Regularised discretisations obtained from first-kind Fredholm operator equations

George Fikioris

School of Electrical and Computer Engineering, National Technical University, Zografou, Greece

Abstract

Judicious discretisations of certain first-kind Fredholm operator equations are tantamount to Fredholm infinite-matrix equations of the second kind. We give detailed explanations for the occurrence of this interesting and useful phenomenon and carefully examine the behaviour of the discretisations, especially with respect to matrix condition numbers. Our main tool is a certain equivalence of the discretisations—within the context of finite matrices—to discretisations of Fredholm operator equations of the second kind.

1 | INTRODUCTION

Consider a linear operator equation whose operator \( G \) can be split into two parts, \( G = G_1 + G_2 \), where the first part \( G_1 \) is invertible, with a known inverse \( G_1^{-1} \). This is the original equation,

\[
GX = Y \quad \text{or} \quad (G_1 + G_2)X = Y. \tag{1}
\]

Application of the ‘regulariser’ \( G_1^{-1} \) yields the transformed equation,

\[
(I + G_1^{-1}G_2)X = G_1^{-1}Y, \tag{2}
\]

in which \( I \) is the identity operator of the underlying infinite-dimensional complex Hilbert space. The transformation of \( (1) \) into \( (2) \) is the method of analytical regularisation (MAR), also called semi-inversion. Use of \( (2) \) instead of \( (1) \) is generally advantageous when \( G_1 \) and \( G_2 \) are such that the following statements are both true.

(i) Equation \( (1) \) is a Fredholm integral equation of the first kind; and

(ii) the operator \( G_1^{-1}G_2 \) in \( (2) \) is compact, so that \( (2) \) is a Fredholm integral equation of the second kind.

The advantages are well-discussed in the computational electromagnetics literature, see especially Nosich’s review articles [1] and [2]. For example, Fredholm theory applies to \( (2) \) but not to \( (1) \). Furthermore, \( (1) \)—but not \( (2) \)—often leads to large matrix condition numbers, which increase with matrix dimension.

While it is generally beneficial to discretise \( (2) \) instead of \( (1) \), this article focuses on a case where it makes no essential difference whether \( (1) \) or \( (2) \) is discretised, even when the operators are such that both \( (i) \) and \( (ii) \) are satisfied. This is true for specific finite discretisations of \( (1) \) and \( (2) \) that the author calls Discretisation 1 and Discretisation 2, respectively. Thus, it is explained when Discretisations 1 and 2 are equivalent within the context of matrices and linear systems of finite dimension. Then, a number of consequences of this phenomenon for computational electromagnetics are described.

Our two discretisations are not always equivalent; they are if the operators in \( (1) \) satisfy conditions which we explicate in detail. On the other hand, we leave out operator properties that do not pertain to the particular matter under discussion. As an example, our main results regarding equivalence (Theorems 1 and 2) hold whether or not \( G_1^{-1}G_2 \) is compact, and whether or not \( G_1 \) is the most singular part of \( G \) (which is often true in applications). Also, both Discretisations 1 and 2 make use of the eigenfunctions of \( G_1 \). Throughout, it is assumed that these functions are known (but this might be difficult in practice, especially if the eigenfunctions are to be computed numerically).

Parts of this article improve, extend, and clarify certain results that first appeared in the 2001 article [3]. Here, whenever possible, we present results in the form of precise theorems, which are proved in the Appendix. Our results pertain to a number of findings and/or methods existing in the literature.
For example, what is here termed Discretisation 1 has been called ‘a regularised discretisation scheme’ that ‘combines both regularisation (semi-inversion) and discretisation in one single procedure’ and ‘bridges the gap between MAR and conventional method of moments solutions’ [1]. Discretisation 1 has recently been called the ‘Method of Analytical Preconditioning (MAP)’ [1,4] which ‘immediately results in a regularised discretisation scheme (i.e. yields a Fredholm second-kind infinite matrix-operator equation)’ [1]. MAR/MAP is especially useful in the case of canonical-shape inversion [1], and has been used in many problems of computational electromagnetics and photonics, see [1–5] and the many relevant works referred to in [5].

When discussing the theory behind such regularised schemes, however, nearly all of the current literature still deals with infinite matrix equations, see especially [1]. Since the study that follows uses finite matrices, it provides greater clarity, if only because infinite matrix equations are always truncated in practice.

2 | NOTATION: GALERKIN's METHOD

Familiar concepts from linear algebra, operator theory, and functional analysis are used here. \( \langle \Phi, \Psi \rangle \) denotes the inner product of the Hilbert-space elements \( \Phi \) and \( \Psi \), for which \( \langle \Phi, \Psi \rangle = \langle \Psi, \Phi \rangle \) and \( \langle a \Phi, \Psi \rangle = \langle \Phi, a \Psi \rangle = a \langle \Phi, \Psi \rangle \), where \( a \in \mathbb{C} \) and where the overbar denotes the complex conjugate. The symbol \( \| \cdot \| \) is used for norms, both of Hilbert-space elements and of Hilbert-space operators; the meaning should be clear from the context. Symmetric operators (often called formally self-adjoint operators), which are well discussed in [6], play an important role. \( \delta_{ij} \) is the Kronecker delta.

The discretisations that concern us arise by applying Galerkin’s method to (1) and (2), with basis (or expansion) functions to be specified. The testing (or weighting) functions (to use the usual language of the method of moments [7,8]) are the same as the basis functions. Thus, when applied to \( HX = Z \), the Galerkin method with basis functions \( \Phi_0, \Phi_1, \ldots, \Phi_{n-1} \) produces an approximate solution \( X_\approx X_n \) given by:

\[
X_n = \sum_{j=0}^{n-1} x_j \Phi_j, \tag{3}
\]

in which the \( n \)-dependent complex numbers \( x_0, x_1, \ldots, x_{n-1} \) are the solutions to the \( n \times n \) system of linear equations:

\[
\sum_{j=0}^{n-1} \langle H \Phi_j, \Phi_i \rangle x_j = \langle Z, \Phi_i \rangle, \quad i = 0, 1, \ldots, n - 1. \tag{4}
\]

Throughout, it is assumed that systems such as (4) are uniquely solvable. As discussed in [9], what we here call Galerkin’s method is applied to operator equations in both \( L^2 \) and Sobolev-space settings.

3 | EQUIVALENT DISCRETISATIONS IN GENERAL

For the purposes of this paper, the term discretisation is used as follows.

**Definition 3.1** An \( n \times n \) discretisation of a given operator equation is specified by an \( n \times n \) system of linear algebraic equations whose solution is intended to approximate the solution to the operator equation. The former solution is the **discretised solution**.

For example, we say that (4) is an \( n \times n \) discretisation of the operator equation \( HX = Z \). Definition 3.1 does not specify the way in which the discretised solution is supposed to approximate the true solution (i.e. the solution to the operator equation), because the specific way is immaterial to our main results.

In elementary matrix theory, two \( n \times n \) systems are equivalent if their solutions are the same. It is worth rephrasing this concept in terms of what we have called discretisations.

**Definition 3.2** Two \( n \times n \) discretisations (not necessarily of the same operator equation) are **equivalent** if the corresponding discretised solutions are equal. We have **row-scaled-equivalent** discretisations when, for all \( i = 0, 1, \ldots, n - 1 \), equation \( i \) of the first discretisation differs from equation \( i \) of the second only by a nonzero multiplicative factor (which may be \( i \)-dependent).

For example, we obtain a row-scaled equivalent discretisation of (4) if we multiply each equation in (4) by \( d_i \), where \( d_i \neq 0 \). With regard to Definition 3.2, let us stress the following:

(i) In nontrivial equivalent discretisations, the two matrices in Definition 3.1 are not the same. In the case of row-scaled-equivalent discretisations, we obtain the first (or second) matrix by left-multiplying the second (or first) by a diagonal matrix.

(ii) The matrix dimension \( n \) is fixed and finite. In particular, we do not assume that our discretisations are convergent. This is why our equivalence is unrelated to the closeness (or lack of closeness) between the discretisations and the true solutions (i.e. the solutions of the respective operator equations).

(iii) The equality of the discretised solutions assumes no round-off errors, quadrature errors, errors in computations of special functions, etc. This is important, because errors of this sort can affect the respective solutions differently, as customarily measured by the matrix condition number. Thus, one of the two equivalent discretisations may be superior to the other in this regard. In what follows, after establishing MAR-related equivalence, matrix condition numbers are compared.
Our notion of equivalence becomes interesting when it pertains to two different operator equations satisfied by the same physical quantity, as is the case with (1) and (2).

There are many well-known examples of what we here call ‘equivalent discretisations’: (i) With due regard to the proper spaces [9], the Rayleigh-Ritz method is equivalent to Galerkin’s method [8–10]. (ii) Reference [10] notes a number of such equivalences pertaining to Pocklington-type equations. (iii) As explained in [10] and elsewhere, a certain discretisation of Hallén’s equation is equivalent to a discretisation of Pocklington’s equation; in this case, both the operator equations and the basis/testing functions are different. For discussions of the Ritz method is equivalent to Galerkin’s method [8–10]. (ii) Reference [10] notes a number of such ‘equivalent discretisations’: (i) With due regard to the proper spaces [9], the Rayleigh-Ritz method is equivalent to Galerkin’s method [8–10]. (iii) As explained in [10] and elsewhere, a certain discretisation of Hallén’s equation is equivalent to a discretisation of Pocklington-type equations.

4 | EQUIVALENT DISCRETISATIONS RELATED TO THE MAR

This section contains two theorems describing equivalent discretisations of (1) and (2). In both theorems, the basis functions \( \Phi_j \) in the Galerkin method are orthogonal eigenfunctions of \( G_1 \), so that

\[
G_1 \Phi_j = \lambda_j \Phi_j, \tag{5}
\]

with

\[
\langle \Phi_j, \Phi_i \rangle = \| \Phi_i \|^2 \delta_{ij}. \tag{6}
\]

Since we have assumed that \( G_1 \) is invertible, it is also true that

\[
\lambda_j \neq 0. \tag{7}
\]

When (5) and (6) are satisfied, (1) and (4), as well as elementary properties of the inner product, give the following \( n \times n \) system, which defines Discretisation 1.

\[
\sum_{j=0}^{n-1} \left( \lambda_i \| \Phi_i \|^2 \delta_{ij} + \langle G_2 \Phi_j, \Phi_i \rangle \right) x_j = \langle Y, \Phi_i \rangle, \quad i = 0, 1, \ldots, n - 1 \quad \text{(Discretisation 1).} \tag{8}
\]

Similarly, (5), (6), (2) and (4) yield the \( n \times n \) system of Discretisation 2:

\[
\sum_{j=0}^{n-1} \left( \| \Phi_i \|^2 \delta_{ij} + \langle G_2^{-1} G_2 \Phi_j, \Phi_i \rangle \right) x_j = \langle G_2^{-1} Y, \Phi_i \rangle, \quad i = 0, 1, \ldots, n - 1 \quad \text{(Discretisation 2).} \tag{9}
\]

Evidently (8) and (9) are not, in general, equivalent discretisations. The two theorems that follow give conditions on \( G_1 \) that guarantee equivalence. Theorem proofs are provided in the Appendix.

**Theorem 1** Let \( G_1 \) be an invertible, symmetric linear operator possessing \( n \) eigenfunctions \( \Phi_0, \Phi_1, \ldots, \Phi_{n-1} \), corresponding to \( n \) distinct eigenvalues \( \lambda_j \). Then (5)–(7) hold for \( j = 0, 1, \ldots, n - 1 \). For the Galerkin method whose basis functions are the aforementioned \( \Phi_j \), Discretisation 2 is specified by the \( n \times n \) system.

\[
\sum_{j=0}^{n-1} \left( \lambda_i \| \Phi_i \|^2 \delta_{ij} + \frac{1}{\lambda_i} \langle G_2 \Phi_j, \Phi_i \rangle \right) x_j = \frac{1}{\lambda_i} \langle Y, \Phi_i \rangle, \quad i = 0, 1, \ldots, n - 1. \tag{10}
\]

Obviously, the \( i \)th equation in (10) is the corresponding equation in (8) divided by \( \lambda_i \). Accordingly, Discretisations 1 and 2 are row-scaled equivalent.

**Theorem 2** Let \( G_1 \) be an invertible, densely defined linear operator possessing a complete orthogonal set of distinct eigenfunctions \( \Phi_j \), corresponding to distinct eigenvalues \( \lambda_j \). Consider the Galerkin method whose basis functions are any \( n \) of the aforementioned \( \Phi_j \) say \( \Phi_0, \Phi_1, \ldots, \Phi_{n-1} \). Then Discretisation 1 is specified by the \( n \times n \) system.

\[
\sum_{j=0}^{n-1} \left( \lambda_i \| \Phi_i \|^2 \delta_{ij} + \lambda_i \langle G_2^{-1} G_2 \Phi_j, \Phi_i \rangle \right) x_j = \lambda_i \langle G_2^{-1} Y, \Phi_i \rangle, \quad i = 0, 1, \ldots, n - 1. \tag{11}
\]

Obviously, the \( i \)th equation in (11) is the corresponding equation in (9) multiplied by \( \lambda_i \). Accordingly, Discretisations 1 and 2 are row-scaled equivalent.

In both theorems, the matrix of Discretisation 1 equals that of Discretisation 2 left-multiplied by the diagonal matrix of eigenvalues of \( G_1 \), namely

\[
\Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{n-1}). \tag{12}
\]

By (7), \( \Lambda \) is non-singular.

In Theorem 1, the key assumption is that \( G_1 \) is symmetric; as a result, all eigenvalues of \( G_1 \) are necessarily real. In Theorem 2, the key assumptions are that the \( n \) basis functions are taken from a ‘large’ set (i.e. a set of infinite cardinality) of

\footnote{These assumptions equivalently can be stated as follows: (i) \( G_1 \) is densely defined; (ii) Equations (5)–(7) hold for \( j = 0, 1, 2, \ldots \), with \( \lambda_j \) and \( \Phi_j \) distinct; and (iii) \( 0, \ldots, n \)-natural.}
eigenfunctions; and that this set is capable of representing all Hilbert-space elements via convergent Fourier series. Therefore, the conditions in Theorems 1 and 2 differ in a significant manner. In particular, the assumptions of Theorem 2 do not imply real eigenvalues, nor a symmetric $G_1$.

The assumptions in Theorem 2 are in a sense milder, but it is often true that the symmetric assumption (of Theorem 1) is easier to establish. For example, all self-adjoint operators are symmetric [6]. If a particular $G_1$ has been shown to be symmetric, then we can establish row-scaled equivalence without separately proving that the eigenfunctions/basis functions are orthogonal: note that (6) is a conclusion in Theorem 1, not an assumption. In Theorem 2, however, eigenfunction orthogonality is not guaranteed by the other assumptions of the theorem and must be shown independently before one applies the theorem.

It is worthy to mention the case of integral equations with Hilbert-Schmidt kernels [12], most often encountered in an $L^2$ setting. In this case, it is immediately apparent whether the self-adjoint assumption is satisfied or not, as explained in [12]. See [13] for a simple example in electrostatics.

### 5 | CONSEQUENCES

We now demonstrate some practical consequences of our equivalence results:

(i) If one of the two discretisations has a known property, the equivalent discretisation automatically shares this property. Most notably, in the case where $G_1^{-1}G_2$ is compact, classical Fredholm theory tells us that Discretisation 2 converges (in the norm of the Hilbert space). It follows immediately that Discretisation 1 also converges, and that the limits of the two discretisations are identical.

(ii) Suppose, still, that $G_1^{-1}G_2$ is compact. Besides converging to the same limit, the two equivalent methods necessarily have identical convergence speeds. Note that the notion of convergence speed pertains only to discretisations that are finite (or to truncations of infinite discretisations), that is, to finite matrices.

(iii) Incidentally (because we know of no such case pertaining to the MAR), divergence of one of the two equivalent methods automatically implies divergence of the other. Divergences do arise in the context of certain popular integral equations of electromagnetics [13,14].

(iv) In a hypothetical computer with perfect hardware and software, Discretisations 1 and 2 will give exactly the same results. Therefore, neither of the two discretisations can be used to ‘validate’ the other—with the exception of errors such as round-off, quadrature errors, errors in computations of special functions, etc.

(v) *A priori* knowledge of equivalence may save us time and effort. The MAR starts with (1) so—as long as the orthogonal eigenfunctions $\Phi_j$ of $G_1$ are known—we can simply apply Galerkin’s method, with the $\Phi_j$ as basis functions, directly to (1). That is, we prefer Discretisation 1; there is no need to apply $G_1^{-1}$ and derive (2).

Let us stress that the application of $G_1^{-1}$ is an operator shorthand for a sequence of calculations that might be tedious. An example of this can be found in [3], which considers the simple-looking case where (1) is a one-dimensional, first-kind equation with a logarithmically singular kernel. Such equations arise frequently in electromagnetics (e.g. they govern the perfectly conducting flat or curved strip [1–3], the linear antenna [11,14], and certain classical equations of electrostatics [13]), and have been the subject of tutorial discussions (see the relevant works cited in [3]). The second-kind counterpart (2) of this equation—which is obtained by applying Carleman’s formula [15]—is a complicated operator equation containing a Cauchy principal-value integral. Although this operator equation is in fact an integral equation, it is barely recognizable as such, as one can see from (20) and (22) of [3].

### 6 | CONDITION NUMBERS

This section compares the condition numbers of the matrices in the two equivalent discretisations (under the conditions of Theorem 1, but similar conclusions can be formulated for case of Theorem 2). The condition number $\kappa$ depends on the matrix norm used, so let us first emphasize that the underlying matrix norm can be any $p$-norm, that is, any matrix norm induced by the $L_p$ vector norm [16]. Within any $p$-norm, it is always true [16, p. 382] that $\kappa \geq 1$. Also, the $\kappa(D)$ of any non-singular diagonal matrix $D = \text{diag}(d_0, d_1, \ldots, d_{n-1})$ is given by [16, pp. 358, 332]

$$
\kappa(D) = \kappa(D^{-1}) = \frac{\max_i |d_i|}{\min_j |d_j|} \geq 1.
$$

(13)

In (13), the equality $\kappa = 1$ occurs if and only if all magnitudes $|d_i|$ are equal, so that $D$ is unitary.

After these clarifications, we resume our study of equivalent discretisations. In the case $G_2 = 0$, the transformed ‘equation’ (2) is actually an explicit expression for the solution $X$ of the original equation (1). Consequently the MAR is intrinsically a perturbation method, with $G_2$ being the perturbing operator, and $G_1^{-1}Y$ being the unperturbed solution. The theorem that follows relates the condition numbers of the two equivalent methods subject to small perturbations. More precisely, it assumes that $G_2$ is a bounded linear operator whose norm is small.

---

This is especially true of the densely-defined assumption, which is ‘not serious’ [6, p. 486]. The proof of Theorem 2, which uses the adjoint operator $G_1^*$, utilizes the densely-defined assumption only to guarantee that $G_1^*$ exists.

This is one reason why we formulated Theorems 1 and 2 in terms of conditions on $G_1$ rather than conditions on $G_1^{-1}$—even if it is $G_1^{-1}$ that appears in (9). (The main reason, of course, is that (1) — and not (2) — is the equation we start out with.)
As is apparent from (4), the $\kappa$ of an $n \times n$ matrix associated with Galerkin's method depends on the normalisation of the basis functions, that is, on the choice of the norms $\|\Phi_j\|$. For a pre-specified normalisation, we denote by $\kappa_1$ and $\kappa_2$ the matrix condition numbers associated with Discretisations 1 and 2, respectively. Our theorem is as follows.

**Theorem 3** Let the conditions of Theorem 1 be satisfied. Assume, additionally, that $G_2$ is a bounded operator. Then for all sufficiently small $\|G_2\|$ the following two statements hold:

$$\|\Phi_j\| = 1 \ (j = 0, 1, \ldots, n - 1) \Rightarrow \kappa_2 \leq \kappa_1. \quad (14)$$

$$\|\Phi_j\| = \frac{1}{\sqrt{|\lambda_j|}} \ (j = 0, 1, \ldots, n - 1) \Rightarrow \kappa_1 \leq \kappa_2. \quad (15)$$

For small perturbations, therefore, neither of the two equivalent methods is superior (i.e. neither has a smaller condition number), because the relation between $\kappa_1$ and $\kappa_2$ depends on the normalisation of the basis functions—provided, of course, that the conditions of Theorem 3 hold. To the extent that it is natural to use orthonormal basis functions, (14) of Theorem 3 can be construed as a trivial advantage of Discretisation 2, in line with the usual rule that discretising second-kind equations is preferable.

For the unperturbed equations with $G_2 = 0$, the normalisations in (14) and (15) are simple and intuitive: The choice in (14) means that the matrix in Discretisation 2 is the identity matrix, see (10), which has the smallest possible condition number, so that $1 = \kappa_2 \leq \kappa_1$. By the same token, the choice in (15) means that the matrix of Discretisation 1 is the unitary diagonal matrix with elements $\lambda_j/|\lambda_j|$, see (8); in this case, (13) implies $\kappa_1 = 1$, so that $1 = \kappa_1 \leq \kappa_2$.

For the perturbed equations with nonzero $G_2$, the essence of the theorem (whose detailed proof is in the Appendix) is that the aforesaid condition-number inequalities are preserved when the perturbing operator $G_2$ is 'small'.

From a practical point of view, if we choose to apply Discretisation 1 (for instance, in accordance with (v) of Section 5), we can use (15) to normalise our basis functions. This amounts to using basis functions $\Phi_j'$, where

$$\Phi_j' = \frac{1}{\|\Phi_j\|\sqrt{|\lambda_j|}} \Phi_j, \quad j = 0, 1, \ldots, n - 1. \quad (16)$$

## 7 CONCLUSION

We demonstrated a straightforward equivalence between two seemingly different discretisations. One results directly from the first-kind equation (1), while the other from its second-kind counterpart (2). Both discretisations arise from the application of the Galerkin's method having orthogonal eigenfunctions of $G_1$ (the eigenfunctions are assumed to be known) as basis functions. Theorems 1 and 2 give conditions on $G_1$ that suffice to ensure equivalence. It is important that the equivalence pertains to finite discretisations, involving matrices of finite (but arbitrary) dimension $n$.

Our results can be viewed in a number of different ways. First, we provide an exception to the usual rule that discretising second-kind equations is preferable (but see the 'trivial advantage' arising from Theorem 3). Second, we better comprehend how a regularised discretisation scheme can arise directly from a first-kind Fredholm operator equation. Third, our theorems can help in theoretically justifying the application of such a scheme in a particular problem (and our normalisation (16) can eliminate the 'trivial advantage'). Fourth, we can better predict the actual behaviour of our scheme if implemented, especially with respect to finite-matrix condition numbers. We thus expect our results, which are theoretical, to be useful for a number of computational-electromagnetic problems.

## ACKNOWLEDGEMENT

We thank Nikolaos L. Tsitsas for comments that led to much improvement of this manuscript.

## CONFLICT OF INTEREST

There is no conflict of interest.

## ORCID

George Fikioris @ https://orcid.org/0000-0002-0987-8835

## REFERENCES

1. Nosich, A.I.: Method of analytical regularisation in computational photonics. Radio Sci. 51, 1421–1430 (2016)
2. Nosich, A.I.: The method of analytical regularisation in wave-scattering and eigenvalue problems: Foundations and review of solutions. IEEE Antenna. Propag. Mag. 41(3), 34–49 (1999)
3. Fikioris, G.: A note on the method of analytical regularisation. IEEE Antenna. Propag. Mag. 43(2), 34–40 (2001)
4. Lucido, M., Santomassimo, C., Panariello, G.: The method of analytical preconditioning in the analysis of the propagation in dielectric waveguides with wedges. J. Lightwave Techn. 36(14), 2925–2932 (2018)
5. Lucido, M., Panariello, G., Schettino, F.: Scattering by a zero-thickness PEC disk: a new analytically regularising procedure based on Helmholtz decomposition and Galerkin method. Radio Sci. 52(1), 2–14 (2017)
6. Naylor, A.W., Sell, G.R.: Linear Operator Theory in Engineering and Science. Springer, New York, NY (1982)
7. Dudley, D.: Mathematical Foundations for Electromagnetic Theory. IEEE Press, Piscataway, NJ (1994)
8. Harrington, R.F.: Field Computations by Moment Methods. IEEE Press, Piscataway, NJ (1993)
9. Kress, R.: Linear Integral Equations, 3rd ed. Springer, New York, NY (2014)
10. Jones, D.S.: Acoustic and Electromagnetic Waves. Oxford University Press, Oxford (1986)
11. Jones, D.S.: Note on the integral equation for a straight wire antenna. IEEE Proc. Pt. H. 128(2), 114–116 (1981)
12. Staßgolding, L., Holst, M.: Green’s Functions and Boundary-value Problems, 3rd ed. Wiley, New York, NY (2011)
13. Fikioris, G., et al.: Unphysical moment-method solutions of an approximate integral equation of electrostatics. IEEE Antenna. Propag. Mag. 142–153 (2017). https://doi.org/10.1109/MAP.2017.2686318
Since (9), replace the two inner products by twice applying the identity (A4) (for \( \Psi = G_2 \Phi \)) and \( \Psi = \text{Y} \). This yields the expression (10) for Discretisation 2, which is what we wanted to show.

**Proof of Theorem 2**

Since \( G_1 \) is densely defined, it possesses an adjoint \( G_1^* \) [6, pp. 527, 486] for which
\[
\langle G_1 \Phi, \Psi \rangle = \langle \Phi, G_1^* \Psi \rangle \quad (A5)
\]

for all \( \Phi \) and \( \Psi \). Let us show that the eigenvalues and eigenfunctions of the adjoint are found from
\[
G_1^* \Phi_j = \overline{\lambda_j} \Phi_j, \quad j = 0, 1, 2, \ldots \quad (A6)
\]

To prove (A6), we can show [6, p. 306] that \((G_3^* - \overline{\lambda_j})\Phi_j\) is orthogonal to all elements \( \Phi_i \) of our complete orthogonal Hilbert-space basis \( \{\Phi_0, \Phi_1, \ldots\} \). This, in turn, can be deduced from (5) and (6), and the complex conjugate of (A5) as follows:
\[
\langle (G_1^* - \overline{\lambda_j}), \Phi_i \rangle = \langle G_1^* \Phi_i, \Phi_i \rangle - \bar{\lambda_j} \langle \Phi_i, \Phi_i \rangle = \langle \Phi_i, \overline{\lambda_j} \Phi_i \rangle - \bar{\lambda_j} \langle \Phi_i, \Phi_i \rangle = 0.
\]

Now for any element \( \Psi \), we have
\[
\langle \Psi, \Phi_i \rangle = \langle G_1 G_1^{-1} \Psi, \Phi_i \rangle = \langle G_1^{-1} \Psi, G_1^* \Phi_i \rangle
\]
\[
= \langle G_1^{-1} \Psi, \overline{\lambda_j} \Phi_i \rangle = \lambda_j \langle G_1^{-1} \Psi, \Phi_i \rangle.
\]

We arrive at (11) by transforming the two inner products in (8) using the identity (A8) for \( \Psi = G_2 \Phi \) and for \( \Psi = \text{Y} \).

**Proof of Theorem 3**

In this proof, all matrices are assumed to be \( n \times n \) and non-singular without explicit mention. \( I \) is the identity matrix. \( \| \|_p \) and \( \kappa(\mathbf{E}) = \| \mathbf{E} \|_p \| \mathbf{E}^{-1} \|_p \) denote the \( p \)-norm and the associated condition number of the (non-singular) matrix \( \mathbf{E} \). It is true that
\[
1 \leq \kappa(\mathbf{E}) = \kappa(\mathbf{E}^{-1}) \quad (A9)
\]

and
\[
\kappa(\mathbf{FE}) \leq \kappa(\mathbf{F}) \kappa(\mathbf{E}), \quad (A10)
\]

see [16, pp. 382–383]. Any \( p \)-norm is a continuous function of the matrix elements \( E_{ij} \). Therefore, the condition number \( \kappa(\mathbf{E}) \) is continuous in the same sense.

Our proof of Theorem 3 makes use of the two Lemmas that follow.

**Lemma 1** If \( \kappa(\mathbf{F}) = 1 \), then \( \kappa(\mathbf{FE}) = \kappa(\mathbf{EF}) = \kappa(\mathbf{E}) \).

**Proof** Equation (A10) yields \( \kappa(\mathbf{FE}) \leq \kappa(\mathbf{E}) \). We have \( \kappa(\mathbf{F}^{-1}) = 1 \) by (A9), so \( \mathbf{E} = \mathbf{F}^{-1}(\mathbf{FE}) \) similarly yields \( \kappa(\mathbf{E}) \leq \kappa(\mathbf{FE}) \). It follows that \( \kappa(\mathbf{FE}) = \kappa(\mathbf{E}) \). The proof of \( \kappa(\mathbf{EF}) = \kappa(\mathbf{E}) \) is analogous. \( \square \)

**Lemma 2** Let \( \mathbf{F} \) be non-singular and let \( \kappa(\mathbf{F}) = 1 \). Then for any matrix \( \mathbf{E} \) whose norm \( \| \mathbf{E} \|_p \) is sufficiently small we have
\[ \kappa(F_1 + F^{-1}E) \leq \kappa(FF_1 + E). \quad (A11) \]

**Proof** Suppose first that \( \kappa(F) > 1 \). Lemma 1 gives
\[ \kappa(F_1) = 1 < \kappa(F) = \kappa(FF_1). \quad (A12) \]

When \( E = 0 \), therefore, (A11) holds as a strict inequality. By continuity of \( \kappa \), this strict inequality must continue to hold for sufficiently small \( \|E\|_p \), completing our proof for \( \kappa(F) > 1 \).

When \( \kappa(F) = 1 \), (A11) holds trivially (as an equality and for all \( E \)) by Lemma 1. \( \square \)

Armed with Lemma 2, we now prove our theorem.

**Proof of Theorem 3** When \( \|\Phi\| = 1 \), Theorem 1 and (12) tell us that \( \kappa_1 = \kappa(\Lambda + A) \) and \( \kappa_2 = \kappa(I + \Lambda^{-1}A) \), where \( A \) is the matrix with elements \( A_{ij} = \langle G_2 \Phi_j, \Phi_i \rangle \). Using elementary properties of the inner product and the Hilbert-space norm [6, pp. 273, 248], we see that
\[ |A_{ij}| \leq \|G_2\| \|\Phi_j\| \leq \|G_2\| \|\Phi_i\| = \|G_2\| \quad \to 0, \quad (A13) \]
so that \( \|A\|_p \to 0 \) by continuity of the \( p \)-norm. Application of Lemma 2 with \( F_1 = I, F = \Lambda \) and \( E = A \) gives the desired inequality \( \kappa_2 \leq \kappa_1 \).

Let us turn to the case \( \|\Phi\| = |\lambda_j|^{-1/2} \). Theorem 1 and (12) now tell us that \( \kappa_1 = \kappa(F_1 + \Lambda B) \) and \( \kappa_2 = \kappa(\Lambda^{-1}F_1 + B) \), where the elements of \( B \) are \( B_{ij} = |\lambda_i|^{-1} \langle G_2 \Phi_j, \Phi_i \rangle \) and where
\[ F_1 = \text{diag}(\lambda_0, 1, \ldots, 1) \quad (A14) \]
is unitary and diagonal. By (13) we have \( \kappa(F_1) = 1 \). Since
\[ |B_{ij}| \leq |\lambda_i|^{-1} \|G_2\| \|\Phi_j\| \|\Phi_i\| = \frac{\|G_2\|}{\sqrt{|\lambda_i|}} \to 0, \quad (A15) \]
we have \( \|B\|_p \to 0 \). Application of Lemma 2 with \( F = \Lambda^{-1} \) and \( E = B \) now gives \( \kappa_1 \leq \kappa_2 \). \( \square \)