



HAL
open science

A simple master Theorem for discrete divide and conquer recurrences

Olivier Garet

► **To cite this version:**

Olivier Garet. A simple master Theorem for discrete divide and conquer recurrences. 2019. hal-02049382v1

HAL Id: hal-02049382

<https://hal.science/hal-02049382v1>

Preprint submitted on 26 Feb 2019 (v1), last revised 15 Feb 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A SIMPLE MASTER THEOREM FOR DISCRETE DIVIDE AND CONQUER RECURRENCES

OLIVIER GARET

ABSTRACT. The aim of this note is to provide a Master Theorem for some discrete divide and conquer recurrences:

$$X_n = a_n + \sum_{j=1}^m b_j X_{\lfloor p_j n \rfloor},$$

where the p_i 's belong to $(0, 1)$. The main novelty of this work is there is no assumption of regularity or monotonicity for (a_n) . Then, this result can be applied to various sequences of random variables $(a_n)_{n \geq 0}$, for example such that $\sup_{n \geq 1} \mathbb{E}(|a_n|) < +\infty$.

1. INTRODUCTION

Divide-and-conquer methods are widely used in Computer Science. The analysis of the cost of the algorithm naturally leads to divide-and-conquer recurrences. The methods to study these recurrences are popularized as “Master theorems” in the litterature of Computer Science. See e.g. the reference books by Cormen et al [3] or Goodrich and Tamassia [6].

In the sequel, we consider sequences $(X_n)_{n \geq 0}$ that are defined by $X_0 = a_0$, then

$$(1) \quad X_n = a_n + \sum_{j=1}^m b_j X_{\lfloor p_j n \rfloor},$$

where the p_i 's belong to $(0, 1)$ and $\lfloor x \rfloor$ denotes the only $n \in \mathbb{Z}$ such that $x - n \in [0, 1)$.

Of course, in Computer Science, a_n and X_n represent computation times and are therefore positive. However, the case of negative a_n and X_n can be of theoretical interest.

In the litterature of Computer Science, (a_n) is supposed to be deterministic. Nevertheless, in the context of randomized algorithm, eventually involving Monte-Carlo simulation, it is natural to consider the case of a random (a_n) and observe the fluctuations of the computation time.

One of the most general results in the field of Computer Science is due to Akra and Bazzi [1]. They do not seek for an exact asymptotic limit, focusing of the order of the fluctuations. Their methods rely on classical real analysis.

The mathematical litterature is more focused on exact methods, that rely on generating functions. The first paper in this spirit is Erdős et al [5], which solved the case $a_n = 0$ with the help of renewal equations. Tauberian theorems lead to simpler proofs of their result, see e.g. Choimet and Queffelec [2]. Recent results

2010 *Mathematics Subject Classification.* 11B37,68W40,60F15.

Key words and phrases. Divide-and-conquer recurrence, Dirichlet series, Tauberian theorem.

by Drmota and Szpankowski [4]) also rely on Tauberian theorems and some other tools in complex analysis. They request some assumptions of monotonicity.

If one wants to cover the case of a random (a_n) , the sequence (a_n) obviously can not be supposed to be monotonic. Quite surprisingly, we did not find in the litterature any theorem of this kind, computing an exact limit without making some assumption of monotonicity.

Let us precise the assumptions: we assume that the b_i 's and the p_i 's are positive numbers with $\sum_{j=1}^m b_j > 1$ and such that there exists j, ℓ with $\frac{\log p_j}{\log p_\ell} \notin \mathbb{Q}$. The rational case, which is not considered here, is also of great interest in Computer Science – see e.g. Roura [7] or Drmota and Szpankowski [4].

It is known that the general growth of (X_n) is governed by the value of the positive root s_0 for the equation

$$\sum_{j=1}^m b_j p_j^s = 1.$$

As said before, the originality of the present paper lies in the assumption on the (a_n) : under the assumption that

$$\sum_{n=1}^{+\infty} \frac{|a_n|}{n^{s_0+1}} < +\infty,$$

we prove that the sequence $\frac{X_n}{n^{s_0}}$ admits a limit L when n tends to infinity and give a fairly simple closed expression for it.

As we will see, this allow to apply our Theorem to a large class of random variables. Then, the limit L is a random variable, which appears as the sum of a random series.

If we specialize to the case where the (a_n) are independent, then one can easily control the random fluctuations of L .

2. THE DETERMINISTIC THEOREM

Theorem 1. *Let $m \geq 1$, (b_1, \dots, b_m) be a family of non-negative numbers and (p_1, \dots, p_m) be a family of non-negative numbers in $(0, 1)$ such that*

- *there exists j, ℓ with $\frac{\log p_j}{\log p_\ell} \notin \mathbb{Q}$;*
- *$\sum_{j=1}^m b_j > 1$.*

We denote by s_0 the positive root s_0 for the equation

$$\sum_{j=1}^m b_j p_j^s = 1.$$

Then, there exist a sequence $(\ell_j)_{j \geq 0}$ such that for every sequence $(a_n)_{n \geq 0}$ with

$$\sum_{n=1}^{+\infty} \frac{|a_n|}{n^{s_0+1}} < +\infty,$$

then the sequence $(X_n)_{n \geq 0}$ defined by $X_0 = a_0$ and the recursion (1) satisfies

$$\lim_{n \rightarrow +\infty} \frac{X_n}{n^{s_0}} = \sum_{n=0}^{+\infty} \ell_n a_n.$$

Proof. We denote by $L_n(a)$ the value of X_n corresponding to the recursion (1) for some sequence a .

The recursion equation. Let n_0 be a non-negative integer and suppose first that $a_n = 0$ for $n > n_0$.

For $n > n_0$, we have $X(n) = \sum_{j=1}^m b_j X(\lfloor p_j n \rfloor)$.

We can choose C such that $|X_k| \leq Ck^{s_0}$ for $0 < k \leq n_1 = \max(n_0, 1/p_1, \dots, 1/p_m)$. Then, it follows by natural induction that $|X_k| \leq Xk^{s_0}$ for each $k \in \mathbb{N}^*$. In the sequel, we put $X(t) = X(\lfloor t \rfloor)$ to simplify some notation. Now define

$$(2) \quad \phi(s) = s \int_{n_0+1}^{+\infty} \frac{X(t)}{t^{s+1}} dt$$

for $s \in \mathbb{C}$ with $\operatorname{Re}(s) > s_0$. The recursion Equation leads to

$$\begin{aligned} \phi(s) &= s \int_{n_0+1}^{+\infty} \sum_{j=1}^m b_j \frac{X(p_j x)}{t^{s+1}} dt = s \sum_{j=1}^m b_j p_j^s \int_{(n_0+1)p_j}^{+\infty} \frac{X(t)}{t^{s+1}} dt \\ &= \left(\sum_{j=1}^m b_j p_j^s \right) \phi(s) + s \sum_{j=1}^m b_j p_j^s \int_{(n_0+1)p_j}^{n_0+1} \frac{X(t)}{t^{s+1}} dt. \end{aligned}$$

So we have

$$(3) \quad \phi(s) = \frac{P(s)}{1 - \sum_{j=1}^m b_j p_j^s}, \quad \text{with } P(s) = s \sum_{j=1}^m b_j p_j^s \int_{(n_0+1)p_j}^{n_0+1} \frac{X(t)}{t^{s+1}} dt$$

Tauberian magic. Suppose here that $a = I_{n_0}$ with $I_{n_0}(i) = \mathbb{1}_{i \leq n_0}$. By natural induction, it is easy to see that $(X_n)_{n \geq 0}$ is non-decreasing.

Following the reasoning by Choimet and Queffelec [2], we see that $1 - \sum_{j=1}^m b_j p_j^s$ does not vanish for $s \in \mathbb{C}$ with $\operatorname{Re}(s) \geq s_0$ and $s \neq s_0$. It follows that for

$$c = \operatorname{Res}_{s_0} \phi = \frac{P(s_0)}{-\sum_{j=1}^m b_j p_j^{s_0} \log(p_j)},$$

the map $s \mapsto \phi(s) - \frac{c}{s-s_0}$ is holomorphic on $\{s \in \mathbb{C}; \operatorname{Re}(s) \geq s_0\}$.

Now note $b(x) = \sum_{n_0 < n \leq x} X_n - X_{n-1}$. The Abel transformation gives

$$\sum_{n=n_0+1}^{+\infty} \frac{X_n - X_{n-1}}{n^s} = s \int_{n_0+1}^{+\infty} \frac{b(t)}{t^{s+1}} dt$$

Since $b(t) = X(t) - X_{n_0}$, we have

$$\sum_{n=n_0+1}^{+\infty} \frac{X_n - X_{n-1}}{n^s} = s \int_{n_0+1}^{+\infty} \frac{X(t)}{t^{s+1}} dt - \frac{X_{n_0}}{(n_0+1)^s} = \phi(s) - \frac{X_{n_0}}{(n_0+1)^s}.$$

Since $(X_n)_{n \geq 0}$ is non-decreasing, the sequence $(X_n - X_{n-1})_{n > n_0}$ is non-negative, so the Wiener-Ikehara Theorem for series applies: since $b(t) = O(t^{s_0})$ when $t \rightarrow +\infty$, we get $b(x) \sim \frac{c}{s_0} x^{s_0}$, so

$$\lim_{n \rightarrow +\infty} \frac{L_n(I_{n_0})}{n^{s_0}} = \frac{\sum_{j=1}^m b_j p_j^{s_0} \int_{(n_0+1)p_j}^{n_0+1} \frac{L_t(I_{n_0})}{t^{s_0+1}} dt}{\sum_{j=1}^m b_j p_j^{s_0} \log(1/p_j)}.$$

For $n_0 = 0$, we have

$$\begin{aligned} \ell_0 &= \lim_{n \rightarrow +\infty} \frac{L_n(\delta_0)}{n^{s_0}} = \lim_{n \rightarrow +\infty} \frac{L_n(I_0)}{n^{s_0}} = \frac{1}{\sum_{j=1}^m b_j p_j^{s_0} \log(1/p_j)} \sum_{j=1}^m b_j p_j^{s_0} \int_{p_j}^1 \frac{1}{t^{s_0+1}} dt \\ &= \frac{1}{\sum_{j=1}^m b_j p_j^{s_0} \log(1/p_j)} \sum_{j=1}^m b_j p_j^{s_0} \frac{p_j^{-s_0} - 1}{s_0} \end{aligned}$$

Note that this equality and the related convergence form the result by Erdős et al [5].

Let $n_0 \geq 1$. Since $\delta_{n_0} = I_{n_0} - I_{n_0-1}$, it follows that

$$L_n(\delta_{n_0})n^{-s_0} = L_n(I_{n_0})n^{-s_0} - L_n(I_{n_0-1})n^{-s_0}$$

has a limit when n tends to infinity. Let us denote it by ℓ_{n_0} .

To compute it, take $a = \delta_{n_0}$ and consider again the associated ϕ . From (2), we get $\ell_{s_0} = \frac{1}{s_0} \lim_{s \rightarrow s_0^+} \frac{1}{s_0} (s - s_0) \phi(s)$. On the other side, Equation (3) is still valid, with

$$\begin{aligned} P(s) &= s \sum_{j=1}^m b_j p_j^s \int_{(n_0+1)p_j}^{n_0+1} \frac{X(t)}{t^{s+1}} dt \\ &= s \sum_{j=1}^m b_j p_j^s \int_{\max(n_0, (n_0+1)p_j)}^{n_0+1} \frac{1}{t^{s+1}} dt, \end{aligned}$$

also

$$\frac{1}{s_0} (s - s_0) \phi(s) = -\frac{s}{s_0} \frac{s_0 - s}{1 - \sum_{j=1}^m b_j p_j^s} \sum_{j=1}^m b_j p_j^s \int_{\max(n_0, (n_0+1)p_j)}^{n_0+1} \frac{1}{t^{s+1}} dt$$

and

$$\ell_{n_0} = \frac{1}{\sum_{j=1}^m b_j p_j^{s_0} \log(1/p_j)} \sum_{j=1}^m b_j p_j^{s_0} \int_{\max(n_0, (n_0+1)p_j)}^{n_0+1} \frac{1}{t^{s_0+1}} dt.$$

With the supplementary assumption that $p_i \leq 1/2$ for each i , we simply have

$$\ell_{n_0} = \frac{1}{\sum_{j=1}^m b_j p_j^{s_0} \log(1/p_j)} \sum_{j=1}^m b_j p_j^{s_0} \int_{n_0}^{n_0+1} \frac{1}{t^{s_0+1}} dt.$$

The general case. For $n, j \geq 0$, we note $K_n^j = L_n(\delta_j)$. It is obvious that $K_n^j = 0$ for $n < j$ and $K_j^j = 1$. It easily follows by natural induction on n that $0 \leq K_n^j \leq \frac{K_n^0}{K_j^0}$. Now, the affine nature of the recursion gives

$$X_n = \sum_{j=0}^n K_n^j a_j$$

For each $j \geq 0$, we have $\lim_{n \rightarrow +\infty} \frac{K_n^j}{n^{s_0}} = \ell_j$. Moreover, for each $j, n \geq 0$, we have

$$\left| \frac{K_n^j a_j}{n^{s_0}} \right| \leq \frac{K_n^0}{n^{s_0}} \frac{|a_j|}{K_j^0} \leq M \frac{|a_j|}{(j+1)^{s_0}}$$

and by the Weierstrass criterion,

$$\lim_{n \rightarrow +\infty} \frac{X_n}{n^{s_0}} = \sum_{j=0}^{+\infty} \ell_j a_j.$$

□

3. APPLICATION TO SEQUENCES OF RANDOM VARIABLES

We give below some applications of Theorem 1 to sequences of random variables.

Convergence.

Theorem 2. *Assume that the p_i 's, the b_i 's and s_0 fulfill the assumptions of Theorem 1 and (a_n) is a sequence of random variables. Under each of the following sets of supplementary assumptions, the sequence $(X_n)_{n \geq 0}$ defined by $X_0 = a_0$ and the recursion (1) is such that $\frac{X_n}{n^{s_0}}$ almost surely converges to some random variable:*

- (A) *The (a_n) are integrable random variables with $C = \sup_{n \geq 1} \mathbb{E}|a_n| < +\infty$.*
- (B) *$\sum_{j=1}^m b_j p_j > 1$ and the a_n are random variables following the Cauchy $\mathcal{C}(0, 1)$ Law.*

Proof. (A) If the (a_n) are integrable random variables with $C = \sup_{n \geq 1} \mathbb{E}|a_n| < +\infty$, then $\mathbb{E}(\sum_{n=1}^{+\infty} \frac{|a_n|}{n^{s_0+1}}) \leq C\zeta(s_0 + 1) < +\infty$, so $\sum_{n=1}^{+\infty} \frac{|a_n|}{n^{s_0+1}} < +\infty$ almost surely, which gives the almost sure behavior of $\frac{X_n}{n^{s_0}}$.

(B) If $\sum_{j=1}^m b_j p_j > 1$ and the a_n are random variables following the Cauchy $\mathcal{C}(0, 1)$ Law, then the condition $\sum_{j=1}^m b_j > 1$ implies that $s_0 > 1$. Let fix η with $s_0 > \eta > 1$. Then $\mathbb{P}(|a_n| > n^\eta) = O(n^{-\eta})$ and $\sum_{n=1}^{+\infty} \mathbb{P}(|a_n| > n^\eta) < +\infty$, so by the Borel-Cantelli Lemma, for almost every ω , there exists $n_0(\omega)$ with $|a_n(\omega)| \leq n^\eta$ for $n \geq n_0(\omega)$, which gives the convergence of $\sum_{n \geq 1} \frac{|a_n|}{n^{s_0+1}}$ and our Master Theorem still applies.

□

Exponential moments.

Theorem 3. *Assume that the p_i 's, the b_i 's and s_0 fulfill the assumptions of Theorem 1 and (a_n) is a sequence of independent random variables. The sequence $(X_n)_{n \geq 0}$ is defined by $X_0 = a_0$ and the recursion (1).*

- *If there exists a distribution μ with exponential moments such that $|a_n|$ is stochastically dominated by μ^{*n} for each $n \geq 0$, then $|X_n|$ as exponential moments for each n .*
- *If $s_0 > 1$ (or equivalently $\sum_{j=1}^m b_j p_j > 1$) and there exists a distribution μ with exponential moments such that $|a_n|$ is stochastically dominated by μ for each $n \geq 0$, then $\frac{X_n}{n^{s_0}} \rightarrow L$ a.s. where $|L|$ has exponential moments.*
- *If $s_0 > 2$ (or equivalently $\sum_{j=1}^m b_j p_j^2 > 1$) and there exists a distribution μ with exponential moments such that $|a_n|$ is stochastically dominated by μ^{*n} for each $n \geq 0$, then $\frac{X_n}{n^{s_0}} \rightarrow L$ a.s. where $|L|$ has exponential moments.*

Proof. We begin with an easy lemma:

Lemma 1. *Let X be a random variable with $\mathbb{E}(e^{\alpha X}) < +\infty$ and Y a random variable following the exponential law $\mathcal{E}(\alpha)$ Then, for $a = \frac{1}{\alpha} \log \mathbb{E}(e^{\alpha X_1})$, we have the stochastic domination $X \prec Y + a$.*

Proof. We just have to prove that for $t \in \mathbb{R}$, $\mathbb{P}(X \geq t) \leq \mathbb{P}(Y + a \geq t)$, or equivalently $\mathbb{P}(X \geq t) \leq \mathbb{P}(Y \geq t - a)$. For $t \leq a$, we have $\mathbb{P}(X \geq t) \leq 1 = \mathbb{P}(Y \geq t - a)$. For $t \geq a$, the Markov inequality gives

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}e^{\alpha X}}{e^{\alpha t}} = \frac{e^{\alpha a}}{e^{\alpha t}} = \exp(-\alpha(t - a)) = \mathbb{P}(Y \geq t - a).$$

This completes the proof. \square

Now, we have a and α such that for each $n \geq 1$

$$|a_n| \prec \mu^{*n} \prec (\delta_a * \mathcal{E}(\alpha))^{*n} = \delta_{na} * \Gamma(n, \theta).$$

Let $(Z_n)_{n \geq 0}$ be a sequence of independent variables with $Z_n \sim \Gamma(n, \theta)$, where $\Gamma(a, \gamma)$ is the Law with the density

$$x \mapsto \frac{\gamma^a}{\Gamma(a)} x^{a-1} e^{-\gamma x} \mathbb{1}_{]0, +\infty[}(x).$$

$\frac{|X_n|}{n^{s_0}}$ is stochastically dominated by

$$M \sum_{j=0}^n \frac{ja + Z_j}{(j+1)^{s_0}},$$

so for $t < 1/\alpha$, we have

$$\begin{aligned} \mathbb{E}(e^{t \frac{|X_n|}{n^{s_0}}}) &\leq \exp(Ma \sum_{j=1}^{n+1} j^{-s_0}) \prod_{j=0}^n \mathbb{E} \exp\left(\frac{tZ_j}{(j+1)^{s_0}}\right) \\ &\leq \exp(Ma \sum_{j=1}^{n+1} j^{-s_0}) \prod_{j=0}^n \left(1 - \frac{\alpha t}{(j+1)^{s_0}}\right)^{-j} \end{aligned}$$

When j is large enough, $(1 - \frac{\alpha t}{(j+1)^{s_0}})^{-j} \leq \exp(\frac{\alpha t}{j^{s_0-1}})$, which gives the existence of an exponential moment for $s_0 > 2$.

The proof in the case $|a_n| \prec \mu$ and $s_0 > 1$ is similar. \square

As an example of domination by μ^{*n} , we can think about the case where a recursive function called with parameter n requires n simulations with an acceptance-rejection method. Then, a_n appears as the sum of n independent variables following a geometric distribution $\mu = \mathcal{G}(p)$.

REFERENCES

- [1] Mohamad Akra and Louay Bazzi. On the solution of linear recurrence equations. *Comput. Optim. Appl.*, 10(2):195–210, 1998.
- [2] D. Choimet and H. Queffélec. *Twelve landmarks of twentieth-century analysis*. Cambridge University Press, New York, 2015. Illustrated by Michaël Monerau, Translated from the 2009 French original by Danièle Gibbons and Greg Gibbons, With a foreword by Gilles Godefroy.
- [3] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Stein. *Introduction to Algorithms*. PHI Learning, 2010.
- [4] Michael Drmota and Wojciech Szpankowski. A master theorem for discrete divide and conquer recurrences. *J. ACM*, 60(3):Art. 16, 49, 2013.
- [5] P. Erdős, A. Hildebrand, A. Odlyzko, P. Pudaite, and B. Reznick. The asymptotic behavior of a family of sequences. *Pacific J. Math.*, 126(2):227–241, 1987.

- [6] Michael T. Goodrich and Roberto Tamassia. *Algorithm Design: Foundations, Analysis, and Internet Examples*. Wiley, 2002.
- [7] Salvador Roura. Improved master theorems for divide-and-conquer recurrences. *J. ACM*, 48(2):170–205, 2001.

UNIVERSITÉ DE LORRAINE, INSTITUT ÉLIE CARTAN DE LORRAINE, UMR 7502, VANDOEUVRE-LÈS-NANCY, F-54506, FRANCE, AND, CNRS, INSTITUT ÉLIE CARTAN DE LORRAINE, UMR 7502, VANDOEUVRE-LÈS-NANCY, F-54506, FRANCE,
E-mail address: `Olivier.Garet@univ-lorraine.fr`