PHYSARUM CAN COMPUTE SHORTEST PATHS: A SHORT PROOF

VINCENZO BONIFACI

ABSTRACT. The purpose of this note is to give a short proof that a standard model for the *Physarum polycephalum* slime mold correctly computes the shortest path in an undirected weighted graph [V. Bonifaci, K. Mehlhorn, G. Varma, *Physarum can compute shortest paths*, in Proc. of the 23rd ACM-SIAM Symposium on Discrete Algorithms, pp. 233–240, SIAM 2012].

1. INTRODUCTION

Physarum polycephalum is a unicellular slime mold that shows remarkable computing capabilities, including the ability to compute shortest paths [6]. Tero, Kobayashi and Nakagaki [8] proposed a mathematical model for the dynamics of the Physarum, which in the computer simulations converges to the shortest path on any input graph. The convergence of this model to the shortest path has been analytically proven by Bonifaci, Mehlhorn and Varma [2, 3]. The model can therefore be seen as a *natural algorithm* for computing shortest paths [4]. The purpose of this note is to give a short proof of the convergence property; in other words, a short proof of the algorithm's correctness.

1.1. **Model of computation.** Let $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the set of real, nonnegative real, and positive real numbers, respectively, and for any two sets F and I let F^I denote the set of vectors whose components have values in F and are indexed by I. Let G be an undirected graph with node set N, edge set E, length labels $l \in \mathbb{R}_{++}^E$ and two distinguished nodes $s_0, s_1 \in N$. In our discussion, $x \in \mathbb{R}_+^E$ will be a state vector representing the diameters of the tubular channels of the Physarum (edges of the graph). The value x_e is called the *capacity* of edge e. The nodes s_0 and s_1 represent the location of two food sources. Physarum's dynamical system is described by the system of differential equations [8]

$$(1) \qquad \dot{x} = |q(x,l)| - x.$$

Equation (1) is called the *evolution equation*, as it determines the dynamics of the system over time. It is a compact representation of a system of ordinary differential equations, one for every edge of the graph; the absolute value operator $|\cdot|$ is applied componentwise. The vector $q \in \mathbb{R}^E$, known as the *current flow*, is determined by the capacity and length of the edges, as follows. Without loss of generality, assume that $N = \{1, 2, \ldots, n\}$, $E = \{1, 2, \ldots, m\}$ and assume an arbitrary orientation of the edges. Let $A = (a_{ve})_{v \in N, e \in M}$ be the incidence matrix of G under this orientation, that is,

$$a_{ve} \stackrel{\text{def}}{=} \begin{cases} +1 & \text{if } v \text{ is the tail of } e \\ -1 & \text{if } v \text{ is the head of } e \\ 0 & \text{otherwise.} \end{cases}$$

Then q is defined as the unit-value flow from s_0 to s_1 of minimum energy, that is, as the unique optimal solution to the following continuous quadratic optimization problem, related to *Thomson's principle* from physics [1, Theorem IX.2]:

$$(2) \qquad \min q^T R q$$
$$\text{s.t. } Aq = b.$$

Key words and phrases. Natural algorithm; shortest path; Thomson's principle.

Here, $R \overset{\text{def}}{=} \text{diag}(l/x) \in \mathbb{R}^{E \times E}$ is the diagonal matrix with value $r_e \overset{\text{def}}{=} l_e/x_e$ for the e-th element of the main diagonal, and $b \in \mathbb{R}^N$ is the vector defined by

$$b_v \overset{\text{def}}{=} \begin{cases} 0 & \text{if } v \notin \{s_0, s_1\}, \\ +1 & \text{if } v = s_0, \\ -1 & \text{if } v = s_1. \end{cases}$$

The value r_e is called the *resistance* of edge e, while b_v is called the *divergence* of q at v. Node s_0 is called the *source*, node s_1 the *sink*. The quantity $\eta \overset{\text{def}}{=} q^T R q$ is the *energy*; the quantity $b_{s_0} = 1$ is the *value* of the flow q.

The optimality conditions for (2) imply that there exist values $p_1, \ldots, p_n \in \mathbb{R}$ (*potentials*) that satisfy *Ohm's law* [1, Section II.1]:

$$(3) \qquad\qquad q_e = (p_u - p_v)/r_e, \qquad \text{whenever edge } e \text{ is oriented from } u \text{ to } v.$$

By the *conservation of energy* principle, the total energy equals the difference between the source and sink potentials, times the value of the flow [1, Corollary IX.4]:

$$(4) \qquad\qquad \eta = (p_{s_0} - p_{s_1})b_{s_0} = p_{s_0} - p_{s_1}.$$

1.2. **Assumptions on the input data.** We assume that the length labels l and the initial conditions $x(0)$ satisfy the following:

    a. each s_0-s_1 path in G has a distinct overall length;
    b. all initial capacities are nonzero:

$$(5) \qquad\qquad\qquad\qquad\qquad\qquad x(0) > 0;$$

    c. the initially minimum capacity cut is the source cut, and it has unit capacity:

$$(6) \qquad\qquad \mathbf{1}_S^T \cdot x(0) \geq \mathbf{1}_0^T \cdot x(0) = 1, \qquad \text{for any } s_0\text{-}s_1 \text{ cut } S,$$

    where $\mathbf{1}_S$ is the characteristic vector of the set of edges in the cut S, and $\mathbf{1}_0$ is the characteristic vector of the set of edges incident to the source.

Condition (a) states that the configuration of lengths is nondegenerate. It may be weakened or eliminated altogether, at the cost of weakened claims on the convergence properties.

Condition (b) is without loss of generality: edges with $x_e(0) = 0$ never change their state under the dynamics; they may be removed from the graph, with no effect on the dynamics. Edges used by the system at any finite time, or in the limit, are always included in the initial set of edges.

Condition (c) can be ensured by normalizing the capacity of each edge by the capacity of the minimum capacity edge, and by adding an edge $e_0 = \{s_0', s_0\}$ with length (say) 1, initial capacity 1 and redefining the source node to be s_0'; this clearly does not affect the structure of shortest paths in G. For an analysis that does not rely on conditions (a) and (c), we refer the reader to [3].

Last but not least, we observe that if at any time the roles of s_0 and s_1 are reversed, that is, s_1 becomes the source and s_0 the sink, the dynamics (1) are unaffected. The model is therefore symmetrical with respect to the two terminals s_0 and s_1.

1.3. **Comparison with the previous proof.** The goal of this paper was to give a simpler and more readable proof than the one in [3]. For example, the previous proof required a rather technical lemma based on Kirchhoff's Matrix-Tree Theorem, which is not needed here, and the potential function used for the previous convergence proof was nonlinear and not everywhere differentiable, while here we use a linear potential function. On the other hand, a small advantage of the previous proof is that it does not require the input normalization (c) discussed in the previous paragraph.

2. Basic properties

Lemma 2.1 (Gronwall)**.** *Let $\alpha, \beta \in \mathbb{R}$ and let f be a continuous differentiable real function on $[0, \infty)$. If $\alpha f(t) \leq \dot{f}(t) \leq \beta f(t)$ for all $t \geq 0$, then*

$$(7) \qquad\qquad f(0)e^{\alpha t} \leq f(t) \leq f(0)e^{\beta t} \qquad \text{for all } t \geq 0.$$

Proof.

$$\frac{d}{dt}\frac{f}{e^{\beta t}} = \frac{\dot{f}e^{\beta t} - \beta f e^{\beta t}}{e^{2\beta t}} \leq 0 \text{ implies } \frac{f(t)}{e^{\beta t}} \leq \frac{f(0)}{e^{\beta 0}} = f(0).$$

A similar calculation establishes $f(t) \geq f(0)e^{\alpha t}$. $\qquad\square$

Lemma 2.2. *Let $x = x(t)$ be the solution of (1) under the initial conditions $x(0) > 0$. The following properties hold at any time $t \geq 0$:*

    a. $x > 0$;
    b. $\mathbf{1}_S^T \cdot x \geq \mathbf{1}_0^T \cdot x = 1$.

Moreover, the following properties hold at any time $t > t_0$, where t_0 is some finite time:

    c. $x < 2$.

Proof. (a.) Let $e \in E$ be any edge. Since $|q_e| \geq 0$, by the evolution equation (1) we have $\dot{x}_e = |q_e| - x_e \geq -x_e$. Gronwall's lemma with $\alpha = -1$ yields $x_e(t) \geq x_e(0)\exp(-t) > 0$, by (5).

(b.) Summing (1) for the edges in cut S yields

$$\mathbf{1}_S^T \cdot \dot{x} = \mathbf{1}_S^T \cdot |q| - \mathbf{1}_S^T \cdot x \geq 1 - \mathbf{1}_S^T \cdot x,$$

where the inequality follows from the fact that a total current of 1 is sent from s_0 to s_1. Gronwall's lemma applied to $\mathbf{1}_S^T \cdot x - 1$ yields

$$(8) \qquad \mathbf{1}_S^T \cdot x \geq 1 + \left(\mathbf{1}_S^T \cdot x(0) - 1\right)\exp(-t) \geq 1.$$

On the other hand, for the source cut one has $\mathbf{1}_0^T \cdot |q| = \mathbf{1}_0^T \cdot q = 1$ (since currents can only leave the source), so that

$$\mathbf{1}_0^T \cdot \dot{x} = 1 - \mathbf{1}_0^T \cdot x.$$

Gronwall's lemma applied with $\alpha = \beta = -1$ therefore yields, together with (6),

$$(9) \qquad \mathbf{1}_0^T \cdot x = 1 + \left(\mathbf{1}_0^T \cdot x(0) - 1\right)\exp(-t) = 1,$$

and the claim follows by comparing (8) and (9).

(c.) Since $\dot{x}_e = |q_e| - x_e \leq 1 - x_e$, again by Gronwall's lemma we obtain

$$x_e(t) \leq 1 + (x_e(0) - 1)\exp(-t),$$

which yields the claim as long as $t_0 \geq \ln\max_e x_e(0)$. $\qquad\square$

An *equilibrium point* of (1) is a vector $x \in \mathbb{R}_+^E$ such that $\dot{x} = 0$. Our assumptions imply that there are a finite number of equilibrium points: each equilibrium corresponds to an s_0-s_1 path of the network, and vice versa, as we now show.

Lemma 2.3. *If $x = \mathbf{1}_P$ for some s_0-s_1 path P, then x is an equilibrium point. Conversely, if x is an equilibrium point, then $x = \mathbf{1}_P$ for some s_0-s_1 path P.*

Proof. If $x = \mathbf{1}_P$, then the current flow is nonzero only on the path P, and since the value of the flow is 1, $|q| = \mathbf{1}_P = x$ and hence $\dot{x} = |q| - x = 0$.

Conversely, if x is an equilibrium point, $|q_e| = x_e$ for any edge $e \in E$. For the remainder of the lemma it is convenient to assume that edges are oriented consistently with the current flow, so that $q \geq 0$ and $q_e = x_e$ for all e (in general, such an orientation may change with time). Whenever x_e is nonzero, dividing by x_e yields, by (3), $l_e = p_u - p_v$ whenever e is oriented from u to v. Summing along any s_0-s_1 directed path P with nonzero x_e values yields $\mathbf{1}_P \cdot l = p_{s_0} - p_{s_1}$. However, we assumed that distinct s_0-s_1 paths have distinct lengths, that is, distinct values of $\mathbf{1}_P \cdot l$. Therefore there is exactly one directed path P such that $x_e > 0$ for all $e \in P$. So the support of x is a subgraph that is acyclic, contains exactly one path from s_0 to s_1, and has no source or sink other than s_0 and s_1; in other words, it is a path. The claim follows using Lemma 2.2(b). $\qquad\square$

3. Convergence to the shortest path

Recall that, by (4),

$$\eta = \sum_{e \in E} r_e q_e^2 = q^T R q = p_{s_0} - p_{s_1}, \tag{10}$$

and let

$$V \stackrel{\text{def}}{=} l^T x = \sum_{e \in E} l_e x_e = \sum_{e \in E} r_e x_e^2 = x^T R x. \tag{11}$$

Here η is the energy dissipated by the system, as well as the potential difference between source and sink. The quantity V has no immediate interpretation; however, it captures in some sense the "infrastructural cost" of the system. While η may decrease or increase during the evolution of the system, we will show that $\eta \leq V$ and that V is always decreasing, except on equilibrium points.

Lemma 3.1. $\eta \leq V$.

Proof. To see the inequality, consider any flow f of maximum value subject to the constraint that $0 \leq f \leq x$. The minimum capacity of a source-sink cut is 1 at any time, by Lemma 2.2(b). Therefore, by the Max Flow-Min Cut Theorem, the value of the flow f must be 1. Then by (2),

$$\eta = q^T R q \leq f^T R f \leq x^T R x = V. \quad \square$$

Lemma 3.2. V *is a Lyapunov function for* (1)*; in other words, it is continuous (as a function of* x*) and satisfies*

(i) $V \geq 0$;
(ii) $\dot{V} \leq 0$.

Moreover, $\dot{V} = 0$ *if and only if* $\dot{x} = 0$.

Proof. V is continuous and nonnegative by construction. Moreover,

$$\begin{aligned}
\dot{V} = l^T \dot{x} = l^T (|q| - x) && \text{by (1)}, \\
= x^T R |q| - x^T R x && \text{by } l = Rx, \\
= (x^T R^{1/2}) \cdot (R^{1/2} |q|) - x^T R x && \\
\leq (x^T R x)^{1/2} \cdot (q^T R q)^{1/2} - x^T R x && \text{by Cauchy-Schwarz [7]}, \\
\leq (x^T R x)^{1/2} \cdot (x^T R x)^{1/2} - x^T R x && \text{by Lemma 3.1}. \\
= 0.
\end{aligned}$$

Observe that $\dot{V} = 0$ is possible only when equality holds in the Cauchy-Schwarz inequality. This, in turn, implies that the two vectors $R^{1/2} x$ and $R^{1/2} |q|$ are parallel, that is, $|q| = \lambda x$ for some $\lambda \in \mathbb{R}$. However, by Lemma 2.2(b), the capacity of the source cut is 1 and, by (2), the sum of the currents across the source cut is 1. Therefore, $\lambda = 1$ and $\dot{x} = |q| - x = 0$. $\square$

Corollary 3.3. *As* $t \to \infty$*,* $x(t)$ *approaches an equilibrium point of* (1)*, and* $\eta(t)$ *approaches the length of the corresponding* s_0*-*s_1 *path.*

Proof. The existence of a Lyapunov function V implies [5, Corollary 2.6.5] that $x(t)$ approaches the set $\{x \in \mathbb{R}_+^E : \dot{V} = 0\}$, which by Lemma 3.2 is the same as the set $\{x \in \mathbb{R}_+^E : \dot{x} = 0\}$. Since this set consists of isolated points (Lemma 2.3), $x(t)$ must approach one of those points, say the point $\mathbf{1}_P$ for some s_0-s_1 path P. When $x = \mathbf{1}_P$, one has $\eta = V = \mathbf{1}_P^T \cdot l$. $\square$

Theorem 3.4. *As* $t \to \infty$*,* $\eta(t)$ *converges to* $\mathbf{1}_{P^*}^T \cdot l$*, where* P^* *is the minimum length path from* s_0 *to* s_1*, and* $x(t)$ *converges to* $\mathbf{1}_{P^*}$*.*

Proof. By Corollary 3.3, it suffices to prove the first part of the claim. Suppose by contradiction η converges to the length of some other s_0-s_1 path $P \neq P^*$. Let $\delta = (\mathbf{1}_P - \mathbf{1}_{P^*})^T \cdot l/2 > 0$ and

$W = \sum_{e \in P^*} l_e \ln x_e$. Then for all sufficiently large t, we may assume $\eta \geq \mathbf{1}_P^T l - \delta = \mathbf{1}_{P^*}^T l + \delta$, and therefore

$$\dot{W} = \sum_{e \in P^*} \frac{l_e}{x_e} \dot{x}_e \qquad \text{by the chain rule of differentiation,}$$

$$= \sum_{e \in P^*} \frac{l_e}{x_e}(|q_e| - x_e) \qquad \text{by (1),}$$

$$= \sum_{e \in P^*} |p_u - p_v| - \sum_{e \in P^*} l_e \qquad \text{by (3),}$$

$$\geq \eta - \mathbf{1}_{P^*}^T l \qquad \text{by triangle inequality and (4),}$$

$$\geq \delta, \qquad \text{by assumption,}$$

but this implies $W \to +\infty$, which is a contradiction since Lemma 2.2(c) implies that $W < (\ln 2) \cdot \sum_{e \in P^*} l_e$ for all sufficiently large t. $\qquad \square$

References

[1] B. Bollobás. *Modern Graph Theory*. Springer, New York, 1998.

[2] V. Bonifaci, K. Mehlhorn, and G. Varma. Physarum can compute shortest paths. In Y. Rabani, editor, *Proc. of the 23rd ACM-SIAM Symposium on Discrete Algorithms*, pages 233–240. SIAM, 2012.

[3] V. Bonifaci, K. Mehlhorn, and G. Varma. Physarum can compute shortest paths. *Journal of Theoretical Biology*, 309:121–133, 2012. doi:10.1016/j.jtbi.2012.06.017.

[4] B. Chazelle. Natural algorithms. In C. Mathieu, editor, *Proc. of the 20th ACM-SIAM Symposium on Discrete Algorithms*, pages 422–431, Philadelphia, PA, 2009. SIAM.

[5] J. B. LaSalle. *The Stability of Dynamical Systems*. SIAM, 1976.

[6] T. Nakagaki, H. Yamada, and Á. Tóth. Maze-solving by an amoeboid organism. *Nature*, 407:470, 2000.

[7] J. Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, 2004.

[8] A. Tero, R. Kobayashi, and T. Nakagaki. A mathematical model for adaptive transport network in path finding by true slime mold. *Journal of Theoretical Biology*, 244:553–564, 2007.

Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Rome, Italy

E-mail address: `vincenzo.bonifaci@iasi.cnr.it`