BALANCING REGULAR MATRIX PENCILS*

DAMIEN LEMONNIER† AND PAUL VAN DOOREN‡

Abstract. In this paper we present a new diagonal balancing technique for regular matrix pencils $\lambda B - A$, which aims at reducing the sensitivity of the corresponding generalized eigenvalues. It is inspired from the balancing technique of a square matrix A and has a comparable complexity. Upon convergence, the diagonally scaled pencil has row and column norms that are balanced in a precise sense. We also show that balancing a pencil boils down to making it closer to some standardized normal pencil. We give numerical examples illustrating that the sensitivity of generalized eigenvalues of a pencil may significantly improve after balancing.

Key words. generalized eigenvalues, normal pencils, balancing

AMS subject classifications. 15A18, 15A22, 65F15, 65F35

1. Introduction. A matrix A with a norm that is several orders of magnitude larger than the modulus of its eigenvalues typically has eigenvalues that are sensitive to perturbations in the entries of A. It is shown in [4] that the Frobenius norm of a matrix can then often be reduced via a diagonal scaling of the type $D^{-1}AD$. Such a scaling can be performed in exact arithmetic if the diagonal elements are constrained to be integer powers of the base of the finite precision arithmetic (typically 2 or 10). As a consequence the eigenvalues do not change, but their sensitivity can significantly be reduced. Such a diagonal scaling is therefore typically used before running any eigenvalue algorithm.

In this paper we introduce a similar scaling method for square pencils $\lambda B - A$ with a determinant $\det(\lambda B - A)$ that is not identically zero for all values of λ. For such pencils – which are called *regular* – one can define generalized eigenvalues via the zeros of the polynomial $\det(\lambda B - A)$. Our scaling method can be viewed as a natural extension of the balancing algorithm of [4] to regular matrix pencils and is aimed at reducing the sensitivity of the generalized eigenvalues of the pencil. This new method differs from that of Ward [7] whose aim it is to make the pencil entries have magnitudes as close to unity as possible, whereas our aim is to make the pencil as close as possible to some standardized normal pencil.

We first recall the classical balancing method for matrices and some of its properties. We then introduce the new balancing method for pencils and derive its analogous properties. We briefly discuss the complexity of the algorithm and finally give some numerical results illustrating the performance of the new scaling method.

2. Normal matrices and balancing. Normal matrices are known to have orthogonal eigenvectors and hence well conditioned eigenvalues [4]. Therefore if one has to compute eigenvalues of an arbitrary $n \times n$ matrix A, it is recommended to make it *closer* to a normal matrix by an error free transformation. Diagonal *scaling* transformations with positive diagonal elements that are integer powers of the base can be performed exactly since they only amount to integer changes in the exponents of the matrix entries. And in order to preserve the eigenvalues one performs diagonal similarities $D^{-1}AD$.

*This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

†Univ. Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium (lemonnier@csam.ucl.ac.be).

‡Univ. Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium (vdooren@csam.ucl.ac.be)

The basic question is thus how to characterize a diagonal scaling $D^{-1}AD$ that makes a matrix closer to a normal matrix. For this we consider two equivalent characterizations of normal matrices. A matrix A is normal iff
(1) A has *orthogonal* eigenvectors or, equivalently, its Schur form A_S :

$$A_S := U^*AU, \quad U^*U = I_n \tag{2.1}$$

is a diagonal matrix Λ_A
(2) the so-called *defect from normality*

$$\gamma(A) := \sum_{i=1}^{n} \sigma_i^2 - \sum_{i=1}^{n} |\lambda_i|^2 \tag{2.2}$$

is zero, where σ_i and λ_i are the singular values and the eigenvalues of A, respectively.

The defect from normality $\gamma(A)$ is always non-negative [2], which easily follows from (2.1) and the fact $\gamma(A) = \gamma(A_S)$, since unitary similarities do not change the eigenvalues nor the singular values of a matrix. Let $Orb(A)$ denote the *orbit* of A, i.e. the set of matrices similar to A. Then $\gamma(A)$ is the *minimum squared distance* between A and any normal matrix similar to A.

THEOREM 2.1. *The optimization problem*

$$\inf_T \|T^{-1}AT\|_F \tag{2.3}$$

has a normal matrix N_a in the closure of the orbit of A as solution. If A is diagonalizable then there exists a bounded T such that $N_a = T^{-1}AT$, otherwise T is unbounded.

Proof. Use the (complex) Schur decomposition $A_S = U^*AU$ and choose a unitary matrix Q such that the matrix $R := U^*TQ$ is triangular. Since unitary transformations do not change the Frobenius norm, the above minimization is then equivalent to

$$\inf_R \|R^{-1}A_SR\|_F,$$

which has the diagonal part Λ_A of A_S as solution. The transformation matrix R will be bounded for a diagonalizable matrix A, and it will be unbounded otherwise (see also [5] for more details). ∎

It then follows that $\sum_{i=1}^n \sigma_i^2$ is the Frobenius norm squared of A_S and therefore also the sum of the entries squared of A_S, while $\sum_{i=1}^n |\lambda_i|^2$ is just the sum of the diagonal entries squared of A_S. A diagonal scaling $D^{-1}AD$, on the other hand, does not change the λ_i's but does modify the σ_i's. So one can reduce the gap γ by scaling A in order to diminish its Frobenius norm. This is exactly what the balancing algorithm [4] does : it solves

$$\inf_D \ \|D^{-1}AD\|_F. \tag{2.4}$$

Let e_i denote the i-th unit vector, then it is shown in [4] that the optimal scaling is achieved when $D^{-1}AD$ satisfies

$$\|(D^{-1}AD)e_i\|_2^2 \ = \ \|e_i^T(D^{-1}AD)\|_2^2 \qquad \forall\ i = 1 \ldots n \tag{2.5}$$

and an algorithm is provided for computing an approximate solution D with elements that are powers of the base of the finite precision arithmetic. Each step of that algorithm decreases the Frobenius norm of the scaled matrix and hence also the distance to the normal matrices with the same spectrum as A.

The aim of this paper is to generalize these balancing ideas to regular matrix pencils $(\lambda B - A)$. In other words we will try to answer the following questions : (1) what is the property of regular pencils that is equivalent to normality in the standard eigenvalue problem, and (2) how to scale an arbitrary pencil so that it gets as close as possible to achieving this property?

3. Normal pencils. We first recall a definition of normal pencils, given in [1].

DEFINITION 3.1. *An $n \times n$ complex regular pencil $\lambda B - A$ is said to be normal if it has orthogonal right and left eigenvectors, i.e. if it has a decomposition of the form*

$$U_l^*(\lambda B - A)U_r = \lambda \Lambda_B - \Lambda_A,$$

where U_l, U_r are unitary, and Λ_A, Λ_B are diagonal.

In order to relate this to a defect we recall the definition of generalized singular values of two square matrices A and B.

DEFINITION 3.2. *The right (resp. left) singular values σ_{ri} (resp. σ_{li}) of $\lambda B - A$ are defined to be the generalized eigenvalues of $\lambda B^T B - A^T A$ (resp. $\lambda B B^T - A A^T$).*

Since the invertibility of B is not essential in these definitions, we first make the simplifying assumption that B is invertible. It then follows easily that

$$\sigma_{ri} = \sigma_i(AB^{-1}), \quad \sigma_{li} = \sigma_i(B^{-1}A).$$

When B is invertible, it is shown in [1] that the pencil $\lambda B - A$ is normal iff both AB^{-1} and $B^{-1}A$ are normal. A good candidate for the *defect from normality of a regular pencil* $\lambda B - A$ appears then to be

$$\Gamma(A, B) := \sum_{i=1}^{n} \sigma_{ri}^2 + \sum_{i=1}^{n} \sigma_{li}^2 - 2\sum_{i=1}^{n} |\lambda_i|^2$$

where λ_i are the generalized eigenvalues of the pencil. Clearly $\Gamma(A, B) = \gamma(AB^{-1}) + \gamma(B^{-1}A)$, which is always positive and is zero iff both AB^{-1} and $B^{-1}A$ are normal, and hence iff the pencil $\lambda B - A$ is normal.

If B is not invertible, we need another "defect from normality" function since $\Gamma(A, B)$ is then the difference between two infinite quantities. We can then consider a transformed pencil

$$\lambda \widehat{B} - \widehat{A} := \lambda(cB - sA) - (sB + cA), \quad c^2 + s^2 = 1. \tag{3.1}$$

It is well-known (see e.g. [1]) that for a regular pencil $\lambda B - A$ there always exists a choice (c, s) for which $\widehat{B}$ is invertible. Since the above transformation does *not* affect the left and right eigenvectors of a pencil, it follows that

$$\lambda \widehat{B} - \widehat{A} \text{ is normal } \iff \lambda B - A \text{ is normal.}$$

Rather than minimizing $\Gamma(A, B)$ one can thus minimize $\Gamma(\widehat{A}, \widehat{B})$ which will reach a minimum when both $\lambda \widehat{B} - \widehat{A}$ and $\lambda B - A$ are normal pencils. Notice however that the value of this defect then changes even though normality is preserved. Without loss of generality, we assume from now on that B is invertible.

But orthogonality of the left and right eigenvectors is not sufficient to guarantee a low sensitivity of the generalized eigenvalues of a regular pencil because eigenvalues can now be arbitrarily large or small, irrespective of the norm of A and B. Let x_i and y_i be respectively the right and left eigenvectors of a given eigenvalue λ_i :

$$Ax_i = \lambda_i Bx_i, \quad y_i^* A = \lambda_i y_i^* B,$$

and define the corresponding "Rayleigh" components :

$$\alpha_i := y_i^* Ax_i/(\|y_i\|_2\|x_i\|_2), \quad \beta_i := y_i^* Bx_i/(\|y_i\|_2\|x_i\|_2), \quad \lambda_i = \alpha_i/\beta_i.$$

In [6] it is shown that a perturbation in A and B of relative size ϵ :

$$\|\delta A\|_2 \leq \epsilon\|A\|_2, \quad \|\delta B\|_2 \leq \epsilon\|B\|_2, \tag{3.2}$$

may yield a perturbed eigenvalue $\tilde{\lambda}_i$, but that the chordal distance :

$$\chi(\lambda_i, \tilde{\lambda}_i) \ := \ \frac{|\alpha_i\tilde{\beta}_i - \tilde{\alpha}_i\beta_i|}{\sqrt{|\alpha_i|^2 + |\beta_i|^2}\sqrt{|\tilde{\alpha}_i|^2 + |\tilde{\beta}_i|^2}} \tag{3.3}$$

between the original and the perturbed eigenvalue is bounded by :

$$\chi(\lambda_i, \tilde{\lambda}_i) \ \leq \ \epsilon\frac{(\|A\|_2^2 + \|B\|_2^2)^{1/2}}{(|\alpha_i|^2 + |\beta_i|^2)^{1/2}} \ + \ O(\epsilon^2),$$

and that there exist perturbations δA and δB, for which this bound is met. The quantity

$$\kappa(\lambda_i) := \ \frac{(\|A\|_2^2 + \|B\|_2^2)^{1/2}}{(|\alpha_i|^2 + |\beta_i|^2)^{1/2}}, \tag{3.4}$$

is thus a valid relative condition number for λ_i in the sense that it measures how a perturbation of relative size ϵ in A and B affects λ_i in the (intrinsically relative) chordal metric. The reason why such a "relative" metric is to be preferred for pencils is linked to the fact that eigenvalues are now given by ratios of computed quantities (see [6] for more details).

When using the QZ-algorithm to compute the generalized eigenvalues of the pencil $\lambda B - A$ one obtains the so-called Schur form of this pencil :

$$A_S := Q^*AZ, \ B_S := Q^*BZ, \quad Q^*Q = I_n, \ Z^*Z = I_n, \tag{3.5}$$

where A_S and B_S are both upper triangular. This algorithm typically induces errors δA and δB in A and B that are of the order of (3.2), where ϵ is the machine accuracy of the computer. Since the orthogonal transformations Q and Z do not affect the quantities used in the definition (3.4) of $\kappa(\lambda_i)$, we can as well analyze the effect of perturbations in the coordinate system of the Schur form. The right and left eigenvectors x_i, y_i can then be normalized as follows :

$$x_i := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{i-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad y_i := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \eta_{i+1} \\ \vdots \\ \eta_n \end{bmatrix}.$$

If we denote the diagonal entries of the triangular matrices A_S, B_S by a_{ii}, b_{ii}, respectively, we then obtain the equalities :

$$\alpha_i = a_{ii}/n_i, \quad \beta_i = b_{ii}/n_i, \quad n_i := (\|y_i\|_2 \|x_i\|_2) \geq 1.$$

Since

$$\|A\|_2 = \|A_S\|_2 \geq \max_i |a_{ii}|, \quad \|B\|_2 = \|B_S\|_2 \geq \max_i |b_{ii}|$$

we finally obtain the inequality

$$\kappa(\lambda_i) \geq \frac{(\max_i |a_{ii}|^2 + \max_i |b_{ii}|^2)^{1/2}}{(|a_{ii}|^2 + |b_{ii}|^2)^{1/2}} \tag{3.6}$$

with *equality* holding only for a normal pencil since then x_i and y_i have norm 1 and A_S and B_S are diagonal. But normal pencils can still have a quantity $\kappa(\lambda_i)$ that can be very large if the pairs (a_{ii}, b_{ii}) vary a lot in norm. This is not the case for the following subclass of normal pencils.

DEFINITION 3.3. *A regular pencil $\lambda B - A$ is standard normal if there exist unitary transformations U_l, U_r and diagonal matrices Λ_A, Λ_B, such that for some real $\gamma \neq 0$*

$$U_l^*(\lambda B - A)U_r = \lambda\Lambda_B - \Lambda_A, \quad |\Lambda_A|^2 + |\Lambda_B|^2 \;=\; \gamma^2 I.$$

For this class of pencils we obviously have

$$1 \leq \kappa(\lambda_i) = \frac{(\max_i |a_{ii}|^2 + \max_i |b_{ii}|^2)^{1/2}}{(|a_{ii}|^2 + |b_{ii}|^2)^{1/2}} \leq \sqrt{2}, \tag{3.7}$$

with the lower bound $\kappa(\lambda_i) = 1$ met for each i in the particular case where $\Lambda_A = \alpha I$ and $\Lambda_B = \beta I$. Obviously the class of standard normal pencils is nearly optimal in terms of eigenvalue sensitivity since $\kappa(\lambda_i) \leq \sqrt{2}$ for each eigenvalue λ_i.

The following theorem explains which pencils can be transformed to standard normal form using left and right transformations.

THEOREM 3.4. *Every regular pencil with a full set of right and left eigenvectors can be transformed into a standard normal form*

$$T_l^{-1}(\lambda B - A)T_r = \lambda\Lambda_B - \Lambda_A, \quad |\Lambda_A|^2 + |\Lambda_B|^2 \;=\; \gamma^2 I \quad \text{with } \gamma \in \mathbb{R}_0.$$

Proof. If $\lambda B - A$ has a full set of right and left eigenvectors x_i, y_i, then putting x_i as the columns of T_r and y_i^* as the rows of T_l^{-1} will diagonalize $T_l^{-1}(\lambda B - A)T_r$. A simple additional diagonal scaling – which can be absorbed in either T_r or T_l – will ensure that moreover $|\Lambda_A|^2 + |\Lambda_B|^2 \;=\; \gamma^2 I$, for some arbitrary real positive γ. ∎

REMARK 3.1. *For non diagonalizable (regular) pencils, the theorem remains valid in the limit, but then T_l, T_r are unbounded. In this case we have that $\lambda\Lambda_B - \Lambda_A$ belongs to the closure of the orbit of $\lambda B - A$ under left and right transformations T_l, T_r [5].*

4. Balancing pencils. We now look for scaling transformations that make a given pencil get closer to a normal one. We could use a scaling of the type

$$D_l^{-1}(\lambda B - A)D_r \tag{4.1}$$

where D_r, D_l are *real positive diagonal scaling matrices*. This does not modify the generalized eigenvalues of the pencil, but the defect from normality $\Gamma(A,B)$ becomes now

$$\Gamma(D_l^{-1}AD_r, D_l^{-1}BD_r) = \sum_{i=1}^{n}\sigma_i^2(D_l^{-1}AB^{-1}D_l) + \sum_{i=1}^{n}\sigma_i^2(D_r^{-1}B^{-1}AD_r) - 2\sum_{i=1}^{n}|\lambda_i|^2.$$

It follows from Section 2 that the optimal D_r, D_l are solutions of

$$\inf_{D_r}\ \|D_r^{-1}B^{-1}AD_r\|_F, \quad \inf_{D_l}\ \|D_l^{-1}AB^{-1}D_l\|_F.$$

But such an approach would require to invert the matrix B (at least implicitly) and it is unclear how to proceed when B is singular.

We now define a new optimization problem inspired from Theorem 3.4 that avoids the inversion of B. It uses the so-called Frobenius inner product for regular pencils defined in a $2n^2$-dimensional space of two $n \times n$ complex matrices :

$$\langle \lambda B_1 - A_1\ ,\ \lambda B_2 - A_2\rangle_F\ :=\ \mathrm{tr}(A_1A_2^* + B_1B_2^*)\ .$$

It follows then that $\|\lambda B - A\|_F^2\ := \langle \lambda B - A, \lambda B - A\rangle_F = \ \|A\|_F^2 + \|B\|_F^2$ where $\|.\|_F$ denotes the usual Frobenius matrix norm.

THEOREM 4.1. *The optimization problem*

$$\inf_{\det(T_l^{-1}T_r)=1}\ \|T_l^{-1}(\lambda B - A)T_r\|_F \tag{4.2}$$

has a standard normal pencil as solution. If $\lambda B - A$ is diagonalizable then $T_r,\ T_l$ have a bounded solution, otherwise they are unbounded.

Proof. Using the Schur decomposition $\lambda B_S - A_S = Q^*(\lambda B - A)Z$ we define triangular matrices $R_r := Z^*T_rQ_r$ and $R_l := Q^*T_lQ_l$ where Q_r are Q_l are chosen to be unitary and $\det Q_l^*Q_r = 1$. Since unitary transformations do not change the Frobenius norm, the above minimization is then equivalent to

$$\inf_{\det(R_l^{-1}R_r)=1}\ \|R_l^{-1}(\lambda B_S - A_S)R_r\|_F$$

where now all matrices are upper triangular. Moreover, if we factor $R_r = D_rU_r$ and $R_l = D_lU_l$, where U_r and U_l are unit upper triangular and D_r and D_l are diagonal, then the problem splits in two subproblems. Clearly U_r and U_l only affect the elements above the diagonal of $\|R_l^{-1}(\lambda B_S - A_S)R_r\|_F$ and these can all be put equal to zero if the pencil is diagonalizable (e.g. when there are no repeated eigenvalues). In such a case the problem reduces further to

$$\inf_{\det(D_l^{-1}D_r)=1}\ \|D_l^{-1}(\lambda\Lambda_B - \Lambda_A)D_r\|_F$$

which is easily solved using a Lagrange multiplier approach. The solution

$$D_l^{-2}D_r^2(\Lambda_B^*\Lambda_B + \Lambda_A^*\Lambda_A) = \gamma^2 I, \quad \gamma^{2n} = \det(\Lambda_B^*\Lambda_B + \Lambda_A^*\Lambda_A)$$

is equivalent to the condition that $D_l^{-1}(\lambda\Lambda_B - \Lambda_A)D_r$ is a standard normal pencil. If the pencil is not diagonalizable, it is still possible to find unbounded diagonal scalings $D_r,\ D_l$ that will make the elements that are above the diagonal in the Schur form tend to zero. ∎

The above theorem suggests to use the same minimization problem but now restricted to *real positive diagonal scaling matrices* :

$$\inf_{\det(D_l^{-1}D_r)=1} \|D_l^{-1}(\lambda B - A)D_r\|_F \tag{4.3}$$

as a technique to balance regular pencils. We will show that this has a unique minimum that is attained when

$$\|(D_l^{-1}AD_r)e_j\|_2^2 + \|(D_l^{-1}BD_r)e_j\|_2^2 = \|e_i^T(D_l^{-1}AD_r)\|_2^2 + \|e_i^T(D_l^{-1}BD_r)\|_2^2 = \gamma^2$$

for all i's and j's. This leads to the following generalization of (2.5).

DEFINITION 4.2. *An $n \times n$ regular complex pencil $\lambda B - A$ is said to be balanced if*

$$\|Ae_j\|_2^2 + \|Be_j\|_2^2 = \|e_i^T A\|_2^2 + \|e_i^T B\|_2^2 = \gamma^2, \quad \forall\ i, j. \tag{4.4}$$

The following theorem proves that every balanced pencil can be seen as the solution of an optimization problem very similar to (4.3).

THEOREM 4.3. *A regular pencil $D_l^{-1}(\lambda B - A)D_r$ with real positive diagonal scalings D_l, D_r, is balanced if and only if it is a solution of*

$$\inf_{\det(D_l^{-1}D_r)=c} \|D_l^{-1}(\lambda B - A)D_r\|_F.$$

Proof. Denote the i-th diagonal entry of D_r and D_l by d_{ri} and d_{li}, respectively, and let $a_{ij},\ b_{ij}$ be the entries of the matrices $A,\ B$. We want to minimize

$$\inf_{d_{li},\ d_{rj}} \sum_{i,j=1}^{n} (|a_{ij}|^2 + |b_{ij}|^2)(\frac{d_{rj}}{d_{li}})^2, \quad \text{where } \ (\frac{\prod d_{lk}}{\prod d_{rk}})^2 = c^2.$$

With the change variables $d_{ri}^2 = \exp(u_{ri})$ and $d_{li}^2 = \exp(-u_{li})$ and when putting $m_{ij} := |a_{ij}|^2 + |b_{ij}|^2$, this becomes

$$\inf_{u_{li},\ u_{rj}} \sum_{i,j=1}^{n} m_{ij}\exp(u_{li} + u_{rj}), \quad \text{where } \ \sum_k (u_{lk} + u_{rk}) = 2\ln c.$$

This is a *convex* minimization problem with a linear constraint. Its solution can be found via the use of a Lagrange multiplier Γ :

$$\inf_{u_{li},\ u_{rj}} \sum_{i,j=1}^{n} m_{ij}\exp(u_{li} + u_{rj}) + \Gamma\left(2\ln c - \sum_k (u_{lk} + u_{rk})\right).$$

This unconstrained minimization has therefore a minimum iff the first order conditions are satisfied. These are $\sum_k (u_{lk} + u_{rk}) = 2\ln c$ and

$$\sum_{i=1}^{n} m_{ij}\exp(u_{li} + u_{rj}) = \sum_{j=1}^{n} m_{ij}\exp(u_{li} + u_{rj}) = \Gamma, \quad \forall\ i, j.$$

Putting $\Gamma = \gamma^2$ and rephrasing it in the original variables, this amounts to

$$\|e_i^T(D_l^{-1}AD_r)\|_2^2 + \|e_i^T(D_l^{-1}BD_r)\|_2^2 \;=\; \|(D_l^{-1}AD_r)e_j\|_2^2 + \|(D_l^{-1}BD_r)e_j\|_2^2 = \gamma^2$$

for all i, j. The optimal pencil $D_l^{-1}(\lambda B - A)D_r$ is therefore balanced. The converse statement is easily checked in a similar manner. ∎

REMARK 4.1. *Notice that if the pencil $\lambda B - A$ can be permuted to a block triangular pencil, then so can the matrix M with elements m_{ij}. One then easily checks that the scalings of the above theorem can be unbounded for this so-called* reducible case*. This case is typically excluded in the scaling problem, since then the generalized eigenvalue problem can de deflated to smaller dimensional ones [7]. When such permutations do not exist, the scaling problem has a bounded solution.*

REMARK 4.2. *The above theorem does not prove that the diagonal scaling procedure will always improve the sensitivity of the eigenvalue problem but the bound (3.4) for $\kappa(\lambda_i)$ suggests that this will be the case. We will illustrate by numerical experiments that the scaling typically improves the sensitivity of the eigenvalues.*

REMARK 4.3. *The above theorem also allows to choose the parameter γ in (4.4) since modifying the constant c in the condition $\det(D_l^{-1}D_r) = c$, automatically scales all the column and row norms. This is used in the numerical method described below.*

5. Numerical method. In order to balance a pencil, we will use a very simple method rather than using convex optimization techniques. This method consists in alternatively updating D_r and D_l such that the compound matrices $\begin{bmatrix} A \\ B \end{bmatrix} D_r$ and $D_l^{-1}\begin{bmatrix} A & B \end{bmatrix}$ have column norms and row norms equal to 1, respectively. By doing so we converge linearly to a balanced pencil with $\gamma = 1$ in (4.4). The proposed method is essentially a "coordinate descent" method where one alternates between computing the optimum in the "coordinates" of D_r and D_l. The convergence is slow but when we restrict ourselves to powers of the base (2 or 10) for the diagonal elements of D_r and D_l, stagnation typically occurs after two or three updates of both D_r and D_l. Each joint update of D_l and Dr in fact requires only $4n^2$ floating point operations if one uses the matrix M with elements $m_{ij} := |a_{ij}|^2 + |b_{ij}|^2 : 2n^2$ to compute the row and column norms and $2n^2$ to perform the two scalings. (A MATLAB code is given in the Appendix for the base 2). The scaling procedure has therefore a marginal cost in comparison to the eigenvalue computation. As in the standard eigenvalue problem one has to test also if there exist permutations that reduce the pencil to a block triangular form so that lower dimensional eigenvalue problems can be isolated. Such a procedure is needed to guarantee that the diagonal scaling will remain bounded but the complexity is also quadratic in n (see [7]).

6. Numerical examples. In the tables below we have compared the precision of the computed eigenvalues without scaling, after applying our proposed scaling procedure and after applying Ward's method [7], which is currently implemented in LAPACK. We considered in Table 6.1 randomly generated diagonalizable pencils $T_l^{-1}(\lambda\Lambda_B - \Lambda_A)T_r$ (where $\lambda\Lambda_B - \Lambda_A$ is in standard normal form), in Table 6.2 randomly generated non diagonalizable pencils $T_l^{-1}(\lambda J_B - J_A)T_r$ (where $J_B^{-1}J_A$ is in Jordan normal form), and in Table 6.3 pencils with elements of strongly varying order of magnitude. We used normally distributed random numbers for the free elements of $\Lambda_A, \Lambda_B, J_A$ and J_B. We imposed the normalization in $\lambda\Lambda_B - \Lambda_A$ by choosing Λ_B to satisfy $\Lambda_B^2 + \Lambda_A^2 = \gamma^2 I$, and the Jordan structure in $\lambda J_B - J_A$ by choosing some repeated consecutive elements on the diagonals of J_A and J_B and assigning

corresponding off-diagonal 1's in J_A. The condition number of the random matrices T_l and T_r was controlled by taking the k-th power of normally distributed random numbers $r_{i,j}$ as their elements. A larger power k then typically yields a larger condition number for the transformation. For these experiments we used the QZ-algorithm [3] applied to different pencils of size $n = 10$. We computed the chordal distances $c_i := \chi(\lambda_i, \tilde{\lambda}_i)$ for all eigenvalues λ_i and compared in each table the quantities $c := \|[c_1, \ldots, c_n]\|_2$ for the original pencil (c_{orig}), for the balanced pencil constructed by our algorithm (c_{bal}), and for the balanced pencil using Ward's method (c_{ward}). In Tables 6.1 and 6.2 we also give the condition numbers $\kappa(T_r)$ and $\kappa(T_l)$.

In Table 6.1 we focus on diagonalizable pencils. When $\kappa(T_r) = \kappa(T_l) = 1$ we observe that balancing does not improve the precision of the calculated eigenvalues, but otherwise it does in general significantly improve the accuracy of the calculated eigenvalues. Recall also that we restrict the diagonal elements of the balancing transformations D_r, D_l to be powers of two. From the table it appears that the proposed balancing algorithm has a positive effect on the precision of the computed eigenvalues. The comparison factor c_{ward}/c_{bal} shows that in general the new method outperforms Ward's algorithm.

TABLE 6.1
Comparison for randomly generated diagonalizable pencils

$n = 10$	c_{orig}	c_{bal}	c_{ward}	c_{ward}/c_{bal}
$\kappa(T_r) = 3.27e+07, \kappa(T_l) = 2.58e+11$	3.01e-03	7.00e-13	2.61e-09	3.72+03
$\kappa(T_r) = 8.24e+12, \kappa(T_l) = 4.21e+10$	3.69e-01	3.20e-12	1.00e-09	3.12e+02
$\kappa(T_r) = 6.81e+08, \kappa(T_l) = 1.75e+07$	7.81e-09	8.84e-14	1.01e-11	1.15e+02
$\kappa(T_r) = 1.06e+07, \kappa(T_l) = 7.82e+08$	1.56e-07	4.90e-13	4.16e-13	8.50e-01
$\kappa(T_r) = 1.46e+05, \kappa(T_l) = 4.08e+05$	2.67e-10	3.52e-15	3.92e-15	1.12e+00
$\kappa(T_r) = 1.92e+03, \kappa(T_l) = 7.72e+02$	6.78e-13	3.04e-15	2.07e-14	6.08e+00
$\kappa(T_r) = 3.95e+01, \kappa(T_l) = 1.75e+01$	2.23e-15	2.20e-15	6.52e-15	2.97e+00
$\kappa(T_r) = 1.00e+00, \kappa(T_l) = 1.00e+00$	4.79e-16	4.79e-16	4.94e-14	1.03e+02

In Table 6.2 we look at non-diagonalizable pencils. We imposed the first example to have two Jordan blocks of size 2, and the second example to have one Jordan block of size 3. The eigenvalue sensitivity is in principle infinite and the calculated eigenvalues have little in common with the true eigenvalues. The table shows that both balancing algorithms do not alter the precision of the computed eigenvalues. In

TABLE 6.2
Randomly generated non-diagonalizable pencils

$n = 10$	c_{orig}	c_{bal}	c_{ward}	c_{ward}/c_{bal}
$\kappa(T_r) = 1.15e+09, \kappa(T_l) = 3.27e+09$	4.88e-01	4.88e-01	4.88e-01	1.00e+00
$\kappa(T_r) = 4.68e+02, \kappa(T_l) = 4.79e+03$	1.30e-01	1.30e-01	1.30e-01	1.00e+00

Table 6.3 we look at pencils with entries of strongly varying size : the largest ones are of the order of 1, the smallest ones are much smaller. Ward's method tries to make the size of these elements equal and in doing so, it applies a scaling that often

deteriorates the sensitivity instead of improving it. The new method, on the other hand, usually significantly improves the sensitivity.

TABLE 6.3
Pencils with elements of strongly varying order of magnitude

c_{orig}	c_{bal}	c_{ward}	c_{ward}/c_{bal}
4.38e-10	4.30e-15	1.02e-05	2.37e+09
1.25e-13	1.90e-15	1.92e-03	1.01e+12
9.16e-12	6.13e-16	1.17e-10	1.92e+05

7. Acknowledgment. We would like to acknowledge the help of Yurii Nesterov in Theorem 4.3 who pointed out that this was a convex optimization problem. We also thank Daniel Kressner who sent us a 3 by 3 example from his thesis for which Ward's scaling significantly deteriorates the sensitivity of the computed eigenvalues. The examples in Table 6.3 are inspired from this.

8. Conclusion. In this paper we presented a new balancing method for matrix pencils. From the point of view of the sensitivity of the eigenvalues we showed that the standard normal pencils are near optimal and that they can be viewed as a natural extension of normal matrices. A diagonal balancing method was then proposed that makes a given pencil as close as possible to a standard normal one. Moreover we showed that the complexity of the new method is comparable to that of the classical balancing of matrices. We also gave numerical evidence that the accuracy of computed generalized eigenvalues may significantly improve after balancing a pencil and that the method often outperforms the method of Ward implemented in LAPACK.

REFERENCES

[1] J.-P. Charlier and P. Van Dooren. A Jacobi-like algorithm for computing the generalized Schur form of a regular pencil. *J. Comput. Appl. Math.*, 27(1-2):17–36, 1989. Reprinted in *Parallel algorithms for numerical linear algebra*, 17–36, North-Holland, Amsterdam, 1990.

[2] R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press, Cambridge, 1991.

[3] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973.

[4] B. N. Parlett and C. Reinsch. Balancing a matrix for calculation of eigenvalues and eigenvectors. *Numer. Math.*, 13:293–304, 1969.

[5] A. Pokrzywa. On perturbations and the equivalence orbit of a matrix pencil.. *Linear Algebra Appl.* 82:99–121, 1986.

[6] G. W. Stewart. Perturbation theory for the generalized eigenvalue problem. In *Recent Advances in Numerical Analysis*, Eds. de Boor C. and Golub G., Academic Press, New York, 1978, 193-206.

[7] R. C. Ward. Balancing the generalized eigenvalue problem. *SIAM J. Sci. Stat. Comput.*, 2:141–152, 1981.

Appendix.

```
function [Dl, Dr, iter] = baleig(A,B,max_iter)

% Performs two-sided scaling Dl\A*Dr, Dl\B*Dr in order to improve
% the sensitivity of generalized eigenvalues. The diagonal matrices
% Dl and Dr are constrained to powers of 2 and are computed iteratively
% until the number of iterations max_iter is met or until the norms are
% between 1/2 and 2. Convergence is often reached after 2 or 3 steps.
% The diagonals of the scaling matrices are returned in Dl and Dr
% and so is iter, the number of iterations steps used by the method.

n=size(A,1); Dl=ones(1,n); Dr=ones(1,n); M=abs(A).^2+abs(B).^2;

for iter=1:max_iter,
    emax=0;emin=0;
    for i=1:n;
        % scale the rows of M to have approximate row sum 1
        d=sum(M(i,:));e=-round(log2(abs(d))/2);
        M(i,:)=pow2(M(i,:),2*e);
        % apply the square root scaling also to A, B and Dl
        Dl(i)=pow2(Dl(i),-e);
        if e > emax, emax=e; end; if e < emin, emin=e; end
    end
    for i=1:n;
        % scale the columns of M to have approximate column sum 1
        d=sum(M(:,i));e=-round(log2(abs(d))/2);
        M(:,i)=pow2(M(:,i),2*e);
        % apply the square root scaling also to A, B and Dr
        Dr(i)=pow2(Dr(i),e);
        if e > emax, emax=e; end; if e < emin, emin=e; end
    end
    % Stop if norms are all between 1/2 and 2
    if (emax<=emin+2), break; end
end
```