Discrete Mathematics and Theoretical Computer Science DMTCS vol. **14:2**, 2012, 271–284

On the algebraic numbers computable by some generalized Ehrenfest urns

Marie Albenque[1†] and Lucas Gerin[2‡]

[1] *LIX, CNRS, UMR 7161, École Polytechnique, France*
[2] MODAL'X, *Université Paris-Ouest*

received 29th April 2011, revised 4th June 2012, 20th November 2012, accepted 4th December 2012.

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 is changed according to the number of black balls among them.The limiting behaviour of the composition of the urn when both the time and the number of balls tend to infinity is investigated and the proportion of black balls is shown to converge to an algebraic number. We prove also that, surprisingly enough, not every algebraic number can be "computed" this way.

Keywords: population protocols, distributed computing, approximation of Markov chains, Ehrenfest

1 Introduction

1.1 Context and motivations

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 objects 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.

This work is motivated by recent studies in *population protocols* (see [AR07] for a detailed introduction), which 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 [AR07] for a formal definition).

†Email: `albenque@lix.polytechnique.fr`, supported by the European project ExploreMaps – ERC StG 208471.
‡Email: `lgerin@u-paris10.fr`, supported by ANR MAGNUM

Let us be more precise. We are given a finite set S of *states*, a *transition rule* $\phi : S^2 \to S$, and $n \geq 2$ identical *agents*, which may be at any moment in one of the $\mathrm{card}(S)$ possible states. A *population protocol* associated to ϕ is a dynamical system $(\sigma_t)_{t\in\mathbb{N}}$ on S^n where at each time step two agents are chosen, and their states updated. Updating is made according to ϕ: if x, y, respectively in states e, f, are chosen, then both their state is turned into $\phi(e, f)$. Population protocols are usually designed to compute predicates over a set of boolean variables. A different but related question is addressed in this article: we use population protocols to compute real numbers instead of predicates.

1.2 Model

To compute numbers with population protocols, we rely on the classical formalism of stochastic urn models, where each ball stands for an agent.

Let us describe more precisely the model we deal with. We fix an integer $k \geq 1$, a real number $x_0 \in [0, 1]$ and a *rule* $f : \{0, \ldots, k\} \to \{\text{black,white}\}$. At time 0, there are n balls in the urn and each of them is randomly colored in black with probability x_0 or in white with probability $1 - x_0$, independently from the $n - 1$ others. The triplet (k, f, x_0) is referred to as the *rule* of the urn.

At each time unit, k balls are picked randomly, uniformly and independently from the past. Let i be the number of black balls among them, then all the k balls are recolored in the color $f(i)$ and put back in the urn. This model is a generalization of the famous Ehrenfest urn model (see for instance [Dur91]) which corresponds 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).

Definition 1 *The real number α is said to be* computable *if there exists a rule (k, f, x_0) such that, for any $\varepsilon > 0$, there exists $c > 0$ such that*

$$\mathbb{P}\left(\left|X^{(n)}_{\lfloor cn\rfloor} - \alpha\right| \geq \varepsilon\right) \overset{n\to\infty}{\to} 0,$$

where $X^{(n)}_\ell$ is the proportion of black balls at time ℓ, and $\lfloor k\rfloor$ stands, as usual, for the largest integer smaller than k.

Roughly speaking, it means that if we start with a large number n of balls and wait for a linear time in n (larger than cn), the proportion of black balls is with high probability close to α.

Before stating our main results, let us first discuss some important features of our model.

- The results we aim for are about the asymptotic properties of population protocols, when the size of the population grows to infinity. From this perspective, the choice of picking agents uniformly at random appears to be the natural generalization of fairness to large (or even infinite) populations.

- It is important to point out the strong assumption that the k balls are all turned 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.

- Because the number we compute corresponds to the proportion of black balls, it is sufficient to consider only two possible states for the balls. This is why we focus on this case.

1.3 Description of the results

The results of this paper are twofold. We first describe in Section 2 the asymptotic behavior of the proportion of black balls for any fixed rule. In Proposition 3, the process $(X^{(n)}_{\lfloor nt \rfloor})_{t\geq 0}$ is shown to be close to the solution of an ordinary differential equation. We then characterize in Theorem 4 the numbers computable by a given rule as the roots of a polynomial related to this differential equation.

Section 3 is devoted to studying the set of computable numbers. This relies on a combinatorial analysis of the latter polynomial. In Theorem 5, we show that the set of computable numbers is dense in $[0,1]$. On the other hand, we prove in Proposition 7 that, surprisingly enough, rational numbers are not computable (except for a few explicitly exhibited ones).

We conclude this introduction by comparing our results with related works. In [BCC+09], the black/white case with $k = 2$ has already been handled but with differences in the approach and the statements of the results. The main difference is that the authors of [BCC+09] had to assume that the initial proportion x_0 of black balls is close to the number α one wants to compute[(i)] (as a counterpart, the main result of [BCC+09] gives a interesting and precise description of the fluctuations of $X^{(n)}_k$ around its mean, for k large). The techniques developed in the present paper allow us to free ourselves from this restrictive assumption.

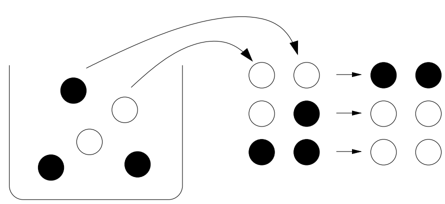
Our main results (Theorem 5 and Proposition 7) may seem surprising as we might expect that any algebraic number would be computable in our setting. This should be put in perspective with a simultaneous and interesting result by Bournez, Fraigniaud and Koegler [BFK12]. They study stochastic population protocols in which agents are only picked in pairs but the number of possible colors is arbitrarily large and the new colors of the two balls may be different. In this different setting, any algebraic number is computable. Note that, as in [BCC+09], the results in [BFK12] hold only with the assumption that the initial proportion of black balls is close to an equilibrium.

Let us also 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 [CS08] in a somewhat different context; they study some qualitative properties of the differential equation in order to discuss the stability of the underlying protocol.

Before going through details, we present now a simple introductory example.

1.4 Heuristic : the example of $(3-\sqrt{5})/2$

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



Recall that $X^{(n)}_\ell$ denotes the proportion of black balls at time ℓ. The sequence $(X^{(n)}_\ell)$ 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

(i) Indeed, the hypothesis "$|b_n(x) - b(x)| \to 0$" in ([BCC+09], Th.2) implies that "$1 - 2(X^{(n)}_0)^2 \frac{n}{n-1} + X^{(n)}_0 \frac{2}{n-1} \to 0$" and thus the initial proportion has to go to α.

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:

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

To get a hint for the asymptotic behavior of $p^{(n)}$, let us compute the conditional expectation of the increments of $X_\ell^{(n)}$:

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

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 $(3-\sqrt{5})/2 \approx 0.382\ldots$, which is the only root of the polynomial $2(1-X)^2-2X(1-X)-2X^2$ in $[0,1]$. We let the balls "compute" $(3-\sqrt{5})/2$.

2 Limiting behavior of urns

We study in this section the sequence $\mathbf{X}^{(n)} := (X_\ell^{(n)})_{\ell\geq 0}$ of the proportions of black balls in the urn. As mentioned above, this sequence is a Markov chain with state space $\left\{\frac{0}{n},\frac{1}{n},\ldots,\frac{n}{n}\right\}$. We denote by E_f the set

$$E_f = \{0 \le i \le k; f(i) = \text{ black}\},$$

recall that the triplet (k,f,x_0) (or equivalently, the triplet (k,E_f,x_0)) is the *rule* of the urn.

Following the heuristic of Section 1.4, we associate to the rule (k,f) the polynomial $b=b_f$ defined by

$$\begin{aligned}
b(y) &= \sum_{i\in E_f}\binom{k}{i}(k-i)y^i(1-y)^{k-i}+\sum_{i\notin E_f}\binom{k}{i}(0-i)y^i(1-y)^{k-i}\\
&= \sum_{i\in E_f}\binom{k}{i}ky^i(1-y)^{k-i}-ky\\
&= k\mathbb{P}\left(\mathcal{B}_{k,y}\in E_f\right)-ky,
\end{aligned}$$

where $\mathcal{B}_{k,y}$ is a binomial random variable with parameters (k,y). In the above example ($k=2$ and $E_f=\{0\}$), this gives

$$2\mathbb{P}\left(\mathcal{B}_{2,y}=0\right)-2y = 2(1-y)^2-2y,$$

which is indeed equal to the polynomial in (1).

The meaning of $b(y)$ can be understood as follows: when n is large, picking k balls uniformly among n balls, a proportion y of them being black, almost amounts to perform k times an experiment with a probability y of success. The quantity $b(y)$ then represents the expectation of the evolution of the number of black balls, after putting back the k recolored ones, as stated in the following lemma:

Lemma 2 *For $y = c/n$ with $c \in \{0, 1, 2, \ldots, n\}$, set*

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

The map $y \mapsto nb^{(n)}(y)$ converges uniformly to $b(y)$ on the set $\{0, 1/n, 2/n, \ldots, n/n\}$. More precisely, for n big enough,

$$\max_{0 \le c \le n} \left| nb^{(n)}(c/n) - b(c/n) \right| \le 5k^3/\sqrt{n}.$$

Proof: The probability that i balls are black when k balls are picked in a urn that contains n balls among which ny are black is equal to: $\binom{ny}{i}\binom{n-ny}{k-i}\binom{n}{k}^{-1}$, indeed to pick exactly i black balls, we need to pick i black balls among the ny black balls and $n-i$ white balls among the $n - ny$ white balls. For $y \in \{0/n, n/n\}$ there is nothing to prove since $nb^{(n)}(0) - b(0) = nb^{(n)}(1) - b(1) = 0$. For any $y \in \{1/n, 2/n, \ldots, (n-1)/n\}$ we write

$$\begin{aligned} nb^{(n)}(y) - b(y) &= \sum_{i=0}^{k} \left(k\mathbf{1}_{i \in E_f} - i\right) \left[\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{n}{k}} - \binom{k}{i} y^i (1-y)^{k-i} \right] \\ &= \sum_{i=0}^{k} \left(k\mathbf{1}_{i \in E_f} - i\right) \binom{k}{i} y^i (1-y)^{k-i} \left(\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{k}{i}\binom{n}{k} y^i (1-y)^{k-i}} - 1 \right) \end{aligned}$$

Let us handle the last term :

$$1 - \frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{k}{i}\binom{n}{k} y^i (1-y)^{k-i}} = 1 - \frac{\binom{ny}{i} i!}{n^i y^i} \frac{\binom{n-ny}{k-i}(k-i)!}{n^{k-i}(1-y)^{k-i}} \frac{n^k}{k!\binom{n}{k}}. \tag{2}$$

To show that (2) goes to zero (and hence, so does $nb^{(n)}(y) - b(y)$), recall that, when j is fixed and m goes to infinity, $\binom{m}{j} \sim e^{-j} m^j / j!$: the last three terms in (2) converge to one.

In order to prove that the convergence is uniform, a little more work is needed. First, recall that a consequence of the Stirling formula is that, for any integers $j \le m$,

$$\exp\big(-\frac{2j^2}{m}\big) \le \frac{\binom{m}{j}}{m^j/j!} \le 1. \tag{3}$$

Plugging this in (2) gives that

$$1 - \exp(-2k^2/n) \le 1 - \frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{k}{i}\binom{n}{k} y^i (1-y)^{k-i}} \le 1 - \exp(-\frac{2i^2}{ny}) \exp(-\frac{2(k-i)^2}{n-ny}),$$

and thus
$$\left|1-\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{k}{i}\binom{n}{k}y^i(1-y)^{k-i}}\right| \leq 1-\exp\left(-2k^2(\tfrac{1}{ny}+\tfrac{1}{n(1-y)})\right).$$

If $1/\sqrt{n} \leq y \leq 1-1/\sqrt{n}$, then this last quantity is less than $4k^2/\sqrt{n}$. Then, for any $y \in [1/\sqrt{n}; 1-1/\sqrt{n}]$, one has

$$\begin{aligned}|nb^{(n)}(y)-b(y)| &\leq \sum_{i=0}^{k}\left|k\mathbf{1}_{i\in E_f}-i\right|\binom{k}{i}y^i(1-y)^{k-i}\left|1-\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{k}{i}\binom{n}{k}y^i(1-y)^{k-i}}\right| \\ &\leq k\times 4k^2/\sqrt{n}\times\sum_{i=0}^{k}\binom{k}{i}y^i(1-y)^{k-i} = 4k^3/\sqrt{n}.\end{aligned}$$

We now deal with the case where $y < 1/\sqrt{n}$ and prove that both $nb^{(n)}(y)$ and $b(y)$ are close to $k\mathbf{1}_{0\in E_f}(1-y)^k$. We assume here that $n > 4k^2$ and therefore $y < 1/2k$.

$$nb^{(n)}(y)-k\mathbf{1}_{0\in E_f}(1-y)^k = k\mathbf{1}_{0\in E_f}\left(\frac{\binom{ny}{0}\binom{n-ny}{k-0}}{\binom{n}{k}}-(1-y)^k\right)+\sum_{i=1}^{k}\left(k\mathbf{1}_{i\in E_f}-i\right)\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{n}{k}}.$$

Then, using again (3) and inequalities $1-ky \leq (1-y)^k \leq 1-\frac{ky}{2}$ and $1-x \leq e^{-x} \leq 1-\frac{x}{2}$ for $0 \leq x \leq 1/2$,

$$\begin{aligned}|nb^{(n)}(y)-k\mathbf{1}_{0\in E_f}(1-y)^k| &\leq k\mathbf{1}_{0\in E_f}\left|\frac{\binom{n-ny}{k}}{\binom{n}{k}}-(1-y)^k\right|+k\sum_{i=1}^{k}\frac{\binom{ny}{i}\binom{n-ny}{k-i}}{\binom{n}{k}} \\ &\leq k\mathbf{1}_{0\in E_f}\left|\frac{\binom{n-ny}{k}}{\binom{n}{k}}-(1-y)^k\right|+k\left(1-\frac{\binom{n-ny}{k}}{\binom{n}{k}}\right) \\ &\leq k(1-y)^k|\exp(-2k^2/(n-ny))-1|+k(1-(1-y)^k e^{-2k^2/(n-ny)}) \\ &\leq k\times 1\times\frac{4k^2}{n}+k\left(1-(1-ky)(1-4k^2/n)\right) \\ &\leq 8k^3/n+k^2y-\frac{4yk^4}{n} \leq \frac{5k^2}{\sqrt{n}}.\end{aligned}$$

On the other hand,

$$\begin{aligned}|b(y)-k\mathbf{1}_{0\in E_f}(1-y)^k| &= |\sum_{i=1}^{k}\left(k\mathbf{1}_{i\in E_f}-i\right)\binom{k}{i}y^i(1-y)^{k-i}| \\ &\leq k(1-(1-y)^k) \leq k^2y \leq k^2/\sqrt{n}.\end{aligned}$$

This proves that for any $y < 1/\sqrt{n}$,

$$|nb^{(n)}(y)-b(y)| \leq 5k^2/\sqrt{n}+k^2/\sqrt{n} \leq 6k^2/\sqrt{n},$$

and concludes the case $y < 1/\sqrt{n}$. The case $y > 1 - 1/\sqrt{n}$ is symmetric. $\square$

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

$$x' = b(x) \quad \text{and} \quad x(0) = x_0 \tag{4}$$

(recall that x_0 is the initial proportion of black balls in the urn). First notice that since

$$b(0) = k\mathbb{P}\left(\mathcal{B}_{k,0} \in E_f\right) = k\mathbf{1}_{0\in E_f} \geq 0 \quad \text{and} \quad b(1) = k\mathbb{P}\left(\mathcal{B}_{k,k}k \in E_f\right) - k = k\mathbf{1}_{k\in E_f} - k \leq 0, \tag{5}$$

this maximal solution x_t actually remains in the interval $[0,1]$. To describe the asymptotic behavior of the sequence $\mathbf{X}^{(n)}$, we speed it up by a factor n, by setting $x_n(t) = X^{(n)}_{\lfloor nt\rfloor}$ and prove that $x_n(t)$ is well approximated by $x(t)$ when n is big:

Proposition 3 *For each rule and for any real numbers $t_0, \varepsilon > 0$, there exist $A, B > 0$ such that, for each $n \geq 1$,*

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

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 [SW95]), we only outline the main ideas of the proof.

For sake of conciseness we set $X_k := X_k^{(n)}$ 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 $X_{\lfloor nt\rfloor}$ as:

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

Then

$$\begin{aligned} X_{\lfloor nt\rfloor} - x_t = M_{\lfloor nt\rfloor} &+ \int_0^{\lfloor nt\rfloor/n} nb^{(n)}(X_{\lfloor ns\rfloor}) - b(X_{\lfloor ns\rfloor})ds \\ &+ \int_0^{\lfloor nt\rfloor/n} b(X_{\lfloor ns\rfloor}) - b(x_s)ds + (X_0 - x_0) + (x_{\lfloor nt\rfloor/n} - x_t), \end{aligned}$$

since $x_{\lfloor nt\rfloor/n} = x_0 + \int_0^{\lfloor nt\rfloor/n} b(x_s)ds$. Now, in order to bound, $f(t) := \sup_{s\leq t} |x_n(s) - x(s)|$, we write

$$\begin{aligned} f(t) \leq \sup_{s\leq t} |M_{\lfloor ns\rfloor}| + \int_0^{\lfloor nt\rfloor/n} |nb^{(n)}(x_n(s)) - b(x_n(s))|ds + \int_0^{\lfloor nt\rfloor/n} |b(x_n(s)) - b(x(s))|ds \\ + (X_0 - x_0) + \sup_{s\leq t} |x_{\lfloor ns\rfloor/n} - x_s|. \end{aligned}$$

The probability for the first term to be large can be bounded with a concentration inequality for martingales while the second term is bounded thanks to Lemma 2. The probability that the two last terms are greater than ε is as small as desired. Since the third 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 ([SW95], p.76-84) or ([DN08], p.45-46) for details. □

The Cauchy-Lipschitz 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 (4) is monotonous, and converges to a root of b. If $b(x_0) \geq 0$ (resp. < 0) then the solution starting from x_0 converges to the smallest (resp. largest) root of b greater (resp. smaller) than x_0, denoted by α. We gather this observation with Proposition 3 to obtain our main result:

Theorem 4 *Assume that $b(x_0) \geq 0$ (resp. < 0) and let α be the smallest (resp. largest) root of b no smaller (resp. no greater) than x_0. For any $\varepsilon > 0$, there exist some constants $c > 0$ and $A, B > 0$ such that for each n*

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

In particular it implies that the rule (k, E_f, x_0) computes *the number α, as defined in Definition 1.*

As regards a more quantitative aspect on time and space complexity, we point out the recent article [AB11] in which this question is discussed for the case of $k = 2$ and $\alpha = 1/\sqrt{2}$ (but the method extends to other situations).

Proof of Theorem 4: First note that such an α always exists by (5). 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 3 gives the desired bound for the right-hand side. □

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 is algebraic. Moreover we have the following properties:

Theorem 5 *The set $\mathcal{L}$*

(i) *is symmetric with respect to $1/2$;*

(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, x_0) a be rule computing to α. We denote by $E^\star$ the set defined by:

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

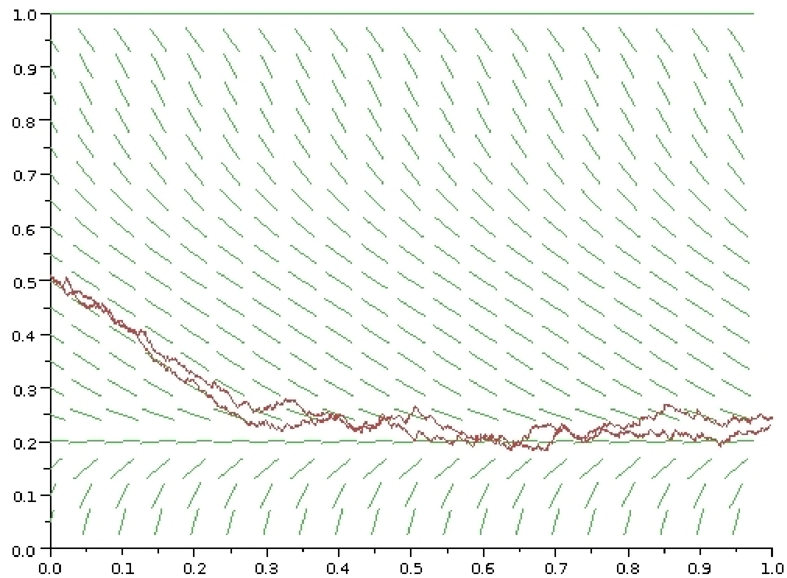


Fig. 1: Two simulations of $(X^{(n)})$, with $n = 2000$ balls up to time 2000, with the flow of the corresponding ODE. Here, $k = 8$, $E = \{0, 4, 6, 8\}$ and $x_0 = 0.5$. The polynomial b is equal to $8\left((1-x)^8 + 70x^4(1-x)^4 + 28x^6(1-x)^2 + x^8\right) - 8x$ and the corresponding α is approximately equal to 0.2079.

Let $b^\star$ be the polynomial associated to the new rule $(k, E^\star, 1 - x_0)$, we have

$$b^\star(1-\alpha) = k\mathbb{P}\left(\mathcal{B}_{k,1-\alpha} \in E^\star\right) - k(1-\alpha) \tag{6}$$
$$= -k\mathbb{P}\left(\mathcal{B}_{k,\alpha} \in E\right) + k\alpha. \tag{7}$$

Hence, $1 - \alpha$ is a root of $b^\star$. One checks easily that if the solution of the ODE $y' = b(y), y(0) = y_0$ converges to α, then the solution of $y' = b^\star(y), y(0) = 1 - y_0$ 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, x_0) such that the associated ODE converges to α. In particular it is necessary that:

$$\mathbb{P}(\mathcal{B}_{k,\alpha} \in E) = \alpha. \tag{8}$$

Fix for now the integer k, and consider the set

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

The proof relies on the following lemma:

Lemma 6 *For any* $0 < \delta < 1/2$*, there exists* $\lambda > 0$ *such that for any integer* k *and* $x \in (\delta, 1-\delta)$*,*

$$|\mathbb{P}\left(\mathcal{B}_{k,x} \equiv 0, 1, \dots, a-1 \,(\mathrm{mod}\, b)\,\right) - a/b| \le e^{-\lambda k}. \tag{9}$$

Proof of Lemma 6: 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 $\mathbf{X} = (X_\ell)_{\ell \ge 0}$ on $\{0, 1, \dots, b-1\}$ starting from $X_0 = 0$ and with probability transitions

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

starting from $X_0 = 0$. It is clear that this Markov chain admits as unique stationary measure the uniform measure π over $\{0, 1, \dots, b-1\}$. By the general coupling inequality (see [Lin92] Chap.I.2.), the desired quantity is smaller than

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

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

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

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 $e^{-\lambda k} < \varepsilon$ and $x_0 = 0.5$. The solution of $y' = b(y)$ starting from 0.5 converges to a root of b. By Lemma 6, such a solution belongs to $(a/b \pm \varepsilon)$ (see Figure 2).

(iii) Fix $k \ge 1$ 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 7 *Let* $x = p/q$ *be a rational number such that* $\gcd(p,q) = 1$ *and* $q \ge 4$ *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 are respectively computed by the rules $(1, \emptyset, 0.5)$, $(1, \{0,1\}, 0.5)$, $(2, \{1\}, 0.5)$, $(3, \{0,3\}, 0.5)$ and $(3, \{1,2\}, 0.5)$.

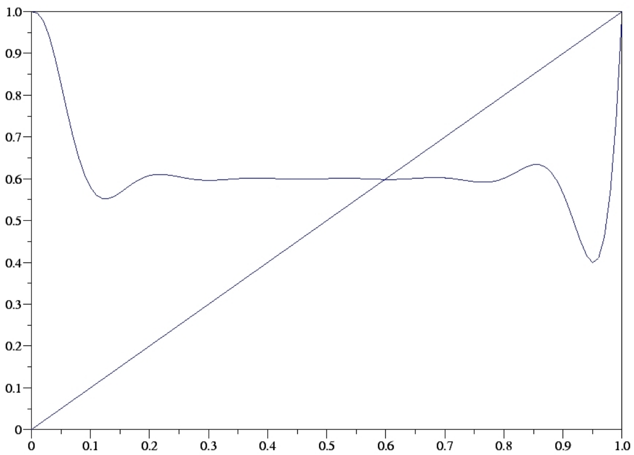


Fig. 2: A plot of the maps $x \mapsto x$ and $x \mapsto \mathbb{P}(\mathcal{B}_{k,x} \in E_{a,b})$, for $k = 30$, $a/b = 3/5$.

We proceed by contradiction. Let $x = p/q$ such that $\gcd(p, q) = 1$ and $q \geq 4$ and assume that $p/q \in \mathcal{L}$. Since it implies that $1 - p/q \in \mathcal{L}$, we can assume without loss of generality that $p \geq 3$. Let (k, E, x_0) be one of the rules that admits p/q as a solution, we can hence write

$$\sum_{i \in E} \binom{k}{i} p^i (q-p)^{k-i} = pq^{k-1}. \tag{10}$$

We now use some well-chosen reductions modulo p to deduce from this relation that $k \equiv 1 \pmod{p^n}$ for every n, which leads to a contradiction (take for example n equal to k). Reduction of (10) modulo p implies that $\mathbf{1}_{0 \in E} q^k \equiv 0 \pmod{p}$, which yields $0 \notin E$, since $\gcd(p, q) = 1$. We go one step further, reducing (10) modulo p^2 and dividing by p leads to the relation :

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

in which the left-hand side can be simplified into $\mathbf{1}_{1 \in E} k q^{k-1} \pmod{p}$. Since $\gcd(p, q) = 1$, we obtain $\mathbf{1}_{1 \in E} k \equiv 1 \pmod{p}$, from which we readily deduce that $1 \in E$ and $k \equiv 1 \pmod{p}$.

We now proceed by induction to show that $k \equiv 1 \pmod{p^n}$, for every $n \leq k$. The following lemma will be useful:

Lemma 8 *Assume $k \equiv 1 \pmod{p^{n-1}}$, with $2 \leq n \leq k$. Then for any $2 < i \leq n$, $\binom{k}{i} \equiv 0 \pmod{p^{n-i+1}}$. Moreover if $p \not\equiv 2 \pmod{4}$, the result is also true for $i = 2$.*

Proof: It is enough to prove the lemma for p being a power of a prime number, otherwise writing the decomposition of $p = \prod p_i^{\alpha_i}$ into a product of prime numbers and applying the result for each of the term gives the result. We start with the classical relation:

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

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

$$\gcd(i, p^{n-1}) \gcd(i-1, p^{n-1}) \binom{k}{i} \equiv 0 \pmod{p^{n-1}}. \tag{11}$$

Now, for $i \geq 4$, we have that $i \leq 2^{i-2} \leq p^{i-2}$, and hence $\gcd(i, p^{n-1}) = \gcd(i, p^{i-2})$, since we assume that p is a power of a prime. Similarly $\gcd(i-1, p^{n-1}) = \gcd(i-1, p^{i-2})$. Now, since $\gcd(i, i-1) = 1$,

$$\gcd(i, p^{i-2}) \gcd(i-1, p^{i-2}) \leq p^{i-2}.$$

Then (11) yields the desired result. The cases $i = 2$ and $i = 3$ are dealt with a direct computation. □

To continue the proof of Proposition 7, we need to proceed differently depending on p being or not equal to 2 modulo 4. Assume first that $p \not\equiv 2 \pmod 4$ and that we proved $k \equiv 1 \pmod{p^{n-1}}$ for some $n \geq 2$. To carry on the recursion, we write the reduction of (10) modulo p^{n+1},

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

Lemma 8 implies that each term in the sum of the l.h.s. vanishes. Expanding the remaining term $kp(q-p)^{k-1}$ and dividing both sides by p gives

$$(1-k)q^{k-1} + \sum_{i=1}^{k-1} (k-i) \binom{k}{i} p^i q^{k-1-i} \equiv 0 \pmod{p^n}$$
$$(1-k)q^{k-1} \equiv 0 \pmod{p^n}.$$

This proves by induction that $k \equiv 1 \pmod{p^n}$ for any n, which leads to a contradiction and concludes the proof in this case.

Assume now that $p \equiv 2 \pmod 4$ and $k \equiv 1 \pmod{p^{n-1}}$ for some $n \geq 2$. We first observe that

$$q^k = (p+q-p)^k = \sum_{i \in E} \binom{k}{i} p^i (q-p)^{k-i} + \sum_{i \notin E} \binom{k}{i} p^i (q-p)^{k-i}.$$

Thus (10) can be written as

$$\sum_{i \notin E} \binom{k}{i} p^i (q-p)^{k-i} = q^k - pq^{k-1}. \tag{12}$$

Taking the reduction of the latter equation modulo p^{n+1} and using that $\binom{k}{i} p^i \equiv 0 \pmod{p^{n+1}}$ when $i > 2$ gives

$$(q-p)^k + \mathbf{1}_{2 \notin E} \binom{k}{2} p^2 (q-p)^{k-2} \equiv q^k - pq^{k-1} \pmod{p^{n+1}}$$
$$q^k - kq^{k-1}p + \binom{k}{2} q^{k-2} p^2 + \mathbf{1}_{2 \notin E} \binom{k}{2} p^2 (q-p)^{k-2} \equiv q^k - pq^{k-1} \pmod{p^{n+1}}$$
$$(1-k)pq^{k-1} + \binom{k}{2} q^{k-2} p^2 + \mathbf{1}_{2 \notin E} \binom{k}{2} p^2 (q-p)^{k-2} \equiv 0 \pmod{p^{n+1}}. \tag{13}$$

We focus on the last term of the left-hand side and write:

$$\begin{aligned}
\mathbf{1}_{2\notin E}\binom{k}{2}p^2(q-p)^{k-2} &= \mathbf{1}_{2\notin E}\binom{k}{2}p^2\sum_{i=0}^{k-2}\binom{k-2}{i}p^iq^{k-2-i}\\
&= \mathbf{1}_{2\notin E}\binom{k}{2}p^2q^{k-2} + \Big(k(k-1)p^2\Big)\mathbf{1}_{2\notin E}\frac{p}{2}\sum_{i=1}^{k-2}\binom{k-2}{i}p^{i-1}q^{k-2-i}\\
&\equiv \mathbf{1}_{2\notin E}\binom{k}{2}p^2q^{k-2} \pmod{p^{n+1}},
\end{aligned}$$

since $p/2 \in \mathbb{N}$ and $(k-1)p^2 \equiv 0 \pmod{p^{n+1}}$.

Then (13) divided by p can be written

$$(1-k)q^{k-1} + (1+\mathbf{1}_{2\notin E})\binom{k}{2}q^{k-2}p \equiv 0 \pmod{p^n}.$$

If $2 \notin E$, the second term disappears and we are left with $(1-k)q^{k-1} \equiv 0 \pmod{p^n}$ and hence $k \equiv 1 \pmod{p^n}$, otherwise we get

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

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

These 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 $x \otimes y$?

As already mentioned, it is proved in [BFK12] 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 ; [BFK12] suggests that this should be the case. It also would be interesting to study whether adding more colors but still requiring that the k balls all turn into the same color has a bigger computational power.

Acknowledgements

We would like to thank O.Bournez and J.Cohen for some very interesting discussions during the preparation of [BCC+09] that raised our interest for the subject and for showing us a preliminary version of [BFK12].

References

[AAE06] D. Angluin, J. Aspnes, and D. Eisenstat. Stably computable predicates are semilinear. In *Proceedings of PODC'06 :* Principles Of Distributed Computing, 2006.

[AB11] G. Aupy and O. Bournez. On the number of binary-minded individuals required to compute $1/\sqrt{2}$. *Theoretical Computer Science*, 412(22):2262–2267, 2011.

[AR07] J. Aspnes and E. Ruppert. An introduction to population protocols. *Bulletin of the European Association for Theoretical Computer Science*, 93:98–117, 2007.

[BCC$^+$09] O. Bournez, Ph. Chassaing, J. Cohen, L. Gerin, and X. Koegler. On the convergence of population protocols when population goes to infinity. *Applied Mathematics and Computation*, 215(4):1340–1350, 2009.

[BFK12] O. Bournez, P. Fraigniaud, and X. Koegler. Computing with large populations. In *Proceedings of MFCS'12 :* Mathematical Foundations of Computer Science, 2012.

[CS08] I. Chatzigiannakis and P. Spirakis. The dynamics of probabilistic population protocols. In *Proceedings of DISC'08 :* International Symposium on Distributed Computing, 2008.

[DN08] R. Darling and J. Norris. Differential equation approximations for markov chains. *Probability Surveys*, (5), 2008.

[Dur91] R. Durrett. *Probability, Theory and Examples*. The Wadsworth & Brooks/Cole Statistics/Probability Series, 1991.

[Lin92] L. Lindvall. *Lectures on the coupling method*. John Wiley & Sons, 1992.

[SW95] A. Shwartz and A. Weiss. *Large deviations for performance analysis*. Chapman & Hall, 1995.