



**HAL**  
open science

# An inequality involving nonnegative matrices and increasing functions

Jean-Luc Gouzé

► **To cite this version:**

Jean-Luc Gouzé. An inequality involving nonnegative matrices and increasing functions. [Research Report] RR-1330, INRIA. 1990, pp.6. inria-00075230

**HAL Id: inria-00075230**

**<https://inria.hal.science/inria-00075230>**

Submitted on 24 May 2006

**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.

# IRIA

UNITÉ DE RECHERCHE  
IRIA-SOPHIA ANTIPOLIS

Institut National  
de Recherche  
en Informatique  
et en Automatique

Domaine de Voluceau  
Rocquencourt  
B.P.105  
78153 Le Chesnay Cedex  
France  
Tél.: (1) 39 63 55 11

## Rapports de Recherche

N° 1330

*Programme 5*  
*Automatique, Productique,*  
*Traitement du Signal et des Données*

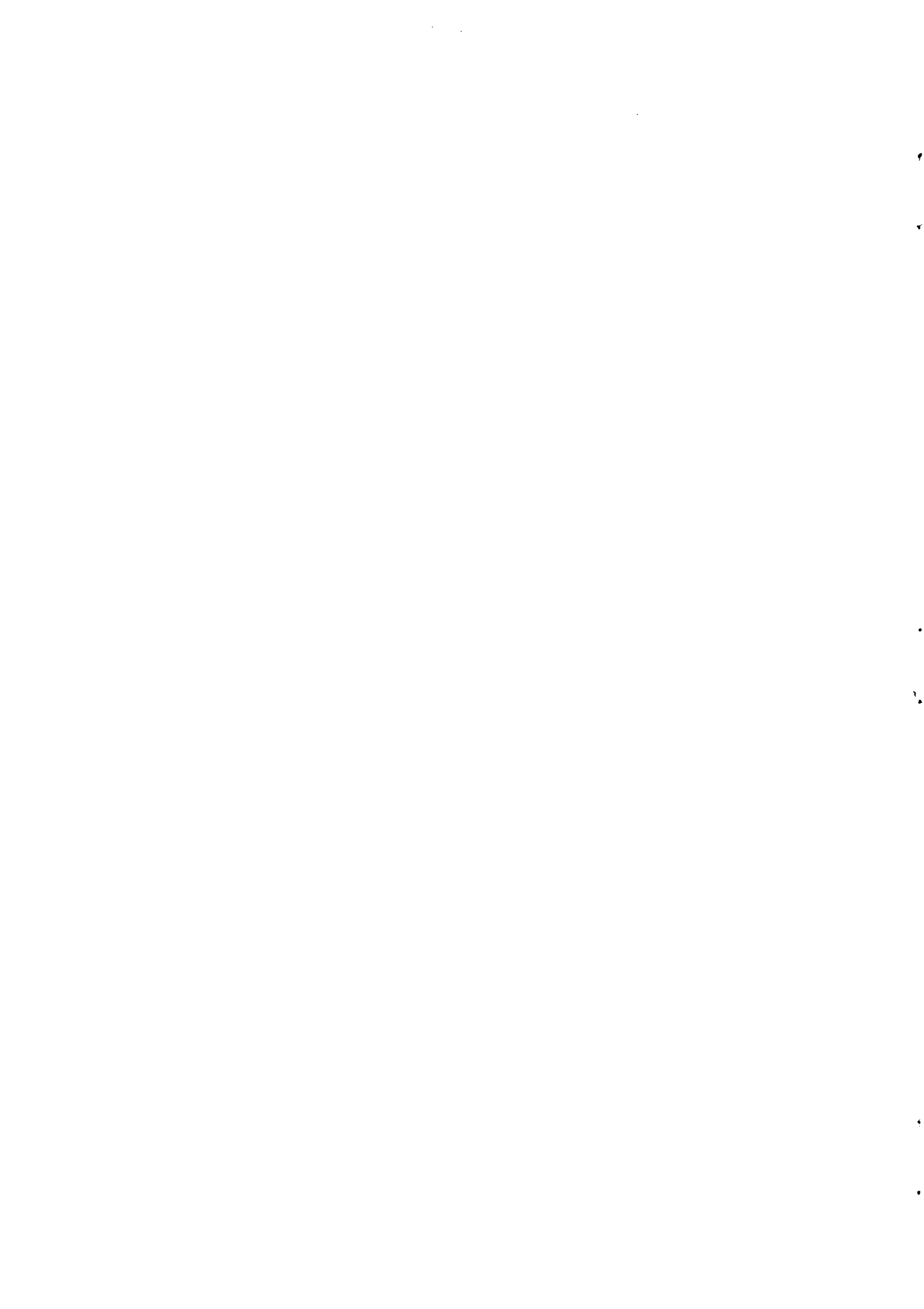
### AN INEQUALITY INVOLVING NONNEGATIVE MATRICES AND INCREASING FUNCTIONS

Jean-Luc GOUZÉ

Octobre 1990



\* R R . 1 3 3 8 \*



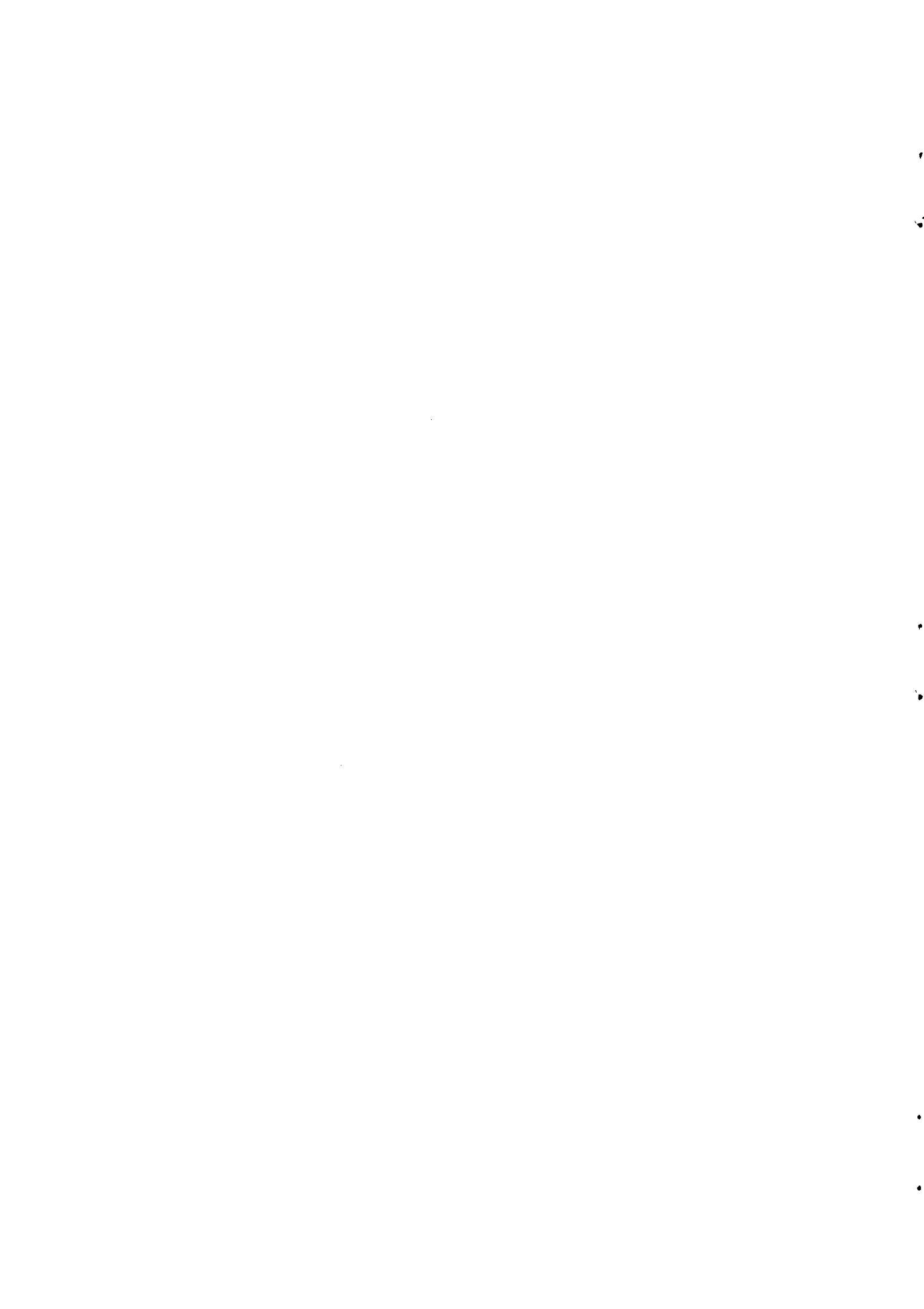
# An inequality involving nonnegative matrices and increasing functions

Jean-Luc Gouzé  
INRIA  
BP 109  
Sophia-Antipolis, 06561 Valbonne Cedex  
France

**Abstract.** Let  $\tilde{f}(x)$  denote the vector  $(f(x_1), \dots, f(x_n))^t$ ; if  $A$  is a doubly stochastic matrix and  $f$  a strictly increasing function, we demonstrate that  $x^t A \tilde{f}(x) - x^t \tilde{f}(x) \leq 0$  for all vector  $x$ ; this inequality has applications in the study of non-linear differential systems.

## Une inégalité “entropique” entre matrices positives et fonctions monotones

**Résumé :** Notons par  $\tilde{f}(x)$  le vecteur  $(f(x_1), \dots, f(x_n))^t$ ; si  $A$  est une matrice positive doublement stochastique et si  $f$  est une fonction réelle strictement croissante, nous démontrons que  $x^t A \tilde{f}(x) - x^t \tilde{f}(x) \leq 0$  pour tout  $x$ ; cette inégalité a des applications pour la construction de fonctions de Lyapounov pour des systèmes différentiels non linéaires.



## 1 Introduction and main result

We want here to prove an inequality involving nonnegative matrices (particularly stochastic and doubly stochastic matrices) and strictly increasing functions. This inequality is by itself interesting, and can also be used to construct Lyapunov functions for non-linear differential  $n$ -dimensional systems, and consequently to prove convergence towards equilibria. For some particular cases, this inequality has been proved in the framework of probabilistic problems ([5]) or of convergence of chemical differential systems ([3]); it can be interpreted as the increasing character of something looking like entropy (cf. [6]).

We first state the result; let  $f$  a real function of a real variable, defined in some real subset  $D$ , and strictly increasing. If  $x$  is a vector in  $\mathbf{R}^n$ , we denote by  $\tilde{f}(x)$  the vector  $(f(x_1), \dots, f(x_n))^t$ . If  $A$  is a real  $(n \times n)$  matrix, the non-oriented graph of  $A$  is partitioned into  $p$  connected classes, corresponding to  $p$  submatrices  $A_i$ .  $A$  is called doubly stochastic if it is a nonnegative matrix verifying  $A\mathbf{1} = \mathbf{1}$  and  $A^t\mathbf{1} = \mathbf{1}$  ( $\mathbf{1}$  is the vector with all the components equal to 1). ( cf. [1]).

**Theorem 1.1** *Let  $A$  a doubly stochastic matrix and*

$$\phi_A(x) = x^t A \tilde{f}(x) - x^t \tilde{f}(x)$$

*Then*

- $\phi_A(x) \leq 0$  for all  $x \in D^n$
- $\phi_A(x) = 0 \Leftrightarrow x = \sum_{i=1}^p \lambda_i \mathbf{1}_i$   
*where  $\mathbf{1}_i$  is the vector having the components corresponding to the vertices of the graph of  $A_i$  equal to 1, and the others equal to 0, and the  $\lambda_i$  are real.*

We remark that  $\phi_A(x) = x^t (A - I) \tilde{f}(x)$ , and that  $M = A - I$  is a singular off-diagonal nonnegative matrix such that  $M\mathbf{1} = 0, M^t\mathbf{1} = 0$  (cf. [1]). We deduce the following corollary:

**Corollary 1** *Let  $M$  a singular off-diagonal nonnegative matrix such that  $M\mathbf{1} = 0, M^t\mathbf{1} = 0$  and*

$$\phi'_M(x) = x^t M \tilde{f}(x)$$

*Then*

- $\phi'_M(x) \leq 0$  for all  $x \in D^n$
- $\phi'_M(x) = 0 \Leftrightarrow x = \sum_{i=1}^p \lambda_i \mathbf{1}_i$   
 where  $\mathbf{1}_i$  is the vector having the components corresponding to the vertices of the graph of  $M$ ; equal to 1, and the others equal to 0, and the  $\lambda_i$  are real.

To prove this corollary, we choose  $\alpha > 0$  such that  $M + \alpha I$  is a nonnegative matrix; then  $\frac{1}{\alpha}M + I$  is a doubly stochastic matrix and we can apply the theorem to obtain the result.

It is now easy to deduce a new corollary that extend the application of the theorem; if  $k$  is a vector with each component non equal to zero, we denote by  $x/k$  the vector  $(x_1/k_1, \dots, x_n/k_n)$ , and by  $(\text{diag } k)$  the diagonal matrix with diagonal  $k$ .

**Corollary 2** *Let  $B$  a nonnegative matrix verifying  $B\mathbf{1} = \mathbf{1}$  and  $B^t k = k$  with  $k > 0$  (that is  $k_i > 0$  for all  $i$ ). Let:*

$$\phi''_B(x) = x^t B \tilde{f}(x/k) - x^t \tilde{f}(x/k)$$

Then

- $\phi''_B(x) \leq 0$  for all  $(x/k) \in D^n$
- $\phi''_B(x) = 0 \Leftrightarrow x/k = \sum_{i=1}^p \lambda_i \mathbf{1}_i$   
 where  $\mathbf{1}_i$  is the vector having the components corresponding to the vertices of the graph of  $B$ ; equal to 1, and the others equal to 0, and the  $\lambda_i$  are real.

Indeed, if we let  $z = x/k$ , it is easy to see that  $\phi''_B(x) = \phi'_{(\text{diag } k)(B-I)}(z)$  and we can apply the corollary to  $M = (\text{diag } k)(B - I)$ .

Let us remark that if  $B$  is a stochastic matrix ( $B\mathbf{1} = \mathbf{1}$ ) without transitory states, it implies ([1]) the existence of  $k > 0$  such that  $B^t k = k$ .

We now demonstrate the inequality before giving applications.

## 2 Demonstration of the inequality:

We first prove the theorem when  $A$  is a permutation matrix. The problem is therefore to study:

$$\phi_P(x) = x^t P \tilde{f}(x) - x^t \tilde{f}(x)$$

where  $P$  is a  $n \times n$  permutation matrix. We can, in a unique way, decompose  $P$ :

$$P = \sum_{i=1}^p P_i^0 \quad (1)$$

where  $P_i^0$  is a cycle of length  $l_i$  (with  $1 \leq l_i \leq n$ ), that is to say a cyclic permutation on  $l_i$  elements, and the null application for the other elements. Of course:

$$\sum_{i=1}^p l_i = n$$

and there is no intersection between the  $p$  classes. This is equivalent to say that the (non-oriented) graph is made of  $p$  disconnected classes, each one being a cycle with matrix  $P_i^0$ . The problem is now reduced to study:

$$\phi_{P_i}(y) = y^t P_i \bar{f}(y) - y^t \bar{f}(y)$$

where  $P_i$  is a cycle on  $l_i$  elements and  $y$  a vector of dimension  $l_i$ . Without any restriction, we can always suppose that the cycle is:

$$y_1 \rightarrow y_2 \rightarrow \dots \rightarrow y_{l_i} \rightarrow y_1$$

We can now demonstrate (cf. [3]):

**Lemma 1**

$$\phi_{P_i}(y) \leq 0 \text{ for all } y$$

and ( $\lambda$  being a real)

$$\phi_{P_i}(y) = 0 \Leftrightarrow y = \lambda \mathbf{1}$$

*Proof:*

$$\phi_{P_i}(y) = \sum_{j=1}^{l_i} (y_j f(y_{j+1}) - y_j f(y_j))$$

where by convention  $y_{l_i+1} = y_1$ . We prove the lemma by induction on  $l_i$ ; if  $l_i = 1$ , the result is trivial; if  $l_i = 2$ , then :

$$\phi_{P_i}(y) = (y_1 - y_2)(f(y_2) - f(y_1))$$

and the result is true because  $f(x)$  is a strictly increasing function. We suppose that the result is true for  $l_i - 1$ ; we can also suppose without any restriction of the problem that  $y_{l_i} \geq y_j$  ( $j = 1, \dots, l_i - 1$ ). Then:

$$\begin{aligned} \phi_{P_i}(y) &= [y_1 f(y_2) - y_1 f(y_1) + \dots + y_{l_i-1} f(y_1) - y_{l_i-1} f(y_{l_i-1})] \\ &\quad + (y_{l_i-1} - y_{l_i})(f(y_{l_i}) - f(y_1)) \end{aligned}$$



The first part of the right member is nonpositive because of the induction hypothesis, and the second part because  $f$  is strictly increasing. If  $\phi(y)$  is zero, then  $y_1 = y_2 = \dots = y_{i,-1}$  because of the induction hypothesis, and moreover  $y_{i,-1} = y_i$ , or  $y_1 = y_i$ , so finally  $y = \lambda \mathbf{1}$ .

It is now a simple matter to obtain the result for  $\phi_P$ :

**Lemma 2**

$$\phi_P(y) \leq 0 \text{ for all } y \in D^n$$

and  $(\lambda_i \text{ being real})$

$$\phi_P(y) = 0 \Leftrightarrow y = \sum_{i=1}^p \lambda_i \mathbf{1}_i$$

where  $\mathbf{1}_i$  is a vector having the components corresponding to the vertices of the graph of  $P_i^0$  equal to 1, and the others equal to 0.

Indeed, it is enough to apply the preceding lemma to each cycle of the permutation by using (1).

Take now a doubly stochastic matrix  $A$ ; by a theorem of Birkhoff (cf. [1]), we know that  $A$  is a convex combination of permutation matrices; we deduce that  $\phi_A(y)$  is a convex combination of  $\phi_P(y)$ , and therefore nonpositive.

If  $\phi_A(y)$  is zero, consider a connected class  $A_c$  of the graph of  $A$ ; then  $\phi_{A_c}$  is also zero;  $A_c$  has a convex decomposition into permutations  $P$ , and each  $\phi_P(y)$  must be zero; moreover, the graphs of these permutations realise a (non-disjoint) covering of the graph of  $A_c$ , and therefore we must have equality for all components belonging to  $A_c$ . The theorem is therefore demonstrated.

### 3 Applications

- We can first make explicit the function  $f$  to obtain useful inequalities. If we take  $f(x) = x$ , we obtain, by using the corollary 1, the result that  $(B - I)\text{diag } k + \text{diag } k(B^t - I)$  is semidefinite negative if  $B$  is a stochastic matrix; this result is simple to obtain directly ([1]).

Take now  $f(x) = e^x$ ; we obtain:

$$x^t B e^{x/k} - x^t e^{x/k} \leq 0$$

in particular  $x^t M e^x \leq 0$  for all singular off-diagonal nonnegative matrix  $M$  such that  $M\mathbf{1} = 0$  and  $M^t\mathbf{1} = 0$ .

Let us take now  $f(p) = \ln p$ , ( $p > 0$ ), we obtain:

$$p^t(B - I)(\ln p - \ln k) \leq 0 \quad (2)$$

for all  $p > 0$  (cf. [5]).

- The main interest of this inequality for dynamical systems is to enable one to construct Lyapunov functions or auxiliary functions that decrease along the trajectories of the system (cf. [4]). Take the differential  $n$ -dimensional non-linear system:

$$\dot{y} = M \tilde{f}(y)$$

where  $M$  is a matrix verifying the hypotheses of the corollary 2 and  $f$  is  $C^1$  (and always strictly increasing). Then the norm function  $V(y) = \|y\|^2$  is decreasing along the trajectories because its derivative with respect to time is  $2\phi'_M(y)$ , and this expression cancels if and only if  $y$  is an equilibrium for the differential system. We deduce that all the trajectories of the system are bounded and converge towards the set of equilibria. For example, the inequality (2) has been used in [3] (in a concealed way) to show convergence to equilibrium of chemical systems. It is also possible to use this inequality (2) to obtain results of convergence for Lotka-Volterra differential systems in population dynamics ([2]).

- Take  $f(p) = \ln p$  for  $p > 0$ . The inequality given by the theorem is :

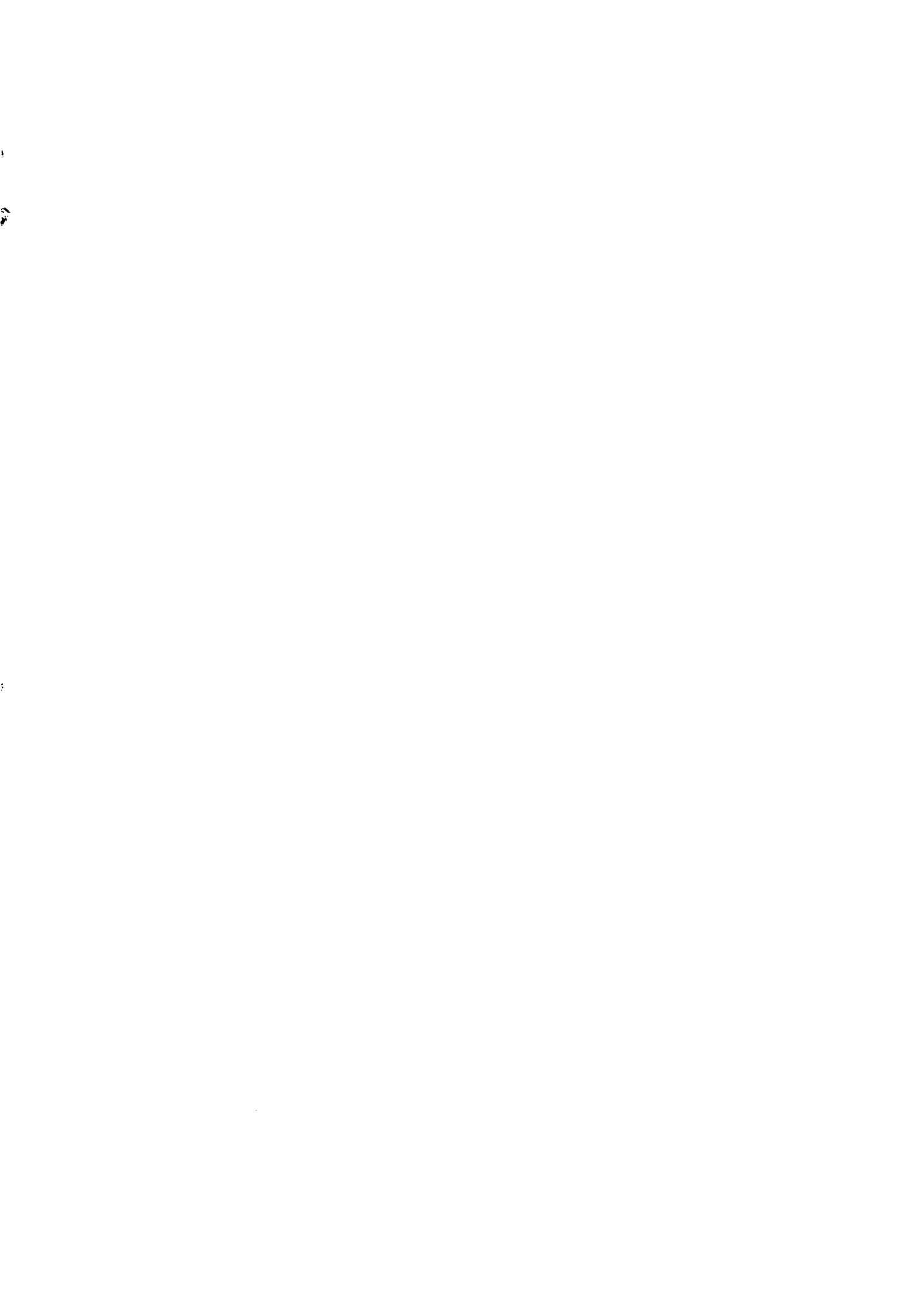
$$p^t A \ln p - p^t \ln p \leq 0$$

If the  $p_i$  are a probability distribution for a discrete finite system, the entropy of such a system is  $-p^t \ln p$  ([6]). It seems to us that one could see the preceding inequality as an "entropic inequality" saying that something that looks like entropy always increases when the system is equitably mixed.

## References

- [1] A. Berman and R.J. Plemmons, 1979. Nonnegative matrices in the mathematical sciences, Academic Press, New-York
- [2] J.L. Gouzé 1990. Global behaviour of  $n$ -dimensional Lotka-Volterra systems. Rapport de Recherche Inria 1350.

- [3] F. Horn and R. Jackson 1972. General mass action kinetics. Arch. Rational Mech. Anal., 47: 81-116
- [4] N. Rouche et J.Mahwin, 1973. Equations différentielles ordinaires, tomes 1 et 2, Masson, Paris
- [5] J. Schnakenberg, 1976. Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Modern Physics,48 :571-585
- [6] P. Walters 1982. An introduction to ergodic theory, Springer-Verlag, New-York



**ISSN 0249 - 6399**