# Towards an Algorithmic Theory of Stability

Mehran Mesbahi and George P. Papavassilopoulos
Department of Electrical Engineering-Systems
University of Southern California, CA 90089-2563
{mesbahi/yorgos}@bode.usc.edu

## Abstract

Let $\delta_\Sigma$ be a measure of the relative stability of a *stable* dynamical system $\Sigma$. Let $\tau_{\mathcal{A}(\Sigma)}$ be a measure of the computational efficiency of a particular algorithm $\mathcal{A}$ which *verifies* the stability property of $\Sigma$. For two representative cases of $\Sigma$, we demonstrate the existence of particular measure $\delta_\Sigma$ and algorithm $\mathcal{A}$ such that,

$$\delta_\Sigma \, \tau_{\mathcal{A}(\Sigma)} \;= c$$

where $c$ depends possibly on the dimension of the system $\Sigma$ and parameters specific to the algorithm $\mathcal{A}$, but independent of any other system characteristics. In particular, given $\Sigma$ and $\mathcal{A}$, one can determine $\delta_\Sigma$ by *measuring* $\tau_{\mathcal{A}(\Sigma)}$.

**Keywords:** Linear Matrix Inequalities; Stability Theory; Interior Point Methods; Conjugate Gradient Method

## 1. Introduction

The field of control and system theory on one hand, and computational complexity on the other, are generally not considered by the researcher of either field to have much in common. Recently, some control and system theorists have begun a serious study of control problems from the computational complexity point of view, e.g., classifying control problems in terms of the complexity class which they belong to [3], [9], [11], [14], [15]. This line of research is concerned with problems of determining whether a control problem is for example NP-complete, etc. These studies in principle convey the idea that the corresponding system problem, whether it is analysis or synthesis, is computationally *difficult*. One major issue which we believe has not been considered in this direction is the role that the theoretical studies of the computational efficiency of algorithms can play in analyzing systems problems that *can be solved efficiently*. Given a control or system problem that we can solve by means of an algorithm in a reasonable time (for example in time which is proportional to a polynomial of the dimension of the system), what does the *running time* of the algorithm disclose about some of the characteristics of the system under study. In this avenue, suppose that one wants to examine the stability properties of a certain dynamical system and we use an algorithm to check whether the system is stable or not. Thus we use an algorithm which accepts as input, a description of the system (e.g., in terms of matrices), and produces as an output "yes" or "no," indicating respectively, whether the system is stable or not stable. Suppose furthermore that the time required for the termination of this algorithm is proportional to the dimension of the system *and* another parameter denoted by $\xi$. We like to show that for certain particular problems in systems and control theory, there exist algorithms for which the corresponding $\xi$ can be viewed as a certain measure of *robustness*, e.g., stability margins or the amount of relative stability.

The major obstacle for the realization of this program is that up to very recently, our understanding of the *total* effort required by an algorithm to solve continuous optimization problems was rather limited. Since many control and system problems can be formulated in terms of optimization problems, our knowledge of the *total* computational effort of the algorithms that we have used to solve these problems was also insufficient for the development of the above approach. The advent of the interior point methods [6], [10], [13], and their precise complexity analysis for a wide array of convex optimization problems, including the linear matrix inequalities, has opened up a very promising avenue for the realization of the afro-mentioned program. The purpose of the present paper is to demonstrate that the computational efficiency of the conjugate gradient and interior point methods depend on, and can convey information about, the robustness properties of two important problems in system stability.

The paper is organized as follows. First, certain basic facts pertaining to the conjugate gradient and the interior point methods are provided in Section 2 which have direct relevance to our subsequent presentation. In Section 3, we demonstrate that the conjugate gradient method not only solves the Lyapunov equation in order to verify the stability property of a linear time invariant system, but also its running time depends on a suitable notion of relative stability. In Section 4, the same idea is elaborated on, this time using the interior point method for cases pertaining to the absolute stability problem. In Sections 3 and 4, we *do not* suggest that the algorithms that we use to solve the corresponding problems are in any sense *optimal*. For example, optimality does not hold in regards to the conjugate gradient method for solving the Lyapunov equation, since much more efficient algorithms are presently available for this purpose. The paper is then concluded with a brief after-thought on the motivations and the implications of the results presented in the paper.

First, a few words on the notation. We use $\otimes$ to denote Kronecker product, **vec** to indicate the operation on the matrices which stacks up the column of the matrix from left to right (and makes a long vector out of the matrix). $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ are used to designate eigenvalues of $A$ with the minimum and the maximum real part, respectively (in the case where more than one eigenvalue is a candidate for the minimum and the maximum, any one of them is chosen arbitrarily); $|\lambda|_{\min}(A)$ and $|\lambda|_{\max}(A)$ serve the similar purpose for the minimum and maximum of the eigenvalues of $A$ in absolute value. The notation herm$(A)$ denotes the hermitian part of the matrix $A$, i.e., $\frac{A+A'}{2}$, and for two symmetric matrices $A$ and $B$, $A > B$, indicates that $A - B$ is positive definite. As noted previously, $\tau_{\mathcal{A}(\Sigma)}$ designates the running time of the algorithm $\mathcal{A}$ that verifies the stability properties of the dynamical system $\Sigma$; when the latter task is done only approximately, e.g., the the solution of the Lyapunov equation is found within an $\epsilon$ approximation, then we still use $\tau_{\mathcal{A}(\Sigma)}$ instead of a more accurate notation $\tau_{\mathcal{A}(\Sigma)}(\epsilon)$ to represent the termination time. Finally, $f(n) = O(g(n))$ $(f(n) = \Omega(g(n)))$ indicates that there exist positive constants $c$ and $m$ such that $0 \leq f(n) \leq cg(n)$ $(0 \leq cg(n) \leq f(n))$, for all $n \geq m$.

## 2. IPMs and CGMs

In this section, we briefly provide some basic facts regarding the numerical algorithms which are the cornerstones of the two main theorems of the paper: the conjugate gradient method (cgm), and the interior point method (ipm). Although there are many variants of the ipms discussed in the literature, we shall focus mainly on the barrier method as presented by Renegar [13], and we shall use "ipm" synonymously with the "barrier method" as discussed in that reference. Our very brief discussion of the conjugate gradient algorithm also follows what is presented in another paper of Renegar [12], restricted to the $n$-dimensional Euclidean space.

Given a positive definite (symmetric) matrix $A \in R^{n\times n}$, a vector $b \in R^n$, an initial point $x_0$, in order to find the solution of $x^*$ of the equation $Ax = b$, the cgm produces the iterates $\{x_i\}_{i\geq 1}$, where $x_i$ is the optimal solution of the problem:

$$\min ||x^* - x||_A \quad (2.1)$$
$$\text{s.t.} \quad x - x_0 \in \text{Span}\,\{Ae_0, A^2e_0, \ldots, A^ie_0\} \quad (2.2)$$

where $e = x^* - x_0$, and $||y||_A = (y'Ay)^{1/2}$ for all $y \in R^n$, i.e., $||.||_A$ is the norm induced by $A$.

As noted in [12], one can prove that if $0 \leq \epsilon < 1$, the iterates of the cgm satisfy

$$||x^* - x_i||_A \leq \epsilon ||x^* - x_0||_A \quad (2.3)$$

within

$$O(\rho(\sqrt{A}) \log_2 \frac{2}{\epsilon}) \quad (2.4)$$

where $\rho(\sqrt{A}) = \lambda_{\max}(\sqrt{A})/\lambda_{\min}(\sqrt{A})$. This on the other hand implies that if $i = \Omega(\rho(\sqrt{A} \log_2 \frac{2}{\epsilon})$, then $||x^* - x_i||_A \leq \epsilon||x^* - x_0||_A$. In other words, in order to obtain a (relative) $\epsilon$-approximation of the solution $x^*$ in terms of the norm induced by $A$, the cgm terminates in time which is proportional to the condition number $\rho(\sqrt{A})$.

In the case that the matrix $A$ is not positive definite and $A'A$ is not singular, one can consider solving $A'Ax = A'b$ and the above statements still hold with $A$ replaced by $A'A$.

In Section 3 we shall use the property (2.3)–(2.4) of the cgm for solving a system of linear equation, to demonstrate that its running time is inversely proportional to certain measure of relative stability, when used for solving the Lyapunov equation.

Our discussion of the ipm that follows is also very brief. We shall use few terms which might not be familiar to readers who do not follow the ipm literature. Since a complete treatment of the subject will take us too far from the main idea of the paper, we refer the reader to the references [10] and [13] for much more information on the ipms.

Given an $n \times n$ symmetric matrix $C$ and a real number $\nu$, the ipm that is considered in this paper can be applied to problems of the form: find $X$ such that,

$$\text{Trace}\; CX = \nu \quad (2.5)$$
$$X \in D_f \cap L + W \quad (2.6)$$

That is, given the matrix $C$ and the real number $\nu$, the method produces as an output the matrix $X$

which belongs to $D_f \cap L + W$, where,

1. $D_f$ is an open, convex subset of the symmetric matrices, which is the domain of a "non-degenerate self–concordant barrier" with parameter $K$. The set of such functions is denoted by $F(K)$. What makes the functions in $F(K)$ special in the ipm theory is that the complexity of the ipm for solving problems of the form (2.5)–(2.6) can be shown to depend interestingly on $K$. $K$ on the other hand, is only depended on the dimension of the space where the problem is formulated in.

2. $L$ is a closed subspace of the space of symmetric matrices.

3. $W \in D_f$.

Let $\nu_{\inf}$ and $\nu_{\sup}$ denote the infimum and the supremum of the Trace $\{CX\}$ subject to the constraint that $X \in D_f \cap L + W$. Then it can be shown that if $\nu_{\inf} < \nu < \nu_{\sup}$, and we use the barrier method to solve (2.5)–(2.6), it produces the iterates $\{X_i\}_{i \geq 1}$ such that after

$$i = O(K \log_2(2 + K + \frac{\nu_{\sup} - \nu_{\inf}}{\min\{\nu_{\sup} - \nu, \nu - \nu_{\inf}\}})) \quad (2.7)$$

iterations, Trace$\{CX_i\} = \nu$. In other words, the solution of the problem (2.5)–(2.6) is at hand in time which is bounded by (2.7).

Each iteration of the barrier method considered above involves solving a linear equation, which for the purpose of our discussion, is considered to depend only on the dimension of the system $n$ (we use for example Gaussian elimination for solving these equations). Thus conclude in order to solve problems of the form (2.5)–(2.6), for any particular value of $\nu$, the computational efficiency of the barrier method discussed in [13] is of the order of the magnitude of $K$ (which depends on the dimension of the system), and some particular combination of $\nu_{\sup}$, $\nu$, and $\nu_{\inf}$. As it becomes evident in Section 4, using a careful reformulation of the original systems and control problem, $\nu_{\sup}$ and $\nu_{\inf}$ can in fact be used to *reveal* information about the relative stability of the corresponding systems problem.

## 3. Stability of Linear Systems

Our first example comes from a most basic stability question in systems and control theory; studying the stability property of a linear time invariant system. Given a matrix $A \in R^{n \times n}$, determine whether the origin is the globally asymptotically stable equilibrium point of $\Sigma_1$:

$$\Sigma_1 : \quad \dot{x} = A\,x \quad (3.1)$$

In particular, one is interested to know whether the trajectories of $\Sigma_1$ goes back to the origin if it is *disturbed* by any non-zero initial condition. This property on the other hand is equivalent to the matrix $A$ being Hurwitz, i.e., all eigenvalues of $A$ should have negative real parts [4].

As it is well known, Lyapunov [7] proved in 1892 that the origin is the global asymptotically stable equilibrium point of the (3.1) if and only if, given a matrix $Q > 0$, there exists a matrix $P > 0$ such that,

$$A'P + PA = -Q \quad (3.2)$$

Suppose that the matrix $A$ is in fact Hurwitz and consider solving (3.2) using Kronecker products, as first proposed by Bellman [2]. Our goal is to demonstrate that the *running time of the conjugate gradient method for solving the Lyapunov equation conveys an estimate of the relative stability of A.*

Let us rewrite (3.2) as

$$B\,(\mathbf{vec}\ P) = -\mathbf{vec}\ Q \quad (3.3)$$

where

$$B := I \otimes A' + A' \otimes I \in R^{n^2 \times n^2}$$

Each eigenvalue of $B$ is the sum of a pair of the eigenvalues of $A$. Consequently $|\lambda|_{\max}(B) = 2|\mathrm{Re}\{\lambda_{\min}(A)\}|$ and $|\lambda|_{\min}(B) = 2|\mathrm{Re}\{\lambda_{\max}(A)\}|$. We also note that $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ are real numbers. In particular, $|\lambda|_{\max}(B) = |\lambda|_{\max}(A + A')$ and $|\lambda|_{\min}(B) = |\lambda|_{\min}(A + A')$.

In order to use the cgm we consider solving the system of linear equation

$$B'B(\mathbf{vec}\ P) = -B'(\mathbf{vec}\ Q)$$

As noted in [12], since the origin is not in the spectrum of $B'B$ ($A$ is assumed to be Hurwitz), if

$$i = \Omega(\varrho \log \frac{2}{\epsilon})$$

then,

$$||P^* - P_i||_{B'B} \leq \epsilon ||P^* - P_0||_{B'B}$$

where,

$$\begin{aligned} \varrho &:= \lambda_{\max}(\sqrt{B'B})/\lambda_{\min}(\sqrt{B'B}) \\ &= |\lambda|_{\max}(A + A')/|\lambda|_{\min}(A + A') \quad (3.4) \end{aligned}$$

More specifically, in order to obtain an $\epsilon$-approximation of the *certificate* of stability $P^*$ in the $B'B$ norm, the computational efficiency of the cgm is in the order of the magnitude of $\varrho$. Hence,

$$\tau_{\text{cgm}(\Sigma_1)}\, 1/\varrho = c$$

where *$c$ is a constant.*

A closer look at the quantity $1/\varrho$ reveals that it can be viewed as *a measure* of the relative stability of the system $\Sigma_1$, that is if

$$\delta_{\Sigma_1} := 1/\varrho$$

then,

$$\delta^{\Sigma_1} = |\lambda|_{\min}(A+A')/|\lambda|_{\max}(A+A') = 2|\bar{\Delta}|/||A+A'||_2$$

where

$$\bar{\Delta} := \{\inf \Delta \in R : A + \Delta I \text{ is not Hurwitz}\}. \quad (3.5)$$

Consequently we have proved the following theorem.

**Theorem 3.1** *Given the system $\Sigma_1$, there is an algorithm $\mathcal{A}$ and a measure of relative stability $\delta_{\Sigma_1}$, such that*

$$\delta_{\Sigma_1} \tau_{\mathcal{A}(\Sigma_1)} = c$$

*for some constant $c$.*

## 4. Absolute Stability Problem

Our second example comes from another fundamental stability problem, this time from nonlinear systems theory; studying the absolute stability problem of the Lur'e type systems [8], [16]. Given the matrices $A \in R^{n\times n}$, $B \in R^{n\times m}$, $C \in R^{p\times n}$, and $D \in R^{p\times m}$, determine whether the origin is the globally asymptotically stable equilibrium of $\Sigma_2$:

$$\Sigma_2 : \qquad \dot{x} = A\,x + B\,u \qquad (4.1)$$
$$y = C\,x + D\,u \qquad (4.2)$$

where it is know that:

1. $u = -\Phi(y)$, for some $\Phi$ belonging to the sector $[0, 1/k]$, for some real number $k$, i.e., $y'\Phi(y) \leq k||\Phi(y)||^2$, for all $y$ [16].
2. The pairs $(A, B)$ and $(A, C)$ are controllable and observable, respectively.
3. The matrix $A$ is Hurwitz. As it is known, via the Positive Real (PR) Lemma [1], [5], [8], [17], one is able to prove that the system $\Sigma_2$ is absolutely stable if,

$$\frac{1}{2}(H(jw) + H^*(jw)) + kI > 0 \quad \forall w \in [0, \infty] \qquad (4.3)$$

or equivalently if,

$$\text{herm}\{\begin{pmatrix} -P & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D + kI \end{pmatrix}\} > 0 \, (4.4)$$
$$P > 0 \, (4.5)$$

is a feasible set of linear matrix inequalities. Suppose that the system $\Sigma_2$ is in fact absolutely stable and that one wants to verify this by utilizing a numerical algorithm. Before we state that main result, let us define two measures of relative stability for $\Sigma_2$.

The first measure of relative stability is in fact a lower bound on the gain margin of $\Sigma_2$, defined by the following semi-definite program:

$$\delta^1_{\Sigma_2} := \sup \lambda \qquad (4.6)$$

such that,

$$\text{herm}\{\begin{pmatrix} -P & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} A & B \\ C & \lambda^{-1}kI + D \end{pmatrix}\} > 0$$
$$P > 0$$

That is, $\delta^1_{\Sigma_2}$ is the maximum factor by which the constant $1/k$ (recall that the nonlinearity belongs to the sector $[0, 1/k]$) can be increased such that $\Sigma_2$ is guaranteed to remain stable.

Define $E := \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Our next measure of the relative stability of $\Sigma_2$ is defined as follows:

$$\delta^2_{\Sigma_2} := \inf_{\Delta} ||\Delta||/||E|| \qquad (4.7)$$

such that there *does not* exist a matrix $P$ that satisfies the following set of linear matrix inequalities,

$$\text{herm}\{\begin{pmatrix} -P & 0 \\ 0 & I \end{pmatrix}(E + \Delta)\} > 0 \qquad (4.8)$$
$$P > 0 \qquad (4.9)$$

In other words, $\delta^2_{\Sigma_2}$ is the minimum relative perturbation of the the quadruple $(A, B, C, D)$, such that for a given constant $k$, the dynamical system $\Sigma_2$ can not be shown to be absolutely stable via the PR Lemma.

We now present the main result of this section, and then provide an sketch of its proof. The complete proof can be found in the journal version of this paper.

**Theorem 4.1** *Given the system $\Sigma_2$, there is an algorithm $\mathcal{A}$ such that for the relative stability measures $\delta^1_{\Sigma_2}$ and $\delta^2_{\Sigma_2}$,*

$$\delta^1_{\Sigma_2} \tau_{\mathcal{A}(\Sigma_2)} = c_1$$

*and*

$$\delta^2_{\Sigma_2}\tau_{\mathcal{A}(\Sigma_2)} = c_2$$

*for some constants $c_1$ and $c_2$.*

*Proof:* The main idea of the proof is to apply the barrier method, whose properties were discussed in Section 2, to a careful reformulation of the feasibility problem (4.4)–(4.5). In this case, one constructs the linear map $\Upsilon$ on the space of symmetric matrices, such that (4.4)–(4.5) is feasible if and only if,

$$\Upsilon(X) < 0 \qquad (4.10)$$

$$X > 0 \qquad (4.11)$$

is feasible. Then the following semi-definite program is considered

$$\inf\ t \qquad (4.12)$$

$$\text{s.t. } \Upsilon(X) < t(\Upsilon(\tilde{X}) + I) \qquad (4.13)$$

$$X > 0 \qquad (4.14)$$

$$||X|| < 1 \qquad (4.15)$$

$$-1 < t < 2 \qquad (4.16)$$

where $\tilde{X} > 0$ is chosen a-priori. It is then shown that if $t = 0$, then (4.10)–(4.11) is feasible, which implies that (4.4)–(4.5) is feasible. In view of the property (2.7) it can be demonstrated that the computational efficiency of the barrier method is proportional to $k$ and the value of $t_{\text{inf}}$ in the semi-definite program (4.12)–(4.16). This value on the other hand can easily be shown to be inversely proportional to $1/\delta^1_{\Sigma_2}$ and $1/\delta^2_{\Sigma_2}$, and hence,

$$\tau_{\mathcal{A}(\Sigma_2)}\delta^i_{\Sigma_2} = O(K) = c_i; \quad i = 1, 2 \qquad (4.17)$$

□

# 5. Conclusion

The main thesis of this paper is that there is a very close relationship between stability analysis of dynamical systems on one hand, and the theoretical studies on the efficiency of certain numerical algorithms. In particular, we have demonstrated that for two basic, but very important stability problems, the efficiency of the conjugate gradient and the interior point methods, can convey certain information about the relative stability of the corresponding systems. This phenomena can in principle be used to give an *algorithmic definition* of the relative stability of a dynamical system. The objection to this approach would be that the stability properties of a system is in principle coordinate free, and thus, should not depend on a particular algorithm. Nevertheless, since at the present time, we are far from obtaining *optimal* algorithms for solving stability problems (e.g., linear matrix inequalities), a *machine independent theory of stability* is far from its realization. Moreover, in order to check the stability of a dynamical system, an algorithm *has* to be introduced (on some particular model of computation), and thereby, one can argue, that stability properties can be viewed with the running time of that algorithm as our frame of reference. The contribution of the paper is thus to demonstrate that the above approach can in fact be adopted for two very important problems in systems analysis. their *relative* stability.

# Acknowledgment

The research of the authors has been supported in part by the National Science Foundation grant CCR-9222734. The authors gratefully acknowledge fruitful discussions with Michael G. Safonov, Len Adleman, and James Renegar.

# References

[1] B. D. O. Anderson and J. B. Moore. Algebraic structure of generalized positive real matrices. *SIAM Journal of Control and Optimization*, 6(4), 1968.

[2] R. Bellman. Kronecker products and the second method of lyapunov. *Mathematische Nachrichten*, 17–19, 1959.

[3] V. Blondel and J. N. Tsitsiklis. Np-hardness of some linear control design problems. In *Proceedings of the 34th Conference on Decision and Control*, pages 2910–2115, New Orleans, LA, December 1995. IEEE Press.

[4] C-T. Chen. *Linear System Theory and Design*. Holt, Rinehart and Winston, Inc., New York, 1984.

[5] R. E. Kalman. On a new characterization of linear passive systems. In *Proceedings of the first Allerton Conference on circuit and system theory*, pages 456–470, 1963.

[6] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[7] M. Lyapunov. General stability problem of motion. *Annals of Mathematics*, 1949. photo-reproduction of the French translation of 1907.

[8] K. S. Narendera and J. H. Taylor. *Frequency Domain Criteria for Absolute Stability.* Academic Press, New York, 1973.

[9] A. Nemirovskii. Several np-hard problems arising in robust stability analysis. *Mathematics of Control, Signals, and Systems*, 6:99–105, 1993.

[10] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming.* SIAM, Philadelphia, 1994.

[11] C. H. Papadimitriou and J. N. Tsitsiklis. Intractable problems in control theory. *SIAM Journal of Control and Optimization*, 24(4):639–654, 1986.

[12] J. Renegar. Condition numbers, the barrier method and the conjugate gradient method. Technical report, School of Operations Research and Industrial Engineering, Cornell University, Itacha, New York, 1994.

[13] J. Renegar. Linear programming, complexity theory and elementary functional analysis. Technical report, School of Operations Research and Industrial Engineering, Cornell University, Itacha, New York, 1994.

[14] E. Sontag. From linear to nonlinear: some complexity comparisons. In *Proceedings of the 34th Conference on Decision and Control*, pages 2916–2920, New Orleans, LA, December 1995. IEEE Press.

[15] J. E. Tierno and J. C. Doyle. Control problems and the polynomial time hierarchy. In *Proceedings of the 34th Conference on Decision and Control*, pages 2927–2931, New Orleans, LA, December 1995. IEEE Press.

[16] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, second edition, 1993.

[17] V. A. Yakubovich. Solution of certain matrix inequalities encountered in nonlinear control theory. *Soviet Mathematics*, 5:652–656, 1964.