Watersheds, waterfalls, on edge or node weighted graphs

Fernand Meyer
Centre de Morphologie Mathématique
Mines-ParisTech

2012 February 29
The watershed transform is one of the major image segmentation tools [4], used in the community of mathematical morphology and beyond. If the watershed is a successful concept, there is another side of the coin: a number of definitions and algorithms coexist, claiming to construct a watershed line or catchment basins, although they obviously are not equivalent. We have presented how the idea was conceptualized and implemented as algorithms or hardware solutions in a brief note:” The watershed concept and its use in segmentation: a brief history” (arXiv:1202.0216v1), which contains an extensive bibliography. See also [25] for an extensive review on the watershed concepts and construction modes.
The present work studies the topography of a relief defined on a node or edge weighted graph with morphological tools. The catchment basin of a minimum is defined as the sets of points it is possible to reach by a non descending path starting from this minimum. The watershed zone, i.e. the points linked with two distinct minima through a non ascending path, is more and more reduced as one considers steeper paths.
Reminders on weighted graphs
Distances on node or edge weighted graphs
Adjunctions between nodes and edges
The flooding adjunction.
Invariants of the flooding and closing adjunction
Flooding graph
Paths of steepest descent and k-steep graphs
The scissor operator, minimum spanning forests and watershed partitions
Lexicographic distances and SKIZ
The waterfall hierarchy.
Emergence and role of the minimum spanning tree
Discussion and conclusion
Reminder on adjunctions

The following section contains a reminder on adjunctions linking erosions and dilations by pairs and from which openings and closings are derived [26],[12],[11]
Preamble on adjunctions

Let $\mathcal{T}$ a complete totally ordered lattice, and $\mathcal{D}, \mathcal{E}$ be arbitrary sets
$O$: the smallest element and $\Omega$: the largest element of $\mathcal{T}$
$\text{Fun}(\mathcal{D}, \mathcal{T})$: the image defined on the support $\mathcal{D}$ with value in $\mathcal{T}$
$\text{Fun}(\mathcal{E}, \mathcal{T})$: the image defined on the support $\mathcal{E}$ with value in $\mathcal{T}$

Let $f$ be a function of $\text{Fun}(\mathcal{D}, \mathcal{T})$ and $g$ be a function of $\text{Fun}(\mathcal{E}, \mathcal{T})$
$\alpha: \text{Fun}(\mathcal{D}, \mathcal{T}) \rightarrow \text{Fun}(\mathcal{E}, \mathcal{T})$ and
$\beta: \text{Fun}(\mathcal{E}, \mathcal{T}) \rightarrow \text{Fun}(\mathcal{D}, \mathcal{T})$ be two operators

**Definition**

$\alpha$ and $\beta$ form an adjunction if and only if:
for any $f$ in $\text{Fun}(\mathcal{D}, \mathcal{T})$ and $g$ in $\text{Fun}(\mathcal{E}, \mathcal{T})$:
$\alpha f < g \iff f < \beta g$
Erosions and dilations

Theorem

If \((\alpha, \beta)\) form an adjunction, then \(\alpha\) is a dilation (it commutes with the supremum of functions in \(\text{Fun}(\mathcal{D}, \mathcal{T})\)) and \(\beta\) is an erosion (it commutes with the infimum of functions in \(\text{Fun}(\mathcal{E}, \mathcal{T})\)).

Proof: Let \(f\) be a function of \(\text{Fun}(\mathcal{D}, \mathcal{T})\) and \((g_i)\) functions of \(\text{Fun}(\mathcal{E}, \mathcal{T})\). Then

\[ f < \beta \bigwedge_i g_i \iff \alpha f < \bigwedge_i g_i \iff \forall i: \alpha f < g_i \iff \forall i: f < \beta g_i \iff f < \bigwedge_i \beta g_i \]

Reading these equivalences from left to right and replacing \(f\) by \(\beta \bigwedge_i g_i\) implies that \(\beta \bigwedge_i g_i < \bigwedge_i \beta g_i\).

Reading these equivalences from right to left and replacing \(f\) by \(\bigwedge_i \beta g_i\) implies that \(\bigwedge_i \beta g_i < \beta \bigwedge_i g_i\).

establishing that \(\beta \bigwedge_i g_i = \bigwedge_i \beta g_i\).
Similarly we show that $\delta$ is a dilation, i.e. commutes with the union.

**Remark:** Calling $O$ a constant function equal to $O$, we have for any $f : O < \beta f$ implying $\alpha O < f$. This relation is true for all $f$, indicating that $\alpha O = O$.

Similarly, we show that $\beta \Omega = \Omega$, where $\Omega$ is the constant function equal to $\Omega$. 
Lemma

\( \alpha \) and \( \beta \) are increasing operators.

Proof.

If \( f < g \) then \( \beta(f \wedge g) = \beta(f) = \beta(f) \wedge \beta(g) < \beta(g) \)
From one operator to the other

From the relation $\alpha f < g \Leftrightarrow f < \beta g$ one derives expressions of one operator in terms of the other one.
If $\alpha$ is known, $\beta$ may be expressed as $\beta g = \bigvee \{ f \mid \alpha f < g \}$.
Inversely if $\beta$ is known, then $\alpha f = \bigwedge \{ g \mid f < \beta g \}$.
Opening and closing

Lemma

The operator $\beta \alpha$ is a closing: increasing, extensive and idempotent. Similarly the operator $\alpha \beta$ is an opening: increasing, anti-extensive and idempotent.

Proof: Openings and closings, obtained by composition of increasing operators, are increasing.
By adjunction we obtain $\alpha f < \alpha f \Rightarrow f < \beta \alpha f$ showing that the closing is extensive and $\alpha \beta f < f \Leftarrow \beta f < \beta f$ showing that the opening is anti-extensive.
In particular, applying an opening to $\alpha f$ yields $\alpha f > \alpha \beta \alpha f$.
On the other hand, $\alpha$ being increasing, and the closing extensive: $f < \beta \alpha f \Rightarrow \alpha f < \alpha \beta \alpha f$ showing that $\alpha f = \alpha \beta \alpha f$.
Applying $\beta$ on both sides yields $\beta \alpha f = \beta \alpha \beta \alpha f$.  

The family of invariants of an opening or a closing

We call $\text{Inv}(\gamma)$ and $\text{Inv}(\varphi)$ the family of invariants of $\gamma$ and $\varphi$. 
The family of invariants of an opening

Lemma

The family of invariants of an opening is closed by union. And the family of closings is closed by intersection.

Proof: Let us prove it for openings; the result for closings being obtained by duality. Suppose that $g_1$ and $g_2$ are invariant by the opening $\gamma$: $\gamma(g_1) = g_1$ and $\gamma(g_2) = g_2$.

Then $\gamma(g_1 \lor g_2) \leq g_1 \lor g_2 = \gamma(g_1) \lor \gamma(g_2)$ by antiextensivity

And $\gamma(g_1 \lor g_2) \geq \gamma(g_1) \lor \gamma(g_2)$ as $\gamma$ is increasing.
Invariants of an opening or a closing

- $\alpha f \in \text{Inv}(\gamma)$ as $\alpha f = \alpha \beta \alpha f = (\alpha \beta)\alpha f$
- $\beta f \in \text{Inv}(\varphi)$ as $\beta f = \beta \alpha \beta f = (\beta \alpha)\beta f$
* The erosion has no inverse as \( \alpha \beta g \leq g \).

* The opening is the pseudo inverse of the erosion: \( \alpha \beta g \) is the smallest set having the same erosion as \( g \).

**Proof:** Suppose that \( f \) verifies \( \beta f = \beta g \) which implies \( \alpha \beta f = \alpha \beta g \). If \( f \leq \alpha \beta g \), we have \( \alpha \beta f \leq f \leq \alpha \beta g \) showing that \( f = \alpha \beta g \).

* if \( g \in \text{Inv}(\gamma) \), we have \( \gamma g = \alpha \beta g = g \): on \( \text{Inv}(\gamma) \), \( \alpha = \beta^{-1} \).
* The dilation has no inverse as $\beta \alpha g \geq g$.

* The closing is the pseudo inverse of the dilation: $\beta \alpha g$ is the largest set having the same dilation as $g$.

* If $g \in \text{Inv}(\varphi)$, we have $\varphi g = \beta \alpha g = g$: on $\text{Inv}(\varphi)$, $\beta = \alpha^{-1}$
Reminder on graphs
A \textit{non oriented graph} $G = [N, E] : N =$ vertices or nodes ; $E =$ edges : an edge $u \in E =$ pair of vertices (see [1],[10])

A \textit{chain} of length $n$ is a sequence of $n$ edges $L = \{u_1, u_2, \ldots, u_n\}$, such that each edge $u_i$ of the sequence ($2 \leq i \leq n - 1$) shares one extremity with the edge $u_{i-1}$ ($u_{i-1} \neq u_i$), and the other extremity with $u_{i+1}$ ($u_{i+1} \neq u_i$).

A \textit{path} between two nodes $x$ and $y$ is a sequence of nodes ($n_1 = x, n_2, \ldots, n_k = y$) such that two successive nodes $n_i$ and $n_{i+1}$ are linked by an edge.

A \textit{cycle} is a chain or a path whose extremities coincide.

A \textit{cocycle} is the set of all edges with one extremity in a subset $Y$ and the other in the complementary set $\overline{Y}$. 
The subgraph spanning a set $A \subset N : G_A = [A, E_A]$, where $E_A$ are the edges linking two nodes of $A$.

The partial graph associated to the edges $E' \subset E$ is $G' = [N, E']$

For **contracting** an edge $(i, j)$ in a graph $G$, one suppresses this edge $u$ and its two extremities are merged into a unique node $k$. All edges incident to $i$ or to $j$ become edges incident to the new node $k$.

**Contraction** :

If $H$ is a subgraph of $G$, The operator $\kappa : (G, H) \rightarrow G'$ contracts all edges of $H$ in $G$
A connected graph is a graph where each pair of nodes is connected by a path.

A tree is a connected graph without cycle.

A spanning tree is a tree containing all nodes.

A forest is a collection of trees.

Labelling a graph = extracting the maximal connected subgraphs.
In a graph $[N, E]$, $N$ represents the nodes, $E$ represents the edges and $\mathcal{T}$ is the set of weights of edges or nodes. ($O$ being the lowest weight and $\Omega$ the largest)

Edges and nodes may be weighted: $e_{ij}$ is the weight of the edge $(i,j)$ and $n_i$ the weight of the node $i$.

The following weight distributions are possible:

$(-, \Diamond)$: no weights

$(-, n)$: weights on the nodes

$(e, \Diamond)$: weights on the edges

$(e, n)$: weights on edges and nodes

Various types of graphs are met in image processing.
Which types of graphs?

In "pixel graphs" the nodes are the pixels and the edges connect neighboring pixels. The weights of the pixels are their value and the weights of edges may be for instance a gradient value computed between their extremities.
Which types of graphs?

We will work with "neighborhood graphs" where the nodes are the catchment basins and the edges connect neighboring bassins. The edges are weighted by a dissimilarity measure between adjacent catchment basins; the simplest being the altitude of the path-point between two basins.
The gabriel graph

Gabriel and Delaunay graphs permit to define neighborhood relations between the sets and nodes of a population.

Two nodes $x$ and $y$ are linked by an edge if there exists no other node in the disk of diameter $(xy)$. Node weights will express features of the nodes, edge weights relationships between adjacent nodes.
Gabriel graph modeling a lymphnode

The nuclei of a lymphnode are segmented and their Gabriel graph constructed.

- 1.1 : histological section.
- 1.2 : marking the nuclei.
- 2.1 : catchment basins of the nuclei on the original image.
- 2.2 : node weighted graph: the node weights represent the nucleus type.
Minimum spanning trees and forests in an edge weighted graph

A minimum spanning tree is a spanning tree for which the sum of the edges is minimal.

A spanning forest is a collection of trees spanning all nodes.

In a minimum spanning forest, the sum of the edges is minimal. The minimum minimorum spanning forest (MSF) has only nodes. With an additional constraint, one gets particular forests:

- a MSF with a fixed number of trees, useful in hierarchical segmentation [18]
- a MSF where each tree is rooted in predefined nodes, useful in marker based segmentation.
A subgraph $G'$ of an edge weighted graph $G$ is a flat zone, if any two nodes of $G'$ are connected by a chain of uniform altitude.

A subgraph $G'$ of a graph $G$ is a regional minimum if $G'$ is a flat zone and all edges in its cocycle have a higher altitude.

$\mu_e : G \to \mu_e G :$ extracts all regional minima of the graph $G$
A subgraph $G'$ of a node weighted graph $G$ is a flat zone, if any two nodes of $G'$ are connected by a path where all nodes have the same altitude.

A subgraph $G'$ of a graph $G$ is a regional minimum if $G'$ is a flat zone and all neighboring nodes have a higher altitude

$$\mu_n : G \rightarrow \mu_n G : \text{extracts all regional minima of the graph } G$$
**Contracting or expanding graphs**

**Contraction:**
For *contracting* an edge \((i, j)\) in a graph \(G\), one suppresses this edge \(u\) and its two extremities are merged into a unique node \(k\). All edges incident to \(i\) or to \(j\) become edges incident to the new node \(k\) and keep their weights.

If \(H\) is a subgraph of \(G\), the operator \(\kappa : (G, H) \rightarrow G'\) contracts all edges of \(H\) in \(G\).

**Expansion:**
The operator \(\cdot : (−, n) \rightarrow \cdot\) \(G\) creates for each isolated regional minimum \(i\) a dummy node with the same weight, linked by an edge with \(i\).
Extracting partial graphs

The following operators suppress a subset of the edges in a graph:
\( \chi : (\neg, \diamond) \rightarrow \chi G \) keeps for each node only one adjacent edges.
\( \downarrow : (e, \diamond) \rightarrow \downarrow G \) keeps for each node only its lowest adjacent edges.
\( \downarrow : (\neg, n) \rightarrow \downarrow G \) keeps only the edges linking a node with its lowest adjacent nodes.

These operators may be concatenated:
\( \chi \downarrow G \) keeps for each node only one lowest adjacent edge.
\( \chi \downarrow G \) keeps for each node only one edge linking it with its lowest adjacent nodes.
Distances on a graph

Case of edge weighed graphs
Constructing distances on an edge weighted graph.

Distances on an edge weighted graph have chains as support:

1) Definition of the weight of a chain, as a measure derived from the edge weights of the chain elements (example: sum, maximum, etc.)

2) Comparison of two chains by their weight. The chain with the smallest weight is called the shortest.

The distance $d(x, y)$ between two nodes $x$ and $y$ of a graph is $\infty$ if there is no chain linking these two nodes and equal to the weight of the shortest chain if such a chain exists.

Given three nodes $(x, y, z)$ the concatenation of the shortest chain $\pi_{xy}$ between $x$ and $y$ and the shortest chain $\pi_{yz}$ between $y$ and $z$ is a chain $\pi_{xz}$ between $x$ and $z$, whose weight is smaller or equal to the weight of the shortest chain between $x$ and $z$. To each distance corresponds a particular triangular inequality: $d(x, z) \leq weight(\pi_{xy} \triangleright \pi_{yz})$ where $\pi_{xy} \triangleright \pi_{yz}$ represents the concatenation of both chains.
Distance on an edge weighted graph based on the length of the shortest chain

**Length of a chain:** The length of a chain between two nodes $x$ and $y$ is defined as the sum of the weights of its edges.

**Distance:** The distance $d(x, y)$ between two nodes $x$ and $y$ is the minimal length of all chains between $x$ and $y$. If there is no chain between them, the distance is equal to $\infty$.

**Triangular inequality:** For $(x, y, z)$: $d(x, z) \leq d(x, y) + d(y, z)$
The weights are assigned to the edges, and represent their altitudes. **Altitude of a chain:** The altitude of a chain is equal to the highest weight of the edges along the chain. **Flooding distance between two nodes:** The flooding distance \( \text{fldist}(x, y) \) between nodes \( x \) and \( y \) is equal to the minimal altitude of all chains between \( x \) and \( y \). During a flooding process, in which a source is placed at location \( x \), the flood would proceed along this chain of minimal highest altitude to reach the pixel \( y \). If there is no chain between them, the level distance is equal to \( \infty \). **Triangular inequality:** For \((x, y, z)\): \( d(x, z) \leq d(x, y) \lor d(y, z) \): ultrametric inequality
The flooding distance is an ultrametric distance

An ultrametric distance verifies
* reflexivity: \( d(x, x) = 0 \)
* symmetry: \( d(x, y) = d(y, x) \)
* ultrametric inequality: for all \( x, y, z \): \( d(x, y) \leq \max\{d(x, z), d(z, y)\} \):
  the lowest lake containing both \( x \) and \( y \) is lower or equal than the lowest lake containing \( x \), \( y \) and \( z \).
Distances on a graph: sum and maximum of the edge weights

The shortest chain (sum of weights of the edges) between $x$ and $y$ is a red line and has a length of 4.
The lowest chain (maximal weight of the edges) between $x$ and $y$ is a red line and a maximal weight of 2. A flooding between $x$ and $y$ would follow this chain.
For this particular example, the shortest and lowest paths are identical.
The lexicographic distance or the cumulative effort for passing the highest edges

**Toughness of a chain:** We call toughness of a chain the decreasing list of altitudes of the highest edges met along this chain.

In fig. A, along the bold chain between x and y, the highest edge rises at the altitude 3. After crossing it, the highest edge on the remaining chain rises at 2. After crossing it, one is at destination. Hence the toughness of the chain from x to y is [3, 2]. Distances are compared in a lexicographic order. The shortest chain is the red chain in fig. B, it has two edges with weight 2 and its toughness is [2, 2].
The lexicographic distance: a more formal definition

The lexicographic length $\Lambda(A)$ of a chain $A = e_{12}e_{23}...e_{n-1n}$ is constructed as follows. Following the chain from the origin towards its end, one records the highest valuation of the chain, let it be $\lambda_1$, then again the highest valuation $\lambda_2$ on the remaining part of the chain and so on until the end is reached. One gets like that a series of decreasing values: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

The lexicographic distance between two nodes $x$ and $y$ will be equal to the shortest lexicographic length of all chains between these two nodes and is written $\text{lexdist}(x, y)$.

Remark: The lexicographic length of a never increasing path (NAP), is simply equal to the series of weights of its edges, as it is the highest edge along the lowest path between $x$ and $y$. Later we introduce shortest path algorithms whose geodesics are NAPs.
The lexicographic distance of depth $k$ between two nodes $x$ and $y$ is written $\text{lexdist}_k(x, y)$ and obtained by retaining only the $k$ first edges in $\text{lexdist}(x, y)$.

**Remark:** $\text{lexdist}_1(x, y)$ is the same as the ultrametric flooding distance as it is the highest edge on the lowest path between $x$ and $y$. 
Distances on a graph

Case of node weighed graphs
Distances on a node weighted graph have paths as support:
1) Definition of the "length" of a path, as a measure derived from the node weights of the path elements (example: sum, maximum, etc.)
2) Comparison of two paths by their length. The path with the smallest length is called the shortest.
The distance \(d(x, y)\) between two nodes \(x\) and \(y\) of a graph is \(\infty\) if there is no path linking these two nodes and equal to the length of the shortest path if such a path exists.
Given three nodes \((x, y, z)\) the concatenation of the shortest path \(\pi_{xy}\) between \(x\) and \(y\) and the shortest path \(\pi_{yz}\) between \(y\) and \(z\) is a path \(\pi_{xz}\) between \(x\) and \(z\), whose length is smaller or equal to the length of the shortest path between \(x\) and \(z\). To each distance corresponds a particular triangular inequality: \(d(x, z) \leq \text{length}(\pi_{xy} \circlearrowright \pi_{yz})\) where \(\pi_{xy} \circlearrowright \pi_{yz}\) represents the concatenation of both paths.
The weights are assigned to the nodes, and represent their altitudes. **Altitude of a path:** The altitude of a path is equal to the highest weight of the nodes along the path. **Flooding distance between two nodes:** The flooding distance $\text{fldist}(x, y)$ between nodes $x$ and $y$ is equal to the minimal altitude of all paths between $x$ and $y$. 
Distance on a node weighted graph based on the cost for travelling along the cheapest path

The weights are assigned to the nodes and not to the edges. Each node may be considered as a town where a toll has to be paid.

**Cost of a path:** The cost of a path is equal to the sum of the tolls to be paid in all towns encountered along the path (including or not one or both ends).

**Cost between two nodes:** The cost for reaching node $y$ from node $x$ is equal to the minimal cost of all paths between $x$ and $y$. We write $\text{tolldist}(x, y)$. If there is no path between them, the cost is equal to $\infty$. 

Illustration of the cheapest path

In this figure, the cheapest chain between $x$ and $y$ is in red and the total toll to pay is $1 + 1 + 2 + 2 = 6$
Two adjunctions between edges and nodes on weighted graphs
Two adjunctions between edges and nodes

**Definition**

We define two operators between edges and nodes:
- an erosion $[\varepsilon_{en} n]_{ij} = n_i \land n_j$ and its adjunct dilation $[\delta_{ne} e]_i = \bigvee e_{ik}$ (\(k\) neighbors of \(i\))
- a dilation $[\delta_{en} n]_{ij} = n_i \lor n_j$ and its adjunct erosion $[\varepsilon_{ne} e]_i = \bigwedge e_{ik}$ (\(k\) neighbors of \(i\))

**Lemma**

The operators we defined are pairwise adjunct or dual operators:
- $\varepsilon_{ne}$ and $\delta_{en}$ are adjunct operators
- $\varepsilon_{en}$ and $\delta_{ne}$ are adjunct operators
- $\varepsilon_{ne}$ and $\delta_{ne}$ are dual operators
- $\varepsilon_{en}$ and $\delta_{en}$ are dual operators
Let us prove that $\delta_{en}$ and $\varepsilon_{ne}$ are adjunct operators
If $G = [e, n]$ and $\overline{G} = [\overline{e}, \overline{n}]$ are two graphs with the same nodes and edges, but with different valuations on the edges and the nodes then
$\delta_{en} n \leq \overline{e} \iff \forall i, j : n_i \lor n_j \leq \overline{e}_{ij} \iff \forall i, j : n_i \leq \overline{e}_{ij} \iff \forall i, j : n_i \leq \bigwedge \overline{e}_{ij}$

$(j$ neighbors of $i)$
which establishes that $\delta_{en}$ and $\varepsilon_{en}$ are adjunct operators.

Let us prove that $\varepsilon_{en}$ and $\delta_{en}$ are dual operators
$[\varepsilon_{en}(-n)]_{ij} = -n_i \land -n_j = -(n_i \lor n_j) = -[\delta_{en} n]_{ij}$
Hence $\delta_{en} n = -[\varepsilon_{en}(-n)]$
The flooding adjunction

The dilation $[\delta_{en}n]_{ij} = n_i \lor n_j$ and its adjunct erosion
$[\varepsilon_{ne}e]_i = \bigwedge e_{ik}$ have a particular meaning in terms of flooding.

If $n_i$ and $n_j$ represent the altitudes of the nodes $i$ and $j$, the lowest flood covering $i$ and $j$ has the altitude $[\delta_{en}n]_{ij} = n_i \lor n_j$

If $i$ represents a catchment basin, $e_{ik}$ the altitude of the pass points with the neighboring basin $k$, then the highest level of flooding without overflow through an adjacent edge is $[\varepsilon_{ne}e]_i = \bigwedge e_{ik}$. (k neighbors of $i$)
The waterfall flooding is completely specified if one knows the level of flood in each catchment basin.

- The waterfall flooding fills each catchment basin up to its lowest pass point.
- In terms of graphs: the flooding in a node is equal to the weight of its lowest adjacent edge.
- It is an erosion between node weights and edge weights:

\[
[\varepsilon_{ne}e]_i = \bigwedge e_{ik} \quad (k \text{ neighbors of } i)
\]
The waterfall flooding is completely specified if one knows the level of flood in each catchment basin.

- The waterfall flooding fills each catchment basin up to its lowest pass point.
- In terms of graphs: the flooding in a node is equal to the weight of its lowest adjacent edge.
- It is an erosion between node weights and edge weights:

\[
[\varepsilon_{ne} e]_i = \bigwedge e_{ik} \quad (k \text{ neighbors of } i)
\]
Erosion between edges and nodes

erosion from edges to nodes
Dilation between nodes and edges

If $n_i$ and $n_j$ represent the altitudes of the nodes $i$ and $j$, the lowest flood covering $i$ and $j$ has the altitude $[\delta_{en}]_{ij} = n_i \lor n_j$
Erosion between edges and edges / between nodes and nodes

By concatenation of operators between edges and nodes we obtain:

- an adjunction between nodes and nodes: \((\varepsilon_{ne}\varepsilon_{en}, \delta_{ne}\delta_{en}) = (\varepsilon_n, \delta_n)\)
- an adjunction between edges and edges: \((\varepsilon_{en}\varepsilon_{ne}, \delta_{en}\delta_{ne}) = (\varepsilon_e, \delta_e)\)
Opening and closing
As $\varepsilon_{ne}$ and $\delta_{en}$ are adjunct operators, the operator $\varphi_n = \varepsilon_{ne}\delta_{en}$ is a closing on $n$ and $\gamma_e = \delta_{en}\varepsilon_{ne}$ is an opening on $e$.

Similarly the operator $\varphi_e = \varepsilon_{en}\delta_{ne}$ is a closing on $e$ and $\gamma_n = \delta_{ne}\varepsilon_{en}$ is an opening on $n$.

In the sequel, the flooding adjunction will play a key role, in particular the associated opening $\gamma_e$ and closing $\varphi_n$. 
The opening $\gamma_e$
Illustration in one dimension of the opening $\gamma_e$

$\gamma_e = \delta_{en}\epsilon_{ne}$: on the left, the result of the erosion, filling each basin to its lowest pass point. On the right the subsequent dilation. Some pass points have a reduced altitude: the amount of reduction is indicated in red. These pass points are those which are not the lowest pass points of a catchment basin.
Illustration on an edge weighted tree of the opening $\gamma_e$

\[\gamma_e = \delta_{en} \varepsilon_{ne}: \text{From left to right: 1) an edge weighted graph, in the centre, 2) the result of the erosion } \varepsilon_{ne}, \text{ 3) the subsequent dilation produces an opening. The edges in red are those whose weight has been reduced by the opening. These edges are not the lowest edges of one or their extremities.}\]
The invariants of the opening $\gamma_e$

Two possibilities exist for an edge $(i, j)$ with a weight $\lambda$ :

- the edge $(i, j)$ has lower neighboring edges at each extremity. Hence $\varepsilon_{ne}(i) < \lambda$ and $\varepsilon_{ne}(j) < \lambda$ ; hence $\delta_{en}\varepsilon_{ne}(i, j) = \varepsilon_{en}(i) \lor \varepsilon_{en}(j) < \lambda$

- the edge $(i, j)$ is the lowest edge of the extremity $i$. Then $\varepsilon_{ne}(i) = \lambda$ and $\varepsilon_{ne}(j) \leq \lambda$ ; hence $\delta_{en}\varepsilon_{ne}(i, j) = \varepsilon_{en}(i) \lor \varepsilon_{en}(j) = \lambda$

**Conclusion:** the edges invariant by the opening $\gamma_e$ are the edges which are the lowest edge of one of their extremities. All edges with lower adjacent edges at their extremities have their weight lowered by the opening $\gamma_e$
Extracting from an edge weighted graph a partial graph invariant by the opening $\gamma_e$

The relation $\varepsilon_{ne} = \varepsilon_{ne}(\delta_{en}\varepsilon_{ne})$ shows that all edges which are not invariant by the opening $\gamma_e = \delta_{en}\varepsilon_{ne}$ play no role in the erosion. As a matter of fact, if the opening lowers the valuation of an edge $(i,j)$ and if the subsequent erosion $\varepsilon_{ne}$ is not modified, it means that the presence of this edge with its weight plays no role in the erosion $\varepsilon_{ne}$. 
The waterfall graph, partial graph of the lowest adjacent edges

Suppressing in an arbitrary graph $g$ all edges which are not invariant by the opening $\gamma_e$ produces a graph $g'$, invariant for the opening $\gamma_e$ and called waterfall graph. The operator keeping for each node only its lowest adjacent edges is written $\downarrow : (e, \Diamond) \rightarrow \downarrow G$

Properties:

- $\downarrow G$ spans all the nodes (each node has at least one lowest neighboring edge; such edges are invariant by $\gamma_e$)
- And $\varepsilon_{ne}(G) = \varepsilon_{ne}(\downarrow G)$ ($\varepsilon_{ne}$ assigns to each node the weight of its lowest adjacent edge, which is the same in $G$ as in $\downarrow G$)
Illustration on a planar edge weighted graph of the opening $\gamma_e$. 

- **erosion**
- **opening**
- **waterfall graph = edges invariant by opening**
Erosion from edges to nodes on the graph \( \downarrow G \)

A: Initial graph \( G \)
B: Erosion \( \varepsilon_{ne} \) from the edges to the nodes on \( G \)
C: The graph \( \downarrow G \)
D: Erosion \( \varepsilon_{ne} \) from the edges to the nodes on \( \downarrow G \): \( \varepsilon_{ne}(G) = \varepsilon_{ne}(\downarrow G) \)
Erosion from edges to nodes on the graph $\downarrow G$

Another illustration of $\varepsilon_{ne}(G) = \varepsilon_{ne}(\downarrow G)$

erosion on the RAG
erosion on the waterfall graph
The invariants of the opening $\gamma_e$
The invariants of the opening $\gamma_e$

If $(e_1, \diamond)$ and $(e_2, \diamond)$ are two edge weight distributions which are invariant for $\gamma_e$, then $(e_1 \lor e_2, \diamond)$ also is invariant for $\gamma_e$.

An arbitrary graph may be transformed into a flooding graph:
* For an arbitrary edge weight distribution $(e, \diamond)$: $(\gamma_e e, \diamond) \in \text{Inv}(\gamma_e)$
* For an arbitrary node weight distribution $(-, n)$: $(\delta_{en} n, n) \in \text{Inv}(\gamma_e)$ as

$$\gamma_e \delta_{en} n = \delta_{en} \varepsilon_{ne} \delta_{en} n = \delta_{en} n$$
The invariants of the opening $\gamma_e$

For a connected graph $g = (e, \diamond) \in \text{Inv}(\gamma_e)$:

- any partial spanning graph belongs to $\text{Inv}(\gamma_e)$
- any subgraph belongs to $\text{Inv}(\gamma_e)$

In particular $\downarrow : G = (e, \diamond) \rightarrow \downarrow G$ containing for each node only its lowest adjacent edges and $\chi \downarrow G$ keeping only one lowest adjacent edge for each node, both belong to $\text{Inv}(\gamma_e)$
The regional minima of the opening $\gamma_e$

**Theorem**

If $G = (e, \diamond) \in \text{Inv}(\gamma_e)$ and $m = (e, \diamond)$ is the subgraph of its regional minima, then $\epsilon_{ne} m = (\epsilon_{ne} e)$ is the subgraph of the regional minima of the graph $\epsilon_{ne} G = (\epsilon_{ne} e)$

**Proof:** A regional minimum $m_k$ of the graph $G = (e, \diamond) \in \text{Inv}(\gamma_e)$ is a plateau of edges with altitude $\lambda$, with all external edges having a weight $> \lambda$. If a node $i$ belongs to this regional minimum, its adjacent edges have a weight $\geq \lambda$ but it has at least one neighboring edge with weight $\lambda$: hence $\epsilon_{ne}(i) = \lambda$. Consider now an edge $(s, t)$ outside the regional minimum, with the node $s$ inside and the node $t$ outside the minimum. Then $e_{st} > \lambda$. As $G = (e, \diamond) \in \text{Inv}(\gamma_e)$, the edge $(s, t)$ is then one of the lowest edges of the nodes $t$ : thus $\epsilon_{ne}(t) = e_{st} > \lambda$. This shows that the nodes spanned by the regional minimum $m_k$ form a regional minimum of the graph $\epsilon_{ne} G = (\epsilon_{ne} e)$
Inverse of $\varepsilon_{ne}$ on the invariants of the opening $\gamma_e$

On $\text{Inv}(\gamma_e)$: $\delta_{en}\varepsilon_{ne} = \text{Identity}$ showing that on $\text{Inv}(\gamma_e) := \delta_{en} = \varepsilon_{ne}^{-1}$
The closing $\phi_n$
The closing $\varphi_n$ is obtained by a dilation $\delta_{en}$ of the node weights followed by an erosion $\varepsilon_{ne}$. One remarks on the following figure that the node weights remain the same, except the isolated regional minima, which take the weight of their lowest neighboring node. We give the proof in the next slide.
The invariants of the closing $\varphi_n$

Two possibilities exist for a node $i$ with a weight $\lambda$:

- the node $i$ is an isolated regional minimum. Then $\delta_{en}$ assigns to all edges adjacent to $i$ a weight bigger than $\lambda$. The subsequent erosion $\varepsilon_{ne}$ assigns to $i$ the smallest of these weights.

- the node $i$ has a neighbor $j$ with a weight $\mu \leq \lambda$. Then $\delta_{en}$ assigns to the edge $(i, j)$ the weight $\lambda$; whatever the weight of the other adjacent edges. The subsequent erosion $\varepsilon_{ne}$ assigns to $i$ the smallest of these weights, that is $\lambda$.

The closing $\varphi_n$ replaces each isolated node constituting a regional minimum by its lowest neighboring node and leaves all other nodes unchanged.
The invariants of the closing $\phi_n$
The invariants of the closing $\varphi_n$

A node weighted graph is invariant for the closing $\varphi_n$ iff it does not contain isolated regional minima.

For an arbitrary node weight distribution $g = (\neg, n) : \neg g$ creates for each isolated regional minimum $i$ a dummy node with the same weight linked by an edge with $i$. Hence $\neg g \in \text{Inv}(\varphi_n)$
The invariants of the closing $\varphi_n$

If $(-, n_1)$ and $(-, n_1)$ are two node weight distributions which are invariant for $\varphi_n$, then $(-, n_1 \wedge n_2)$ also is invariant for $\varphi_n$.
For an arbitrary node weight distribution $(-, n) : (-, \varphi_n n) \in \text{Inv}(\varphi_n)$
For an arbitrary edge weight distribution $(e, \diamond) : (e, \varepsilon_{ne} e) \in \text{Inv}(\varphi_n)$ as
$\varphi_n \varepsilon_{ne} e = \varepsilon_{ne} \delta_{en} \varepsilon_{ne} e = \varepsilon_{ne} e$
The invariants of the closing $\varphi_n$

For a graph $g = (e, \diamond) \in \text{Inv}(\varphi_n)$: any partial or subgraph which does not create an isolated regional minimum also belongs to $\text{Inv}(\varphi_n)$

Each node $i$ belonging to a regional minimum of $g$ is linked with at least another node $j$ in this minimum through an edge with the same weight. This edge is one of the lowest adjacent edges of $i$ and links $i$ with one of its lowest neighboring nodes.

For this reason after pruning, the graphs $\downarrow g$ and $\downarrow g$ still belong to $\text{Inv}(\varphi_n)$. 
The regional minima of the closing $\varphi_n$

**Theorem**

If $G = (-, n) \in \text{Inv}(\varphi_n)$ and $m = (-, n)$ is the subgraph of its regional minima, then $\delta_{en} m = (\delta_{en} n, \diamond)$ is the subgraph of the regional minima of the graph $\delta_{en} G = (\delta_{en} n, \diamond)$

**Proof:** A regional minimum $m_i$ of a graph $G = (-, n) \in \text{Inv}(\varphi_n)$ is a plateau of pixels with altitude $\lambda$, containing at least two nodes (there are no isolated regional minima in $\text{Inv}(\varphi_n)$). All internal edges of the plateau get the valuation $\lambda$ by $\delta_{en} n$. If an edge $(i, j)$ has the extremity $i$ in the minimum and the extremity $j$ outside, then $\delta_{en} n(i, j) > \lambda$. Hence, for the graph $(\delta_{en} n, \diamond)$, the edges spanning the nodes of $m_i$ form a regional minimum.
Inverse of $\delta_{en}$ on the invariants of the closing $\varphi_n$

On $\text{Inv}(\varphi_n) : \varepsilon_{ne} \delta_{en} = \text{Identity}$ showing that on $\text{Inv}(\varphi_n) : \varepsilon_{ne} = \delta_{en}^{-1}$
The flooding graphs
The flooding graph

**Definition**

An edge and node weighted spanning graph $G = [N, E]$ is a flooding graph iff its weight distribution $(n, e)$ verify the relations:

- $\delta_{en}n = e$
- $\epsilon_{ne}e = n$

**Corollary**

For a flooding graph weight distribution $(n, e)$:

- $n \in \text{Inv}(\varphi_n)$
- $e \in \text{Inv}(\gamma_e)$

**Proof.**

$e = \delta_{en}n = \delta_{en}\epsilon_{ne}e = \gamma_e e$ and $n = \epsilon_{ne}e = \epsilon_{ne}\delta_{en}n = \varphi_n n$
Properties of the flooding graph

As $G$ is invariant by $\gamma_e$, all its edges are the lowest edge of one of their extremities.

As $G$ is invariant by $\varphi_n$, it has no isolated regional minimum.
In a flooding graph, $\varepsilon_{ne} e = n$, hence all edges adjacent to a node have weights which are higher or equal than this node and at least one of them has the same weight.

On the other hand, each node $i$ has at least one neighbor $j$ which is lower or equal (otherwise it would be an isolated regional minimum). The weight $e_{ij}$ verifies $e_{ij} = \delta_{en} n(i, j) = n_i \vee n_j = n_i$. This shows that the edges linking a node with lower or equal nodes have the same weight than this node.

In particular, the edges linking a node $i$ to its lowest neighboring nodes belong to the lowest adjacent edges of this node and have the same weight: $\downarrow G \subset \downarrow G$ and $\downarrow\downarrow G = \downarrow\downarrow G = \downarrow G$.
Constructing flooding graphs.

If $G = (e, \diamond) \in \lnv(\gamma_e)$ then $(e, \varepsilon_{ne} e)$ is a flooding graph (since $e = \gamma_e e = \delta_{en} \varepsilon_{ne} e = \delta_{en} n$)

If $G = (-, n) \in \lnv(\varphi_n)$ then $(\delta_{en} n, n)$ is a flooding graph (since $n = \varphi_n n = \varepsilon_{ne} \delta_{en} n = \varepsilon_{ne} e$)
Deriving flooding graphs from ordinary graphs

If $G = (e, \diamondsuit)$ is an arbitrary edge weighted graph, $\downarrow (e, \diamondsuit) \in \text{Inv}(\gamma_e)$, as the edges lowered by $\gamma_e$, have been suppressed, the other edges keeping their weights. Recall that $\varepsilon_{ne}e = \varepsilon_{ne} \downarrow e$. The derived flooding graph simply is $(\downarrow e, \varepsilon_{ne} \downarrow e)$.

If $G = (-, n)$ is an arbitrary node weighted graph, $(-, \rightarrow n) \in \text{Inv}(\varphi_n)$, as isolated regional minima, if any, have been duplicated. The derived flooding graph simply is $(\delta_{en} \rightarrow n, \rightarrow n)$. 
Suppressing edges in a flooding graph, but leaving at least one lower neighboring edge for each node (like that, no isolated regional minima are created) produces a partial graph which also is a flooding graph, with the same distribution of weights on the nodes and on the remaining edges.
Regional minima of flooding graphs
Regional minima of a flooding graph

We proved earlier these theorems:

- If $G = (e, \diamond) \in \text{Inv}(\gamma_e)$ and $m = (e, \diamond)$ is the subgraph of its regional minima, then $\varepsilon_{ne}m = (\varepsilon_{ne}e)$ is the subgraph of the regional minima of the graph $\varepsilon_{ne}G = (\varepsilon_{ne}e)$.

- If $G = (\varepsilon_{en}n, \varepsilon_{en}e) \in \text{Inv}(\varphi_n)$ and $m = (\varepsilon_{en}n, \varepsilon_{en}e)$ is the subgraph of its regional minima, then $\delta_{en}m = (\delta_{en}n, \diamond)$ is the subgraph of the regional minima of the graph $\delta_{en}G = (\delta_{en}n, \diamond)$.

As in a flooding graph $n = \varepsilon_{ne}e$ and $e = \delta_{en}n$ we derive:

**Theorem**

*If $G$ is a flooding graph with the weight distribution $(e, n)$, then the node weighted graph $(\varepsilon_{en}n, \varepsilon_{en}e)$ and the edge weighted graph $(e, \diamond)$ have the same regional minima subgraph*
The regional minima on the edge or node graph within the flooding graph

regional minima on the edges

regional minima on the nodes
Labeling the regional minima on the edge or node graph within the flooding graph

As the minima are identical on the nodes or the edges of a flooding graph, it is possible to assign the same labels to nodes or to edges.

labeling the regional minima on the nodes and/or the edges
Paths of steepest descent and catchment basins
The catchment basins of the minima

Lemma

From each node outside a regional minimum starts a never ascending path to a regional minimum in a flooding graph.

Proof: Any node $i$ outside a regional minimum has a lower neighboring node, if it does not belong to a plateau. Otherwise it belongs to a plateau, containing somewhere a node $j$ with a lower neighboring node outside, as the plateau is not a regional minimum. The plateau being connected, there exists a path of constant altitude in the plateau between $i$ and $j$. Following this path, it is possible, starting at node $i$ to reach the lower node $k$. This shows that for each node there exists a never ascending path to a lower neighboring node. Taking this new node as starting node, a still lower node may be reached. The process may be repeated until a regional minimum is reached.

Thanks to this lemma, it is possible to define the catchment basins of the minima.
The catchment basins of the minima

| Definition |
|------------|
| The catchment basin of a minimum \( m \) is the set of all nodes from which starts a never ascending path towards \( m \). |

As in a flooding graph, each node and its lower neighboring edges have the same weight, each node in such a never ascending path is followed by an edge with the same weight except the last one belonging to a regional minimum. For this reason the catchment basins based on the node weights or on the edge weights are identical.

**The watershed zones** are the nodes belonging to more than one catchment basin.

**The restricted catchment basins** are the nodes which belong to only one catchment basin: it is the difference between the catchment basin of a minimum \( m \) and the union of catchment basins of all other minima.

From a node in a restricted catchment basin, there exists a unique non ascending path towards a unique regional minimum.
M-flooding graphs

The catchment basins rely entirely on the non ascending paths of the graph reaching a regional minimum. The altitude of the regional minimum has no importance. If we consider only the end points of such a path, the fact that the minimum is an isolated node or not has no importance either. For this reason, we may relax the definition of the flooding graphs for which \( \delta_{en} n = e \) and \( \varepsilon_{ne} e = n \) for all edges and nodes.

Consider a flooding graph where all non regional minima nodes have a positive weight. Assigning to all minima a weight 0 does not invalidate the relation \( \delta_{en} n = e \). The relation \( \varepsilon_{ne} e = n \) also remains true, except for isolated regional minima. We call M-flooding graphs, the node and edge weighted graphs verifying \( \delta_{en} n = e \) everywhere and \( \varepsilon_{ne} e = n \) for all nodes which are not isolated regional minima.

In what follows we consider NAP which end with a node in a regional minimum. Whether this minimum is isolated or not has no importance.
Extending the restricted cathment basins and reducing the watershed zone: steep, steeper, steepest flooding graphs
Catchment basins and segmentation.

The watershed transform is mainly used for segmentation, with the aim to create a partition representing precisely the extension of each object. Large watershed zones are ambiguous as they separate restricted catchment basins without precise localisation of the contour separating them. For obtaining precise segmentations, it is important to reduce these zones and even suppress them completely if possible. Reducing the number of never ascending paths from each node to a regional minimum would help constraining the construction of the watershed partition. Ideally, if only one such path remains for each node outside the regional minima, the solution would be unique, the restricted catchment occupy the whole space and the watershed zones be empty.

In this section we show how to reduce the number of paths and keep only those which have some degree of steepness. Increasing the steepness reduces the number of paths.
Flooding tracks

In a flooding graph each edge is the lowest edge of one of its extremities and has the same weigh: we call such a pair made of a node and adjacent edge with the same weight **flooding pair**.

Two couples \((i, ij)\) and \((j, jk)\) are chained if the node in the second couple is an extremity of the edge of the first couple.

A **flooding track** is a list of chained couples of never increasing weight. The **lexicographic weight** of a flooding track is the list of never increasing weights of its pairs.

Two flooding tracks may then be compared by comparing their lexicographic weights using the **lexicographic order relation**.
Pruning the flooding graph to get steeper paths.

We define a pruning operator $\downarrow^k$ operating on a flooding graph $G$. The pruning $\downarrow^k$ considers each node outside the regional minima and suppresses its adjacent edges, if they are not the highest edge of a flooding track verifying:

- their lexicographic weight is minimal.
- their length is $k$. It may be shorter if its last couple belongs to a regional minimum.

After pruning, each node outside the regional minima is the origin of one or several $k$-steep flooding tracks (remark that the pruning only suppresses the highest edge of the track). If there are several of them, they have the same weights. We call them the $k$-steep adjacent paths to the node $i$. We say that the graph $\downarrow^k G$ has a $k$-steepness or is $k$-steep.
As the steepness degree increases, the pruning becomes more and more severe, producing a decreasing series of partial graphs, hence for $k > l$: $\downarrow^k G \subset \downarrow^l G$. Furthermore $\downarrow^k \downarrow^l G = \downarrow^l \downarrow^k G = \downarrow^{k \lor l} G$.

The pruning $\downarrow^1$ does nothing as each edge is the lowest edge of one of its extremities in any graph invariant by $\gamma_e$. On the contrary, $\downarrow^1 = \downarrow$ transforms an arbitrary graph into a graph invariant by $\gamma_e$.

The pruning $\downarrow^2$ keeps for each node $i$ the adjacent edges which are followed by a second couple of minimal weight. These edges are those linking $i$ with one of its lowest neighboring nodes.
A morphological characterization of k-steep graphs
Eroding k-steep graphs

Consider a k-steep graph $\downarrow^k G = G = (e, n)$, with $k \geq 2$. We define the erosion $\varepsilon G = (\varepsilon_e e, \varepsilon_n n)$, where $\varepsilon_e = \varepsilon_{en} \varepsilon_{ne}$ and $\varepsilon_n = \varepsilon_{ne} \varepsilon_{en}$. $\varepsilon^{(1)} G = \varepsilon G$ and $\varepsilon^{(m)} G = \varepsilon \varepsilon^{(m-1)} G$.

It does not change the catchment basins if we assign to the regional minima the weight 0, whereas all other nodes and edges have weights $> 0$ (like that as soon the erosion assigns the value 0 to a node or an edge, they remain equal to 0 for all subsequent erosions).
Eroding k-steep graphs

Consider a k-steep lexicographic track $\tau$ of a flooding graph $G$ made of a series of non increasing flooding pairs $(i_1, e_1), (i_2, e_2), \ldots, (i_k, e_k)$. Node and edge of each flooding pair have the same weights. As $\tau$ has a minimal lexicographic weight, each of its flooding pairs, except the first one constitutes one of the lowest adjacent flooding pair of the previous flooding pair. For this reason the erosion $\varepsilon G$ assigns to the edge $e_h$ the weight of the adjacent edge $e_{h+1}$ and to the node $i_h$ the weight of the adjacent node $i_{h+1}$. In other worksd successive erosions $(\varepsilon_e e, \varepsilon_n n)$ let glide the value of each pair upwards in the track. If this track is of length $k$, after $k - 1$ erosion, the weight of the ultimate pair will have reached the first one. If such a track is of length $l < k$, it ends with a pair in a regional minimum, with the value 0. Successive erosions $(\varepsilon_e e, \varepsilon_n n)$ also let glide the value of each pair upwards in the track and the last pair, with value 0 also moves upwards and reaches the pair $(i, ij)$ after $l - 1$ erosions. During the next erosions, the value of the pair $(i_1, e_1)$ remains stable and equal to 0.
Eroding k-steep graphs

Consider a k-steep flooding graph $G = \downarrow^k G$. During the $k - 1$ successive erosions, the values of the flooding pairs glide upwards along the k-steep lexicographic path. For the erosion $l < k$, the $(l + 1)th$ pair $(s, st)$ has reached the pair $(i, ij)$. Hence the edge $ij$ remains one of the lowest adjacent edge of the node $i$ and both share identical weights. But $(s, st)$ belongs to the flooding graph and verifies $n_s = e_{ne} e(s)$ and $n_s = e_{ne} e(s)$ and so does $(i, ij)$. This shows that the graph $\varepsilon(l) G$ still is a flooding graph.

**Theorem**

For an ordinary flooding graph $G$, $\downarrow^k G$ is a k-steep flooding graph and for $l < k$, $\varepsilon(l) \downarrow^k G$ still is a flooding graph.
Eroding flooding graphs

Consider a flooding graph $G = (e, n)$. The erosion $\varepsilon_n = \varepsilon_{ne}\varepsilon_{en}$ enlarges the regional minima. Hence if $G$ is invariant by $\varphi_n$ it is still the case for the graph $(e, \varepsilon_n n)$.

On the contrary, after the erosion $\varepsilon_e = \varepsilon_{en}\varepsilon_{ne}$, the graph $(\varepsilon_e e, n)$ is not necessarily invariant for $\gamma_e$. One has to prune the edges which are not the lowest edges of one of their extremities: $\downarrow \varepsilon_e e$.

Repeating the operator $\zeta G = (\downarrow \varepsilon_e e, \varepsilon_n n)$ produces a decreasing series of partial graphs $\zeta^{(n)} G = \zeta \zeta^{(n-1)} G$ we will now characterize.
Two equivalent modes of pruning.

**Theorem**

For $m \geq 1$, and defining $\zeta^{(0)} = \text{identity}$, the operators $\zeta$ applied to $\zeta^{(m-1)} G$ and $\downarrow^{m+1}$ applied to $\downarrow^m G$ keep and discard the same edges. The operators $\zeta^{(k)}$ and $\downarrow^{k+1}$ produce partial spanning graphs with the same edges and nodes.

**Proof:** We prove it by induction.

a) $m = 1$: $\zeta$ and $\downarrow^2$ select the edges linking a node with its lowest node (this covers the case where the edge belongs to a regional minimum)
Two equivalent modes of pruning.

b) We suppose that $\zeta^{(m-1)} G$ and $\downarrow^m G$ have the same edges and show that it is still the case for $m = m + 1$.

Consider a couple $(i, ij)$ of $\zeta^{(m-1)} G$. It belongs to a $m - 1$-steepest track. The second pair $(j, jk)$ of this track also belongs to a $(m - 1)$—steepest track and holds the weight of the lowest pair of the track. The operator $\zeta$ applied to $\zeta^{(m-1)} G$ assigns this weight to the edge $ij$ but not necessarily to the node $i$ if there exists another pair $(i, il)$ with a lower weight after this erosion. In this case the edge $ij$ is discarded by the operator $\zeta$. But it is also discarded by the operator $\downarrow^{m+1} G$, as no $(m + 1)$-steepest path passes through $ij$. 
The operator $\downarrow^m G$ is of theoretical interest but of poor practical value, as it is based on a neighborhood of size $m$. On the contrary the operator $\zeta$ is purely local and uses a neighborhood of size 1: an erosion from node to node, an erosion from edge to edge and the suppression of any edge which is not the smallest edge of one of its extremities. Repeating $m$ times the operator $\zeta$ produces a graph which has the same edges as the operator $\downarrow^{m+1} G$. It is then sufficient to restore the original weights of edges and nodes of the graph $G$ onto the graph $\zeta^{(m)} G$ to obtain the same result as $\downarrow^{m+1} G$. 
A graph $G$ is a $k$-steepest graph if and only if for each $l < k$, $\varepsilon^{(l)} G$ is a flooding graph.

**Proof:** If a graph $G$ is a $k$-steepest graph, then $G = \downarrow^k G$ and we have already established that for $l < k$, $\varepsilon^{(l)} \downarrow^k G$ is a flooding graph.

Inversely suppose that for each $l < k$, $\varepsilon^{(l)} G$ is a flooding graph, then $\varepsilon^{(l)} G$ has the same edges as $G$. As a matter of fact $\varepsilon^{(l)} G$ is invariant by $\gamma_e$ and each edge is the lowest edge of a node: $\downarrow \varepsilon^{(l)} G = \varepsilon^{(l)} G$.

But $\varepsilon^{(l)} G = \varepsilon \varepsilon^{(l-1)} G$ and $\downarrow \varepsilon \varepsilon^{(l-1)} G = (\downarrow \varepsilon_e e, \varepsilon_n n) \varepsilon^{(l-1)} G = \zeta \varepsilon^{(l-1)} G$ (considering $G = \varepsilon^{(0)} G$). It follows that for $l < k$, we have $\varepsilon^{(l)} G = \zeta^{(l)} G$.

The operator $\zeta$ applied to $k$ times to $G$ never suppresses an edge from $G$. But we know that $\zeta^{(k-1)} G$ and $\downarrow^k G$ have the same nodes and edges, showing that indeed $G$ is a $k$-steepest graph.
Deriving an algorithm for delineating the catchment basins
It is now possible to imagine an algorithm associated to the operator $\zeta$. If $G$ is a flooding graph, we assign a distinct label to each regional minimum and a weight 0. The operator $\zeta$ propagates the weights upwards along the steepest tracks as it is repeated. In parallel, we also propagate the labels: every time a pair $(i,j)$ takes the weight 0 from a pair containing a labeled node $j$, the node $i$ takes the label of $j$. Like that, as the labeled zones with weight 0 expand, their labels also expand.

We now have to consider two cases. In the first case, we construct restricted catchment basins separated by watershed zones. In the second, we create a partition by assigning each node to one and only one basin, at the price of arbitrary choices.
Catchment basins and watershed zones

During the successive operations $\zeta$, every time a pair $(i, ij)$ takes the weight 0 from a pair containing the node $j$, the node $i$ takes the label of $j$ provided the pair $(i, ij)$ is unique. If there are two equivalent pairs $(i, ij)$ and $(i, il)$, it means that there exist two minimal lexicographic tracks starting at $i$ towards one or two distinct minima. If $j$ and $k$ hold the same label, this label is assigned to $i$. If on the contrary they are distinct, we have 2 possibilities:

- we assign to $i$ a label $Z$, indicating that it belongs to a watershed zone. The upstream of $i$ also will get this same label $Z$. The regions with label $Z$ belong to the watershed zone and the other labeled regions are restricted catchment basins.
- we assign to $i$ one of the labels (either randomly, or by applying an additional rule for braking the ties), producing a partition in catchment basins with an empty watershed zone. The result is not unique and depends on the succession of choices which have been made.
On the left a flooding graph where the minima have weights equal to 0 and their labels are indicated by distinct colors. The next two figures show the propagation of the weights and of the labels as the operator $\zeta$ has been applied twice in a sequence.
Starting with a node or edge weighted graph we have shown how to extract from it a flooding graph where nodes and edges are weighted. The operators $\downarrow^m G$ extract from the graph $G$ partial graphs which are steeper and steeper flooding graphs. The series of partial graphs $\downarrow^m G$ is decreasing with $m$.

The operator $\zeta$ permits an iterative construction of $\downarrow^m G$, using only small local neighborhood transformations.
The scissor operator and the watershed partitions

In the previous section we introduced pruning operators which extract from a flooding graph k-steep flooding graphs. These operators do not make any choice among the k-steep flooding graphs, they take them all.

We now introduce operators aiming at creating partitions: they extract minimum spanning forests from the flooding graph by pruning. Contrarily to the preceding operators, they do make arbitrary choices among equivalent edges to be suppressed.
The flooding pairs

The previous section has shown how to extract from a node or edge weighted graph partial graphs which are flooding graphs. In a flooding graph, each node has at least one adjacent edge with the same weight. We consider here the nodes and edges outside the regional minima. There exists at least one (in general several) one to one correspondence between each node outside a regional minimum and one of its adjacent edges with the same weight. For each such one to one correspondence, we call flooding pair, the couple of node and edge which have been associated. They hold the same weight.
The flooding pairs

Let us show how to construct such a one to one correspondance. If an edge is the lowest adjacent edge of only one of its extremities, then they form a flooding pair; this is in particular the case if the other extremity of the edge has a lower weight.

If on the contrary, it has the same weight as both its extremities, it belongs to a plateau. As this plateau is not a regional minimum, there exists a pair of neighboring nodes, a node \(s\) inside the plateau and a lower node \(t\), outside. This edge \((s, t)\) forms a flooding pair with the node \(s\). As the plateau is connected, there exists a tree spanning its nodes, having \(s\) as root. There exists a unique path between \(s\) and each node \(i\) of the plateau. The last edge on this path before reaching \(i\) and \(i\) itself form a flooding pair.
The scissor operator creates a partition of the nodes

We have shown that there exists at least one (in general several) one to one correspondance between each node outside a regional minimum and one of its adjacent edges with the same weight. For each such one to one correspondance, we call flooding pair, the couple of node and edge which have been associated. To each such one to one correspondance we associate a pruning operator \( \chi \), called scissor. This operator suppresses all edges outside a regional minimum which do not form a flooding pair with one of their extremities.

After applying the operator \( \chi \) to a flooding graph, there exists one and only one path from each node to a regional minimum: the catchment basins of the minima partition the nodes, and the watershed zones are empty.
The drainage minimum spanning forest
The drainage minimum spanning forest

After contracting all edges in the regional minima of a flooding graph and applying the scissor operator $\chi$ one gets a spanning forest, where each tree is rooted in a minimum:

- the resulting graph is a partition where each connected component contains a regional minimum node: there exists a path linking each node outside the minima with a minimum.
- each connected component is a tree as the number of nodes is equal to the number of edges (the number of flooding pairs) plus one (the regional minimum node).
- it is a minimum spanning forest: each node is linked to its tree through one of its lowest neighboring edge. The total weight of each tree is thus equal to the total sum of the nodes outside the regional minima. This total weight is independent of the particular choice made by $\chi$. 
The drainage minimum spanning forest

Expanding again the regional minima and replacing them by a MST of each regional minimum creates again a minimum spanning forest, identical to the preceding one outside the minima. Its total weight is computed as follows:

- if $M_i$ is a regional minimum with $n$ nodes of weight $\lambda_i$, the weight of its MST is equal to $(n - 1) \times \lambda_i$;
- each other node contributes by its weight, which is also the weight of its lowest adjacent edge.

Each particular forest is based on a particular scissor operator $\chi$ and of a particular MST in each regional minimum. We call $\phi$ the operator which extracts from a flooding graph such a MSF (minimum spanning forest).
The catchment basins

\[ G = \text{a flooding graph}, \ C = \phi G \text{ a minimum spanning forest:} \]

- Each node belongs to a regional minimum or is the origin of a unique never ascending path leading to a regional minimum.
- Each tree of \( C \) contains a unique regional minimum; its nodes form the catchment basin of this minimum.
Catchment basins of increasing steepness

$G$ = a flooding graph, $\downarrow^m G$ still is a flooding graph and $C_m = \phi \downarrow^m G$ a minimum spanning forest of steepness $m$.

As for $m > l : \downarrow^m G \subset \downarrow^l G$, any minimum spanning forest $C_m = \phi \downarrow^m G$ is also a minimum spanning forest of $\downarrow^l G$.

For increasing values of $m$, the number of forests of steepness $m$ decreases.
Illustration
We chose as topographic surface the distance function expressed on the nodes of a hexagonal grid to two binary connected sets, encoded with the value 0. The edge weights are obtained by the dilation $\delta_{en}$. Like that the lowest neighboring edges of a node connects it with its neighbors with equal or lower weights. In the following figure, the edges with the same weight have the same color (cyan = 4, magenta = 3, green = 2).
The flooding graph of a distance function

As the minima have 2 nodes each, the pixel graph is invariant by $\varphi_n$. The dilation $\delta_{en}$ assigns weights to the edges and creates a flooding graph.
The minimum spanning forest by keeping one lowest neighboring edge for each node

The scissor \( \chi \) leaves one lowest neighboring edge for each node. There exists a huge number of choices for \( \chi \). The following illustration shows a particular scissor \( \chi \) producing an unexpected partition in two catchment basins (the catchment basins should be the half plane separated by the mediatrix of both binary sets, as illustrated in the previous slide).
Two geodesics for the flooding ultrametric distance, leading to an unexpected partition.
Looking one node further

The operator $\downarrow^2 G$ leaves only the edges linking a node to its lowest neighbors. The following figure shows the remaining edges. The green zone represents a restricted catchment basin. The yellow zone is an extended catchment basin containing the watershed zone.
The minimum spanning forest by keeping one edge towards a lowest neighboring node

The pruning $\chi \downarrow = \chi \downarrow^2$ leaves for each node one edge towards a lowest neighboring node. The following figure shows one such solution and the resulting partition.
The catchment basins, skeletons by zone of influence for lexicographic distance functions
The operator $\downarrow^k G$ prunes the flooding graph and leaves only NAPs with a steepness equal to $k$. From each node of the graph starts a NAP whose $k$ first edges have a minimal lexicographic weight. If one follows such a path, the next $k$ edges following each node as one goes downwards along the path also has a minimal lexicographic weight. This shows that each NAP is a geodesic line for a lexicographic distance function $\text{lexdist}_k$ we define below.

After the pruning $\downarrow^k$ some nodes belong to two or more catchment basins. In order to obtain a partition, one applies the scissor operator $\chi$ leaving for each node only one lowest adjacent edge. Like that the thick watershed zones are suppressed and the catchment basins form a partition.
Constructing a watershed partition as the skeleton by zones of influence of the minima

It is possible to obtain the same result, if one labels the regional minima and computes for the other nodes the shortest lexicographic distance \( \text{lexdist}_k \) to the minima. If one applies a greedy algorithm, one may in addition propagate the labels of the minima all along the geodesics and construct a partition of the space. If a node is at the same lexicographic distance of two nodes, a greedy algorithm will arbitrarily assign to it one of the labels of the minima. This is quite similar to the operator \( \chi \) which also does arbitrary choices.
Defining a lexicographic distance along non ascending paths.

Consider a flooding graph. In a NAP, each node except the last one forms with the following edge a flooding pair, i.e. they have the same weights. And these weights are decreasing as one follows the NAP downwards. The $k$ first values, starting from the top, may be considered as a lexicographic distance of depth $k$. In what follows we give a precise definition of such distances.

The shortest distances and their geodesics may be computed with classical algorithms. Propagating the labels of the minima along the geodesics during their construction constructs the zones of influence of the minima, i.e a partition into catchment basins. The solution is not necessarily unique. However the number of solutions decreases with the depth of the lexicographic distance which is considered.
Lexicographic distances of depth $k$
Comparing NAPs with the lexicographic order.

Let $S$ be the set of sequences of edge or node weights, i.e. elements of $\mathcal{T}$. Let $S_k$ be the set of sequences with a maximal number $k$ of elements. For a sequence $s \in S_k$, we define the lexicographic weight $w_k(s)$: it is equal to $\infty$ if $s$ is not a NAP and equal to $s$ itself otherwise.

We define an operator $\text{first}_k$ which keeps the $k$ first edges and nodes of any NAP, or the NAP completely if its length is smaller than $k$. The operator $\text{first}_k$ maps any sequence of $S$ into $S_k$.

We define on the NAPs of $S_k$ the usual lexicographic order relation, which we will note $\prec$, such that: $(\lambda_1, \lambda_2, \ldots, \lambda_k) \prec (\mu_1, \mu_2, \ldots, \mu_k)$ if either $\lambda_1 < \mu_1$ or $\lambda_i = \mu_i$ until rank $s$ and $\lambda_{s+1} < \mu_{s+1}$. We define $a \preceq b$ as $a \prec b$ or $a = b$. 

Comparing NAPs with the lexicographic order.

Like that, it is possible to compare any two sequences $s_1$ and $s_2$ of $S_k$ by comparing their weights:

- if $s_1$ and $s_2$ are not NAPs, then $w_k(s_1) = w_k(s_2) = \infty$ and we consider that $s_1 \equiv s_2$ (they are equivalent)
- if one of them, say $s_1$, is a NAP and not the other, then $w_k(s_1) \prec w_k(s_2) = \infty$ and $s_1 \prec s_2$
- if both of them are NAP, then they compare as their weights:
  $$s_1 \prec s_2 \iff \{w_k(s_1) \prec w_k(s_2)\} \quad \text{and} \quad s_1 \equiv s_2 \iff \{w_k(s_1) = w_k(s_2)\}$$

If $s_1$ and $s_2$ are NAPs belonging to $S$, we compare them with:
$$s_1 \prec s_2 \iff w_k[\text{first}_k(s_1)] \prec w_k[\text{first}_k(s_2)]$$
For NAPs of $S$ we define the operator $⊞_k$ called “addition”, which operates as a minimum:

$$a ⊞_k b = \begin{cases} a & \text{if } a \preceq b \\ b & \text{if } b \preceq a \end{cases} \quad \forall a, b \in S$$

The $⊞_k$ operation is associative, commutative and has a neutral element $\infty$ called the zero element: $a ⊞_k \infty = \infty ⊞_k a = a$
The operator $\boxtimes_k$, called “multiplication”, permits to compute the lexicographic length $\Lambda(A)$ of a sequence, obtained by the concatenation of two sequences. Let $a = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $b = (\mu_1, \mu_2, \ldots, \mu_k)$ we will define $a \boxtimes_k b$ by:

- if $a$ or $b$ is not a NAP then $a \boxtimes_k b = \infty$
- if $a$ and $b$ are NAPs and $\lambda_k < \mu_1$ then $a \boxtimes_k b = \infty$
- if $a$ and $b$ are NAPs and $\lambda_k \geq \mu_1$ then $a \boxtimes_k b = w_k \left[ \text{first}_k(a \triangleright b) \right]$, where $a \triangleright b$ is the concatenation of both sequences.

In particular $a \boxtimes_k \infty = \infty \boxtimes_k a = \infty$. so that the zero element is an absorbing element for $\boxtimes_k$. 


Algebraic shortest paths algorithms
A path algebra on a dioïd structure

The operator $\bowtie_k$ is associative and has a neutral element 0 called unit element: $a \bowtie_k 0 = a$. The multiplication is distributive with respect to the addition both to the left and to the right.

The structure $(S, \boxplus_k, \bowtie_k)$ forms a dioïd, on which Gondran and Minoux defined a path algebra [10], where shortest paths algorithms are expressed as solutions of linear systems.
Transposing the dioïd to square matrices

Addition and multiplication of square matrices of size $n$ derives from the laws $\oplus$ and $\otimes_k$ : for $A = (a_{ij})$, $B = (b_{ij})$, $i, j \in [1, n]$ :

$$C = A \oplus B = (c_{ij}) \iff c_{ij} = a_{ij} \oplus b_{ij} \quad \forall i, j$$

$$C = A \otimes_k B = (c_{ij}) \iff c_{ij} = \sum_{1 \leq k \leq n} a_{ik} \otimes_k b_{kj} \quad \forall i, j$$

where $\sum$ is the sum relative to $\oplus$.

As there is no ambiguity, for $\sum_{1 \leq k \leq n} a_{ik} \otimes_k b_{kj}$, we simply write $\sum_{1 \leq k \leq n} a_{ik} b_{kj}$.
Unity and zero matrices

With these two laws, the square matrices also become a dioïd with

\[
\hat{\varepsilon} = \begin{bmatrix}
\varepsilon & \cdots & \varepsilon \\
\varepsilon & \cdots & \varepsilon \\
\varepsilon & \cdots & \varepsilon \\
\varepsilon & \cdots & \varepsilon \\
\end{bmatrix}
\]

zero matrix \( \hat{\varepsilon} = \)

\[
E = \begin{bmatrix}
e & \cdots & \varepsilon \\
e & \cdots & \varepsilon \\
e & \cdots & \varepsilon \\
e & \cdots & \varepsilon \\
\end{bmatrix}
\]

and unity matrix \( E = \)

\[
A^0 = E
\]
If $G = [X, U]$ is a weighted graph with

- a set $X$ of nodes, numbered $i = 1, \ldots, N$.
- a set $U$ of edges $u = (i, j)$ with weights $s_{ij} \in S$.

The incidence matrix $A = (a_{ij})$ of the graph is given by:

$$a_{ij} = \begin{cases} s_{ij} & \text{si} \\ \varepsilon & \text{sinon} \end{cases} \quad (i, j) \in U$$
To each path $\mu = (i_1, i_2, ..., i_k)$ of the graph, one associated its k-lexicographic weight $w(\mu) = s_{i_1i_2} \boxdot_k s_{i_2i_3} \boxdot_k ... \boxdot_k s_{i_{k-1}i_k}$, which is different from $\infty$ only if the path is a NAP.

If $\pi$ is a never increasing sequence, then $w_k[\text{first}_k(\pi)] = \text{first}_k(\pi)$
The shortest paths for the lexicographic distance between any couple of nodes may be computed thanks to $A^n$ or $A^{(n)} = E \boxplus A^1 \boxplus A^2 \boxplus \ldots \ldots A^n$.

**Lemma**

$$A^n_{ij} = \sum_{\mu \in C^{n}_{ij}} w(\mu) \quad \text{and} \quad A^{(n)}_{ij} = \sum_{\mu \in C^{(n)}_{ij}} w(\mu)$$

where:

- $C^{n}_{ij}$ is the family of paths between $i$ and $j$, containing $n + 1$ nodes (not necessarily distinct)
- $C^{(n)}_{ij}$ is the family of paths between $i$ and $j$, containing at most $n + 1$ nodes (not necessarily distinct)

Defining $A' = E \boxplus A$, as $\boxplus$ est idempotent ($a \boxplus a = a \quad \forall a \in S$), we have: $A'^{n} = A^{(n)}$
In a graph with \( N \) nodes, an elementary path has at most \( N \) nodes, separated by \( N - 1 \) edges. Hence, necessarily \( A^{(N)} = A^{(N-1)} \) and \( A^* \), the limit of \( A^{(n)} \) for increasing \( n \), is also equal to \( A^{(N-1)} \).

Thanks to \( A^* \), the computation of the catchment basins is immediate. If \( m_1 \) is a regional minimum, then a node \( i \) belongs to its catchment basin if and only if

\[
A^*_{im_1} \leq \bigoplus_{m_j \neq m_1} A^*_{im_j}
\]

\( A^* \) may be obtained by the successive multiplications:

\[
A, \ A^2 = A \boxtimes_k A, \ A^4 = A^2 \boxtimes_k A^2, \ ... A^{2^i} = A^{2^{i-1}} \boxtimes_k A^{2^{i-1}} \text{ until } 2^i \geq N - 1, \text{ i.e. } i \geq \log (N - 1)
\]
A path algebra on a dioïd structure

\[ A^* \text{ verifies } A^* = E \boxplus A \boxtimes_k A^*. \]

Multiplying by a matrix \( B \):
\[ A^* B = B \boxplus A \boxtimes_k A^* B. \]
Defining \( Y = A^* B \) shows that \( A^* B \) is solution of the equation
\[ Y = B \boxplus A \boxtimes_k Y. \]
Furthermore, it is the smallest solution.
With varying \( B \) it is possible to solve all types of shortest distances:

- The smallest solution of \( Y = E \boxplus A \boxtimes_k Y \) yields \( A^* E = A^* \)

with \( B = \begin{bmatrix} \varepsilon & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \varepsilon \\ \varepsilon & \cdots & \cdots & \varepsilon \end{bmatrix} \), one gets \( A^* B \), the i-th column of the matrix \( A^* \),

that is the distance of all nodes to the node \( i \).
Gondran and Minoux have shown that most of the classical algorithms solving systems of linear equations (Gauss, Gauss-Seidel, etc.) are still valid in this context and correspond often to known shortest paths algorithms defined on graphs. We now give a few examples.
The Jacobi algorithm

Setting \( B = \begin{bmatrix} 0 \\ \varepsilon \\ \vdots \\ \varepsilon \end{bmatrix} \) and \( Y^{(0)} = \begin{bmatrix} \varepsilon \\ \varepsilon \\ \vdots \\ \varepsilon \end{bmatrix} \), the iteration

\[
Y^{(k)} = A \ast Y^{(k-1)} \boxplus B
\]

converges to \( Y^{(N)} = A^\ast \ast B \).
The Gauss Seidel algorithm

$A$ is decomposed as $A = L \oplus E \oplus U$, where $L$ is an inferior triangular matrix, $E$ the unity matrix $E$, and $U$ an upper triangular matrix. The upper part of $L$ and lower part of $U$ have the value $\varepsilon$.

The solution of $Y = A \ast Y \oplus B$ is obtained by the iteration:

$$Y^{(k)} = LY^{(k-1)} \oplus UY^{(k)} \oplus B$$

This algorithm is faster as Jacobi’s algorithm, as the product $UY^{(k)}$ already uses intermediate results, freshly computed during the current iteration $Y^{(k)}$. 
The Jordan algorithm

The Jordan algorithm is used in classical linear algebra for inverting matrices, by successive pivoting. In our case, where the shortest paths are elementary path the algorithms is:

For $k$ from 1 to $N$:

For each $i$ and $j$ from $i$ to $N$, do: $a_{ij} = a_{ij} \oplus a_{ik} * a_{kj}$
Gondran established the algebraic counterpart of the famous shortest path algorithm by Moore-Dijkstra.

**Theorem** (Gondran): Let $\overline{Y} = A^*B$ be the solution of $Y = AY \boxplus B$, for an arbitrary matrix $B$. There exists then an index $i_0$ such that $\overline{y}_{i_0} = \sum b_i$. The smallest $b$ is such solution: $\overline{y}_{i_0} = b_{i_0}$

Each element of $Y = AY \boxplus B$ is computed by

$$y_k = \sum_{j \neq k} a_{kj} \star y_j \boxplus b_k = \sum_{j \neq k, i_0} a_{kj} \star y_j \boxplus a_{ki_0} y_{i_0} \boxplus b_k$$

Suppressing the line and column of rank $i_0$ and taking for $b$ the vector $b_{i_0}^{(1)} = a_{ki_0} y_{i_0} \boxplus b_k$, one gets a new system of size $N - 1$ to solve.

The Moore Dijkstra algorithm can also directly be computed on a flooding graph $G$, as presented below.
Shortest paths algorithms on the graph

- The shortest path algorithm of Dijkstra is first presented \([22]\). We show that for a lexicographic distance of depth 1, it becomes an algorithm for constructing the minimum spanning forest of Prim.
- The core expanding algorithms take advantage of the particular structure of the lexicographic distances. They are faster than the Dijkstra algorithm and better suited to hardware implementations. For a lexicographic distance of depth 2, they produce the same geodesics as the topographic distance.
The shortest path algorithms below are applied on a graph which is invariant by the opening $\gamma_e$ (the regional minima may ad libitum be contracted beforehand). A domain $D$ is used and expanded, containing at each stage of the algorithms the nodes for which the shortest distance to the minima is known. Initially the minima are labeled and put in $D$ with a value 0. We say that a flooding pair $(j, jl)$ is on the boundary of the domain $D$, if $j$ is outside $D$ and $l$ inside $D$. In this case we say that ”$j$ floods $l$” , or ”$l$ is flooded by $j$”. We say that $j$ belongs to the outside boundary $\partial^+ D$ of $D$ and $l$ to the inside boundary $\partial^- D$ of $D$. The domain $D$ is progressively expanded by progressive incorporation of nodes in $\partial^+ D$ belonging to flooding pairs on the boundary of $D$. 
The Moore Dijkstra algorithm

The shortest path algorithm by Moore-Dijkstra constructs the distances of all nodes to the minima in a greedy manner. It uses a domain $D$ which contains at each stage of the algorithm the nodes $i$ for which the shortest distance $\delta^*_k(i)$ and label $\lambda(i)$ is known.

**Initialisation:**
The nodes of the regional minima are labeled and put in the domain $D$.

**Repeat until the domain $D$ contains all nodes:**
For each flooding pair $(j, jl)$ on the boundary of $D$, estimate the shortest path as $\delta_k(j) = e_{jl} \square_k \delta^*_k(l)$.

The node with the lowest estimate is correctly estimated. If the corresponding flooding pair is $(s, st)$:

- $D = D \cup \{s\}$
- $\delta^*_k(s) = \delta_k(s)$
- $\lambda(s) = \lambda(t)$
The node with the lowest estimate is correctly estimated and is introduced in the domain $D$. If is necessarily the shortest path, as any other path would have to leave $D$ through another boundary flooding pair with a higher estimate.
The Moore Dijkstra algorithm may be advantageously controlled by a hierarchical queue structure. Each node, as it gets its estimate, is put into the HQ. As long as the HQ is not empty, the extracted node is among the nodes with the lowest estimate, the one which has been introduced in the HQ first. The HQ has thus the advantage to correctly sequence the treatment in the presence of plateaus: it treats the nodes in the plateaus from their lower boundary inwards. The processing order is proportional to the distance of each node to the lower boundary of the plateau.
The Moore Dijkstra algorithm: case where $k = 1$

The algorithms remain the same but the computations are simplified as $\delta_k(j) = e_{jl} \circ 1 \delta^*_1(l)$ is simply the weight of the node and the edge in the flooding pair $(j, (jl))$. In other terms, the domain $D$ is expanded by introducing into $D$ the flooding pair on the boundary of $D$ with the lowest weight. This corresponds exactly to the algorithm of PRIM for constructing minimum spanning forests. The same algorithm has been used in [16] for constructing the waterfall hierarchy.

This is not surprising, as for $k = 1$, the lexicographic weight of a NAP simply is the weight of the first edge. The distance is in this case the ultrametric flooding distance.

Remark: There is a complete freedom in the choice of the flooding pairs which are introduced at any time into $D$. A huge number of solutions are compatible with this distance, some of them quite unexpected as illustrated by an example given above.
The core expanding shortest distance algorithm

Due to the particular structure of lexicographic distances, it is possible to identify another type of nodes for which the shortest distance may immediately be computed. Let $\partial^- D$ be the set of nodes in the boundary $\partial^- D$ with the lowest valuation $\text{first}_{k-1}(d^*_k)$. For each $t \in \partial^- D$ and for each $s \in \partial^+ D$ flooding $t$ we do $e_{ts} \bigotimes_k \text{first}_{k-1} \delta^*_k(t)$ producing a NAP of length $k$. The value of $\delta^*_k(t) \text{first}_{k-1}$ is simply obtained by appending the weight of $s$ to $\text{first}_{k-1} \delta^*_k(t)$. As $t$ belongs to $\partial^- D$, this value is the smallest possible.

This analysis shows that each node of $\partial^- D$ permits to introduce into $D$, all its neighbors by which it is flooded, whatever their weight. This algorithm is more "active" as Dijkstra’s algorithm, as each node inside $D_k$ is able to label and introduce into $D$ all its flooding neighbors at the same time.
The core expanding shortest distance algorithm

The algorithm is the following:

**Initialisation:**
The nodes of the regional minima are labeled and put in the domain $D$.

**Repeat until the domain $D$ contains all nodes:**
Let $\partial^- D$ be the subdomain of $D$ of nodes flooded by nodes outside $D$ with the lowest valuation $\text{first}_{k-1}(d^*_k)$. For each $t \in \partial^- D$ and $t$ is flooded by $\in \partial^+ D$:

- $D = D \cup \{s\}$
- $\delta_k^*(s) = e_{ts} \boxplus_k \text{first}_{k-1} \delta_k^*(t)$
- $\lambda(s) = \lambda(t)$
The core expanding shortest distance algorithm may also be advantageously controlled by a hierarchical queue structure. Each node, as it is introduced in $D$ is put into the HQ. As long the HQ is not empty, the extracted node is among the nodes with the lowest estimate, the one which has been introduced in the HQ first. This node gives its label and the correct distances to all its neighbors outside $D$ by which it is flooded. If a node extracted from the HQ belongs to a plateau, as it has been introduced in the HQ, it is closer to the lower boundary of the plateau than other nodes which may have been introduced in the HQ later. The algorithm is particularly suitable for a hardware implementation: each node is entered in the HQ only once. On the contrary, with the Moore-Dijkstra algorithm a node may be introduced several times in the HQ, as its estimate may vary before it gets its definitive value.
The core expanding shortest distance algorithm: case where $k = 1$

The domain $\partial^- D$ has been defined as the subdomain of $D$ of nodes flooded by nodes outside $D$ with the lowest valuation $\text{first}_{k-1}(d_k^*)$. For $k = 1$, we have $\text{first}_{k-1}(d_k^*)$ is empty, and $\partial^- D$ occupies the whole domain $\partial^- D$. This means that any node in $\partial^- D$ can be expanded by appending one of the outside node through which it is flooded. The algorithms for constructing graph cuts by J.Cousty find also their place in this context \cite{9}.

We have illustrated this situation earlier, showing that the minimum spanning forests with 0 steepness may be absolutely unexpected. Controling the algorithm with a hierarchical queue limits the anarchy to some extend.
If $k = 2$, then the valuation $\text{first}_{k-1}(d^*_k(i)) = \text{first}_1(d^*_2(i))$ is the valuation of the node $i$ itself. This means that $\partial^- D$ contains the nodes with lowest weight belonging to $\partial^- D$. If $i$ is such a node, it introduces into $D$ each of its neighbors $j$ belonging to $\partial^+ D$, each with a valuation $\text{first}_1(d^*_2(j))$ equal to its weight $n_j$.

If in addition, one uses a HQ for controlling the process, we get the classical algorithm for constructing catchment basins [17],[19].
The classical watershed algorithm

Label the nodes of the minima and them in the domain $D$, each with a priority with a weight.
As long as the HQ is not empty, extract the node $j$ with the highest priority from the HQ:
For each unlabeled neighboring (on the flooding graph) node $i$ of $j$:
* $\text{label}(i) = \text{label}(j)$
* put $i$ in the queue with priority $\nu_i$
As a matter of fact, this algorithm has first been derived from the watershed line, as zone of influence of the minima for the topographic distance, defined below.
The topographic distance

Consider an arbitrary path \( \pi = (x_1, x_2, ..., x_p) \) between two nodes \( x_1 \) and \( x_n \). The weight \( \nu_p \) at node \( x_p \) can be written:

\[
\nu_p = \nu_p - \nu_{p-1} + \nu_{p-1} - \nu_{p-2} + \nu_{p-2} - \nu_{p-3} + \ldots + \nu_2 - \nu_1 + \nu_1
\]

The node \( k - 1 \) is not necessarily the lowest node of node \( k \), therefore \( \nu_{k-1} \geq \varepsilon_n \nu_k \) and \( \nu_k - \nu_{k-1} \leq \nu_k - \varepsilon_n \nu_k \). Replacing each increment \( \nu_k - \nu_{k-1} \) by \( \nu_k - \varepsilon_n \nu_k \) will produce a sum \( \nu_p - \varepsilon_n \nu_p + \nu_{p-1} - \varepsilon_n \nu_{p-1} + \ldots + \nu_2 - \varepsilon_n \nu_2 + \nu_1 \) which is larger than \( \nu_p \). It is called the topographic length of the path \( \pi = (x_1, x_2, ..., x_p) \). The path with the shortest topographic length between two nodes is called the topographic distance between these nodes.
The topographic distance

The path with the shortest topographic length between two nodes is called the topographic distance between these nodes [23], [19]. It will only be equal to $\nu_p$ in the case where the path $(x_1, x_2, ..., x_p)$ precisely is a path of steepest descent, from each node to its lowest neighbor. In other terms the topographic distance and the distance lexicist2 have the same geodesics. If we define the toll to pay along a path $\pi = (x_1, x_2, ..., x_p)$ as $\nu_1$ for the first node and $\nu_i - \epsilon \nu_i$ for the others, then the lowest toll distance for a node $x_p$ to a regional minimum will be $\nu_p$ if there exists a path of steepest descent from $x_p$ to $x_1$. In other terms, $x_p$ and $x_1$ belong to the same catchment basins if one considers the topographic distance. But they also belong to the same catchment basin if one considers the depth 2 lexicographic distance, as, by construction, $x_k$ is the lowest neighbor of $x_{k-1}$. 
Distance on a node weighted graph based on the cost for travelling along the cheapest path

We recall the toll distance presented earlier. The weights are assigned to the nodes and not to the edges. Each node may be considered as a town where a toll has to be paid.

**Cost of a path:** The cost of a path is equal to the sum of the tolls to be paid in all towns encountered along the path (including or not one or both ends).

**Cost between two nodes:** The cost for reaching node $y$ from node $x$ is equal to the minimal cost of all paths between $x$ and $y$. We write $\text{tolldist}(x, y)$. If there is no path between them, the cost is equal to $\infty$. 
Illustration of the cheapest path

In this figure, the shortest chain between $x$ and $y$ is in red and the total toll to pay is $1 + 1 + 2 + 2 = 6$
Finally, any image \( f \) may be considered as the global toll of its pixel graph if one takes:

- as reference nodes the regional minima of the image. Each of them has as toll its altitude.
- as local toll for all other nodes, the difference between their altitude and the altitude or their lowest neighbor: \( g = f - \varepsilon f \)

If in addition, we give a different label to each regional minimum, we may as previously propagate this label along each smallest toll path. We obtain like that a tessellation: to each minimum is ascribed a catchment basin: the set of all nodes which are closer to this minimum then to any other minimum.
Inversely: the catchment basins of the cheapest paths for a distribution of tolls.

Inversely we may chose a number of starting nodes called roots, with a toll to pay and for all other nodes, the toll to pay for reaching or crossing this node. The toll to pay constitutes a topographic surface where each root is a regional minimum. In addition we propagate the labels of the minima along the geodesics of the cheapest distance, and get a partition of the nodes. As a result one get a partition of the space, where each region with a given label is the catchment basin of the root with the same label for this distance function.
Assigning to each node the global toll for reaching it, starting from one of the roots.
Illustration of the cheapest path

For each node we have indicated the local toll value (left value) and the global toll to pay to reach the closest reference node, who shares the same color (label):

- Each reference node became a regional minimum of the graph.
- The local toll of any other node is equal to the difference between its global toll and the global toll of its lowest neighbor.
- Each non reference node got its value and label from one of its lowest neighbors.
- The values of the nodes are computed and the labels propagated along a path of steepest descent.
The catchment basins of this surface are the SKIZ of the minima, both for the topographic distance and for the depth 2 lexicographic distance. Each node is the extremity of a geodesic line, which is the same both for the toll distance and for the lexicographic distance of depth 2 computed on the same topographic surface.
Top down or bottom up
For increasing values of $k$, the domain $\partial^- D$ becomes smaller and smaller, indicating that the number of equivalent catchment basins compatible with a given lexicographic depth is reduced as the value of $k$ becomes bigger. This is in accordance with the fact that with increasing values of $k$, the pruning $\downarrow^k$ becomes more and more severe.
Obtaining catchment basins with $k$ — steepness

After applying the operator $\downarrow^k$ on a flooding graph $G$, there remain only NAP with $k$ steepness.
The same is true if we apply the operator $\zeta^{(k-1)}$ in order to prune edges of $G$ and subsequently restore the initial weights of the edges onto the remaining edges.

If the only NAPs remaining in the graph have a $k$ steepness, any shortest path algorithm with a lower steepness will extract them. In particular, the most simple algorithms for the distances $d_1$ or $d_2$ will extract catchment basins of steepness $k$. 
The following 3 images show respectively the shortest lexicographic distances of depth 1, 2 and 3. If a path is the shortest path for a lexicographic distance of depth $k$, it also is a shortest path for a lexicographic distance of smaller depth. The following three figures present the lexicographic distances of depth 1, 2 and 3. The partition of catchment basins for the distance 3 is also solution for the distance 2 and 1. Similarly the partition for distance 2 is also solution for distance 1. The number of solutions decreases with the lexicographic depth.
Illustration: lexicographic distance of depth 1
Illustration: lexicographic distance of depth 2
Illustration: lexicographic distance of depth 3

Watersheds, waterfalls, on edge or node weighted graphs

2012 February 29
Erosion and pruning

Repeating the operator $ζ G = (\downarrow \varepsilon e, \varepsilon n n)$ produces a decreasing series of partial graphs $ζ^{(n)} G = ζζ^{(n-1)} G$, which are steeper and steeper. In the following figures, we present in red the edge which is not the lowest edge of one of its extremities. After pruning this edge, the operator is applied again. Applying $ζ$ a number $n$ of times is equivalent to constructing the partitions compatible with a SKIZ for a lexicographic distance of depth $n + 1$. In our case, $ζ^{(3)} G$ produces a graph with the same edges as the pruning $χ$ applied to the graph where each edge has been weighted by its lexicographic distance of depth 4 to the nearest minimum, illustrated in the previous figure.
Erosion and pruning
The hierarchy of nested catchment basins
The waterfall hierarchy has been introduced by S. Beucher in order to obtain a multiscale segmentation of an image \([2],[3],[3]\). Given a topographic surface \(S_1\), typically a gradient image of the image to segment, a first watershed transform produces a first partition \(\pi_1\). The waterfall flooding of \(S_1\) consists in flooding each catchment basin up to its lowest neighboring pass point, producing like that a topographic surface with less minima. The watershed segmentation of this surface produces a coarser partition \(\pi_2\) where each region is the union of a number of regions of \(\pi_1\).
Chaining the waterfall floodings

Flooding each catchment basin of $S_1$ up to its lowest pass point produces a new topographic surface $S_2$ which will be submitted to the same treatment as the initial surface $S_1$. Its watershed transform produces a second partition $\pi_2$. The partition $\pi_2$ is coarser than $\pi_1$ as each of its tile is a union of tiles of $\pi_1$. The following figure shows how the flooding of $S_1$ produces $S_2$, which, in the second figure, has also been flooded.

The same process can be repeated several times until a completely flat surface is created. The partitions obtained by the watershed construction on the successive waterfall floodings are coarser and coarser: they form a hierarchy.
The waterfall hierarchy

A 1 dimensional topographic surface is represented through the altitude of its pass points in the figure below. The first flooding assigns weights to the nodes and its watershed construction produces 4 catchment basins, separated by 3 watershed lines. The second flooding has only 2 catchment basins separated by 1 watershed line. The last image orders the watershed lines of the initial image into 3 categories, in cyan the watershed lines which disappeared during the first flooding, in dark blue those which disappeared after the second flooding and in red the one which survived the second flooding.
Let us come back to the watershed on weighted graphs. The watershed of the topographic surface produces a partition $\pi_1$ into catchment basins, represented by its region adjacency graph $RAG_1$. The first flooding floods each catchment basin up to its lowest neighboring pass point. This corresponds to the erosion $\varepsilon_{ne}$ of the graph $RAG_1$. The theory of the watershed on weighted graphs can now be applied on this graph. The resulting watershed appears in form of minimum spanning forest $MSF_1$; each tree of the forest spans a catchment basin of the partition $\pi_2$. The next level of the hierarchy may be represented by a new region adjacency graph $RAG_2$, whose nodes are obtained by contracting each tree of the forest $MSF_1$ in the graph $RAG_1$. Repeating the same treatment to the graph $RAG_2$ produces the next level of the hierarchy, where each tree of the minimum spanning forest $MSF_2$ has been obtained by merging several trees of the $MSF_1$. The process is then repeated until a graph is created with only one regional minimum.
Construction of the level 1 of the hierarchy

We start with an arbitrary node or edge weighted and connected graph $G$. As explained earlier, we extract a graph flooding graph $G'$. For a steepness $k$, we prune $G'$ and get $\downarrow^k G'$. The scissor operator $\chi$ produces a minimum spanning forest $F_1 = \chi \downarrow^k G'$, spanning the finest watershed partition, the lowest level of the hierarchy. The graph representing the second level of the hierarchy is obtained by contracting all edges of the forest $F_1$; each tree becomes one node. The nodes are connected by edges of the graph $G$. The result of this contraction $G'_1 = \kappa(G, F_1)$ is again a connected graph, to which the same treatment as previously can be applied.
Construction of the level 2 of the hierarchy

Only the nodes of the graph $G'_1$ are weighted: $G'_1 = (e_1, \diamond)$. The nodes will be weighted with $\varepsilon_{ne}$ and we get the graph $(e_1, \varepsilon_{ne}e_1)$, which becomes a flooding graph of steepness $k$ in $\downarrow^k (e_1, \varepsilon_{ne}e_1)$. A final scissor operator creates a forest $F_2$ spanning nodes of $G'_1$. $F_2$ represents the level 2 of the hierarchy, and the nodes of $G$ spanned by each of its trees constitutes a catchment basin of level 2.

The contraction $\kappa(G'_1, F_1)$ produces again a connected graph $G'_2 = (e_1, \diamond)$ to which the same treatment may be applied.

This process is repeated until the graph $\kappa(G'_m, F_m)$ contains a unique regional minimum of edges.
Each new minimum spanning forest $F_n$ makes use of new edges of the initial graph. The union of all MSF constructed up to the iteration $m$, $\bigcup_{k \leq m} F_k$, is a minimum spanning forest of the initial graph $G$, which converges to a MST of $G$ as $m$ increases. Hence the union of all edges present in the series $F_n$ is a minimum spanning tree of the graph $G$. The particular minimum spanning tree which emerges depends upon the depth $k$ of the pruning operator $\downarrow^k$, and of the particular choices among alternative solutions made by the pruning operators $\chi$ used at each step. We call the operator extracting the minimum spanning tree $\mu(G)$.

The figure in the next slide presents a RAG, its flooding graph in the first row, and in the second row the MSFs $\bigcup_{k \leq m} F_k$ for $m = 1, 2, 3$. The last one being the MST.
Emergence of a minimum spanning tree

RAG Flooding □ graph
MSF □ level □ 1 MSF □ level □ 2 MSF □ level □ 3

Fernand Meyer Centre de Morphologie Mathématique Mines-ParisTech

Watersheds, waterfalls, on edge or node weighted graphs
2012 February 29 / 201
If $G$ and $G'$ are two flooding trees, $G'$ being a partial tree of $G$ included in $G$, then the pruning operators $\chi$ have less choices for pruning $G'$ as for pruning $G$. For this reason the family of MST derived for $G'$ is included in the family of MST derived for $G$: $\{\mu(G')\} \subset \{\mu(G)\}$.

In particular if one considers the decreasing sequence of minimum spanning forests $\chi \downarrow^k G$, one obtains decreasing families of MST: for $k < l$, we have $\{\mu(\downarrow^l G)\} \subset \{\mu(\downarrow^k G)\}$. The choices are more and more constrained. One gets minimum spanning trees which are steeper and steeper for increasing $k$ in $\downarrow^k G$. 
All MSTs of an edge weighted graph share a fundamental property: between any two nodes of the graph, there exists a unique path in the MST which links them, and this path has a minimal flooding weight. Thanks to this property, replacing the graph $G$ by its MST $T$ permits to construct the catchment basins linked to the lexdist$_1$. This procedure is fast but has no control on the quality of the result if one chooses an MST at random (the one which is produced by the preferred MST extraction algorithm). Fast, as there are as the number of edges is strongly reduced (in a tree: number of edges = number of nodes -1). Limiting, as one has no control on the steepness of the watershed which is ultimately extracted.
In order to get high quality partitions and segmentations, one has to carefully choose the MST. The preceding slide has shown that the MSTs form a decreasing family of trees as their steepness increases. Higher quality segmentations will be obtained for a higher steepness. In order to improve things, one may prune the graph $G$ with the operator $\zeta_k$ before extracting one of its MST, yielding a MST of $k$-steepness. One has then to consider a trade-off between the required speed and the quality of results one targets. This obviously also depends on the type of graphs. Some graphs contain only very few MST; in the extreme case, a graph where all edges have distinct weights, contains a unique MST.
Conclusion
Outcome

Starting with the flooding adjunction, we have introduced the flooding graphs, for which node and edge weights may be deduced one from the other.

Each node weighted or edge weighted graph may be transformed in a flooding graph, showing that there is no superiority in using one or the other, both being equivalent.

We have introduced pruning operators $\downarrow^k$ and $\zeta^{(k)}$ which extract subgraphs of increasing steepness. For an increasing steepness, the number of never ascending paths becomes smaller and smaller. This reduces the watershed zone, where catchment basins overlap.

The scissor operator $\chi$, associating to each node outside the regional minima one and only one edge choses a particular watershed partition. Again, with an increasing steepness, the number of equivalent solutions becomes smaller. Ultimately, for natural image, an infinite steepness leads to a unique solution, as it is not likely that two absolutely identical non ascending paths of infinite steepness connect a node with two distinct minima.
Finally, we have shown that the NAP paths remaining after the pruning $\downarrow^k$ or $\zeta^{(k)}$ are identical with geodesics of lexicographic distances. We have shown how the path algebra may be adapted to lexicographic distances. We have presented the Moore-Dijkstra algorithm for constructing skeletons by zone of influence for lexicographic distances. For the depth 1, the lexicographic distance becomes the flooding distance and the algorithms become classical algorithms for constructing minimum spanning forests. For the depth 2, it is equivalent with the topographic distance. The beneficial effect of hierarchical queue algorithms is highlighted, as it permits to correctly divide plateaus among neighboring catchment basins. We also presented the core expanding algorithms which are particularly efficient for lexicographic distances.
The waterfall hierarchy is obtained by contracting the trees of a first watershed construction and the new graph submitted to a novel watershed construction. The process is iterated until only one region remains. The union of the edges of all forests produces constitute a minimum spanning tree of the initial graph.

MST and MSF defined before only for edge weighted graphs may be extended to node weighted graphs.

MST and MSF may be ordered into nested classes according to their steepness.

The classical order is thus inverted: classically a MST is extracted from the graph as a first step, and a MSF derived from it, by cutting a number of its edges. This way of doing leaves not much control as the quality of the result depends upon the choice of the MST. As MST and MSF may be ordered into nested classes according to their steepness.
The algebraic approach to the watershed presented here sets the stage in which a number of earlier definitions and algorithms may be reinterpreted:

- the waterfall algorithm [3]
- the watershed line defined a zone of influence of the minima for various distances, in particular the flooding distance and the topographic distance [19],[23]. In the case of the flooding distance applied on an edge weighted graph, one finds the graph cuts [9], and the algorithm for constructing a waterfall hierarchy described in [16]
- the role of the hierarchical queues for a correct division of plateaus between catchment basins if one uses myopic distances [17]

We also explored the place of the choice for constructing a watershed partition. The choice becomes more and more restricted as one considers lexicographic distances with increasing depth. The waterfall hierarchy transposed to MST provides a hierarchical decomposition of the MST.
Old algorithms may be reinterpreted, in particular those relying on pruning and labeling graphs \[7, 15, 24, 28\]. The classical algorithm for constructing watersheds derived from the hierarchical queues is a particular case of core expanding algorithm. In order to reduce the number of equivalent watershed partitions one is able to extract from a flooding graph, one may use a mixture of pruning \(\zeta^{(k)}\) up to a given depth, and on the resulting graph apply a myopic algorithm, with a lexicographic depth of 1 or 2.
The watershed transform introduced by S. Beucher and C. Lantuejoul [4] is one of the major image segmentation tools, used in the community of mathematical morphology and beyond. If the watershed is a successful concept, there is another side of the coin: a number of definitions and algorithms coexist, claiming to construct a watershed line or catchment basins, although they obviously are not equivalent. We have presented how the idea was conceptualized and implemented as algorithms or hardware solutions in a brief note:” The watershed concept and its use in segmentation: a brief history” (arXiv:1202.0216v1), which contains an extensive bibliography. Here we give a more restricted list of references.
[1] C. Berge.  
*Graphs.*  
Amsterdam: North Holland, 1985.

[2] S. Beucher.  
*Segmentation d’Images et Morphologie Mathématique.*  
PhD thesis, E.N.S. des Mines de Paris, 1990.

[3] S. Beucher.  
Watershed, hierarchical segmentation and waterfall algorithm.  
*ISMM94 : Mathematical Morphology and its applications to Signal Processing,* pages 69–76, 1994.

[4] S. Beucher and C. Lantuéjoul.  
Use of watersheds in contour detection.  
In *Proc. Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation,* 1979.

[5] S. Beucher and C. Lantuéjoul.  
Use of watersheds in contour detection.
[6] S. Beucher and F. Meyer.
The morphological approach to segmentation: the watershed transformation.
In E. Dougherty, editor, *Mathematical morphology in image processing*, chapter 12, pages 433–481. Marcel Dekker, 1993.

[7] Moga A. Bieniek, A.
An efficient watershed algorithm based on connected components. *Pattern Recognition*, 33(6):907–916, 2000.
cited By (since 1996) 78.

[8] Gunilla Borgefors.
Distance transformations in digital images. *Comput. Vision Graph. Image Process.*, 34:344–371, June 1986.

[9] Jean Cousty, Gilles Bertrand, Laurent Najman, and Michel Couprie.
Watershed cuts: Minimum spanning forests and the drop of water principle.
[10] M. Gondran and M. Minoux.
*Graphes et Algorithmes.*
Eyrolles, 1995.

[11] H. Heijmans.
*Morphological Image Operators, Advances in Electronics and Electron Physics, Supplement 24, Editor-in-Chief P. Hawkes.*
Boston: Academic Press, 1994.

[12] H. Heijmans and C. Ronse.
The algebraic basis of mathematical morphology, part I: Dilations and erosions.
*Comput. Vision, Graphics and Image Processing,* 50:245–295, 1990.

[13] C. Lantuéjoul.
*La squelettisation et son application aux mesures topologiques de mosaïques polycristallines.*
PhD thesis, École nationale supérieure des mines de Paris, 1978.
[14] C. Lantuéjoul and S. Beucher.
On the use of the geodesic metric in image analysis.
*J. Microsc.*, 1981.

[15] F. Lemonnier.
*Architecture Electronique Dédiée aux Algorithmes Rapides de Segmentation Basés sur la Morphologie Mathématique.*
PhD thesis, E.N.S. des Mines de Paris, 1996.

[16] B. Marcotegui and S. Beucher.
Fast implementation of waterfalls based on graphs.
*ISMM05 : Mathematical Morphology and its applications to Signal Processing*, pages 177–186, 2005.

[17] F. Meyer.
Un algorithme optimal de ligne de partage des eaux.
In *Proceedings 8ème Congrès AFCET, Lyon-Villeurbanne*, pages 847–857, 1991.

[18] F. Meyer.
Minimal spanning forests for morphological segmentation.
[19] F. Meyer.
Topographic distance and watershed lines.
*Signal Processing*, pages 113–125, 1994.

[20] F. Meyer and S. Beucher.
Morphological segmentation.
*JVCIP*, 1(1):21–46, Sept. 1990.

[21] Fernand Meyer.
Grey-weighted, ultrametric and lexicographic distances.
In Christian Ronse, Laurent Najman, and Etienne Decenciã¨re, editors, *Mathematical Morphology: 40 Years On*, volume 30 of *Computational Imaging and Vision*, pages 289–298. Springer Netherlands, 2005.

[22] E.F. Moore.
The shortest path through a maze.
[23] Laurent Najman and Michel Schmitt.
Watershed of a continuous function.
*Signal Processing*, 38(1):99 – 112, 1994.
Mathematical Morphology and its Applications to Signal Processing.

[24] D. Noguet.
A massively parallel implementation of the watershed based on cellular automata.
In *Application-Specific Systems, Architectures and Processors, 1997. Proceedings., IEEE International Conference on*, pages 42 –52, jul 1997.

[25] Jos B. T. M. Roerdink and Arnold Meijster.
The watershed transform: Definitions, algorithms and parallelization strategies.
*Fundamenta Informaticae*, 41:187–228, 2001.

[26] J. Serra, editor.
[27] Soille Pierre Vincent, Luc.
Watersheds in digital spaces: An efficient algorithm based on immersion simulations.
*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, 1991.

[28] Iwanowski M. Aswiercz, M.
Fast, parallel watershed algorithm based on path tracing.
*Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6375 LNCS(PART 2):317–324, 2010.