# The Generic Chaining

Michel Talagrand

ISBN 3-540-24518-9

The Appendix was not published in the book.

Therefore, the Appendix is shown on the next pages.

# A. Appendix: Trees and Majorizing Measures

In this appendix we describe different ways to measure the size of a metric space. Some of these ways played an important part in the development of the theory. We will show that they are all equivalent to the functional $\gamma_2(T,d)$. It is possible to consider more general notions corresponding to other functionals considered in the book, but for simplicity we consider only the case of $\gamma_2$.

A *tree* $\mathcal{T}$ of a metric space $(T,d)$ is a *finite* collection of subsets of $T$ with the following two properties.

$$\text{Given } A, B \text{ in } \mathcal{T}\,,\ \text{ if } A \cap B \neq \emptyset\,,\ \text{ then either } A \subset B \text{ or else } B \subset A\,. \tag{A.1}$$

$$\mathcal{T} \text{ has a largest element}\,. \tag{A.2}$$

If $A$, $B \in \mathcal{T}$ and $B \subset A$, $B \neq A$, we say that $B$ is a *child* of $A$ if

$$C \in \mathcal{T}\,,\, B \subset C \subset A \Rightarrow C = B \text{ or } C = A\,. \tag{A.3}$$

We denote by $c(A)$ the number of children of $A$. We will consider only trees with the following property

$$\text{If } A \in \mathcal{T} \text{ and } \ c(A) = 0\,, \text{ then } A \text{ contains exactly one point}\,. \tag{A.4}$$

A *separated* tree is a tree $\mathcal{T}$ such that to each $A$ in $\mathcal{T}$ with $c(A) \geq 1$ is associated an integer $s(A) \in \mathbb{Z}$ with the following properties.

$$\text{If } B_1 \text{ and } B_2 \text{ are children of } A, \text{then } d(B_1, B_2) \geq 4^{-s(A)}\,. \tag{A.5}$$

$$\text{If } B \text{ is a child of } A\,, \text{ then } s(B) > s(A)\,. \tag{A.6}$$

Here of course $d(B_1, B_2) = \inf\{d(x_1, x_2)\,,\, x_1 \in B_1\,,\, x_2 \in B_2\}$. We then define

$$S_{\mathcal{T}} = \{t \in T\ ;\ \{t\} \in \mathcal{T}\}$$

and the *depth* of $\mathcal{T}$,

$$d(\mathcal{T}) = \inf_{t \in S_{\mathcal{T}}} \sum_{t \in A \in \mathcal{T}} 4^{-s(A)} \sqrt{\log c(A)}\,.$$

Here and below we make the convention that the summation does not include the term $A = \{t\}$ (for which $c(A) = 0$). We view $d(\mathcal{T})$ as a “measure of the size" of the separated tree $\mathcal{T}$. We can then measure the size of $T$ by $\sup\{d(\mathcal{T})\,;\, \mathcal{T} \text{ separated tree}\}$.

The notion of tree we just considered is but one of many possible. Let us now consider another (more restrictive) notion. An *organized* tree is a tree $\mathcal{T}$ such that to each $A \in \mathcal{T}$ with $c(A) \geq 1$ are associated $j = j(A) \in \mathbb{Z}$, $t \in T$ and $t_1, \cdots, t_{c(A)} \in B(t, 4^{-j})$ with the properties that

$$1 \leq \ell < \ell' \leq c(A) \Rightarrow d(t_\ell, t_{\ell'}) \geq 4^{-j-1}$$

and that each ball $B(t_\ell, 4^{-j-2})$ contains exactly one child of $A$.

If $B_1$ and $B_2$ are children of $A$, then

$$d(B_1, B_2) \geq 4^{-j(A)-2}\,, \tag{A.7}$$

so that an organized tree is also a separated tree (with $s(A) = j(A) + 2$), but the notion of organized tree is more restrictive. (For example we have no control of the diameter of the children of $A$ in a separated tree.)

We define the depth $d'(\mathcal{T})$ of an organized tree by

$$d'(\mathcal{T}) = \inf_{t \in S_\mathcal{T}} \sum_{t \in A \in \mathcal{T}} 4^{-j(A)} \sqrt{\log c(A)}\ .$$

If we simply view $\mathcal{T}$ as a separated tree using (A.7), then $d(\mathcal{T}) = d'(\mathcal{T})/16$ (where $d(\mathcal{T})$ is the depth of $\mathcal{T}$ as a separated tree). Thus we have shown the following.

**Proposition A.1.** *We have*

$$\sup\{d'(\mathcal{T})\ ;\ \mathcal{T} \text{ organized tree}\} \leq 16 \sup\{d(\mathcal{T})\ ;\ \mathcal{T} \text{ separated tree}\}\ . \tag{A.8}$$

**Proposition A.2.** *We have*

$$\gamma_2(T, d) \leq L \sup\{d'(\mathcal{T})\ ;\ \mathcal{T} \text{ organized tree}\}\ . \tag{A.9}$$

*Proof.* We consider the functional

$$F_n(A) = F(A) = \sup\{d'(\mathcal{T})\ ;\ \mathcal{T} \subset A\,,\, \mathcal{T} \text{ organized tree}\}\,,$$

where we write $\mathcal{T} \subset A$ as a shorthand for "$\forall B \in \mathcal{T}\,,\, B \subset A$". In the course of the proof of Theorem 1.3.1 we have noted that this theorem holds true as soon as (1.31) holds true when $a$ is of the type $r^{-j-1}$. We check this condition when $r = 4$, $\theta(n) = 2^{n/2-2}, \beta = 1$, and $\tau = 1$. Consider $n \geq 0$ and $m = N_{n+1}$. Consider $j \in \mathbb{Z}$, $t \in T$ and $t_1, \cdots, t_m \in B(t, 4^{-j})$ with

$$1 \leq \ell < \ell' \leq m \Rightarrow d(t_\ell, t_{\ell'}) \geq 4^{-j-1}\ .$$

Consider sets $H_\ell \subset B(t_\ell, 4^{-j-2})$ and $c < \min_{\ell \leq m} F(H_\ell)$. Consider, for $\ell \leq m$ a tree $\mathcal{T}_\ell \subset H_\ell$ with $d'(\mathcal{T}_\ell) > c$ and denote by $A_\ell$ its largest element. Then it should be obvious that the tree $\mathcal{T}$ consisting of $U = \bigcup_{\ell \leq m} H_\ell$ (its largest element) and the reunion of the trees $\mathcal{T}_\ell$, $\ell \le m$, is organized (with $j(U) = j$, and $A_1, \cdots, A_m$ as children of $U$). Moreover $S_\mathcal{T} = \bigcup_{\ell \le m} S_{\mathcal{T}_\ell}$.

Consider $t \in S_\mathcal{T}$, and let $\ell$ with $t \in S_{\mathcal{T}_\ell}$. Then

$$\begin{aligned}\sum_{t \in A \in \mathcal{T}} 4^{-j(A)} \sqrt{\log c(A)} &= 4^{-j} \sqrt{\log m} + \sum_{t \in A \in \mathcal{T}_\ell} 4^{-j(A)} \sqrt{\log c(A)} \\ &\ge 4^{-j} \sqrt{\log m} + d'(\mathcal{T}_\ell) \ge 4^{-j} \sqrt{\log m} + c \,.\end{aligned}$$

Since $\sqrt{\log m} \ge 2^{n/2}$, this proves (1.31).

To prove (A.9) we apply Lemma 1.3.3. To control the diameter of $T$, we simply note that if $s$, $t \in T$, and $j$ is the largest integer with $4^{-j} \ge d(s,t)$, then the tree $\mathcal{T}$ consisting of $T$, $\{t\}$, $\{s\}$, is organized with $j = j(T)$ and $c(T) = 2$, so $d'(\mathcal{T}) \ge 4^{-j}\sqrt{\log 2}$. $\square$

**Proposition A.3.** *Given a metric space $(T,d)$ we can find on $T$ a probability measure $\mu$, supported by a countable subset of $T$, and such that*

$$\sup_{t \in T} \int_0^\infty \sqrt{\frac{1}{\log \mu(B(t,\epsilon))}} \mathrm{d}\epsilon \le L\gamma_2(T,d) \,. \tag{A.10}$$

A probability measure $\mu$ on $(T,d)$ such that the left-hand side of (A.10) is usefully small is called a majorizing measure. The (in)famous theory of majorizing measures used the infimum of the left-hand side of (A.10) over all choices of $\mu$ as a measure of the size of the metric space $(T,d)$.

*Proof.* Consider an admissible sequence $(\mathcal{A}_n)$ with

$$\forall t \in T\,, \sum_{n \ge 0} 2^{n/2} \Delta(A_n(t)) \le 2\gamma_2(T,d) \,.$$

Since card $\mathcal{A}_n \le N_n$, there is a probability measure $\mu$ on $T$, supported by a countable set, and satisfying

$$\forall n \ge 1\,, \forall A \in \mathcal{A}_n\,, \mu(A) \ge \frac{1}{2^n N_n} \ge \frac{1}{N_n^2}$$

so that given $t \in T$

$$\begin{aligned}\epsilon > \Delta(A_n(t),d) \Rightarrow \mu(B(t,\epsilon)) &\ge \frac{1}{N_n^2} \\ \Rightarrow \sqrt{\log \frac{1}{\mu(B(t,\epsilon))}} &\le 2^{n/2+1} \,.\end{aligned} \tag{A.11}$$

Now, since $\mu$ is a probability, $\mu(B(t,\epsilon)) = 1$ for $\epsilon > \Delta(T,d)$, and thus $\log(1/\mu(B(t,\epsilon)) = 0$. Thus

$$\begin{aligned}\int_0^\infty \sqrt{\log \frac{1}{\mu(B(t,\epsilon))}}\mathrm{d}\epsilon &= \sum_{n\geq 1}\int_{\Delta(A_n(t))}^{\Delta(A_{n-1}(t))} \sqrt{\log \frac{1}{\mu(B(t,\epsilon))}}\mathrm{d}\epsilon \\ &\leq \sum_{n\geq 1} 2^{n/2+1}\Delta(A_{n-1}(t)) \leq L\gamma_2(T,d)\end{aligned}$$

using (A.11). □

**Proposition A.4.** *If $\mu$ is a probability measure on $T$, (supported by a countable set) and $\mathcal{T}$ is a separated tree on $T$, then*

$$d(\mathcal{T}) \leq L \sup_{t\in T}\int_0^\infty \sqrt{\log \frac{1}{\mu(B(t,\epsilon))}}\mathrm{d}\epsilon\ .$$

This completes the proof that the 4 "measures of the size of $T$" considered in this appendix are indeed equivalent.

*Proof.* The basic observation is as follows. The sets

$$B(C, 4^{-s(A)-1}) = \{x \in T\ ;\ d(x,C) \leq 4^{-s(A)-1}\}$$

are disjoint as $C$ varies over the children of $A$ (as follows from (A.5)). So one of them has measure $\leq c(A)^{-1}$.

We then proceed in the following manner. We start with the largest element $A_0$ of $\mathcal{T}$. We then select a child $A_1$ of $A_0$ with $\mu(B(A_1, 4^{-s(A_0)-1})) \leq 1/c(A_0)$, and a child $A_2$ of $A_1$ with $\mu(B(A_2, 4^{-s(A_1)-1})) \leq 1/c(A_1)$, etc., and continue this construction as long as we can. It ends only when we reach a set of $\mathcal{T}$ that has no child, and hence by (A.4) is reduced to a single point $t$. If $t \in A \in \mathcal{T}$, by construction we have

$$\mu(B(t, 4^{-s(A)-1})) \leq \frac{1}{c(A)}$$

so that

$$4^{-s(A)-2}\sqrt{\log c(A)} \leq \int_{4^{-s(A)-2}}^{4^{-s(A)-1}} \sqrt{\frac{1}{\log \mu(B(t,\epsilon))}}\mathrm{d}\epsilon\ . \tag{A.12}$$

By (A.6) the intervals $]4^{-s(A)-2}, 4^{-s(A)-1}[$ are disjoint for different sets $A$ with $t \in A \in \mathcal{T}$, so summation of the inequalities (A.12) yields

$$\frac{1}{16}d(\mathcal{T}) \leq \sum_{t\in A\in\mathcal{F}} 4^{-s(A)-2}\sqrt{\log c(A)} \leq \int_0^\infty \sqrt{\frac{1}{\log \mu(B(t,\epsilon))}}\mathrm{d}\epsilon\ .$$

□

Springer

http://www.springer.com/978-3-540-24518-6

The Generic Chaining
Upper and Lower Bounds of Stochastic Processes
Talagrand, M.
2005, VIII, 222 p., Hardcover
ISBN: 978-3-540-24518-6