# Appendix A

# Fredholm's Integral Equations of the First Kind

## A.1 Introduction

This appendix encompasses all the theoretical aspects of the Fredholm's integrals equations of the first kind which have not been included in Chapter 4, mainly devoted to the analytical study of spectral recovery and color mappings. Precisely, the main issue tackled here has to do with the development of a complete analytical and numerical solution to the color formation equation, which was shown to be expressed as

$$y\left(\mu\right) = \int_{\lambda_0}^{\lambda_1} S\left(\mu,\lambda\right)\, E\left(\lambda\right)\, R\left(\lambda\right)\, d\lambda\,,\ \ \mu \in [\mu_0, \mu_1] \tag{A.1}$$

The above expression is the *continuous color formation equation*, which describes in a broad way the formation of a multispectral signal and, in particular, of a color signal. By solving this equation it is possible that either $E(\lambda)$ or $R(\lambda)$ functions can be recovered if the sensor sensitivity $S(\mu, \lambda)$ and one of these two functions is known *a priori*. This is usually known as *spectral recovery* and has been extensively dealt in works such as [MW86, TW90, HFD90, MW92, DI93, FC84, Sap98, Sap99, TO00].

The novelty here is that the issue has been generalized in a way that a continuous solution is provided from where several existing approaches can be derived afterwards, depending on the constraints applied in each case, as it is done in Chapter 4. The importance of having a general expression is that the existence, uniqueness and numerical behaviour of the solution – or the lack of it – can be studied in a theoretical and general framework, leaving apart the limitations of each particularization. Furthermore, approximations can be easily derived by choosing some of the methods provided later, accordingly to the available data in each case.

The importance of such a formulation can be better understood if we compare this problem to the one encountered when a process is described by a differential equation. In order to solve a particular case using a set of measured data, it is likely to follow a numerical scheme based on some linearized approximation finally deriving to their difference equations. Nevertheless, if a deeper

knowledge on the process is to be drawn, it is a must to obtain the general differential equation which describes the whole family of cases, rather than just studying the numerical scheme. The same applies here.

Therefore, after some broad properties of general linear operators, this appendix describes the conditions to be fulfilled in order to obtain a proper solution to the recovery problem as well as to illustrate the basics for the analytical solution of the color equation in terms of a *Fredholm's Integral equation of the First Kind* (IFK). This equation is directly derived from Eq. (A.1), where the interior of the integral has been rewritten as the product of a *kernel* function $K$ and a *solution* function $f$, which is unknown and has to be computed from the *data* function $g$, which is measured in a series of wavelengths $\mu_i$ by the imaging sensors. Hence, we get that

$$g(\mu) = \int_{\lambda_0}^{\lambda_1} K(\mu, \lambda)\ f(\lambda)\ d\lambda\,,\ \ \mu \in [\mu_0, \mu_1] \tag{A.2}$$

The solution of this kind of integral equations is illustrated later in this Chapter and it is shown that the recovered function $f$ is expressed as a linear combination of a family of functions which straightforwardly depends on the kernel function $K$. The nature of the solution fully depends on how the spaces where functions $f$ and $g$ span are restricted. This notion is not new and was already formulated by Forsyth in [For90] for the case of color formation. Nevertheless, there it was only shown for a precise case, while here this result is not restricted to some discretization scheme, but states for a wider range of cases, both continuous and discrete.

The last section of the Chapter deals with the methods needed to be applied in order to obtain finite solutions of the former equation in terms of matrices and vectors. These methods are used in Chapter 4 to analyze the issue of finding an expression for the color mapping in the color constancy context. There, the discretization methods were applied to obtain some particular expressions relating our description to those previously obtained by other authors.

In brief, the aim of this Chapter is to provide a useful reference on the IFK issue to easily follow the theoretical exposition about color constancy in Chapter 4 and to be looked up when necessary.

## A.2 General Linear Operator

First of all, let us define some clarifying concepts involved in the theory of integral equations. For further details and proofs the reader should refer to the more extensive works in [Pip91, Kre89, Han00, Win91]. An *operator* $A$ is a mapping between spaces of functions, $X$ and $Y$, that applies a function $\varphi \in X$ onto another function $\psi \in Y$ as follows

$$\begin{aligned} A\colon X &\longrightarrow Y \\ \varphi &\longmapsto \psi = A\,\varphi \end{aligned} \tag{A.3}$$

We can also define an *inverse* operator $A^{-1}$ where

$$\begin{aligned} A^{-1}\colon A(X) &\longrightarrow X \\ \psi &\longmapsto \varphi = A^{-1}\,\psi \end{aligned} \tag{A.4}$$

Note that $A(X) \subseteq Y$, i.e., the image set of $X$ by the mapping $A$ lives within $Y$, but they may not be exactly the same. This way, $A\,A^{-1} = A^{-1}\,A = Id$, which is the *identity* operator. The *existence* and *uniqueness* of a solution for an equation using operators, $A\,\varphi = \psi$, can be equivalently expressed as the *existence* of an *inverse* operator such that $A^{-1}\colon Y \to X$. It is also said that the operator $A$ is *invertible*. Then, the solution can be expressed as $A^{-1}\,g = f$. Notice the subtlety between the solution to the former equation and the inverse operator. In order this operator to be a solution, it is necessary that at least $A(X) = Y$, and not only that $A(X) \subset Y$.

That is, the inverse operator $A^{-1}$ exists if and only if the operator $A$ is *bijective*, i.e.,

- $\forall \psi \in A(X) \subset Y,\ \exists!\, \varphi \in X \mid A\,\varphi = \psi$ (*injectiveness*).
- $A(X) = Y$ (*exhaustiveness*).

The set of functions $A(X) = \{A\,\varphi\colon \varphi \in X\}$ is usually known as the *rank* of the operator $A$. It is important to be aware that from the mere existence of a certain operator $A^{-1}\colon A(X) \to X$ does not follow the existence of the inverse operator $A^{-1}\colon Y \to X$, which provides us with a unique solution to the equation $A\,\varphi = \psi$. Further considerations must be taken into account.

We restrict our study to a very special kind of operators which are known as *linear operators*. An operator $A$ is lineal if it is true that

$$A\,(\alpha\,\varphi_1 + \beta\,\varphi_2) = \alpha\,A\,\varphi_1 + \beta\,A\,\varphi_2 \tag{A.5}$$

with $\alpha, \beta \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in X$, being $X$ and $Y$ two vector spaces. Besides, this sort of operators can have other relevant properties which assure the inversion of the operator such as those of *boundness* and *compactness*.

A space, in general, has a norm if we can define a *norm* for all its elements. For example, over the space of real functions we can define a quadratic norm such as $\|\varphi\|_2^2 = \int_a^b |\varphi(x)|^2\,dx$. Then, a linear operator $A$ between two normed spaces $X$ and $Y$ is *bounded* if $\exists\, C > 0$ so that $\|A\varphi\| \leq C\|\varphi\|$, $\forall \varphi \in X$. That is, an bounded operator over a function which is bounded will always produce another bounded function.

Another key property of linear operators is that of *compactness*. Basically, a set is *compact* if it is neither *open* nor *unbounded*, i.e., it is closed and bounded. Analogously, a linear operator $A\colon X \to Y$, where both $X$ and $Y$ are normed spaces, will be said to be *compact* if $\forall Z \subset X$ *bounded*, the closure of $A(Z)$, $\overline{A(Z)} \subset Y$, is a compact set. In other words, the closure of the image of a bounded subset is also bounded.

At this point, we can set out that for any linear operator $A\colon X \to Y$ being *invertible* is the same as having a *bounded inverse operator* $A^{-1}$. Moreover, any *compact* linear operator $A$ has no bounded inverse unless the rank $A(X)$ has a finite dimension.

On the other hand, in any practical problem it is important to know how well–conditioned a solution is in order to find it by numerical methods. Classically, the well–condition of a problem has been defined as the *existence* and *unicity* of the solution besides a *continuous* dependence of the solution to the data. If any of these properties is not fulfilled the problem is ill–conditioned and the solution will show many numerical difficulties.

Existence and unicity is fulfilled whenever an operator $A$ is bijective. As seen before, this implies the existence of the inverse operator $A^{-1}$. Meanwhile, the third condition is satisfied if and only if the inverse operator $A^{-1}$ is continuous. Therefore, it can be said that the solution of an equation such as $A\,\varphi = \psi$ is well–conditioned. However, in general, for any *compact linear operator* $A\colon X \to Y$, where $X$ and $Y$ are normed space, the solution of the equation $A\,\varphi = \psi$ is *ill–conditioned* unless the space X has finite dimension.

Summarizing, for any compact linear operator $A\colon X \to Y$ the solution of the equation $A\,\varphi = \psi$ exists, is unique and is well–conditioned if and only if both the *solution space* $X$ and the *data space* $A(X) = Y$ have finite dimension. Otherwise, if the condition on $A(X) = Y$ is not fulfilled, which happens whenever the dimension of the two spaces is different, there might be no solution at all. If the data space have infinite dimension, the solution is not unique. Finally, if it is the solution space which has infinite dimension, the solution to the equation will not be stable.

This result is a pretty basic but important result since specifically states why solution and data spaces should be of finite dimension. It must be noticed that it is not a consequence of any lack of data of the measure process which could be solved in a finer version of the equation, but rather it is a consequence of the nature of the equations. This property should be taken into account in any recovery approach or color constancy algorithm.

## A.3 Integral Operators

Let us now restrict our study to the special kind of operators called *integral operators*. An integral operator $K$ is of the form

$$(K\circ)(x) = \int_a^b K(x,y) \circ \, dy \,,\; x \in [a,b] \tag{A.6}$$

where function $K(x,y)$ is usually called *kernel*. This operator is helpful because a first kind Fredholm integral equation can be written as follows

$$g(x) = K\,f(x) = \int_a^b K(x,y)\, f(y)\, dy \tag{A.7}$$

or even simpler

$$g = Kf \tag{A.8}$$

where function $g$ is the *data function* and $f$, the *solution function*.

Because of the linearity property of integrals, $K$ is a *linear operator*. Moreover, if the kernel function $K(x,y)$ is well–behaved (continuous, smooth, etc.) and quadratic bounded[1], then it is a *compact operator*. So following the principal result obtained before, in general, we conclude that the *IFK equations are ill–conditioned problems* unless constraints on the function spaces exist or be imposed.

The only way to guarantee a well–conditioned solution to Eq. (A.8) is to *restrict* the spaces where the solution functions $f$ and the data functions $g$ live to have a finite dimension. As this is not always possible because of the nature

[1] $\int_a^b \int_a^b K^2(x,y)\, dx\, dy < \infty$

of these spaces in practice, we could face an ill–conditioned problem when trying to solve the corresponding IFK equation.

In fact, the straightforward application of usual numerical methods to the matrices obtained by discretization of these sort of equations give rise to numerical answers which may lack of sense due to errors. As a matter of fact, a paradoxical behaviour appears: the finer the discretization, the worse the solution becomes [Win91, Han00]. This should have been considered in works such as those of Sapiro [Sap99], Ho et al. [HFD90], Chang and Hsieh [CH95], Tominaga [TW89, TW90], though no analysis of this nature is done in those papers.

## A.4 Analytical Solution of an IFK

In the previous sections we basically showed one important thing, that is not always feasible to find a well–behaved solution to an IFK equation, in general. This does not mean that the problem is completely unsolvable. We show in this section the basics about how to analytically find a solution to an IFK.

First, we find a solution in the simplest case, that is, when the kernel is a symmetrical function. This case is then used in the general case, i.e., for nonsymmetrical kernels by means of turning any kernel into a symmetrical one. Finally, we get the general expression for a solution of any IFK equation.

The solution functions of an IFK will be written in terms of a set of functions derived from the kernel. This way, it will be completely clear the relation between the solution and the kernel.

### A.4.1 Symmetrical Kernels

In the case where kernel $K(x, y)$ is a symmetrical function, i.e., $K(x, y) = K(y, x)$, the operator $K$ has some especially attractive properties such that of having *eigenfunctions* and *eigenvalues*. These functions can be thought as an analogy to the eigenvectors of matrices as Eq. (A.9) claims.

In the next lines we consider the properties of eigenfunctions and eigenvalues of symmetrical kernels belonging to an IFK. Furthermore, we obtain representations of kernel $K$ in terms of a series of eigenfunctions. We first consider the equation for the eigenfunctions

$$K\,\phi = \lambda\,\phi \tag{A.9}$$

The following facts are known in case $K(x, y)$ is symmetrical, well-behaved, and the quadratic bounded condition holds [Win91]

- Eq. (A.9) has at least one solution $\phi \neq 0$ and this corresponds to a value of $\lambda \neq 0$. The function $\phi$ and the constant $\lambda$ are an eigenfunction and its corresponding eigenvalue of $K$.
- If eigenfunctions $\phi_1$ and $\phi_2$ belong to different eigenvalues $\lambda_1$ and $\lambda_2$, then $\phi_1$ and $\phi_2$ are *orthogonal*, that is, $(\phi_1, \phi_2) = 0$.
- Two different eigenfunctions may belong to the same eigenvalue. However, a particular nonzero eigenvalue can be only associated with a finite number of linearly independent eigenfunctions, which may be orthonormalized.

- Eq. (A.9) has only a finite number of nonzero eigenvalues if and only if $K(x,y)$ is *separable*[2].
- All eigenvalues are *real* and all eigenfunctions may be chosen to be *real*.
- If there are infinity many nonzero eigenvalues $\lambda_i$, then $\lim_{i\to\infty}\lambda_i = 0$, having ordered the eigenvalues so that $|\lambda_i| \geq |\lambda_{i+1}|$ previously.

The existence of eigenfunctions immediately suggests that Eq. (A.9) may be used to expand a function in terms of these eigenfunctions. If for some $f$ we can write $g = Kf$, then we can expand $g$ in terms of eigenfunctions $\phi_i$

$$g = \sum_{i=1}^{\infty}(g,\phi_i)\,\phi_i, \quad \text{where} \quad (g,\phi_i) = \int_a^b g(x)\,\phi_i(x)\,dx \tag{A.10}$$

The kernel $K$ can also be represented, considered as a function of the variable $y$, in a series of eigenfunctions as Eq. (A.10)

$$K(x,y) = \sum_{i=1}^{\infty}(K(x,y),\phi_i(y))\,\phi_i(y) \tag{A.11}$$

$$(K(x,y),\phi_i(y)) = \int_a^b K(x,y)\,\phi_i(y)\,dy = (K\phi_i)(x) = \lambda_i\,\phi_i(x) \tag{A.12}$$

So that

$$K(x,y) = \sum_{i=1}^{\infty}\lambda_i\,\phi_i(x)\,\phi_i(y) \tag{A.13}$$

i.e., there is an identification between the kernel $K$ and its eigenfunctions and eigenvalues.

### A.4.2 Nonsymmetrical Kernels

When removing the assumption of symmetry, relatively little of the preceding can be saved. First, an integral operator may have no eigenvalues at all. Second, it is not always possible to assure that nonzero eigenvalues exist. This is a problem because matrices most of the times have some nonzero eigenvalues. Therefore, as matrices arise in numerical computations, such manipulations may introduce entirely spurious eigenvalues. More problems, despite nonsymmetrical operators may have eigenvalues and eigenfunctions, they may be *complex*. Orthogonality results disappear, and so do expansion theorems, although it is still true that $lim_{i\to\infty}\lambda_i = 0$.

Let us depict the proper solution for a nonsymmetrical kernel turning it into a symmetrical kernel. Because function $g$ is known, we may apply to it any reasonable operator $T$ that we choose and obtain an also known function $h$

$$g = K\,f \Rightarrow T\,g = TK\,f = h \tag{A.14}$$

Thus, the IFK equation becomes $h = (TK)\,f$, where $TK$ is the new operator, $h$ is the data function and $f$ is the unknown. A useful operator to solve the problem is the *transposed operator* $T = K^*$ defined as follows

$$(K^*\circ)\,(y) = \int_a^b K(x,y)\circ dx, \quad y \in [a,b] \tag{A.15}$$

[2] $K$ is separable when $K(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_i(x)\cdot\beta_j(y)$ and $n,m < \infty$.

It is similar to $K$, but the role of the variable $x$ has been interchanged with the variable $y$. Observe that if $K$ is symmetrical, $K = K^*$. Hence, the equation turns into $h = K^*K\,f$, where the kernel $K^*K$ looks like

$$(K^*K\circ)(z) = \int_a^b \left[\int_a^b K(x,z)\,K(x,y)dy\right]\circ dx, \quad z\in[a,b] \tag{A.16}$$

This is a symmetrical operator and has all the properties listed above. Additionally, the nonzero eigenvalues of $K^*K$ are not only *real*, but also *positive*.

It is also interesting to define the analogous operator $KK^*$. It has to be noticed that $KK^* \neq K^*K$, in general. However, $KK^*$ is also symmetrical and its eigenvalues are nonnegative. An important property is that the eigenvalues are exactly the same as those of $K^*K$. Nonetheless, the eigenfunctions of the two operators differ in general.

For the positive eigenvalues $\mu_i$ of the integral operator $K^*K$, we define the set of *singular values* $\sigma_i = +\sqrt{\mu_i}$. If $v_i$ are the eigenfunctions of the operator $K^*K$, thus

$$K^*Kv_i = \sigma_i^2\,v_i \tag{A.17}$$

Now we define the next function

$$u_i = \frac{1}{\sigma_i}\,Kv_i \tag{A.18}$$

Applying the transposed operator $K^*$ to $u_i$ we get that

$$K^*u_i = \frac{1}{\sigma_i}\,K^*Kv_i = \sigma_i\,v_i \tag{A.19}$$

Notice that Eq. (A.18) and Eq. (A.19) are dual relations. If we apply operator $K$ to Eq. (A.19), we obtain

$$KK^*u_i = \sigma_i\,Kv_i = \sigma_i^2\,u_i \tag{A.20}$$

Thus, functions $u_i$ defined by Eq. (A.19) are eigenfunctions of the operator $KK^*$. Moreover, it may be shown that Eq. (A.19) generates all such eigenfunctions. On the other hand, we may start with the eigenfunctions $u_i$ of $KK^*$ and reverse all the steps of the argument to show that the functions $v_i$ defined by Eq. (A.19) are the eigenfunctions of $K^*K$.

Summarizing to this point, we have associated with the original operator $K$ two sets of functions, $v_i$ and $u_i$, which are, respectively, eigenfunctions of $K^*K$ and $KK^*$. The following expressions encompass their properties

$$\begin{array}{rcl|rcl} K^*u_i &=& \sigma_i\,v_i & KK^*u_i &=& \sigma_i^2\,u_i \\ Kv_i &=& \sigma_i\,u_i & K^*Kv_i &=& \sigma_i^2\,v_i \end{array} \tag{A.21}$$

Functions $u_i$ and $v_i$ are called *singular functions* belonging to kernel $K$, and values $\sigma_i$ are referred to as *singular values*. These functions are *orthogonal* and we can assume them to be also *normalized*. A special case to take into account are functions $v$ such that $K^*K\,v = 0$. We include these among the singular functions and agree that the corresponding singular value is *zero*. The same for those that satisfies the expression $KK^*\,u = 0$.

The following equation is also true and relates both operators by means of the scalar product of two functions

$$(Kf_1, f_2) = (f_1, K^* f_2) \tag{A.22}$$

Let us now consider a function $g$ given by $g = Kf$ for some $f$. Suppose that $g$ can be expanded into a series of functions $u_i$ and some of the previous properties are applied

$$\begin{aligned} g = Kf \Leftrightarrow g &= \sum_{i=1}^{\infty} (g, u_i)\, u_i \\ g &= \sum_{i=1}^{\infty} (Kf, u_i)\, u_i = \sum_{i=1}^{\infty} (f, K^* u_i)\, u_i \\ g &= \sum_{i=1}^{\infty} (f, \sigma_i\, v_i)\, u_i = \sum_{i=1}^{\infty} \sigma_i\, (f, v_i)\, u_i \end{aligned} \tag{A.23}$$

So that

$$g = Kf \Leftrightarrow g = \sum_{i=1}^{\infty} \sigma_i\, (f, v_i)\, u_i \tag{A.24}$$

This expression is somewhat suspect because we simply assumed a series expansion. Actually, it may be shown that it is quite valid as long as function $g$ can be written as $g = Kf$ [Kre89, Pip91]. As said in Section A.2, function $g$ can not be any possible function but must belong to the space spanned by $u_i$ functions. Otherwise, the equation $g = Kf$ does not hold. Furthermore, it should be notice that Eq. (A.24) is a generalization of the expansion in Eq. (A.10) found in the case of symmetrical kernels.

This result suggests a possible expansion for an arbitrary kernel $K(x, y)$ as a series of singular functions $u_i$ and $v_i$

$$K(x, y) = \sum_{i=1}^{\infty} (K(x, y), v_i(y))\, v_i(y) \tag{A.25}$$

where

$$(K(x, y), v_i(y)) = \int_a^b K(x, y)\, v_i(y)\, dy = (Kv_i)(x) = \sigma_i\, u_i(x) \tag{A.26}$$

Thus, we obtain the next expansion for the kernel $K$

$$K(x, y) = \sum_{i=1}^{\infty} \sigma_i\, u_i(x)\, v_i(y) \tag{A.27}$$

i.e., the structure of the kernel $K$ is defined by the singular functions and the singular values, and *vice versa*. Both data and solution functions belong to the spaces spanned by the singular functions, which get determined by the kernel, respectively. The contrary is also true. Therefore, both sides of the equation are constrained in order to obtain a solution, which only favors the *existence* of one solution.

### A.4.3 IFK Solution with General Kernels

In the previous section, function $f$ was tacitly supposed to be known in order to write function $g$ as a series. Now we want to find an expression for that function $f$ and, consequently, function $g$ must be known *a priori*. Thus, it is clear that if there exists a solution $f$ so that $g = Kf$, then the function $g$ must be expressed in terms of a series as follows

$$g = \sum_{i=1}^{\infty} \sigma_i \, (f, v_i) \, u_i \tag{A.28}$$

If we also expand the solution function $f$ in terms of the singular functions $v_i$ and, after that, the operator $K$ is applied to obtain the function $g$, we get that

$$f = \sum_{i=1}^{\infty} f_i \, v_i \Rightarrow Kf = \sum_{i=1}^{\infty} f_i \, Kv_i = g \tag{A.29}$$

Using some of the properties in Eq. (A.23) and the expansion of function $g$ in Eq. (A.10), the resulting expression is

$$Kf = \sum_{i=1}^{\infty} f_i \, \sigma_i \, u_i = \sum_{i=1}^{\infty} (g, u_i) \; u_i = g \tag{A.30}$$

Identifying the coefficients of these two series it is clear that

$$g_i = (g, u_i) = f_i \, \sigma_i \Rightarrow f_i = \frac{(g, u_i)}{\sigma_i} \tag{A.31}$$

Hence, formally a solution for the IFK equation is the following expression

$$f = \sum_{i=1}^{\infty} \frac{(g, u_i)}{\sigma_i} \, v_i \tag{A.32}$$

Here, it is self–evident the space to where $f$ belongs is not arbitrary, rather it is spanned by $v_i$ functions. So, in order an IFK to have a solution both functions $g$ and $f$ can not be arbitrary and the kernel $K$ determines the spaces where they live.

Besides, it is at once clear where the problems of IFKs come from. The terms $f_i = \frac{(g,u_i)}{\sigma_i}$ seem to diverge if there are infinitely many singular values $\sigma_i$ that go to zero. However, the theorem of Piccard gives the least conditions to make that this series converges [Kre89, Pip91], that is, the kernel $K$ must be squared integrable, which implies that the coefficients tend to zero, i.e., $\lim_{i\to\infty} f_i = 0$.

We see that the integration with $K$ has a smoothing effect: the higher the components in $f$, the more they are damped in $g$ due to the multiplication by $\sigma_i$. Moreover, Eq. (A.32) shows that the *inverse problem*, that of computing $f$ from $g$, indeed has the opposite effect on the oscillations in $g$, namely, an amplification of $g$'s components $(g, u_i)$ by a factor $\sigma_i^{-1}$.

This, of course, amplifies the high–frequency components. The trouble arises when the right–hand side $g$ is contaminated by errors, for then we will most likely divide small but nonnegligible coefficients $(g, u_i)$ with very small singular values $\sigma_i$, which will lead to a dramatic high–frequency perturbation of the computed solution.

It is important to notice that Eq. (A.32) is just one solution to the IFK. Because of the lack of uniqueness of the solution (infinite dimension of the data space), for any function $h$ such that $Kh = 0$ and being $f$ a solution to the IFK $g = Kf$, the new function $f + h$ is also a solution to the IFK, i.e., $h \in Null(K)$.

## A.5 Discretization of an IFK

The aim of discretizing integral equations lies in turning them into systems of linear equations which can be numerically solved to get approximate solutions. Nevertheless, there is *no exclusive way* of discretizing an IFK and may depend on the sort of information at hand when facing a particular problem. Moreover, some of them are definitely numerically better than others.

### A.5.1 Quadrature Rule

This is the most basic and straightforward approach widely used although it *definitively* does not provide the best solution. Basically, it consists in applying a numerical integration method to the integral equation, that is,

$$\int_0^1 \phi(x)\, dx \approx \sum_{j=1}^{m} \omega_j\, \phi(x_j) \tag{A.33}$$

where $x_j$ are the abscises of the particular quadrature rule and their corresponding weights $\omega_j$.

Using any quadrature rule, it is possible to numerically approximate the integral of the IFK in the following way

$$g(x) = \int_a^b K(x,y)\, f(y)\, dy \approx \sum_{j=1}^{m} \omega_j\, K(x,y_j)\, f(y_j) \tag{A.34}$$

Then, we take the set of *measured points* of the data function $g$, $g(x_i)$, get the next system of equations

$$g(x_i) = \sum_{j=1}^{m} \omega_j\, K(x_i,y_j)\, f(y_j)\,,\ i = 1,\ldots,n \tag{A.35}$$

It is immediately apparent that we now have a matrix equation

$$\mathbf{g}^t = \mathbf{K}\,\mathbf{f}^t$$

where

$$\begin{array}{rcl} (\mathbf{K})_{ij} & = & \omega_j\, K(x_i,y_j) \\ \mathbf{g} & = & (g(x_1),\ldots,g(x_n)) \\ \mathbf{f} & = & (f(y_1),\ldots,f(y_m)) \end{array} \tag{A.36}$$

Then, if the matrix $\mathbf{K}$ has an inverse, we can at once obtain the solution $\mathbf{f}$ as

$$\mathbf{f}^t = \mathbf{K}^{-1}\,\mathbf{g}^t \tag{A.37}$$

In any case, the solution computed in that way is not completely stable in front of perturbations of function $g$, that is, errors that may exist in the

data values. Therefore, if we solved the system by standard numerical analysis techniques the obtained values $f_j$ would be merely some coarse or wrong approximation to the true values $f(y_j)$ [Win91, Han00].

A way to predict how good or bad is the behaviour of the matrix $\mathbf{K}$ against perturbations is to study the *condition number* of the matrix. This index is defined in metric $\|\circ\|_2$ as

$$cond(\mathbf{K}) = \|\mathbf{K}\|_2 \, \|\mathbf{K}^{-1}\|_2 = \frac{\sigma_1}{\sigma_n} \tag{A.38}$$

where $\sigma_1$ and $\sigma_n$ are the largest and the smallest singular values of $\mathbf{K}$, respectively. A big condition number means that the matrix is ill–conditioned. If it is infinite, the matrix is said to be *singular*.

In the case the matrix comes from a discretized IFK, we have that the singular values of $\mathbf{K}$ decay gradually until they stabilize at about $\sigma_1$ times the machine precision, which can be considered as zero [Han00]. Consequently, the condition number $cond(\mathbf{K})$ is approximately the inverse of the machine precision, that is, infinite in practice.

In short, although this is the most straightforward method to approximate the solutions of IFK equations, *it is only useful if a coarse solution is needed*. Furthermore, there is no improvement by changing the quadrature rule because the algorithm has a *natural limit* beyond which it provides unsatisfactory information. Unfortunately, the limit is very problem–dependent. This class of limitations should be considered in any approach involving the inversion of integral operators, such as spectral recovery.

### A.5.2 General Series Expansion

The concept of expanding a function into a series is a pretty common method in mathematics to express a particular element as being part of a space. As explained in Section A.2, the ability at solving an IFK consists in determining a good set of basis functions where to expand the data and the solution functions. Otherwise, it will be impossible to find even an approximation. So any knowledge about how these functions must be, accordingly to any physical fact, should be embedded in the spaces where they live.

#### A Family of Orthonormalized Arbitrary Basis Functions

If we choose an arbitrary set of *orthonormal* basis functions, $\{\psi_i\}_{i=1,\ldots,n}$, defined in the interval of integration $[a, b]$, it is possible to write both functions $g$ and $f$ in the following way

$$g(x) = \sum_{i=1}^{n} g_i \, \psi_i(x) \qquad \text{and} \qquad f(y) = \sum_{j=1}^{n} f_j \, \psi_j(y) \tag{A.39}$$

so that the IFK equation $g = Kf$ formally becomes

$$g(x) = \sum_{i=1}^{n} g_i \, \psi_i(x) = \sum_{j=1}^{n} f_j \int_a^b K(x, y) \, \psi_j(y) \, dy \tag{A.40}$$

Using the orthogonality of the functions, we get that

$$g_i = (g, \psi_i) = \sum_{j=1}^{n} f_j \int_a^b \int_a^b K(x,y)\, \psi_i(x)\, \psi_j(y)\, dxdy \tag{A.41}$$

In matrix notation

$$\mathbf{g}^t = \mathbf{K}\, \mathbf{f}^t$$

where

$$\begin{array}{rcl} (\mathbf{K})_{ij} & = & \int_a^b \int_a^b K(x,y)\, \psi_i(x)\, \psi_j(y)\, dxdy \\ \mathbf{g} & = & (g_1, \ldots, g_n) \\ \mathbf{f} & = & (f_1, \ldots, f_n) \end{array} \tag{A.42}$$

**Two Families of Orthonormalized Basis Functions**

In fact, there is no reason for functions $g$ and $f$ to be expanded by the same set of basis functions. Function $g$ should be expanded in a space which reflects the constraints that helps the IFK to have a solution. Furthermore, any known feature of $f$ should also be contained in its expansion. Usually properties of the basis functions are those of being nonnegative, continuous, smooth, nonsingular, and bounded. Finally, each expansion can have a different number of elements.

By incorporating all these ideas, we can now use two different families of *orthonormal* basis functions, $\{\psi_i\}_{i=1,\ldots,n}$ and $\{\theta_j\}_{j=1,\ldots,m}$, to expand functions $g$ and $f$, respectively

$$g(x) = \sum_{i=1}^{n} g_i\, \psi_i(x) \qquad \text{and} \qquad f(y) = \sum_{j=1}^{m} f_j\, \theta_j(y) \tag{A.43}$$

Thus, the IFK equation is discretized in the following way

$$g(x) = \sum_{i=1}^{n} g_i\, \psi_i(x) = \sum_{j=1}^{m} f_j \int_a^b K(x,y)\, \theta_j(y)\, dy \tag{A.44}$$

so that

$$g_i = (g, \psi_i) = \sum_{j=1}^{m} f_j \int_a^b \int_a^b K(x,y)\, \psi_i(x)\, \theta_j(y)\, dxdy \tag{A.45}$$

The resulting system is no longer squared and the least squares method must be used to compute the inverse matrix. In matrix formulation, the system is

$$\mathbf{g}^t = \mathbf{K}\, \mathbf{f}^t$$

where

$$\begin{array}{rcl} (\mathbf{K})_{ij} & = & \int_a^b \int_a^b K(x,y)\, \psi_i(x)\, \theta_j(y)\, dxdy \\ \mathbf{g} & = & (g_1, \ldots, g_n) \\ \mathbf{f} & = & (f_1, \ldots, f_m) \end{array} \tag{A.46}$$

It is important to retain that it is possible to prove by using the condition number argument that the matrix $\mathbf{K}^t\, \mathbf{K}$ that comes up in the least squares method becomes more and more ill–behaved as additional terms in the series expansions are used [Win91, Han00]. So, in general, the fewer, the better.

### Galerkin Method

This method is based on the previous one but the orthogonality condition has been relaxed and it is only necessary that the basis functions used in the series expansions be linearly independent, allowing the employment of such functions as B–Splines. No longer can the $(g, \psi_i)$ be thought of as Fourier coefficients, rather they are simply numbers that can be calculated. The equation obtained is therefore

$$(g, \psi_i) = \sum_{j=1}^{m} f_j \int_a^b \int_a^b K(x, y)\, \psi_i(x)\, \theta_j(y)\, dxdy \tag{A.47}$$

This equation gives rise to the following linear system

$$\mathbf{g}^t = \mathbf{K}\, \mathbf{f}^t$$

where

$$\begin{array}{rcl} (\mathbf{K})_{ij} & = & \int_a^b \int_a^b K(x, y)\, \psi_i(x)\, \theta_j(y)\, dxdy \\ \mathbf{g} & = & ((g, \psi_1), \ldots, (g, \psi_n)) \\ \mathbf{f} & = & (f_1, \ldots, f_m) \end{array} \tag{A.48}$$

### Collocation Method

If the functions spanning the data function space are *relaxed* to be Dirac $\delta$–functions, the preceding approach turns into the next scheme

$$\psi_i(x) = \delta(x - x_i) \Longrightarrow (g, \psi_i) = g(x_i) \tag{A.49}$$

And the matrix $\mathbf{K}$ can be calculated with a single integration which often is a great saving in computing time

$$\mathbf{g}^t = \mathbf{K}\, \mathbf{f}^t$$

where

$$\begin{array}{rcl} (\mathbf{K})_{ij} & = & \int_a^b K(x_i, y)\, \theta_j(y)\, dy \\ \mathbf{g} & = & (g(x_1), \ldots, g(x_n)) \\ \mathbf{f} & = & (f_1, \ldots, f_m) \end{array} \tag{A.50}$$

## A.5.3 Series of Singular Functions

In the preceding Section basis functions used to expand $f$ and $g$ were chosen primarily based on the behaviour of the data and the expected behaviour of the solution, and the properties of the integral operator $K$ played a secondary role. We also made obvious in Section A.4.2 that the operator $K$ completely determined its singular functions and singular values and was, in turn, completely determined by them. Thus, it seems reasonable to use those functions for the series expansion, if available.

In Section A.4.3 we obtained the following expansion for a solution $f$

$$f(x) = \sum_{k=1}^{\infty} \frac{(g, u_k)}{\sigma_k}\, v_k(x) \tag{A.51}$$

This expression has several practical difficulties in order to compute a solution of an IFK. First, as said, because of the decreasing of the singular values

$\sigma_k$, the coefficients of this expansion are a likely source of error. Nevertheless, the useful information is carried by the first components of the series, which correspond to the biggest singular values. The rest of them only adds noise to the solution. Hence, *truncation* of the series is a very helpful method to remove these errors out. Secondly, the series is an infinite summation and must be finitely approximated. Besides, the integral operator $K$ must be first replaced by a matrix operator $\mathbf{K}$, as shown in previous Sections.

This way, both the matrix operator and its inverse operator can be written in terms of the standard SVD as

$$\begin{array}{rcccl} \mathbf{K} & = & \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^t & = & \sum_{k=1}^{n} \mathbf{u}_k^t\,\sigma_k\,\mathbf{v}_k \\ \mathbf{K}^{-1} & = & \mathbf{V}\,\mathbf{\Sigma}^{-1}\,\mathbf{U}^t & = & \sum_{k=1}^{n} \mathbf{v}_k^t\,\frac{1}{\sigma_k}\,\mathbf{u}_k \end{array} \tag{A.52}$$

where $\mathbf{v}_k$, $\mathbf{u}_k$ are the singular vectors and $\sigma_k$ the singular values of matrix[3] $\mathbf{K}$.

Just let us remind that $\mathbf{u}_k$ is the $k^{th}$ eigenvector of matrix $\mathbf{K}\,\mathbf{K}^t$, and $\mathbf{v}_k$ that of matrix $\mathbf{K}^t\,\mathbf{K}$ [PFTV93]. Both matrices have the same set of eigenvalues, i.e., $\sigma_k^2$. Moreover, vector $\mathbf{v}_k$ corresponds to the $k^{th}$ column of matrix $\mathbf{V}$ and vector $\mathbf{u}_k$ to that of matrix $\mathbf{U}$. Matrix $\mathbf{\Sigma}$ is defined as the diagonal matrix $diag(\sigma_k)$.

An observation should be done at this point. Vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the singular vectors of the matrix $\mathbf{K}$, which is a discretization of the kernel $K$. *They are not, in general, any discretization of the singular functions of kernel $K$, i.e., $u_i$ and $v_i$.* So, we can not directly go from Eq. ($A$.32) to Eq. ($A$.55), but rather first we must find the matrix $\mathbf{K}$ by one of the aforementioned methods.

Putting together both *truncation* and *discretization* allows us to rewrite the solution in terms of singular vectors and values

$$\mathbf{f}^t = \sum_{k=1}^{N} \frac{\mathbf{u}_k \cdot \mathbf{g}^t}{\sigma_k}\,\mathbf{v}_k^t \tag{A.53}$$

where $N \leq n$, and $\mathbf{f}$ and $\mathbf{g}$ are the solution and the data vectors, respectively, which depend on the precise method of obtaining the matrix $\mathbf{K}$. We can then rewrite this solution using matrices as follows

$$\mathbf{f}^t = \left( \sum_{k=1}^{N} \mathbf{v}_k^t\,\frac{1}{\sigma_k}\,\mathbf{u}_k \right) \mathbf{g}^t = \mathbf{K}^+\,\mathbf{g}^t \tag{A.54}$$

where

$$\begin{array}{rcl} (\mathbf{K}^+)_{ij} & = & \sum_{k=1}^{N} \frac{1}{\sigma_k}\,v_{ik}\,u_{jk} \\ \mathbf{g} & = & (g_1, \ldots, g_n) \\ \mathbf{f} & = & (f_1, \ldots, f_m) \end{array} \tag{A.55}$$

It is important to notice that $\mathbf{K}^+ \neq \mathbf{K}^{-1}$ since the matrix $\mathbf{K}^+$ has been obtained from truncation of the first $N$ elements of the general solution, which correspond to those having *the greatest* singular values. Thus, the condition number of the matrix can been reduced and so its singularity. In fact, $cond(\mathbf{K}^+) = \frac{\sigma_1}{\sigma_N}$. Since $\sigma_N \geq \sigma_n$, it is true that $cond(\mathbf{K}^+) \leq cond(\mathbf{K})$, which is an improvement.

[3]Here we consider matrix $\mathbf{K}$ as being squared for simplicity, despite the fact there also exists a SVD for general rectangular matrices.

There are some other methods to get a numerical solution for an IFK which behaves better than that of inverting matrix **K** such as the Tikhonov regularization [Pip91, Kre89, Han00, Win91] and those called POCS (Projection Onto Convex Sets) [SS82, YW82, Mar94, ST96, HAD00]. These methods permit adding restrictions on the solution space, such as continuity, smoothness, boundness, etc. Despite its interest for all those who want to recover spectral functions from color data, we are not going into further details on this subject.

## A.6 Conclusions

This Appendix mainly deals with three issues. First, the nature of the solution of the IFK equation is studied, concluding with the constraints that must be fulfilled in order to attain a solution that exists, be unique, and well–behaved. If any of these conditions is otherwise broken, it is also reported the way that the solution will look like accordingly.

Secondly, a general expression for the solution of an IFK in terms of the singular functions and values of the kernel is put forward. This expression completely describes the nature of any solution to the recovery problem and exhibits the difficulties that may arise when dealing with that type of problem. Any further theoretical study on the subject should take account of this expression.

Finally, in order to obtain a practically computation on the problems described by an IFK, a bunch of methods to discretize the solution of any IFK is proposed. Some of them have been successfully used in Chapter 4 to theoretically derive color mappings in the context of color constancy, as well as in the spectral recovery problem.

In additions, we are fully convinced that more further studies on this class of integral equations may report a deeper insight into the question of the recovery of spectral functions from sets of physical measurements, which should improve the efficiency of those proposed in this Appendix.