# Interpolating Self Consistent Field for Eigenvector Nonlinearities*

Rob Claes[a], Karl Meerbergen[a]

[a]*Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001, Leuven, Belgium*

**Abstract**

One of the most common approaches for solving eigenvalue problems with eigenvector nonlinearities (NEPv) is the Self Consistent Field (SCF) method that uses a vector from the previous iteration to build a zeroth order approximation of the nonlinearity. This approach is often slow and unreliable, which is why most applications use mixing scheme extensions of SCF, that use a linear combination of multiple previous iterates to build the zeroth order approximation. In this paper, we present a method that uses multiple previous iterates to build an interpolating first order approximation of the nonlinearity. It can be shown that this method converges superlinearly to the desired eigenpair.

*Keywords:* nonlinear eigenvalue problem, self consistent field, Secant method, interpolation

## 1. Introduction

We consider the nonlinear eigenvector problem (NEPv)

$$A(v)v = \lambda v \tag{1}$$

where $A : \mathbb{C}^n \to \mathbb{C}^{n\times n}$ is a Hermitian matrix valued function and $(\lambda, v) \in \mathbb{C} \times \mathbb{C}^n$ an eigenpair. The function $A$ is defined such that for any $\alpha \neq 0$ it holds that $A(\alpha v) = A(v)$.

Eigenvector nonlinearities commonly arise in electronic structure calculations in quantum chemistry and physics [1], more specifically in the Hartree-Fock approximations of the Schrödinger and Gross-Pitaevskii equations [2]. More recently, applications of NEPv can be found in the field of data science: spectral clustering [3, 4], balanced graph cuts [5], robust Rayleigh quotient minimization for linear discriminant analysis [6], core-periphery detection in networks [7] and orthogonal canonical correlation analysis [8].

A frequently used method to solve the NEPv is Self Consistent Field (SCF) [2, 9], which is given in the basic version by Algorithm 1. In most applications, the eigenpair with either the smallest or largest eigenvalue is calculated, but SCF can target any eigenvalue of the problem. Even though the simplest version of SCF can be implemented in a few lines of code, its convergence properties can be difficult to

*This work is supported by the Research Foundation Flanders (FWO) Grant G0B7818N and the KU Leuven Research Council.

understand. The convergence of a specific subclass is studied in [10], while more general approaches are taken in [11, 12, 13].

**Algorithm 1:** Standard SCF for NEPv

**Input:** $v_0$

**for** $k = 1, 2, 3, \ldots$ **do**

| Solve $A(v_{k-1})v_k = \lambda_k v_k$ for the eigenpair $(\lambda_k, v_k)$.

**end**

In many applications standard SCF converges very slowly or does not converge at all. Level-shifted SCF is often used for applications for which standard SCF does not converge [14]. The method applies standard SCF to a slightly altered problem $B(v) = A(v) - \sigma vv^*$ with a fixed parameter $\sigma$. It is clear that the altered problem will converge to the same normalized eigenvector with a corresponding eigenvalue that is shifted by the value of $\sigma$. It is possible to find a large enough value of $\sigma$ to make the level-shifted SCF convergent. A major drawback is that fast convergence is only achieved for a narrow range of values for $\sigma$ [15].

The SCF iteration can also be stabilized by the addition of a mixing scheme like DIIS [16]. The general idea of this approach is to alter Algorithm 1 by evaluating $A$ not in $v_{k-1}$ but in a linear combination of previous iterates $\tilde{v} = \sum_{i=k-p}^{k-1} c_i v_i$. The coefficients $c_i$ can be determined by different optimality criteria [17].

In this paper, we introduce a new SCF extension to incorporate more than one iterate in each iteration. The presented approach exhibits superlinear convergence and – due to its resemblance to the root finding method – is called Secant SCF. The remainder of this text is structured as follows. The concept of Secant SCF is introduced in Section 2 together with a derivation of the convergence properties. The reliability of the proposed method is determined by three important numerical aspects that are discussed in Section 3. Finally, a numerical example in Matlab is presented in Section 4.

## 2. Secant SCF

The standard SCF method uses only one previous iterate to build a zeroth order approximation of the eigenvector nonlinearity. Mixing schemes like DIIS use multiple previous iterates, but still use a zeroth order approximation in each iteration. In contrast, we introduce a method that uses multiple previous iterates to build a higher order approximation of the nonlinearity. To this end, we start with two iterations of standard SCF to obtain approximations $(\lambda_1, v_1)$ and $(\lambda_2, v_2)$. Next, the interpolating polynomial of degree 1 in $\lambda$, $A^S(\lambda)$ is constructed such that $A^S(\lambda_i) = A(v_i) - \lambda_i I$ for $i = 1, 2$. While the interpolating polynomial is uniquely defined, it can be expressed in different ways using different basis functions. In the remainder of this paper, we use the Newton polynomial basis as explained in Section 3. The resulting method, called Secant SCF, is given in Algorithm 2. The name originates from the similarities to the univariate root finding

method, Secant. The convergence similarity between our method and the standard root finding method is discussed in section 2.1.

**Algorithm 2:** Secant SCF for NEPv

**input:** $v_0$

**for** $k = 1, 2$ **do**
  Solve $(A(v_{k-1}) - \lambda_k I)\, v_k = 0$ for the eigenpair $(\lambda_k, v_k)$.

**end**

**for** $k = 3, 4, \ldots$ **do**
  Construct $A_k^S(\lambda)$, the polynomial of degree 1 such that $A_k^S(\lambda_i) = A(v_i) - \lambda_i I$ for $i = k-2, k-1$.
  Solve $A_k^S(\lambda_k) v_k = 0$ for the eigenpair $(\lambda_k, v_k)$.

**end**

An intuitive extension of this idea might be to use $k > 2$ previous iterates. However, there are two main reasons to expect no significant improvement. The first reason is based on the similarity between our method and the Secant method. It can be shown that, under mild conditions, the Secant method converges with order of convergence $\varphi \approx 1.62$. The extension of Secant with three previous iterates, Muller's method [18], is known to have an order of convergence of $\varphi \approx 1.84$. Therefore, we do not expect a significant reduction in the amount of iterations when using more than two previous iterates. The second reason originates from the linearization of the polynomial eigenvalue problem: the linearized pencil is of size $(k-1)n$. This means that each iteration becomes significantly more expensive due to the larger size of the linearization. However, the impact of the increased problem size can be reduced by structure exploiting methods like [19].

### *2.1. Convergence rate*

In order to derive the convergence rate of Secant SCF, an important result from perturbation theory is stated in the following Theorem and Corollary.

**Theorem 1.** *[Generalization of [20, Theorems 3.6 and 3.7]] Consider the regular matrix pencil $(A, B)$ for which a $\sigma \in \mathbb{C}$ exists such that $A - \sigma B$ is nonsingular. The matrix $X$ is defined by the Jordan form $J = X^{-1}(A - \sigma B)^{-1} B X$ with $J$ a Jordan matrix. Let $(\tilde{\lambda}, \tilde{x})$ be an estimation of a simple eigenpair with residual $r = A\tilde{x} - \tilde{\lambda} B\tilde{x}$. Then there exists an eigenvalue $\lambda$ of $(A, B)$ such that the following inequality holds:*

$$\left| \frac{1}{\lambda - \sigma} - \frac{1}{\tilde{\lambda} - \sigma} \right| \leq \kappa_2(X) \left\| (A - \sigma B)^{-1} \right\|_2 \|r\|_2 . \tag{2}$$

**Corollary 2.** *Consider de regular matrix pencil $(A, B)$. Let $(\tilde{\lambda}, \tilde{x})$ be an estimation of a simple eigenpair with corresponding residual $r = A\tilde{x} - \tilde{\lambda} B\tilde{x}$. Then there exists an eigenvalue $\lambda$ of $(A, B)$, and a positive constant $\gamma < \infty$ such that the following inequality holds:*

$$\left| \lambda - \tilde{\lambda} \right| \leq \gamma \|r\|_2 . \tag{3}$$

Note that the value of $\gamma$ is dependent on the matrix pencil $(A, B)$ by the condition number $\kappa_2(X)$ of the matrix $X$ defined by the Jordan form $J = X^{-1}(A - \sigma B)^{-1}BX$ for some value of $\sigma$ such that $A - \sigma B$ is nonsingular.

The derivation of the convergence rate of Secant SCF starts somewhat backwards and assumes that for a given problem Secant SCF converges to a solution. We will not prove that Secant SCF converges. But we will prove that if it converges, the convergence rate is superlinear for a simple eigenvalue. Now, define the sequences $\{\lambda_k\}$ and $\{v_k\}$ from the pairs $(\lambda_k, v_k)$ in iteration $k$ of Secant SCF. Assume that the sequences converge to $\lambda_*$ and $v_*$ respectively as $k \to \infty$ such that $A(v_*)v_* = \lambda_* v_*$. Under the assumption that all $\lambda_k$ are distinct, it is possible to define an interpolating entire function $x(\lambda)$ such that $x(\lambda_k) = v_k$ for all $k$ [21]. Now, consider the matrix valued function $\mathcal{A}(\lambda) = A(x(\lambda)) - \lambda I$. Ignoring the first two iterations, Secant SCF can be written as finding the smallest eigenvalue and corresponding eigenvector of

$$A_k^S(\lambda) = \mathcal{A}(\lambda_{k-1}) + (\lambda - \lambda_{k-1})\frac{\mathcal{A}(\lambda_{k-2}) - \mathcal{A}(\lambda_{k-1})}{\lambda_{k-2} - \lambda_{k-1}}. \tag{4}$$

The next step is to define the value $e_k = \lambda_k - \lambda_*$ as the error in iteration $k$. For small values of $e_k$, $\mathcal{A}(\lambda_k)$ can be developed as

$$\begin{aligned} \mathcal{A}(\lambda_k) &= \mathcal{A}(\lambda_* + e_k) && (5)\\ &= \mathcal{A}(\lambda_*) + e_k\mathcal{A}'(\lambda_*) + \frac{e_k^2}{2}\mathcal{A}''(\lambda_*) + \mathcal{O}(e_k^3) && (6) \end{aligned}$$

Next, define the residual for iteration $k$ as

$$r_k = A_k^S(\lambda_*)v_*. \tag{7}$$

This definition of the residual might seem odd, since it does not involve the original eigenvector nonlinearity. It is however very effective at measuring the convergence of the proposed method, since a good approximation of the nonlinearity $A_k^S$ leads to a small residual at the desired solution $(\lambda_*, v_*)$. Using the Taylor expansion (6), the residual can be expressed as

$$\begin{aligned} r_k &= A_k^S(\lambda_*)v_* && (8)\\ &= e_{k-1}\mathcal{A}'(\lambda_*)v_* + \frac{e_{k-1}^2}{2}\mathcal{A}''(\lambda_*)v_* && (9)\\ &+ \frac{\lambda_* - \lambda_{k-1}}{\lambda_{k-2} - \lambda_{k-1}}\left((e_{k-2} - e_{k-1})\mathcal{A}'(\lambda_*) + \frac{1}{2}(e_{k-2}^2 - e_{k-1}^2)\mathcal{A}''(\lambda_*)\right)v_* + \mathcal{O}(e_{k-1}^3) + \mathcal{O}(e_{k-2}^3) \end{aligned}$$

Neglecting higher order terms, $r_k$ is approximated as

$$r_k \approx e_{k-1}\mathcal{A}'(\lambda_*)v_* + \frac{e_{k-1}^2}{2}\mathcal{A}''(\lambda_*)v_* - \frac{e_{k-1}}{e_{k-2}-e_{k-1}}\left((e_{k-2}-e_{k-1})\mathcal{A}'(\lambda_*) + \frac{1}{2}(e_{k-2}^2 - e_{k-1}^2)\mathcal{A}''(\lambda_*)\right)v_* \tag{10}$$

$$= \frac{e_{k-1}^2}{2}\mathcal{A}''(\lambda_*)v_* - \frac{1}{2}e_{k-1}(e_{k-1}+e_{k-2})\mathcal{A}''(\lambda_*)v_* \tag{11}$$

$$= -\frac{1}{2}e_{k-1}e_{k-2}\mathcal{A}''(\lambda_*)v_* \tag{12}$$

In the remainder of this derivation, the following assumption is made.

**Assumption 3.** *The local convergence of Secant SCF, described by the errors $e_k$, is given by*

$$|e_k| \approx \beta\,|e_{k-1}|^p\,, \quad p>1, \beta>0 \tag{13}$$

*for all $k \geq K > 0$.*

This assumptions states that after $K$ iterations Secant SCF reaches local convergence of order $p$. We are now ready to derive the numerical value of $p$ for Secant SCF.

Starting from (12) and taking the 2-norm, leads to

$$\|r_k\|_2 \approx \frac{1}{2}\,|e_{k-1}|\,|e_{k-2}|\,\|\mathcal{A}''(\lambda_*)v_*\|_2\,. \tag{14}$$

Using the inequality of Corollary 2 to replace $\|r_k\|_2$, and by defining $\tau = \frac{1}{2}\gamma\,\|\mathcal{A}''(\lambda_*)v_*\|_2$ we get:

$$|e_k| \leq \tau\,|e_{k-1}|\,|e_{k-2}|\,. \tag{15}$$

By Assumption 3 and defining $\alpha = \left(\frac{\tau}{\beta}\right)^{\frac{1}{p-1}}$:

$$\beta\,|e_{k-1}|^{p-1} \leq \tau\,|e_{k-2}| \tag{16}$$

$$|e_{k-1}| \leq \alpha\,|e_{k-2}|^{\frac{1}{p-1}}\,. \tag{17}$$

By Assumption 3:

$$\beta\,|e_{k-2}|^p \leq \alpha\,|e_{k-2}|^{\frac{1}{p-1}} \tag{18}$$

or

$$\frac{\beta}{\alpha} \leq |e_{k-2}|^{\frac{1}{p-1}-p}\,. \tag{19}$$

It is clear that $\frac{\beta}{\alpha} > 0$ and that $\lim_{k\to\infty}|e_{k-2}| = 0$, which leads to the conclusion that $\frac{1}{p-1} - p \geq 0$ or $p \geq \frac{1+\sqrt{5}}{2} \approx 1.62$ for all $k > K$. The order of convergence for Secant SCF is thus at least the same or higher than for the standard Secant method.

## 3. Numerical implementation

This section contains important remarks on the numerical implementation of the proposed method. The choice of polynomial basis, the selection of the eigenvalues in each iteration, and the stopping criterion are essential for a reliable numerical method.

*Choice of basis*

For theoretical derivations the choice of basis is not relevant, but it plays an important role in the numerical stability of the implementation. Secant SCF will always lead to a series of interpolation points that grow closer and closer to each other. This may lead to numerical instabilities in the calculation of the interpolating polynomial. In this discussion, we consider the monomial basis, Newton polynomials, and Lagrange polynomial formulations as described in [19]. The Lagrange formulation is often preferred due to its simple formulation, but the Newton and monomial formulation leads to a more stable algorithm for problems of the form (21). A reliable calculation of the divided difference coefficients for Newton polynomials is possible since the nonlinearity only appears on the diagonal of the matrix. The numerical experiment shows the difference in convergence between the Newton and monomial basis on one hand and the Lagrange formulation on the other hand.

*Selection of eigenvalues*

In order to obtain reliable convergence, it is important to make an appropriate selection of eigenpairs to continue the iteration. The solutions of interest are the smallest, real eigenvalues and the corresponding eigenvectors of (1). The linearized pencil from the interpolating polynomial generally does not preserve the complete spectrum of (1) which introduces spurious eigenvalues that can be smaller than the smallest eigenvalue of (1). This problem can be alleviated by selecting the eigenvalue nearest to the eigenvalue from the previous iteration. Therefore we propose the following scheme for interpolating SCF: select the smallest real eigenvalue for the first two standard SCF iterations, followed by the eigenvalue nearest to the previous eigenvalue in the following iterations.

*Stopping criterion*

The simplest way to define the stopping criterion for Secant SCF is when the residual of the NEPv is smaller than a given tolerance. However, when the tolerance is set too small, Secant SCF will stagnate, or, in the worst case, start diverging. This diverging phenomenon originates in the calculation of the divided difference coefficient. Close to a solution, the consecutive approximations are very close to each other, meaning that $\lambda_k - \lambda_{k-1} \approx 0$. In this situation, it is better to set a minimum value for the distance between two interpolation points as the stopping criterion.

## 4. Numerical experiments

The discretization of the Gross-Pitaevskii equation for modelling the Bose-Einstein condensation, leads to a NEPv of the form

$$H(v)v = \lambda v, \tag{20}$$

$$H(v) = A_f + \beta \text{Diag}(|v|)^2. \tag{21}$$

It is defined by a Hermitian positive definite matrix $A_f$, the parameter $\beta$ and the complex eigenvector $v$, where $|\cdot|$ defines the element-wise absolute value. This model was first studied in [22] and the discussion on the convergence of SCF for this problem was extended by [15]. For the potential function $f(x, y)$, $A_f$ is given by

$$A_f = \text{Diag}(\tilde{f}) - \frac{1}{2}L - \imath\omega L_\phi \tag{22}$$

where

$$\tilde{f} = h^2[f(x_1, y_1), \ldots, f(x_N, y_1), f(x_1, y_2), \ldots, f(x_N, y_2), \ldots, f(x_N, y_N)]^T \tag{23}$$

with $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$ the interior points of the discretization interval $[-1, 1] \times [-1, 1]$ using $N + 2$ equidistant points in each direction, such that $h = \frac{2}{N+1}$. The matrix $L$ is given by

$$L = D_{2,N} \otimes I_N + I_N \otimes D_{2,N} \tag{24}$$

and $L_\phi$ is given by

$$L_\phi = h\text{Diag}(y_1, \ldots, y_N) \otimes D_N - D_N \otimes (h\text{Diag}(x_1, \ldots, x_N)) \tag{25}$$

with tridiagonal matrices $D_N = \text{tridiag}(-0.5, 0, 0.5)$ and $D_{2,N} = \text{tridiag}(1, -2, 1)$.

For this example, we take the potential function $f(x, y) = \frac{1}{2}(x^2 + y^2)$ and set $\omega = 0.85$, $\beta = 6$ and $N = 25$ which leads to a NEPv of size $625 \times 625$. The starting vector $v_0$ is set to the eigenvector with the smallest eigenvalue of $A_f$. This approach for the starting vector is often chosen in applications, but differs from the strategy in [15, 22]. The eigenvalue selection follows the scheme proposed in Section 3

In this experiment, Secant SCF is compared to standard SCF and SCF with DIIS. An eigenpair is considered converged when its residual norm is smaller than $\varepsilon_{tol} = 10^{-14}$. Figure 1 shows the convergence of standard SCF (Algorithm 1), SCF with DIIS using two previous vectors, and Secant SCF (Algorithm 2). It is clear that standard SCF has a linear convergence pattern, while the extension with DIIS converges faster. The Secant method illustrates its superlinear convergence and converges in the least amount of iterations. In this case, the tolerance is set large enough to be the stopping criterion for Secant SCF. It is also clear that, for this problem, the Newton and monomial basis functions are more reliable than the Lagrange polynomial implementation which stagnates at a residual of approximately $10^{-9}$.

The convergence graph in Figure 1 shows that Secant SCF converges in significantly fewer iterations than standard SCF. This does not always guarantee that Secant SCF also has a lower computational cost. Our

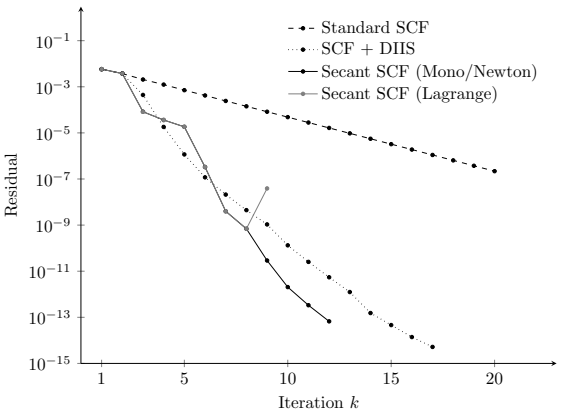


Figure 1: Comparison between standard SCF, SCF+DIIS and Secant SCF.

proof-of-concept implementation in Matlab showed an average time per iteration of 12 ms for standard SCF, 24 ms for Newton Secant SCF, 31 ms for Lagrange Secant SCF, 25 ms for monomial Secant SCF, and 17 ms for SCF with DIIS. For the presented example the total computational cost for Newton and monomial Secant SCF is roughly the same as for SCF with DIIS with an average total computational time of 290 ms. Note that the presented timings should be interpreted in a relative sense because they are all based on proof-of-concept implementations. We leave the discussion regarding an efficient numerical implementation for future research.

## References


[1] A. Szabo, N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, Dover publications, 1996.

[2] Y. Saad, J. T. Chelikowsky, S. M. Shontz, Numerical methods for electronic structure calculations of materials, SIAM Rev. 52 (1) (2010) 3–54.

[3] T. Bühler, M. Hein, Spectral clustering based on the graph $p$-Laplacian, in: Proceedings of the 26th International Conference on Machine Learning, 2009, pp. 81–88.

[4] M. Hein, T. Bühler, An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse PCA, in: Advances in Neural Information Processing Systems 23, 2010, pp. 847–855.

[5] L. Jost, S. Setzer, M. Hein, Nonlinear eigenproblems in data analysis: Balanced graph cuts and the ratioDCA-Prox, in: Extraction of quantifiable information from complex systems, Springer, 2014, pp. 263–279.

[6] Z. Bai, D. Lu, B. Vandereycken, Robust Rayleigh quotient minimization and nonlinear eigenvalue problems, SIAM J. Sci. Comput. 40 (5) (2018) A3495–A3522.

[7] F. Tudisco, D. J. Higham, A nonlinear spectral method for core–periphery detection in networks, SIAM J. Math. Data. Sci. 1 (2) (2019) 269–292.

[8] L. Zhang, L. Wang, Z. Bai, R.-C. Li, A self-consistent-field iteration for orthogonal CCA, IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).

[9] R. M. Martin, Electronic structure: basic theory and practical methods, Cambridge university press, 2020.

[10] C. Yang, W. Gao, J. C. Meza, On the convergence of the self-consistent field iteration for a class of nonlinear eigenvalue problems, SIAM J. Matrix Anal. Appl. 30 (4) (2009) 1773–1788.

[11] Y. Cai, L.-H. Zhang, Z. Bai, R.-C. Li, On an eigenvector-dependent nonlinear eigenvalue problem, Tech. rep., Univ. Texas Arlington (2017).

[12] X. Liu, X. Wang, Z. Wen, Y. Yuan, On the convergence of the self-consistent field iteration in Kohn–Sham density functional theory, SIAM J. Matrix Anal. Appl. 35 (2) (2014) 546–558.

[13] P. Upadhyaya, E. Jarlebring, E. H. Rubensson, A density matrix approach to the convergence of the self-consistent field iteration, Numerical Algebra, Control and Optimization (4) (2019).

[14] E. Cancès, C. L. Bris, On the convergence of SCF algorithms for the Hartree-Fock equations, M2AN, Math. Model. Numer. Anal. (2000).

[15] Z. Bai, R.-C. Li, D. Lu, Sharp estimation of convergence rate for self-consistent field iteration to solve eigenvector-dependent nonlinear eigenvalue problems, SIAM J. Matrix Anal. Appl. 43 (1) (2022) 301–327.

[16] P. Pulay, Convergence acceleration of iterative sequences - the case of SCF iteration, Chemical physics letters 73 (2) (1980) 393–398.

[17] A. J. Garza, G. E. Scuseria, Comparison of self-consistent field convergence acceleration techniques, The Journal of chemical physics 137 (5) (2012) 054110.

[18] D. E. Muller, A method for solving algebraic equations using an automatic computer, Mathematical tables and other aids to computation 10 (56) (1956) 208–215.

[19] R. Van Beeumen, K. Meerbergen, W. Michiels, Compact rational krylov methods for nonlinear eigenvalue problems, SIAM J. Matrix Anal. Appl. 36 (2) (2015) 820–838.

[20] Y. Saad, Numerical methods for large eigenvalue problems, Manchester University Press, 1992.

[21] J.-A. Désidéri, Interpolation of infinite sequences by entire functions, Applied numerical mathematics 58 (12) (2008) 1918–1932.

[22] E. Jarlebring, S. Kvaal, W. Michiels, An inverse iteration method for eigenvalue problems with eigenvector nonlinearities, SIAM J. Sci. Comput. 36 (4) (2014) A1978–A2001.