# Note on the Complexity of Some Eigenvector Problems

Thomas Eiter and Georg Gottlob

Christian Doppler Labor für Expertensyteme
Technische Universität Wien
Paniglgasse 16, 1040 Wien

December 27, 1995

CD-TR 95/89

# Note on the Complexity of Some Eigenvector Problems

**Thomas Eiter and Georg Gottlob**
Information Systems Department, TU Vienna
Paniglgasse 16, A-1040 Wien, Austria
(eiter|gottlob)@dbai.tuwien.ac.at

December 27, 1995

**Abstract**

We consider the computation of eigenvectors $\mathbf{x} = (x_1, \ldots, n)$ over the integers, where each component $x_i$ satisfies $|x_i| \leq c$ for a constant $c$. We address decisional problems in this context, and describe some problems complete for NP, $\Theta_2^p$, $F\Delta_2^p$, $\Delta_2^p$, and $\Sigma_2^p$.

## 1 Introduction and preliminaries

Let $R$ be a ring with 1. Recall that $\lambda \in R$ is an eigenvalue of an $n \times n$ matrix $M = (m_{i,j})$ over $R$ if the equation

$$M \cdot \mathbf{x}^T = \lambda \mathbf{x}^T$$

has nontrivial solutions, i.e., solutions $\mathbf{x} \neq \mathbf{0}$, where $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{x}^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is the transpose of $\mathbf{x}$, and $\mathbf{0} = (0, \ldots, 0)$ is the zero vector; all vectors $\mathbf{x}$ that satisfy this equation are eigenvectors (for the eigenvalue $\lambda$). It is well-known that if $R$ is a field, then the eigenvectors for $\lambda$ form a vector space, whose dimension is the multiplicity of $\lambda$ as root of the characteristic polynomial of $M$.

In this short note, we address the computational complexity of some computational problems on eigenvectors. In particular, we consider bounded eigenvectors over the integers, i.e. $R = \mathcal{Z}$. Recall that the norm of an integer vector $\mathbf{x}$, $||\mathbf{x}||$, is defined by $||\mathbf{x}|| = \sqrt{\sum_{i=1}^{n} x_i^2}$. Vector $\mathbf{x} \in V$ is maximal in a set of vectors $V$ iff $||\mathbf{y}|| \leq ||\mathbf{x}||$, for every $\mathbf{y} \in V$. We say that vector $\mathbf{x}$ is bounded by a number $b \geq 0$ ($b$-bounded) if $|x_i| \leq b$, for every $i = 1, \ldots, n$.

We consider the following problems (vectors and matrices are tacitly assumed to be over the integers), where $c \geq 1$ is a fixed constant:

**P1:** Given an $n \times n$ matrix $M$, an eigenvalue $\lambda$ of $M$, and a real number $K$, does there exist a $c$-bounded eigenvector $\mathbf{x}$ such that $||\mathbf{x}|| \geq K$ ?

**P2:** Given an $n \times n$ matrix $M$, an eigenvalue $\lambda$ of $M$, and an integer $z$, decide if there is any maximal eigenvector $\mathbf{x}$ among the $c$-bounded eigenvectors with $x_1 = z$, i.e., the first component of $\mathbf{x}$ is $z$.

**P3:** Given an $n \times n$ matrix $M$ and an eigenvalue $\lambda$ of $M$, compute the lexicographically first among the maximal $c$-bounded eigenvectors.

**P4:** Given an $n \times n$ matrix $M$, an eigenvalue $\lambda$ of $M$, a subset $I$ of the components, and an integer $z$, does there exist a $c$-bounded $\preceq$-maximal eigenvector $\mathbf{x}$ so that $x_1 = z$, where $\mathbf{x} \preceq \mathbf{y}$ iff $\mathbf{x}$ and $\mathbf{y}$ coincide on the components in $I$ and $\|\mathbf{x}\| \preceq \|\mathbf{y}\|$.

Notice that correct problem instances can be recognized in polynomial time, since deciding whether $\lambda$ is an eigenvalue of $M$ can be done (e.g. using standard Gaussian elimination) in polynomial time.

## 2 Complexity results

For determining the complexity of problems P1–P4, we refer to variants of problems involving the classical satisfiability problem.

Let $C$ be a set of propositional clauses. A truth assignment $\tau$ to the variables of $C$ satisfies $C$, if each clause $C$ contains at least one literal with value *true*. An assignment $\tau$ is *not-all-equal satisfying* (*nae-satisfying*) for $C$, if each clause in $C$ contains two literals that have different value according to $\tau$; clearly, each nae-satisfying assignment of $C$ satisfies $C$ in the standard sense. Moreover, if $\sigma$ is an nae-satisfying assignment, then also the complementary assignment $\overline{\sigma}$, in which each variable has opposite truth value, is nae-satisfying.

Let $C = \{C_1, \ldots, C_m\}$ be an instance of 3SAT, i.e., a set of propositional clauses $C_i = \alpha_{i,1} \vee \alpha_{i,2} \vee \alpha_{i,3}$, $i = 1, \ldots, m$ on variables $X = \{x_1, \ldots, x_n\}$. Then denote by $C'$ the set of the following clauses:

- $x_j \vee x_j^* \vee z_j$ and $x_j \vee x_j^* \vee \neg z_j$, for each $j = 1 \ldots, n$,
- $\alpha_{i,1}^* \vee \alpha_{i,2}^* \vee w_i$ and $\alpha_{3,i}^* \vee x_0 \vee \neg w_i$, for each $i = 1, \ldots, m$

where $x_0$, all $z_j$, all $x_j^*$, and all $w_i$ are fresh variables and $\alpha_{i,j}^* = x_\ell$, if $\alpha_{i,j} = x_\ell$,and $\alpha_{i,j}^* = x_\ell^*$ if $\alpha_{i,j} = \neg x_\ell$.

The following is easily verified. Let $\tau$ be an nae-satisfying assignment for $C'$. If $\tau(x_0) =$ *false*, then $\tau$, restricted to $X$, satisfies $C$; if $\tau(x_0) =$ *true*, then the complementary assignment $\overline{\tau}$, restricted to $X$, satisfies $C$. On the other hand, if an assignment $\sigma$ satisfies $C$, then $\sigma$ is extendible to at least one nae-satisfying assignment of $C'$ in which $x_0 =$ *false*. Thus, we obtain the following.

**Lemma 2.1** *Let $C$ be any instance of 3SAT. The nae-satisfying assignments $\tau$ of $C'$ such that $\tau(x_0) =$ false, correspond on the variables $X$ 1-1 to the satisfying assignments of $C$.*

As a consequence, deciding whether a SAT instance is satisfiable under nae-satisfaction (NAESAT) is NP-hard [1], even if all clauses have size 3 (NAE3SAT). More general, this allows to conclude that checking the validity of certain Quantified Boolean formulas (QBFs), based on the notion of nae-satisfaction is as hard as for the standard notion of satisfaction. An NAESAT instance $C$ on variables $X$ can be seen as a QBF $\Phi = \exists X C$, where $C$ is viewed as a conjunction of its clauses and the quantifier $\exists$ ranges over all truth assignments to $X$. $\Phi$ is valid under nae-satisfaction (briefly, nae-valid) if $C$ is a Yes instance. Accordingly, a QBF $\forall Y \exists X\, C$ where $C$ is in CNF is nae-valid if for every assignment $\sigma$ to $Y$, there is an assignment $\tau$ to $X$ such that the combined assignment $\sigma \cup \tau$ nae-satisfies $C$.

**Lemma 2.2** *Given a QBF $\Phi = \forall Y \exists X\, C$, where $C$ is in CNF, deciding whether it is nae-valid is $\Pi_2^p$-complete, and hard even if $C$ is in 3CNF (all clauses have size 3).*

**Proof.** Membership in $\Pi_2^p$ is easy: A guess for $\tau$ such that no extension of $\tau$ by $\sigma$ does nae-satisfy $C$ can be checked with a call to an NP oracle (the check is in co-NP).

Hardness follows from a straightforward reduction of checking the validity of a QBF of the given form in the standard sense, which utilizes the reduction from 3SAT to NAE3SAT given by the transformation from above. I.e., construct for $C$ the formula $C'$, and consider the QBF $\Psi = \forall Y \exists X \cup X^* \cup \{z_j, x_0, w_i\}\, C'$. This formula is nae-valid iff $\Phi$ is valid in the standard sense. ■

**Theorem 2.3** *Problem P1 is NP-complete, for every fixed $c \geq 1$.*

**Proof.** Membership in NP is clear, since a guess for a suitable eigenvector $\mathbf{x}$ can be verified in polynomial time.

We show hardness by a reduction from NAE3SAT [1], which is, given a 3SAT instance $C = \{C_1, \ldots, C_M\}$ on variables $X = \{x_1, \ldots, x_n\}$, decide whether $C$ is nae-satisfiable.

We will describe the matrix $M$ in terms of the equations emerging for each component $x_i$ from the equation $M \cdot \mathbf{x}^T = \lambda \mathbf{x}^T$; the eigenvalue $\lambda$ is 1.[1] That is, we state for each $x_i$ the equation

$$\sum_{j=1}^{n} m_{i,j} \cdot x_j = x_i$$

Unless stated otherwise, the equation is $1 \cdot x_i = x_i$ (i.e., the $i$-th row of $M$ has 1 at column $i$ and 0 everywhere else).

We construct a $k \times k$ matrix, $k = 2n + m$, as follows. The components $x_1, \ldots, x_n$ of an eigenvector $\mathbf{x} = (x_1, \ldots, x_k)$ are supposed to correspond to the variables $x_1, \ldots, x_n$ of $C$, and the component values to (partial) truth assignments. For each $i = 1, \ldots, n$, we have the equation

$$c \cdot x_{n+i} = x_i; \tag{1}$$

[1] Here and in the other proofs, we take $\lambda = 1$; it is easy to see that 1 is indeed always an eigenvalue of $M$.

any $c$-bounded solution requires that $(x_i, x_{n+i})$ is one of $(0,0)$, $(c,1)$, $(-c,-1)$. Intuitively, this corresponds to a partial truth assignment, where $(c,1)$ means that variable $x_i$ of $C$ is set true, $(-c,-1)$ that $x_i$ is set false, and $(0,0)$ that $x_i$ is undefined.

For each clause $C_i = x_{i_1}^{s_{i,1}} \vee x_{i_2}^{s_{i,2}} \vee x_{i_3}^{s_{i,3}}$ in $C$, where $s_{i,j} \in \{0,1\}$ (as usual $x^0$ is the literal $x$ and $x^1$ the literal $\neg x$), we have the equation

$$(-1)^{s_{i,1}} \cdot x_{i_1} + (-1)^{s_{i,2}} \cdot x_{i_2} + (-1)^{s_{i,3}} \cdot x_{i_3} \quad = \quad x_{2n+i} \tag{2}$$

The value of $x_{2n+i}$ in any $c$-bounded eigenvector $\mathbf{x}$ is either 0 or $\pm c$. If all literals in $C_i$ have assigned a truth value, i.e., $x_{i_j} = \pm c$, then $x_{2n+i}$ must have value $\pm c$; this is only possible if two terms $(-1)^{s_{i,j}} x_{i_j}$ add up to 0, i.e., the literals corresponding literals $L_j$ have opposite value under the truth assignment represented by $\mathbf{x}$.

It should be clear from above how the matrix $M$ is completed. Notice that $\lambda = 1$ is an eigenvalue of $M$ (e.g., set $x_1 = c$ and $x_j = 0$, $j = 2, \ldots, n$, and choose the other components appropriately). Define $K = \sqrt{n(c+1) + m \cdot c}$.

It is easy to see that the instance $M$, $\lambda$, $K$ of P1 is a Yes-instance iff $C$ is nae-satisfiable. This proves the result. ■

*Remark.* Problem P1 remains in NP also if $c$ is not fixed, but a number in the input.

**Theorem 2.4** *Problem P2 is $\Theta_2^p$-complete, for every fixed $c \geq 1$. Hardness holds even for $z = 0$.*

**Proof.** The problem can be solved with a logarithmic number of queries to an NP oracle, by computing first the norm $\delta$ of a maximal $c$-bounded eigenvector in binary search (cf. problem P1; note that $\delta \leq c^2 \cdot n$, and the input size is larger than $n$) and then querying the oracle whether there exists a $c$-bounded eigenvector $\mathbf{x}$ such that $\|\mathbf{x}\| = \delta$ and $x_1 = z$.

Hardness is shown by a reduction from the following problem. Consider the variant of NAE3SAT in which it is known that $C$ is nae-satisfied by a certain assignment $\tau$ in which $\tau(x_1) = \tau(x_2)$; deciding whether there is an nae-satisfying assignment for $C$ in which $\tau(x_1) \neq \tau(x_2)$ is NP-complete. Therefore, by the results in [3], given $k$ instances $C^1, \ldots, C^k$ of this problem (on different variables), deciding whether the number of Yes-instances is even is $\Theta_2^p$-complete.

We reduce this problem to P2 as follows. For each instance $C^i$ on variables $X^i = \{x_1^i, \ldots, x_{n_i}^i\}$, we use the equations of the transformation of NAE3SAT to P1 in the proof of Theorem 2.3;[2] note that an eigenvector of respective norm $K^i$ always exists. Furthermore, we add for $C^i$ equations

$$\begin{aligned} 1 \cdot x_s^i + 1 \cdot x_t^i &= x_t^i \\ 1 \cdot x_{n+1}^i + (-1) \cdot x_{n+2}^i + 2 \cdot x_r^i &= x_s^i \end{aligned}$$

where $x_r^i$, $x_s^i$, and $x_t^i$ are new variables. Clearly, $x_s^i = 0$ in any eigenvector, and $x_t^i = \pm c$ in any maximal eigenvector. Since $x_{n+1}^i$ and $x_{n+2}^i$ have in any eigenvector values from

[2] Here and in the rest of the paper, for better readability we use component names $x_j^i$, $y_i$, etc which can be easily transformed to names $x_1, \ldots, x_n$ as stated in problems P1–P4.

$-1, 0, 1$, $x_r^i$ has either value 0, 1, or $-1$. Moreover, for each variable $x_{2n_i+j}^i$ introduced for the clause $C_j^i$ on the right hand side of equation 2, take a new variable $x'^i_{2n_i+j}$ that is set to the same value, i.e., add the equation

$$1 \cdot x_{2n_i+j}^i \quad = \quad x'^i_{2n_i+j}$$

If $C^i$ is a Yes-instance, then all maximal $c$-bounded eigenvectors $\mathbf{x}$ have $x_r^i = \pm 1$ and norm $\sqrt{n(c+1) + 2 \cdot c \cdot m + c + 1}$; if $C^i$ is not a Yes-instance, then they have $x_r^i = 0$ and norm $\sqrt{n(c+1) + 2 \cdot c \cdot m + c}$.

To combine the instances $C^i$, introduce for each variable $x$ occurring so far fresh variables $p_1, \ldots, p_\ell$, where $\ell = \lceil c \rceil$, that are equated to $x$, i.e., satisfy

$$1 \cdot x \; = \; p_j, \qquad j = 1, \ldots, \ell \tag{3}$$

and set up the equations

$$1 \cdot x_s + 1 \cdot x_t \quad = \quad x_t \tag{4}$$

$$c \cdot x_r^1 + \cdots + c \cdot x_r^k + 1 \cdot x_r \quad = \quad x_s \tag{5}$$

where $x_r$, $x_s$, and $x_t$ are new variables. The value of $x_s$ in any eigenvector is 0, and $x_t = \pm c$ in every maximal $c$-bounded eigenvector; since the value of $x_r^i$ is from $0, -1, 1$, the value of $x_r$ is from $-c, 0, c$.

If the number of Yes-instances among $C^1, \ldots, C^k$ is odd, then $x_r = \pm c$, and if it is even, then $x_r = 0$. (Note that if $\mathbf{x}$ is an eigenvector, so is $(-1)\mathbf{x}$; hence, the $x_r^i$ with value $\neq 0$ can be chosen so that they pairwise cancel.) Indeed, in each maximal eigenvector $x_r^i = 0$ if $C^i$ is not a Yes-instance, and $x_r^i = \pm 1$ if $C^i$ is a Yes-instance. The variable copies $p_i$ assure that for every $C^i$, taking the right $x_r^i$ has a larger global influence on the norm of a maximal eigenvector than $x_r$, which will be set according to whether the number of $x_r^i$'s $\neq 0$ is odd or not.

Thus, deciding whether some maximal $c$-bounded eigenvector exists for $M$, $\lambda = 1$ such that $x_r = 0$, is $\Theta_2^p$-hard. By reordering the components such that $x_r$ is the first, this proves the result. ■

Recall that a truth assignment $\tau$ to the variables $x_1, \ldots, x_n$ is lexicographically smaller than a truth assignment $\sigma$, if $\tau(x_i) = \mathit{false}$ for the least index $i$ such that $\tau(x_i) \neq \sigma(x_i)$.

**Lemma 2.5** *Given an NAE3SAT instance $C$, computing the lexicographically first nae-satisfying assignment of $C$, $\tau^*$, is $F\Delta_2^p$-complete, as is computing $\tau^*(x_i)$ for any given variable $x_i$. This holds even if $C$ is known to be nae-satisfiable.*

**Proof.** Reduce the analogous problem for 3SAT, and using the transformation of a 3SAT instance $C$ to a NAE3SAT instance $C'$ described above Lemma 2.1. Then, order the variables arbitrarily but so that the order starts with $x_0, x_1, \ldots, x_n$. By Lemma 2.1, the lexicographic first nae-satisfying assignment of $C'$ corresponds to the lexicographic first assignment of $C$. Computing the latter, as well as a given bit thereof, is well-known $F\Delta_2^p$-complete [2]. This proves hardness. Membership is clear, as for each component only constantly many values have to be considered, and deciding if a partial vector can be completed to a $c$-bounded eigenvector is obviously in NP. ■

**Theorem 2.6** *Problem P3 is $F\Delta_2^p$-complete, for every fixed $c \geq 1$.*

**Proof.** Membership is straightforward: the norm of a maximal $c$-bounded eigenvector, $\delta$, can be computed in polynomial time with a logarithmic number of calls to an NP-oracle (query instances of P1); computing the lexicographically first eigenvector $\mathbf{x}^*$ of norm $\delta$ can be done with a linear number of NP oracle calls, computing $x_1^*$, $x_2^*$ and so on in order; for each component $x_i$ only a constant number of values are possible, and deciding if a partial vector $x_1, \ldots, x_k$ can be completed to a $c$-bounded eigenvector with norm $\delta$ is in NP.

To show hardness, use the same reduction as in the proof of Theorem 2.3. Under the assumption that $C$ is nae-satisfiable, $\delta$ takes the value $K = n(c+1) + m \cdot n$, which is the largest possible value. Therefore, maximal eigenvectors correspond to nae-satisfying assignments of $C$. In particular, the lexicographically first maximal $c$-bounded eigenvector (in which $x_0 = -c$) corresponds to the lexicographically first nae-satisfying assignment of $C$ (in which $x_0 = \mathit{false}$). ■

**Corollary 2.7** *Let $c \geq 1$ be fixed. Given an $n \times n$ matrix $M$, an eigenvalue $\lambda$ of $M$, and integers $i, z$, deciding if $x_i = z$ for the lexicographically first maximal eigenvector $\mathbf{x}$ among those that are c-bounded is $\Delta_2^p$-complete.*

**Proof.** Deciding any given bit of the lexicographically first satisfying truth assignment is $\Delta_2^p$-complete [2]. Thus, the result follows from the proof of Theorem 2.6. ■

**Theorem 2.8** *Problem P4 is $\Sigma_2^p$-complete, for every fixed $c \geq 1$. Hardness holds for $z = 0$.*

**Proof.** The problem is in $\Sigma_2^p$, since a guess for a $\preceq$-maximal $c$-bounded eigenvector $\mathbf{x}$ can be verified with a single call to an NP oracle (deciding whether a vector $\mathbf{y} \preceq \mathbf{x}$ with $\mathbf{x} \neq \mathbf{y}$ exists is in NP).

Hardness is shown by using a slight variant of the problem in Lemma 2.2. It is assumed that for each assignment $\sigma$ to $Y$, a particular assignment $\tau$ to all variables $X$ exists in which $\tau(x_1) = \tau(x_2)$ such that $\sigma \cup \tau$ nae-satisfies $C$; it is asked whether a $\tau$ exists such that $\sigma \cup \tau$ nae-satisfies $C$ and $\tau(x_1) \neq \tau(x_2)$. This problem is easily shown to be $\Pi_2^p$-complete.

We reduce this problem to the complement of P4. Let the input be the formula $\Phi = \forall Y \exists X\, C$. For the formula $C$, seen as an NAE3SAT instance on variables $Y \cup X$, set up the same equations as in the reduction of NAE3SAT to P1 described in the proof of Theorem 2.3. Furthermore, add the equations

$$1 \cdot x_s + 1 \cdot x_t = x_t \tag{6}$$

$$1 \cdot x_1 + (-1) \cdot x_2 + 2c \cdot x_r \;=\; x_s, \tag{7}$$

where $x_r$, $x_s$, and $x_t$ are new variables; take for each $x_{2n+i}$ that was introduced for clause $C_i$ (cf. eq. (2)) a fresh variable $x'_{2n+i}$ and equate it to $x_{2n+i}$, i.e., set up

$$1 \cdot x_{2n+i} = x'_{2n+i}. \tag{8}$$

Recall that $x_1$ and $x_2$ have in any eigenvector values from $-1, 0, 1$; thus, $x_r$ has value 0 if $x_1 = x_2$ and value $\pm 1$ if $x_1 \neq x_2$. Consider for the moment eigenvectors in which each $y_i$ from $Y$ has value $\pm c$; this represents a truth assignment $\sigma$ to the variables in $Y$, where $\sigma(y_i) = \mathit{true}$ iff $y_i = c$, and $\sigma(y_i) = \mathit{false}$ iff $y_i = -c$. The assumption on $\Phi$ entails the following for every $\preceq$-maximal eigenvector $\mathbf{x}$ corresponding to $\sigma$. If $\sigma$ can not be extended by some $\tau$ to an nae-satisfying assignment for $C$ in which $\tau(x_1) \neq \tau(x_2)$, then $\mathbf{x}$ has norm $\sqrt{n(c+1) + 2 \cdot c \cdot m + c}$ and $x_r = 0$ ($n = |Y \cup X|$); if $\sigma$ is extendible by such a $\tau$, then $\mathbf{x}$ has norm $\sqrt{n(c+1) + 2 \cdot m \cdot c + 1}$ and $x_r = \pm 1$.

Therefore, if we only consider eigenvectors in which each $y_i$ is assigned a truth value and choose for $I$ the set $Y$, a $\preceq$-maximal such vector exists in which $x_r = 0$, precisely if for the assignment to the $y_i$ variables represented by this vector, every assignment $\tau$ to the $x_i$ variables in which $\tau(x_1) \neq \tau(x_2)$ does not nae-satisfy $C$. However, we have also to take into account $\preceq$-maximal eigenvectors in which some $y_i$ might have value 0.

To handle this, we introduce for each variable $y_i$ three new variables $p_i$, $p'_i$ and $p''_i$, and set up the equations

$$\begin{aligned} c \cdot p'_i &= p_i \\ p_i + (-1) \cdot y_i &= p''_i; \end{aligned}$$

for $I$, we choose the set of components $p_i$. The value of $p_i$ is in each eigenvector from $0, -c, c$; consider a $\preceq$-maximal such vector. If $p_i = 0$, then, as easy to see, $y_i$ has value $\pm c$ and $p''_i$ has value $\mp c$. Otherwise, if $p_i = \pm c$, then $y_i = p_i$ must hold. Therefore, each $\preceq$-maximal eigenvector represents a total truth assignment $\sigma$ to the variables $Y$; moreover, it has $x_r = 0$ iff an assignment $\tau$ to the variables $X$ exists, where $\tau(x_1) \neq \tau(x_2)$, such that $\sigma \cup \tau$ nae-satisfies $C$ (in fact, such a $\tau$ is encoded by the vector).

Now reorder the components such that $x_r$ becomes the first. Then, the instance of P4 where $M$ is as obvious, $\lambda = 1$, $I = \{p_i\}$ and $z = 0$, is not a Yes instance iff the formula $\Phi$ is valid under nae-satisfaction. This proves the result. ■

## References


[1] M. Garey and D. S. Johnson. *Computers and Intractability – A Guide to the Theory of NP-Completeness*. W. H. Freeman, New York, 1979.

[2] M. Krentel. The Complexity of Optimization Problems. *Journal of Computer and System Sciences*, 36:490–509, 1988.

[3] K. Wagner. Bounded Query Classes. *SIAM J. Comp.*, 19(5):833–846, 1990.