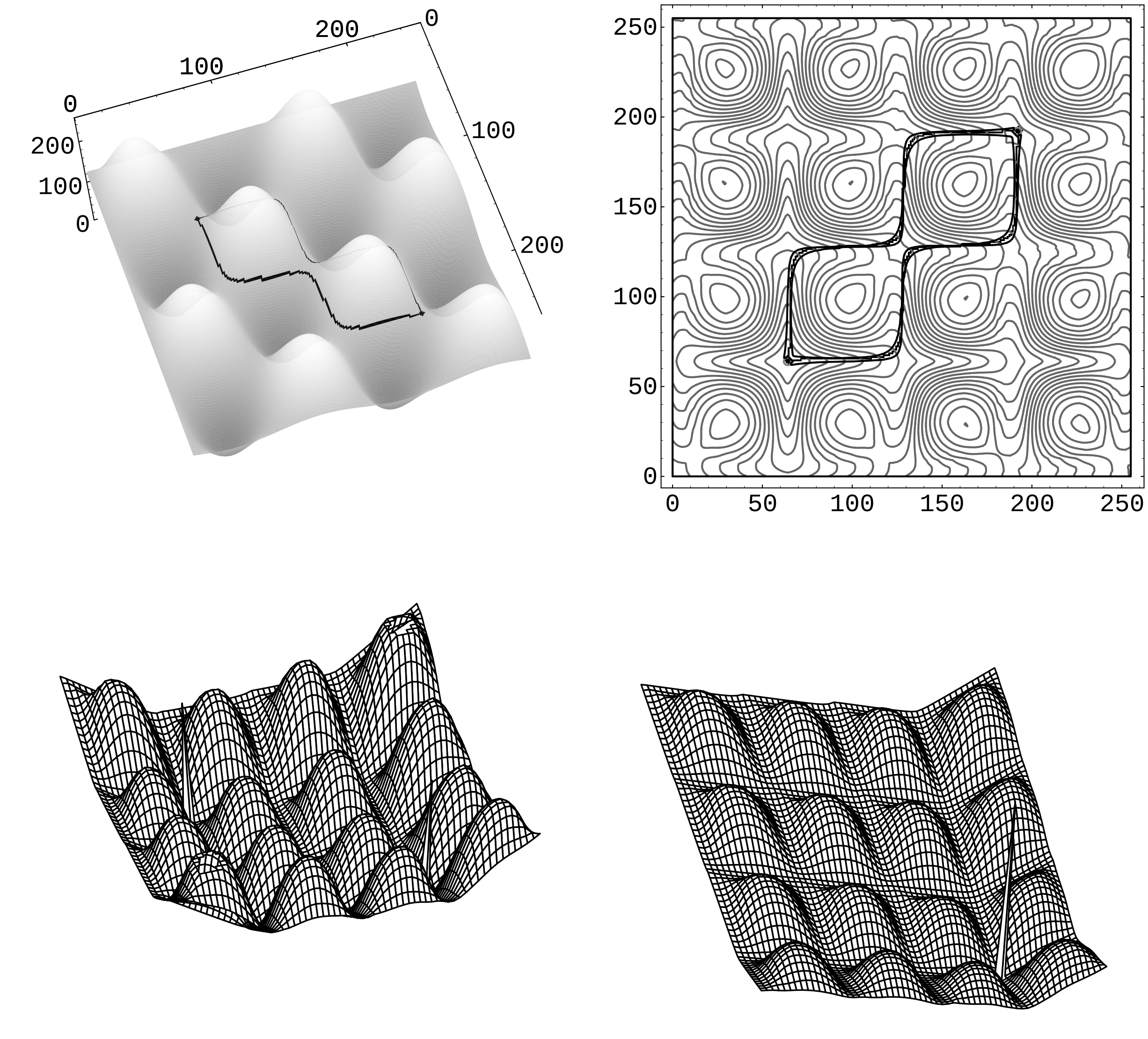

Figure 6.4: a. (Upper-Left) An "egg–box" surface with two minimal geodesics due to symmetry. b. (Upper-Right) Equal geodesic distance contours from the source point, and the minimal level set as black curve. c. (Lower-Left) Sum of the two geodesic distance maps. d. (Lower-Right) The geodesic distance map from the destination point (192,192).