HAL
open science

On the algebraic numbers computable by some generalized Ehrenfest urns

Marie Albenque, Lucas Gerin

To cite this version:

Marie Albenque, Lucas Gerin. On the algebraic numbers computable by some generalized Ehrenfest urns. Discrete Mathematics and Theoretical Computer Science, 2012, Vol. 14 no. 2 (2), pp.271-284. 10.46298/dmtcs.565 . hal-00589621v1

HAL Id: hal-00589621

https://hal.science/hal-00589621v1

Submitted on 29 Apr 2011 (v1), last revised 3 Jul 2017 (v3)

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.

On the algebraic numbers computable by some generalized Ehrenfest urns

Marie Albenque and Lucas Gerin

April 29, 2011

Abstract

This article deals with some stochastic population protocols, motivated by theoretical aspects of distributed computing. We modelize the problem by a large urn of black and white balls from which at every time unit a fixed number of balls are drawn and their colors are changed according to the number of black balls among them. When the time and the number of balls both tend to infinity the proportion of black balls converges to an algebraic number. We prove that, surprisingly enough, not every algebraic number can be "computed" this way.

1 Introduction

The aim of this article is to tackle some questions of distributed computing in theoretical computer science, from a statistical mechanics standpoint. Distributed computing deals with large computing systems using many small processing elements. These small elements are thought as elementary elements in a complex network whose interactions at a low level may be pretty difficult to understand and modelize. There is a clear analogy with statistical mechanics, in which physical systems are well described at a macroscopic level, while molecular-level phenomena seem chaotic.

More precisely this work is motivated by recent studies in *population protocols* (see [2] for a detailed introduction). They are models of decentralized networks consisting of mobile *agents* interacting in pairs. The way agents interact is known (and assumed to be simple) but not their movements. These movements are driven by an "adversary", which picks at each time step two agents according to a process only assumed to be *fair* (roughly speaking, the fairness condition ensures that any possible configuration is eventually attained ; see again [2] for a formal definition).

1.1 Description of our model

Let $\{e_1, \ldots, e_q\}$ be a finite set of *states*, and a *transition rule*

$$\phi : \ \{e_1, \ldots, e_q\}^2 \ \to \{e_1, \ldots, e_q\} \, ,$$

We are also given a set of $n \geq 2$ identical *agents*, which may be at any moment in one of the q possible states. A *population protocol* associated to ϕ is a dynamical system $(\sigma_t)_{t\in\mathbb{N}}$ on $\{e_1, \ldots, e_q\}^n$ where at each time step two agents are chosen, and their states updated. Updating is made according to ϕ : if x, y in states e, f are chosen, their both states are turned into $\phi(x, y)$.

Let $\mathcal{P}$ be a boolean function defined on the states of a population, *i.e.*:

$$\mathcal{P} : \{e_1, \ldots, e_q\}^n \to \{\text{true,false}\}\,.$$

It is a main question in distributed computing to ask wether the population protocol can compute the *boolean expression* $\mathcal{P}$, meaning that for any fair scheme of updating,

$$\{\mathcal{P}(\sigma_0) = \text{ true }\} \Leftrightarrow \{(\sigma_t)_{t\in\mathbb{N}} \text{ converges to } e_1^n\}\,.$$

In this setting the computational power of such systems is now fairly well understood as described in [1].

We adress in this article a different but related question: when the scheme is random, which real numbers can be computed by a population protocol? Let us be more precise; we consider a system of n identical individuals and restrict to the case of two states. We modelize these individuals as indistinguishable black/white balls in a urn. For a given rule

$$f : \{0, \ldots, k\} \to \{\text{black,white}\}\,,$$

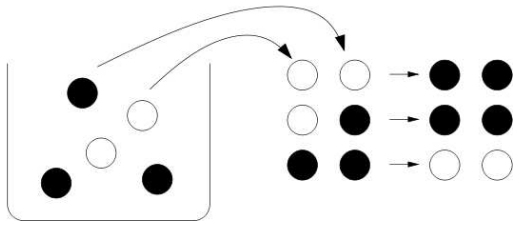
we consider the dynamic of the population driven as follows. We assume that at time 0, there are n balls in the urn and that each of them is randomly colored in black or white with probability 1/2, independently from the $n-1$ others. At each time unit, $k \geq 2$ of the balls are chosen randomly and uniformly (independently from the past). If i of the k balls are black, then the k balls are all recolored in the color $f(i)$, and put back in the urn. Our model is in fact a generalization of the famous Ehrenfest urn model (see for instance [8]) which is the process corresponding to $k = 1$ and $0 \mapsto$ black $1 \mapsto$ white or equivalently at each time step a single ball is picked and its color is changed.

In [4], the case $k = 2$ has already been handled in but with differences in the approach and the statements of the results. The generalization of the model to $k \geq 2$ generates an infinite number of rules and, as we will see, allows to compute an infinite set of numbers. An important feature of our model is that the k balls all turn into the *same* color. From our original setting, this is motivated by the fact that in the complex network there is no hierarchy and not much communication between the agents: when they meet, they instantly all take the same decision. We also stress the fact that our assumption for the choice of the agents is much more restrictive than fairness, but we will see that it enables us to carry on interesting computations. Moreover, this assumption (random and uniform choices of the agents) seems a natural way to extend fairness to very large (and even infinite) populations.

Throughout the process, the number of balls in the urn remains constant, and we are interested in the evolution of the proportion of black balls. In order to understand what we mean by "computing a real number", and before going through details, we present a simple example.

1.2 Heuristic : the example of $(5-\sqrt{17})/2$

Take $k=2$ and consider the function $f: 0 \mapsto$ black ; $1 \mapsto$ black ; $2 \mapsto$ white, as illustrated below



Let $X_\ell^{(n)}$ be the proportion of black balls at time ℓ, then $(X_\ell^{(n)})$ defines a Markov chain on the set $\left\{\frac{0}{n}, \frac{1}{n}, \dots, \frac{n}{n}\right\}$, which admits a unique invariant measure $\pi^{(n)}$. Transition probabilities of the chain X are clearly rational numbers so the components of $\pi^{(n)}$, as a solution of a linear system of rational equations, are rational numbers. Its mean is thus rational. Hence ergodic theorem for Markov chains states that almost surely:

$$\frac{X_1^{(n)} + \cdots + X_\ell^{(n)}}{\ell} \to p^{(n)} := \mathrm{Mean}(\pi^{(n)}) \in \mathbb{Q}.$$

We do not pay attention to the exact expression of this mean, but run rather a non-rigorous computation that gives us a hint for its asymptotic behavior (when n goes large) :

$$\begin{aligned}
\mathbb{E}\left[X_{\ell+1}^{(n)} - X_\ell^{(n)} | X_\ell = x\right] &= +\frac{2}{n}\mathbb{P}\left(\text{both are white}\right) - \frac{1}{n}\mathbb{P}\left(\text{one white, one black}\right) \\
&\quad - \frac{2}{n}\mathbb{P}\left(\text{both are black}\right), \\
&= +\frac{2}{n}\frac{\binom{n-nx}{2}}{\binom{n}{2}} - \frac{1}{n}\frac{nx(n-nx)}{\binom{n}{2}} - \frac{2}{n}\frac{\binom{nx}{2}}{\binom{n}{2}}, \\
&\overset{n\to\infty}{\sim} \frac{1}{n}\left(2(1-x)^2 - x(1-x) - 2x^2\right). \qquad (1)
\end{aligned}$$

Take now ℓ large, our system converges to its stationary regime, and thus we expect the righ-hand term in (1) to vanish. Hence, for large n, $p^{(n)}$ should be close to the irrational number $(5-\sqrt{17})/2 \approx 0.4384\dots$, which is the only root of the polynomial $2(1-X)^2 - X(1-X) - 2X^2$ in $[0,1]$. We let the balls "compute" $(5-\sqrt{17})/2$. The aim of this paper is to give rigorous answers to some questions raised by this simple example:

1. What sense can we give to this double limit (when n, ℓ go simultaneously to infinity)? This is a very usual question in stochastic processes theory, and will be done in the next section with an approximation of $X^{(n)}$ by an ordinary differential equation, after a proper time rescaling (Proposition 2 and Theorem 3).

2. Is any algebraic number computable by such population protocols? In the case $k=2$ these computable numbers are given in [4]. We obtain results on different directions: on one hand the set of numbers is large as it is a dense

subset of $(0,1)$ (Theorem 4) but on the other hand, we prove that almost no rational numbers is computable (Proposition 6).

Let us note that the link between the evolution of some stochastic population protocols and that of an associated ordinary differential equation has been used for the first time by Chatzigiannakis and Spirakis [6] in a somewhat different context ; they study some qualitative properties of the differential equation in order to discuss the stability of the underlying protocol.

Our main results (Theorem 4 and Proposition 6) may seem surprising as we might expect that any algebraic number would be computable in our setting. This should be compared to a recent and nice result by Bournez,Fraigniaud and Koegler [5] which says that, if the agents are picked in pairs but with more than two states, then any algebraic number is computable.

2 Limiting behavior of urns

We now introduce some notations. We denote by E_f the set:

$$E_f = \{i \leq k; f(i) = \text{ black}\},$$

and the pair (k, f) (or equivalently, the pair (k, E_f)) is referred to as the *rule* of the urn. We study the sequence of the proportions of black balls in the urn $\mathbf{X}^{(n)} := (X_\ell^{(n)})_{\ell \geq 0}$ when performing the random experiment described in the introduction. It is clear that $\mathbf{X}^{(n)}$ is a Markov chain with state space $\left\{\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n}{n}\right\}$.

Following the heuristic of Section 1.2, we associate to the rule (k, f) the polynomial $b = b_f$ defined by

$$\begin{aligned} b(y) &= \sum_{i \in E_f} \binom{k}{i} k y^i (1-y)^{k-i} - ky \\ &= k\mathbb{P}\left(\mathcal{B}_{k,y} \in E_f\right) - ky, \end{aligned}$$

where $\mathcal{B}_{k,y}$ is a binomial random variable with parameters (k, y). The quantity $b(y)$ is the generalization to any rule of the polynomial obtained in (1). The meaning of this polynomial can be understood as follows: when n is large, picking k balls uniformly among n balls, a proportion of x of them being black, almost amounts to perform k times an experiment with a probability x of success. The quantity $b(x)$ represents then the expectation of the evolution of the number of black balls, after putting pack the k recolored ones. The following lemma states this idea rigorously:

Lemma 1. *For $y \in \left\{\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n}{n}\right\}$, set*

$$b^{(n)}(y) = \mathbb{E}\left[X_1 - X_0 | X_0 = y\right].$$

The map $y \mapsto b^{(n)}(y)$ converges uniformly to $b(y)$ on the interval $[0,1]$. More precisely, there exists a constant c depending only on (k, f) such that

$$\sup_y \left|b^{(n)}(y) - b(y)\right| \leq c/n.$$

Proof. We simply write:

$$b^{(n)}(y) - b(y) = \sum_{i=0}^{k} (f(i) - i) \left[\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{n}{k}} - \binom{k}{i} y^i (1-y)^{k-i} \right].$$

To show that this quantity converges to zero, it suffices to note that, when j is fixed and m goes to infinity, $\binom{m}{j} \sim m^j/j!$. Stirling formula and convex inequalities permit to prove that the convergence is uniform. □

The rest of the section is devoted to the study of the convergence of $\mathbf{X}^{(n)}$. For that purpose, we define $t \mapsto x_t$ as the unique maximal solution of the ordinary differential equation (ODE) such that

$$x' = b(x) \quad \text{and} \quad x_0 = 1/2. \tag{2}$$

First notice that since

$$b(0) = k\mathbb{P}\left(0 \in E_f\right) \geq 0 \quad \text{and} \quad b(1) = k\mathbb{P}\left(k \in E_f\right) - k \leq 0, \tag{3}$$

this maximal solution x_t actually remains in the interval $[0, 1]$. We can now describe the asymptotic behavior of the sequence $\mathbf{X}^{(n)}$ when n goes to infinity. To do so, we speed up the process $\mathbf{X}^{(n)}$ by a factor n, by setting[1] $x_n(t) = X^{(n)}_{\lfloor nt \rfloor}$ and obtain the following result.

Proposition 2. *For any real numbers $t_0, \varepsilon > 0$,*

$$\mathbb{P}\left(\sup_{t<t_0} |x_n(t) - x_t| > \varepsilon \right) \leq Ae^{-Bn},$$

where $A, B > 0$ may depend on k, f, ε, t_0 but not on n.

Proof. As this proposition can be seen as an instance of the general theory of large deviations for Markov processes, sometimes known as Kurtz's Theorem (see [10]), we only outline the main ideas of the proof.

For sake of conciseness we set $X_k := X^{(n)}_k$ and introduce the classical martingale

$$M_k = X_k - X_0 - \sum_{\ell=0}^{k-1} b^{(n)}(X_\ell).$$

This equation enables to rewrite X as:

$$X_{\lfloor nt \rfloor} = X_0 + M_{\lfloor nt \rfloor} + n \int_0^t b^{(n)}(X_{\lfloor ns \rfloor}) ds.$$

Now, in order to bound, $f(t) := \sup_{s \leq t} |x_n(s) - x_s|$, we write

$$f(t) \leq \sup_{s \leq t} |M_{\lfloor ns \rfloor}| + \int_0^t |nb^{(n)}(x_n(s)) - b(x_n(s))| ds + \int_0^t |b(x_n(s)) - b(x_s)| ds.$$

The probability for the first term to be large can be bounded with a concentration inequality for martingales while the second term is as small as desired thanks to Lemma 1. Since the last term is smaller than $\sup |b'| \int f(s) ds$, an application of the Grönwall Lemma gives a bound for $\mathbb{P}(f(t) > \varepsilon)$. Again, we refer to ([10],p.76-84) or ([7],p.45-46) for details. □

[1]The notation $\lfloor k \rfloor$ stands, as usual, for the smallest integer larger than k.

Cauchy-Lipschitz's Theorem implies that $x'(t)$ is never equal to zero (unless x is constant and equal to a root of b). Hence any solution of (2) is monotonous, and converges to a root of b. If $b(1/2) \geq 0$ (resp. < 0) then the solution starting from $1/2$ converges to the smallest (resp. largest) root of b greater (resp. smaller) than $1/2$, denoted by α. We gather this observation with Proposition 2 to obtain our main result:

Theorem 3. *Assume that $b(1/2) \geq 0$ (resp. < 0) and let α be the smallest (resp. largest) root of b greater (resp. smaller) than $1/2$. For any $\varepsilon > 0$, there exist some constants $c > 0$ and $A, B > 0$ such that*

$$\mathbb{P}\left(\left|X^{(n)}_{\lfloor cn \rfloor} - \alpha\right| \geq \varepsilon\right) \leq Ae^{-Bn}.$$

We say that the rule (k, E) computes *the number α.*

Roughly speaking, it means that if we start with a large number n of balls and wait a linear time, the proportion of black balls is with high probability close to α. As regards a more quantitative aspect on time and space complexity, we point out the recent article [3] in which this question is discussed for the case of $k = 2$ and $\alpha = 1/\sqrt{2}$ (but the method extends to other situations).

of Theorem 3. First note that such an α always exists by (3). Take now c large enough, so that $|x_c - \alpha| \leq \varepsilon/2$. It suffices then to write

$$\mathbb{P}\left(\left|X^{(n)}_{\lfloor cn \rfloor} - \alpha\right| \geq \varepsilon\right) \leq \mathbb{P}\left(\left|X^{(n)}_{\lfloor cn \rfloor} - x_c\right| \geq \varepsilon/2\right).$$

Proposition 2 gives the desired bound for the right-hand side. □

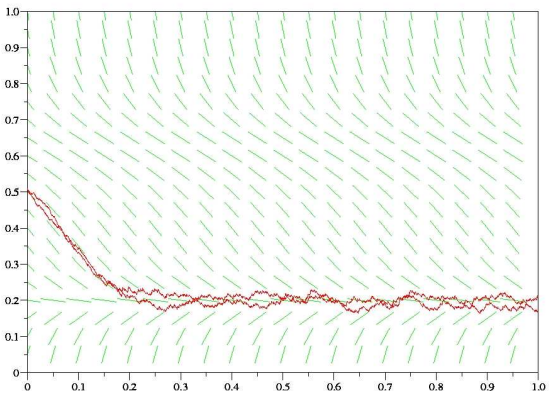


Figure 1: Two simulations of $(X^{(n)})$, with $n = 2000$ balls up to time 2000, with the flow of the corresponding ODE. Here, $k = 8$ and $E = \{0, 4, 5, 8\}$.

3 The set of computable numbers

We give in this section some properties about the set $\mathcal{L}$ of numbers that can be computed by our urns. A first basic observation is that each element of $\mathcal{L}$ is the root of a polynomial b_f and hence are algebraic. Moreover we have the following properties:

Theorem 4. *The set $\mathcal{L}$*

(i) *is symmetric w.r.t. $1/2$;*

(ii) *is dense in $[0,1]$;*

(iii) *contains numbers of any algebraic degree ;*

(iv) *does not contain every algebraic number.*

Proof. **(i)** Let α in $\mathcal{L}$ and (k,E) a be rule converging to α. We denote by $E^\star$ the set defined by:

$$i \in E^\star \Leftrightarrow k-i \notin E, \quad .$$

Let $b^\star$ be the polynomial associated to this new rule $(k,E^\star)$, we have

$$b^\star(1-\alpha) = k\mathbb{P}\left(\mathcal{B}_{k,1-\alpha} \in E^\star\right) - k(1-\alpha) \tag{4}$$
$$= -k\mathbb{P}\left(\mathcal{B}_{k,\alpha} \in E\right) + k\alpha. \tag{5}$$

Hence, $1-\alpha$ is a root of $b^\star$. One checks easily that if the solution of the ODE $y' = b(y), y_0 = 1/2$ converges to α, then the solution of $y' = b^\star(y), y_0 = 1/2$ converges to $1-\alpha$.

(ii) Let a/b be a rational number in $[0,1]$, and ε, δ two positive reals such that

$$(a/b-\varepsilon, a/b+\varepsilon) \subset (\delta, 1-\delta).$$

We are looking for a number $\alpha \in (a/b \pm \varepsilon)$ and a rule (k,E) such that the associated ODE converges to α. In particular it is necessary that:

$$\mathbb{P}(\mathcal{B}_{k,\alpha} \in E) = \alpha. \tag{6}$$

Fix for now the integer k, and consider the set

$$E_{a,b} = \{i \leq k; i \equiv 0,1,2,\ldots,a-1 \mod b\}.$$

The proof relies on the following lemma:

Lemma 5. *For any $0 < \delta < 1/2$, there exists $\lambda > 0$ such that for any integer k and $x \in (\delta, 1-\delta)$,*

$$\left|\mathbb{P}\left(\mathcal{B}_{k,x} \equiv 0,1,\ldots,a-1 \mod b\right) - a/b\right| \leq e^{-\lambda k}. \tag{7}$$

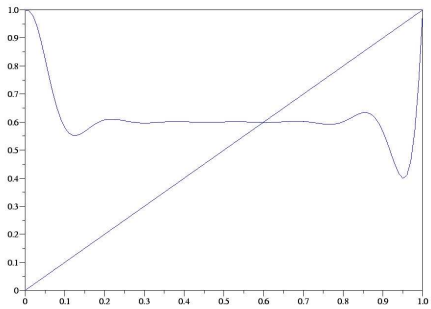


Figure 2: A plot of the maps $x \mapsto x$ and $x \mapsto \mathbb{P}(\mathcal{B}_{k,x} \in E_{a,b})$, for $k = 30$, $a/b = 3/5$.

Proof of Lemma 5. A proof based on linear algebra would give the best constant λ. As we do not need here this exact value, we give a probabilistic and shorter proof. The value modulo $(b-1)$ of a random variable $\mathcal{B}_{k,x}$ is the position at time k of the walk $\mathbf{X} = (X_\ell)_{\ell \geq 0}$ on $\{0, 1, \ldots, b-1\}$ starting from $X_0 = 0$ and with probability transitions

$$\mathbb{P}(X_{\ell+1} = X_\ell + 1 \mod b) = 1 - \mathbb{P}(X_{\ell+1} = X_\ell \mod b) = x.$$

starting from $X_0 = 0$. It is clear that this Markov chain admits as unique stationary measure the uniform measure π over $\{0, 1, \ldots, b-1\}$. By the general coupling inequality (see [9] Chap.I.2.), the desired quantity is smaller than

$$\mathbb{P}\left(X_0 \neq \tilde{X}_0, X_1 \neq \tilde{X}_1, \ldots, X_k \neq \tilde{X}_k\right),$$

where $X, \tilde{X}$ are two i.i.d. copies of $\mathbf{X}$, starting from 0 and π. These two walks meet necessarily if during b successive steps X goes b steps forward while $\tilde{X}$ remains motionless. This occurs with probability $x^b(1-x)^b$, hence

$$|\mathbb{P}\left(\mathcal{B}_{k,x} \equiv 0, 1, \ldots, a-1 \mod b\right) - a/b| \leq \left(1 - x^b(1-x)^b\right)^{\lfloor k/b \rfloor},$$

which decays exponentially in k, provided x is bounded away from 0 and 1. □

Assume k is a multiple of b, this ensures that $0 \in E$ and $1 \notin E$ and thus that neither 0 or 1 is a root of b. So we might as well take a smaller δ such that all the roots of b in the interval $[0, 1]$ belong in fact to $(\delta, 1-\delta)$. Let k be such that $e^{-\lambda k} < \varepsilon$.Any solution of $y' = b(y)$ starting inside $(\delta, 1-\delta)$ converges to a root of b. By Lemma 5, such a solution belongs to $(a/b \pm \varepsilon)$ (see Figure 2).

(iii) Fix $k \geq 1$ and consider the set $E = \{1\}$. The associated polynomial is

$$k\alpha(1-\alpha)^{k-1} - \alpha.$$

Its unique root in $(0, 1)$ is

$$x_0 = 1 - \sqrt[k-1]{1/k},$$

which has algebraic degree $k-1$.

(iv) We give in fact in the next proposition a much stronger result stating that almost no rational numbers belong to the set $\mathcal{L}$.

□

Proposition 6. *Let $x = p/q$ be a rational number such that $p \wedge q = 1$ and $q \geq 4$ then $x \notin \mathcal{L}$.*

Before proving this proposition, observe that the only rational numbers between 0 and 1 that do not satisfy the above conditions are 0, 1, 1/2, 1/3 and 2/3. These numbers all belong to $\mathcal{L}$ and correspond respectively to the rules $(1, \emptyset)$, $(1, \{0, 1\})$. $(2, \{1\})$, $(3, \{0, 3\})$ and $(3, \{1, 2\})$.

We proceed by contradiction. Let $x = p/q$ such that $p \wedge q = 1$ and $q \geq 4$ and assume that $p/q \in \mathcal{L}$. Since it implies that $1 - p/q \in \mathcal{L}$, without loss of generality we can (and will) assume that $p \geq 3$. Let (k, E) be one of the rules that admits p/q as a solution, hence:

$$\sum_{i \in E} \binom{k}{i} p^i (q-p)^{k-i} = pq^{k-1} \tag{8}$$

We now use some well-chosen reductions modulo p to deduce from this relation that $k \equiv 1\,[p^n]$ for every n, which leads to a contradiction (take for example n equal to k). Reduction of (8) modulo p implies that $\mathbf{1}_{0 \in E} q^k \equiv 0\,[p]$. This implies that $0 \notin E$, since $p \wedge q = 1$. We go one step further, reducing (8) modulo p^2 and dividing by p leads to the relation :

$$\mathbf{1}_{1 \in E} k (q-p)^{k-1} \equiv q^{k-1}\,[p]\,.$$

¿From which we readily deduce that $1 \in E$ and $k \equiv 1\,[p]$.

We now proceed by induction to show that for every $n \leq k$, $k \equiv 1[p^n]$. The following lemma would be useful:

Lemma 7. *Assume $k \equiv 1\,[p^{n-1}]$, then for every $2 \leq n \leq k$ and $2 < i \leq n$, $\binom{k}{i} \equiv 0\ [p^{n-i+1}]$. Moreover if $p \not\equiv 2\,[4]$, the result remains true for $i = 2$.*

Proof. We start with the classical relation:

$$i(i-1)\binom{k}{i} = k(k-1)\binom{k-2}{i-2}.$$

Since $k \equiv 1\,[p^{n-1}]$, we get:

$$(i \wedge p)((i-1) \wedge p)\binom{k}{i} \equiv 0\,[p^{n-1}]\,.$$

Now $(i \wedge p)((i-1) \wedge p) < p^2$, so if $i > 3$ the proof is complete, otherwise a simple computation enables to conclude for $i = 3$. The case $i = 2$ and $p \not\equiv 2\,[4]$ is straightforward. □

We need to proceed differently depending on p is or is not equal to 2 modulo 4. Assume first that $p \not\equiv 2\,[4]$ and that we proved $k \equiv 1\,[p^{n-1}]$. To carry on the recursion, we consider the following reductions of (8) modulo p^{n+1}, obtained by successive applications of Lemma 7:

$$kp(q-p)^{k-1} + \sum_{i=2}^{k} \mathbf{1}_{i\in E}\binom{k}{i}p^i(q-p)^{k-i} \equiv q^{k-1}\,[p^{n+1}]$$

$$(1-k)q^{k-1} + \sum_{i\geq 1}^{k-1}(k-i)\binom{k}{i}p^i q^{k-1-i} \equiv 0\ \ [p^n]$$

$$(1-k)q^{k-1} \equiv 0.$$

This concludes the proof in this case. Assume now that $p \equiv 2\,[4]$, it is now convenient to use the fact that the rule $(k, E^\star)$ admits $1 - p/q$ as a solution leading to the relation:

$$\sum_{i\notin E}\binom{k}{i}p^i(q-p)^{k-i} = (q-p)q^{k-1}. \tag{9}$$

Taking the reduction of the latter equation modulo p^{n+1}, we obtain:

$$(1-k)q^{k-1}pq^{k-2}\binom{k}{2}(1+\mathbf{1}_{2\notin E}) \equiv 0\ \ [p^n],$$

and hence

$$(1-k)(q+\frac{pk}{2}) \equiv 0\,[p^n].$$

We conclude the proof by noticing that $(q+\frac{pk}{2})$ is both prime with $p/2$ and odd hence prime with p.

4 Conclusion

We do not claim that our model is realistic for physical systems but it seems to us that our first results raise some interesting theoretical questions in the current research on the computational power of population protocols. Although our model is very general and allows to compute a large set of numbers, some algebraic numbers as "simple" (on a computational point of view) as 1/5 are not computable. A natural question is then to ask if the set $\mathcal{L}$ has a nice structure : has it interesting symmetries? can it be endowed with a certain algebraic structure which is consistent with computability? In other words: does there exist an operation $\otimes$ such that if x and y belongs to $\mathcal{L}$ then a certain combination of their associated rules computes $x \otimes y$? This looks to be an interesting issue but unfortunately we did not get results in this direction.

As already mentioned, it is proved in [5] that any algebraic number is computable for $k = 2$ and $q > 2$ colors, but with the significant difference that the 2 balls may be turned into two different colors. A question remains: what happens in our model with $k > 2$ and two states if we consider more general rules for which the k balls may be recolored differently from each other? With this new model it is possible to

compute any rational number but we still do not know if any algebraic number is computable ; [5] suggests that this should be the case.

Acknowledgements

The second author would like to thank O.Bournez and J.Cohen for some very interesting discussions during the preparation of [4] that raised his interest for the subject and for showing us a preliminary version of [5].

References

[1] D. Angluin, J. Aspnes, D. Eisenstat. Stably computable predicates are semilinear. Proceedings of *Principles Of Distributed Computing (PODC'06)* (2006).

[2] J.Aspnes, E.Ruppert. An introduction to population protocols. *Bulletin of the European Association for Theoretical Computer Science* **93** p.98-117 (2007).

[3] G.Aupy, O.Bournez, On the number of binary-minded individuals required to compute $1/\sqrt{2}$. *Theoretical Computer Science* **412**, 22:2262-2267 (2011).

[4] O.Bournez, Ph.Chassaing, J.Cohen, L.Gerin, X.Koegler. On the convergence of population protocols when population goes to infinity. *Applied Mathematics and Computation* **215** n4 p.1340-1350 (2009).

[5] O.Bournez, P.Fraigniaud, X.Koegler. Computing with Large Populations. *preprint* (2011).

[6] I.Chatzigiannakis, P.Spirakis. The Dynamics of Probabilistic Population Protocols. Proceedings of the *22nd International Symposium on Distributed Computing (DISC '08)* p.498-499 (2008). Full version at `http://arxiv.org/abs/0807.0140`.

[7] R. Darling et J. Norris. Differential equation approximations for Markov chains, *Probability Surveys* **5** (electronic) (2008).

[8] R. Durrett, *Probability, Theory and Examples*, The Wadsworth & Brooks/Cole Statistics/Probability Series (1991).

[9] T.Lindvall. *Lectures on the coupling method*. John Wiley & Sons (1992).

[10] A.Shwartz and A.Weiss. *Large deviations for performance analysis*. Chapman & Hall (1995)