Available online at www.sciencedirect.com
SCIENCE DIRECT
Statistics & Probability Letters 71 (2005) 47–60
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Fragment size distributions in random fragmentations with cutoff

M. Ghorbel[a,b], T. Huillet[a,*]

[a] *Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 5 mail Gay-Lussac, 95031, Neuville sur Oise, France*
[b] *Laboratoire d'Analyse, Géométrie et Applications, CNRS-UMR 7539, Institut Galilée, Université de Paris 13, 93430, Villetaneuse, France*

Received 6 January 2004
Available online 19 November 2004

Abstract

We consider the following fragmentation model with cutoff: a fragment with initial size $x_0 > 1$ splits into $b > 1$ daughter fragments with random sizes, the partition law of which has exchangeable distribution. In subsequent steps, fragmentation proceeds independently for each sub-fragments whose sizes are bigger than some cutoff value $x_c = 1$ only. This process naturally terminates with probability 1. The size of a fragment is the random mass attached to a leaf of a "typical" path of the full (finite) fragmentation tree. The height's law of typical paths is first analyzed, using analytic and renewal processes techniques. We then compute fragments' size limiting distribution ($x_0 \uparrow \infty$), for various senses of a typical path. Next, we exhibit some of its statistical features, essentially in the case of the exchangeable Dirichlet partition model.

MSC: primary 60G57; 62E17; secondary 60K99; 62E15; 62E20

Keywords: Fragmentation models; Random partitions

*Corresponding author.
E-mail address: thierry.huillet@ptm.u-cergy.fr (T. Huillet).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2004.10.021

1. Introduction

A binary version ($b = 2$) of the random fragmentation model with cutoff outlined in the Abstract arises in Computer Science; see Sibuya and Itoh (1987) and references therein for historical background. When branching number $b > 2$, a phase transition on the variance of the number of splitting events of the induced fragmentation tree was observed by Dean and Majumdar (2002). Recently, Mahmoud (2003) and Itoh and Mahmoud (2003) studied analytically the random length of a fragmentation path for various tree-pruning policies, essentially when $b = 2$. Some connections with the Rényi's parking problem were also outlined and studied. The case $b > 2$ was also addressed in Javanian et al. (2003). Subsequently, Janson (2003) observed that renewal processes played some role in these problems.

Incomplete (or one-sided) trees and paths, as those obtained after pruning the full fragmentation tree, have recently become a popular subject. The underlying problem is to define the pruning policy consisting in picking at random one of the b parental mass fractions $(U_1, \ldots, U_b)$ transmitted to each descendant at each step of the fragmentation process.

We first supply a pruning construction which is parametrized by a real parameter β and we call it β-pruning (or equivalently β-size-biased picking).

Applying a β-pruning rule at each step, we are left with a β-path in the full interval tree. The law of the height of β-paths is then studied, mainly in the limit $x_0 \uparrow \infty$.

The size of a fragment is the random mass attached to a leaf of a β-path of the full (finite) fragmentation tree. There are two notions of fragments' size, namely the backward and forward ones; they are the masses attached to a leaf of a β-path just before or after it crosses the cutoff value $x_c = 1$. Using renewal processes techniques, we compute their joint limiting distribution ($x_0 \uparrow \infty$). We then exhibit some of its statistical features, essentially in the case of the exchangeable Dirichlet partition model. Several examples are supplied.

2. Fragmentation models with cutoff

Let us start with generalities on simple fragmentation processes. Consider an interval (or a cube) of length (volume) $x_0 > 0$. At step one, there is a probability $p \in (0, 1)$ to split the interval and so, with probability $\overline{p} := 1 - p$ the initial fragment remains unchanged for ever. If it splits, it splits into $b > 1$ fragments with random sizes, say $(U_1 x_0, \ldots, U_b x_0)$, where $(U_1, \ldots, U_b)$ has exchangeable distribution throughout, implying in particular that each U_k, $k = 1, \ldots, b$ all share the same distribution, say the one of U_1. On each first-generation sub-fragment, the splitting process is then iterated, independently, a property which we refer to in the sequel as its renewal or regenerative structure (see Feller, 1971). We also assume that U_1 has a density $f(u) > 0$ on $(0, 1)$ with total mass 1. If probability p is independent of fragments' sizes at each step, the fragmentation model is called homogeneous and the fragmentation tree is a classical Galton–Watson tree; on the other hand, when the splitting probability is fragments' length dependent, the fragmentation process could be termed heterogeneous.

In this work, a particular such fragmentation process is studied. We call it fragmentation with cutoff. Here, $x_0 > 1$ and the splitting probability at each step is 1 if the size of the fragment to be split is larger than $x_c = 1$, 0 otherwise; such a fragmentation model is not homogeneous. This