Version: October 31, 2011

1 What is multi-dimensional signal analysis?

In order to describe what multi-dimensional signal analysis is, we begin by explaining a few important concepts. These concepts relate to what we mean with a signal, and there are, in fact, at least two ways to think of signal that will be used in this presentation.

One way of thinking about a signal is as a function, e.g., of time (think of an audio signal) or spatial position (think of an image). In this case, the signal f can be described as

$$f\,:\,X \rightarrow Y \tag{1}$$

which is the formal way of saying that f maps any element of the set X to an element in the set Y. The set X is called the *domain* of f. The set Y is called the *codomain* of f and must include $f(x)$ for all $x \in X$ but can, in general be larger. For example, we can described the exponential function as a function from $\mathbb{R}$ to $\mathbb{R}$, even though we know that it maps only to positive real numbers in this case. This means that the codomain of a function is not uniquely defined, but is rather chosen as an initial idea of a sufficiently large set without a further investigation of the detailed properties of the function. The set of elements in Y that are given as $f(x)$ for some $x \in X$ is called the *range* of f or the *image* of X under f. The range of f is a subset of the codomain.

Given that we think of the signal as a function, we can apply the standard machinery of analysing and processing functions, such as convolution operations, derivatives, integrals, and Fourier transforms.

Another way of thinking about signals is to see them as elements of some vector space, the *signal space*. In principle, the signal space can be directly defined from its representation as a function. Typically, functions can be added and they can be multiplied by scalars and, consequently, they form a vector space. Such a signal space can, however, sometimes be too abstract to be useful and instead it is possible to establish a more concrete signal space, e.g., $\mathbb{R}^N$ for some known N. Given the signal space, the standard machinery of operations relevant for vector spaces as defined in linear algebra can be applied. This includes establishing a basis for the signal space, determining coordinates of a particular signal relative to the basis, and transforming coordinates from one basis to another.

In general, we may choose which of the two views that fits best for solving a particular problem. In some cases, it may even be fruitful to keep both representations in mind. However, they lead to various issues that we may have to sort out already now in order to have an idea about where to go next. For example, what do we mean with multi-dimensional signal analysis? In the following sections, this question will be illuminated in different ways, but a more concrete answer will have to wait until later when some of the initial theory is applied to practical problems.

1.1 Multi-dimensional?

From the two different views of a signal, as a function or as a vector, the concept of dimensionality can be defined in a number of different ways:

- Outer dimension
- Inner dimension
- Linear dimension
- Intrinsic dimension
- Degrees of freedom

These are the terms that are used in this presentation, but you may see the same concepts under different names in other texts.

1.1.1 Outer dimension

It was said above that we can see a signal as a function f that maps from a domain X to a codomain Y. In fact, signals are more specific that this rather abstract definition. Initially and additionally, we also assume that $X = \mathbb{R}^N = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$, the set of all ordered N-tuples of real numbers, where each of the real spaces refers either to a temporal or spatial variable. In this case, the *outer dimension* is $= N$, i.e., it specifies how many real variables that are needed to determine a specific element of X, which the signal f then maps to a value in Y.

The typical example of a signal is an audio signal, which we can see as a function of a single variable, time, which is an element in $\mathbb{R}$. This signal has outer dimension $= 1$. An image needs two spatial variables to define the variable in X and, therefore, has outer dimension $= 2$. A video sequence can then be seen as a function of two spatial variables and one temporal variable, and a 3D volume reconstructed from computer tomography data is a function of three spatial variables, i.e., they both have outer dimension $= 3$. In all these examples, the outer dimension varies from one (the audio signal), to two (the image), and to three (the video sequence and the 3D tomography volume).

In the examples that we consider in this presentation, the variables are often, but not always, not in $\mathbb{R}$ but are discrete in the sense that the signal value is only known for a discrete (sampled) set of values of the variables. For example, a digital image has known intensity values only at integer values of the two spatial variables, representing pixel coordinates. In these case, there is often an assumption about an underlying continuous signal function that has been sampled, e.g., more or less according to the sampling theorem. Consequently, by processing the discrete signal in suitable ways we expect to produce results that are discrete versions of the corresponding continuous signal, at least to some degree of approximation. The sampling frequency of the continuous signal, measured in samples per unit of the variable is a measure of resolution of the variable. According to the sampling theorem, the sampling frequency and the resolution are directly related to the frequency content of the continuous signal, at least if we expect its discrete representation is sufficiently accurate to be of practical use.

In addition to being discrete, the signal may also be truncated in its outer dimensions, i.e., have known values only within a specific range of the variables. This is typically the case for an image, where we have a limited range of the two spatial variables given by the width and height of the image. The truncation means that we are able to store the entire signal, as a finite set of samples, in memory and then process it with a finite number of computational steps.

Returning to the domain x of a signal f, the idea of its outer dimension means that X is decomposed into a cartesian product of simpler spaces: $X = X_1 \times X_2 \times \ldots \times X_N$, and typically these spaces are either $\mathbb{R}$ or sampled and truncated versions of $\mathbb{R}$. The outer dimension is then equal to N. Furthermore, in the general case when $N > 1$, we can either think of f as a function of N real variables, or as a function of a single element in $\mathbb{R}^N$. Both these views are consistent with the idea that the function has outer dimension $= N$, and we may choose the one that we feel more comfortable with or that makes it easier to derive certain results. This means that we can write either $f(x_1, x_2)$ or $f(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2)$ to represent a function of outer dimension $= 2$.

From a practical point of view, the outer dimension of a signal is often limited to between one and three. There are, however, examples of signals of even higher outer dimensionality, even though it becomes more and more impractical to store and process such type of signals unless the resolution is further reduced to allow storage and processing of the resulting data sets. Examples of such signals are 4D tomography volumes; representing the 3D volume as a function of time (a signal of outer dimensionality four) and the light field; a signal that represents for each point in 3D space how much light that passes through that point in a certain direction. The latter signal requires 3 variables to define which 3D point is considered, and 2 additional variables to define the direction, i.e, it is a signal of outer dimensionality five. We may even extend the light field signal to represent, in addition to the previous specification, how much light there is for each individual frequency (or wavelength) of the light, and also for a particular polarization orientation, etc. This would then add one or two more outer dimensions to the signal.

With these last examples of signals, it should be clear that the initial assumption that the outer dimensions should refer to either temporal or spatial variables cannot be maintained in general. That said, it should be noted that signals that have temporal and or spatial outer dimensions account for the majority of the practical examples. However, the actual physical interpretation of the outer dimensions is sometimes not that important, it is rather the number of these dimensions, the outer dimension, that is important. For example, signals of outer dimension 3 are sometimes processed in similar ways regardless of whether they represent a video sequence or a 3D tomography volume. Also, if we apply some type of transform on a signal, e.g., the Fourier transform, the result is again a function of the same outer dimension as before the transform was applied. In this case, however, the variables of the transformed function may be related to temporal or spatial frequency. This is, again, an example of why we should not restrict ourselves to only temporal or spatial variables for a signal. Such a transformed signal can be processes in a similar way as we would process the original signal.

The figures 1, 2 and 3 illustrate three examples of signals that have different outer dimensions.

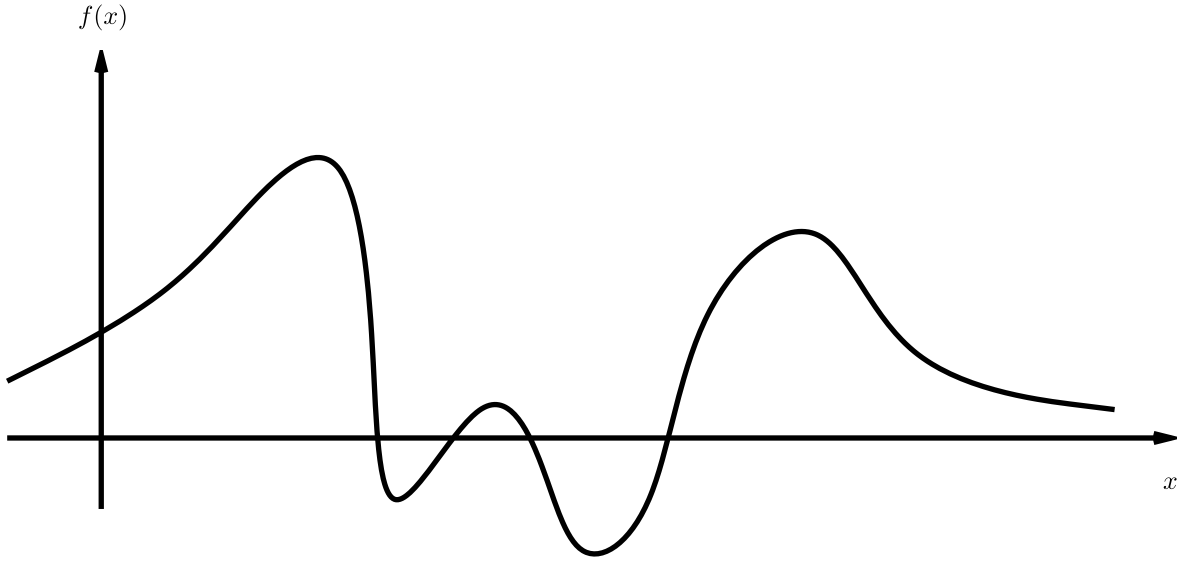

Figure 1: An example of a signal that has outer dimension = 1.

1.1.2 Inner dimension

Having analysed the domain X, let us turn to the codomain Y. Typically, a signal maps X to some type of physical measurement, e.g., in terms of air pressure (an audio signal) or intensity of light (an image). In many cases we can assume that $Y = \mathbb{R}^M$ and in this case we define the *inner dimension* of the signal as $= M$. For example, a colour image typically contains, at each discrete position, information about the intensity of light in the red, green, and blue ranges of the spectrum. In this case we have a signal that maps $\mathbb{R}^2$ to $\mathbb{R}^3$, and it has inner dimension $= 3$.

The restriction to $Y = \mathbb{R}^M$ simplifies the way we think about signals, as mappings from $\mathbb{R}^N$ to $\mathbb{R}^M$. As a consequence, it becomes simpler to organise the storage and processing of our signals. It should be noted that this restriction is fairly general, in the sense that many types of codomains that be represented in terms of $\mathbb{R}^M$ for some inner dimension M. For example, we may consider signals that maps $\mathbb{R}^N$ to $\mathbb{C}$, and then think of $\mathbb{C}$ as equal to $\mathbb{R}^2$ plus a particular algebra. As another example, we may consider signals that map $\mathbb{R}^N$ to symmetric 2×2 matrices, and then represent such matrix as an element of $\mathbb{R}^3$.

The figures 4 and 5 illustrate two examples of signals with different inner dimensions.

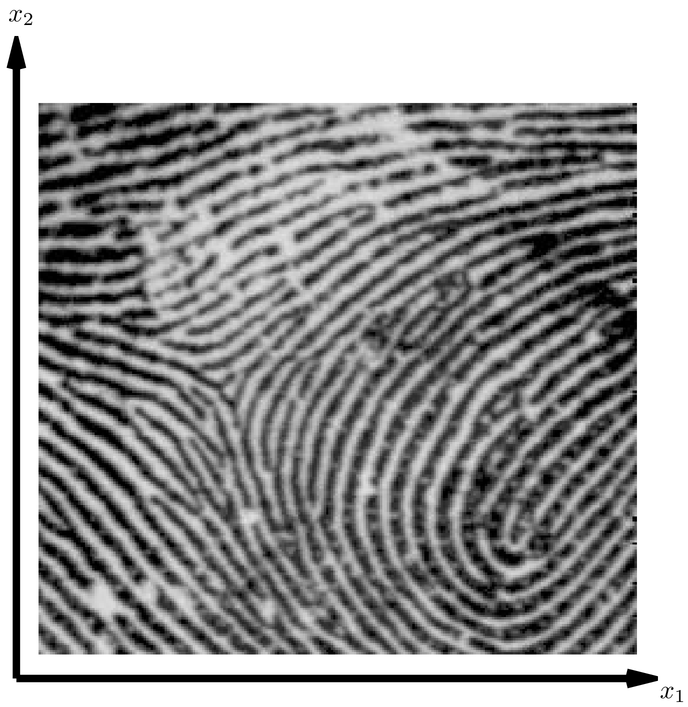

Figure 2: An example of a signal that has outer dimension = 2.

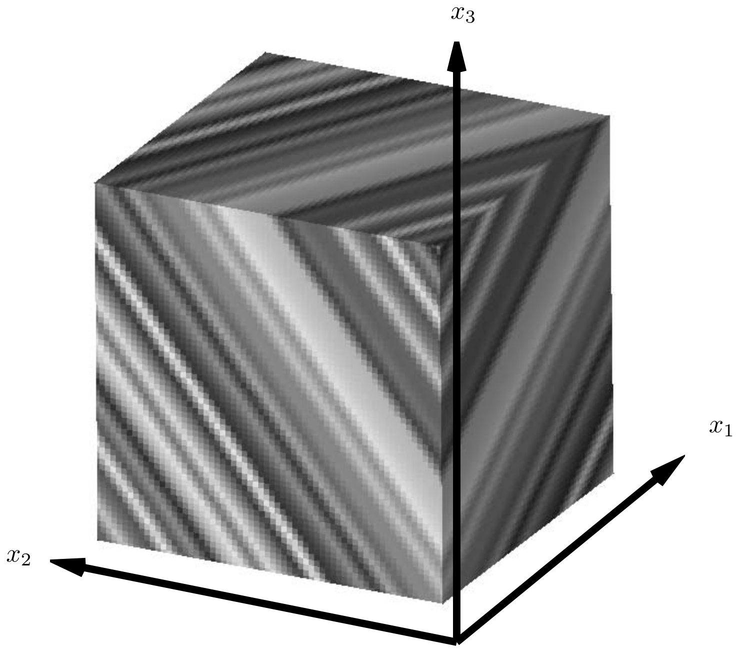

Figure 3: An example of a signal that has outer dimension = 3.

Figure 4: An example of a signal that has inner dimension = 1.

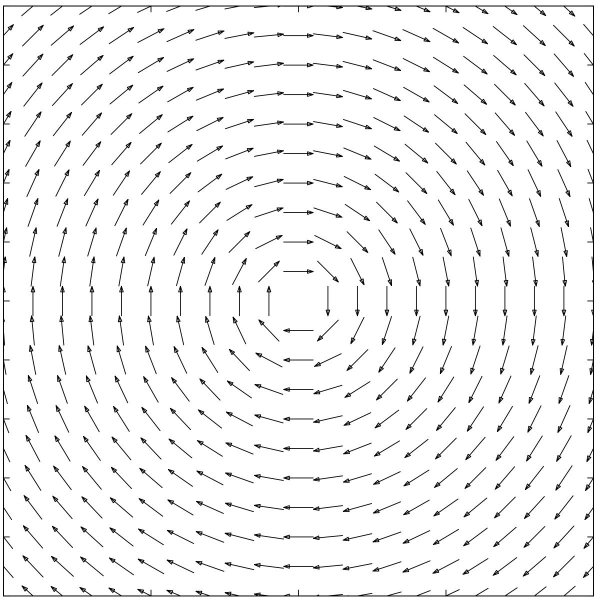

Figure 5: An example of a signal that has inner dimension = 2.

1.1.3 Linear dimension

Let us now change the view of a signal from being a function to being an element of a vector space, the signal space V. The dimension of V is the *linear dimension* of the vector f. Here are some examples of signals that have different linear dimensions.

- An 1024×720 pixel image has 737280 elements in total. This means that we can see the entire image as an element of $\mathbb{R}^{737280}$ and it has a linear dimension of 737280.
- The space of two-variable functions that are sampled in 9×9 points. In this case, the linear dimension is $9 \times 9 = 81$.
- l^2 = the space of square summable functions of discrete variables. In this case, the linear dimension is infinite but enumerable.
- $\mathcal{L}^2$ = the space of square integrable functions on $\mathbb{R}^N$. Also in this case, the linear dimension is infinite but enumerable.

Notice that we can see the second case as related to the first case in the sense that the 9×9 region may be a neighbourhood around some point in the larger image.

Formally, the last two function spaces are characterised also by the outer dimensionality. For example, $\mathcal{L}^2(\mathbb{R}^3)$ is used to denote the space of square integrable functions of three real variables, and this space is distinct from $\mathcal{L}^2(\mathbb{R})$ which consists of single variable functions. Often, however, the outer dimension of a particular signal is known and it is then sufficient to say that the signal is in $\mathcal{L}^2$ to specify that it is square integrable.

1.1.4 Intrinsic dimension

We return to the view of a signal as a function, which then has some outer dimensionality N. In some cases we may find that even though the function needs N real variables in order to produce a value in the codomain, some such functions can be rewritten as functions of lower outer dimension. In this case, there is a transformation $T : X \to Z$ and a function $g : Z \to Y$ such that for every $x \in X$ it is the case that $f(x) = g(z)$ where $z = T(x)$. Furthermore, the outer dimension of g is less than the outer dimension of f.

In order to be useful, we assume that the transformation T and the function g has been chosen such that the outer dimension of g is the smallest possible, i.e., we cannot find a second variable transformation that further reduces the outer dimension. In this case, the outer dimension of g is the *intrinsic dimension* of f. Notice that in general the transformation T and the function g are not unique for particular signal f, even though the intrinsic dimension is unique.

The concept of intrinsic dimension is fairly general in the sense that we have not stated any constraints on T. However, the concept often implies that T is a linear transformation of the variables. In this case the notation *i1D*, *i2D*, etc, is sometimes used for signals of intrinsic dimension 1, or 2, etc.

Example Consider the two-variable function f defined as:

$$f(x_1, x_2) = \sin(3\, x_1 + x_2) \tag{2}$$

By means of the variable transformation

$$y = 3\, x_1 + x_2 \tag{3}$$

f can be rewritten as

$$f(x_1, x_2) = \sin(y). \tag{4}$$

From this formulation it is clear that f only depends on the single variable y. This intrinsic dimension of f is $= 1$. Notice that the variable transformation

$$y' = 6\, x_1 + 2\, x_2 \tag{5}$$

works as well since it gives $f(x_1, x_2) = \sin(y'/2)$. This illustrates the fact that T and g are not unique.

1.1.5 Degrees of freedom

In some cases we have reasons to assume a model for how different observations of the signal may appear. Such an assumption may not be strictly correct for each and every observation, but at least sufficiently correct for a majority of observations and is, therefore, relevant at least from a statistical perspective.

Example Assume that a certain signal of time can be assumed to have only one single frequency component. In order to specify which signal we are considering in a certain case, we then only need to specify three numbers that represent the amplitude A of the signal, its frequency ω, and its phase ϕ:

$$f(t) = A \, \sin(\omega \, t - \phi) \tag{6}$$

This signal has three degrees of freedom. In this case, we again consider the signal f as an element of the signal space V, and ask how many parameters are needed to specify a particular observed signal vector f. The is the reason why the variable t is not included in the degrees of freedom.

Another model of the signal may be that it lies in certain subspace U of V. The dimension of U can then be used as a measure of the degrees of freedom of the signal. This idea will be further discussed in the section on Principal Component Analysis, but it should be emphasised that the concept of degrees of freedom often assumes that it refers to the minimal number of parameters needed to represent a particular observed signal. Furthermore, there is no direct relation between the degrees of freedom of a signal and the dimension of U, other than the latter is an upper limit of the former. For example, the signal may have a single degree of freedom, manifested by a one-variable curve in V that "fills up" the entire V.

1.2 Signal analysis?

The process of analysing a signal means that it is globally or locally represented by some particular set of functions that we have chosen as suitable for this purpose.

1.2.1 Global analysis

When we do a global analysis of a signal, the *entire* signal is represented by means of the special function sets mentioned above. Typically, this set then becomes relatively large in order to represent each and every possible observation of the entire signal. An example of a global analysis it to perform a Fourier transform of the entire signal.

Figure 6 illustrates the entire signal f, i.e., it it has a domain given by the t that are included in the interval in the figure. In a global analysis, we analyse the signal as being a function of this entire domain. This means that also the analysing functions have this domain.

1.2.2 Local analysis

A more common approach to signal analysis is to analyse a restriction of the signal function to a subset of its domain. The reason is that, as already has been mentioned, a global analysis requires a large set of functions in order to be able to describe the entire signal in a useful way. By choosing a suitably small subset of the signal, however, the analysis can often be significantly simplified.

In the typical case, the subset of the domain that is a neighbourhood $\Omega(x_0)$ of some size around a point $x_0 \in X$. This means that the analysing functions only need to worry about what is going in $\Omega(x_0)$, and therefore have $\Omega(x_0)$ as domain.

Even if the local analysis has been simplified by the fact that it deals only with a smaller part of signal, as defined by $\Omega(x_0)$, it also means that it cannot say anything about the signal outside $\Omega(x_0)$. A remedy to this problem is to allow x_0 to be variable, e.g., by moving it along or around the signal domain X. Consequently, the local analysis is made of many different neighbourhoods $\Omega(x_0)$, or even for all possible such neighbourhoods. If the local analysis that is made for each x_0 is simple enough or if the analysis of different neighbourhoods can be made in parallel, the computational cost for this approach is in many cases acceptable.

In the case of local analysis, the signal space V consist of a vector space related to the signal function restricted to $\Omega(x_0)$, not to its entire domain. As a consequence, V can be a vector space of conveniently

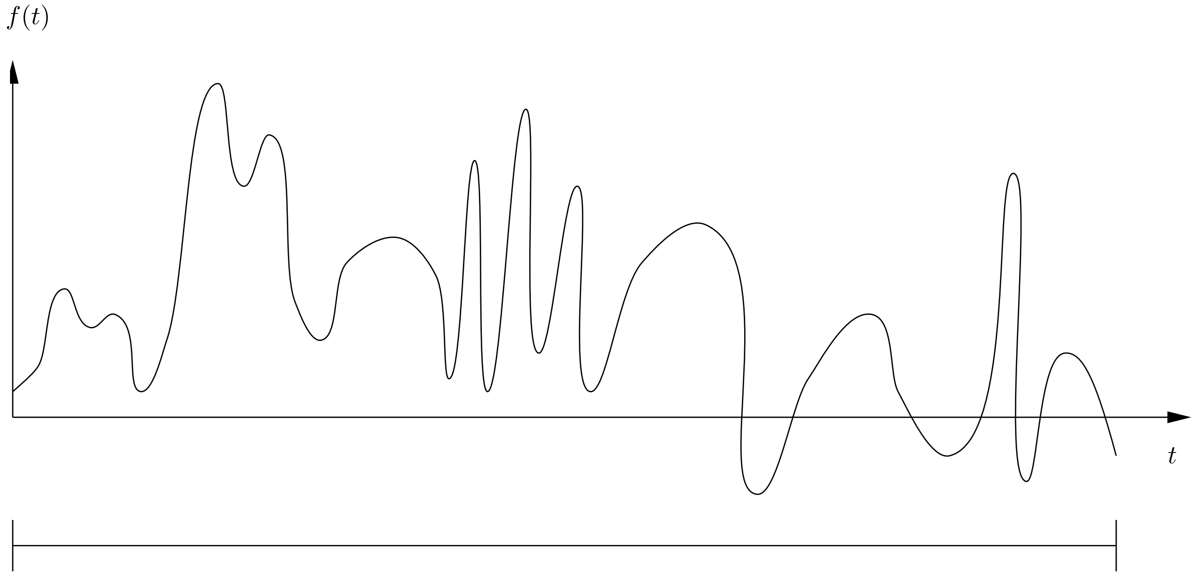

Figure 6: Global analysis: the analysing functions have a domain that is as large as the domain of the signal function.

small dimensionality even if the linear dimension of the entire signal is very large. Compare the first two examples in Section 1.1.3, where we consider an image. If we restrict the analysis of the image to neighbourhoods of 9×9 pixels, the corresponding signal space is of dimension 81, instead of being of dimension 737280 if we consider the entire 1024×720 pixel image. Also, it is likely that for a sufficiently small neighbourhood, the degrees of freedom can be less than these 81 dimensions since it is not likely that we will observe each and every possible element in this 81 dimensional space V with equal probability.

Figure 7 illustrates a signal f that is analysed in a neighbourhood around the point x_0.

1.3 Represents?

In this presentation we will consider two distinct but related ways that a signal is represented by some set of functions.

- We form scalar products between the signal (possibly restricted to some subdomain given by $\Omega(x_0)$) and all the functions that we use to represent the signal. This means that every function in this set of functions produces a scalar, and if the the functions have been chosen in a suitable way these scalars will provide relevant or useful information about the signal. In this case the functions are referred to as *analysing functions*.
- We describe our signal (possibly restricted to some subdomain given by $\Omega(x_0)$) as a linear combination of the representing functions. I this case we need to determine the coefficients in the linear combination that produce the particular observation of the signal function that we have. In this case the functions are referred to as *reconstructing* or *synthesising* functions.

We will soon see that there is a close relation between reconstructing and analysing functions. They are each others duals, and a formal definition of what that means is presented in the following section.

1.4 Which set of functions?

So far, it has been suggested that the set of analysing or reconstructing functions should be suitably chosen. Is it possible to be more specific?

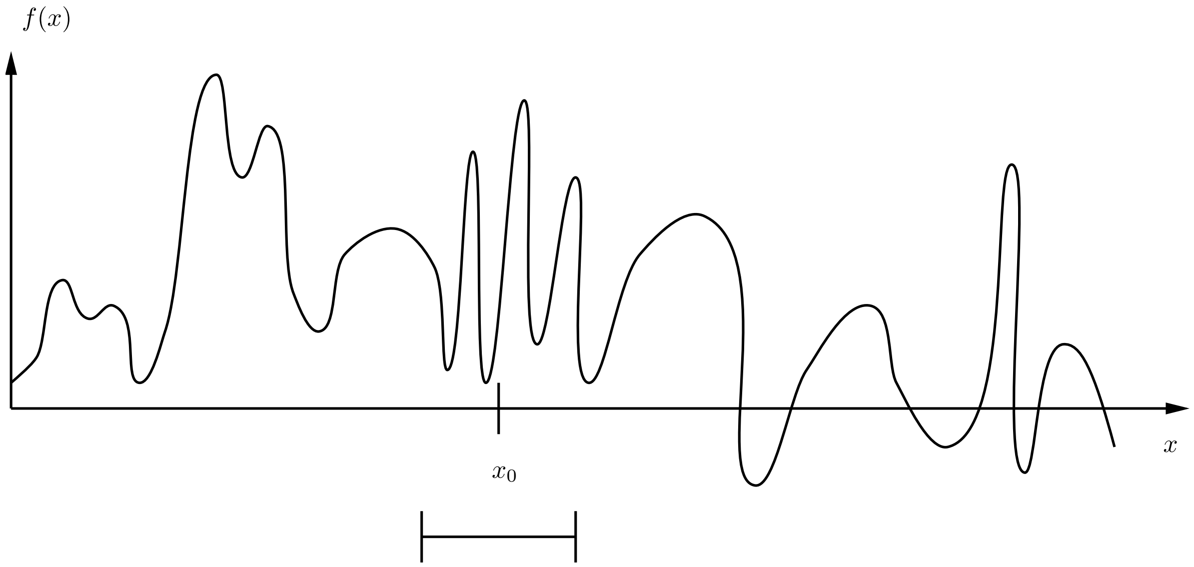

Figure 7: Local analysis: the analysing functions have a domain that is much smaller than the domain of the signal.

At this point in the presentation it suffices to say that exactly which functions we choose for the purpose of representing a signal is very application dependent. There are, however, certain recurrent characters of the representing functions; they fall into either of three cases:

- The representing functions form a basis of the signal space V, i.e., they span V and are linearly independent.
- The representing functions are fewer than what is needed to span V but they are linearly independent. This means that they form a basis of some subspace U of V, or a *subspace basis* of V.
- The functions span V but are too many to be linearly independent. In this case, and with some additional and rather general constraints, they then form a *frame* of V.

In order to complete the characteristic cases, we can consider a set of functions that does not span V, while at the same time they are not linearly independent, a *subspace frame*. This case has, however, not many practical applications and is not considered here.

An important lesson that this presentation aims to provide is that there is no general set of representing function that can be used as a general toolbox for solving any type of problem. It is the properties of the signal at hand and the problem that we try to solve that determine which functions are suitable for analysis or reconstruction of the signal. As an example we have the Fourier transform, which analyses a signal in terms of complex exponential functions of different frequencies. This set of functions is very suitable if we want to understand the global operation of convolving the signal by some filter, but may not be as useful in a local analysis.

Here are some examples of functions that appear frequently in the literature for analysis or reconstruction of signals

- The set of functions of time $t \in \mathbb{R}$ defined as $\delta(t-\tau)$ for all $\tau \in \mathbb{R}$, or the discrete functions of $k \in \mathbb{Z}$ defined as $\delta[k-l]$ for all $l \in \mathbb{Z}$. These are the so-called trivial bases to be discussed later on.
- The set of functions of time $t \in \mathbb{R}$ defined as e^{iut} for all $u \in \mathbb{R}$, or the set of functions of $k \in \mathbb{Z}$ defined as e^{iuk} for all $u \in]-\pi, \pi]$, or the set of P functions of $k = 0, 1, \ldots, P-1$ defined as $e^{2\pi ikl/P}$ for $l = 0, 1, \ldots, P-1$. These bases are used to define various types of Fourier transforms.

- Polynomials: for example in the one-variable case we can use t^n or in two-variable case $x_1^n\, x_2^m$, for some set of integers n and m. Later on we will see that these functions relate to estimation of derivatives.
- The set of functions of time t defined as $\phi(\frac{t-b}{a})$ for $a, b \in \mathbb{R}$ and a suitable choice of function ϕ. These functions are related to the wavelet transform.

2 Some linear algebra

In this section we recapitulate some concepts from linear algebra and then extend that to new ones. In particular, the concept of a dual basis is defined. In the following, V is used to denote a general vector space but keep mind that this is the signal space, and it will be used as such later on.

Let V be an m-dimensional vector space and let $\{\mathbf{e}_k\}$ be a basis for V. An arbitrary vector $\mathbf{v} \in V$ can then be written as

$$\mathbf{v} = c_1\, \mathbf{e}_1 + c_2\, \mathbf{e}_2 + \ldots + c_m\, \mathbf{e}_m \tag{7}$$

where $\{c_k\}$ are the *coordinates* of $\mathbf{v}$ relative to the basis $\{\mathbf{e}_k\}$. From linear algebra we remember that coordinates of a vector relative to a basis are always unique.

A central issue: how can we determine the coordinates of $\mathbf{v}$ relative to the basis $\{\mathbf{e}_k\}$ in the general case?

2.1 Scalar product

We assume that the vector space is complex, i.e., it has $\mathbb{C}$ as the scalar field. The statements in the following presentation can be simplified to the case of a real vector space and these are given in parentheses.

In order to proceed we first need a scalar product for V. In this presentation the scalar product is denoted $\langle\ |\ \rangle$ and it has the following properties

1. $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{R}$ $\qquad$ ($\to \mathbb{R}$ for a real vector space)
2. $\langle \mathbf{v} + \mathbf{u} \mid \mathbf{x} \rangle = \langle \mathbf{v} \mid \mathbf{x} \rangle + \langle \mathbf{u} \mid \mathbf{x} \rangle$
3. $\langle \alpha\, \mathbf{v} \mid \mathbf{u} \rangle = \alpha \langle \mathbf{v} \mid \mathbf{u} \rangle$ for all $\alpha \in \mathbb{C}$ $\qquad$ (for all $\alpha \in \mathbb{R}$)
4. $\langle \mathbf{v} \mid \mathbf{u} \rangle = \overline{\langle \mathbf{u} \mid \mathbf{v} \rangle}$ $\qquad$ ($\langle \mathbf{v} \mid \mathbf{u} \rangle = \langle \mathbf{u} \mid \mathbf{v} \rangle$ for a real vector space)
5. $\langle \mathbf{v} \mid \mathbf{v} \rangle \geq 0$ $\quad$ with equality if and only if $\mathbf{v} = \mathbf{0}$

The properties $2 - 4$ lead to

$$\langle \mathbf{v} \mid \mathbf{u} + \mathbf{x} \rangle = \langle \mathbf{v} \mid \mathbf{u} \rangle + \langle \mathbf{v} \mid \mathbf{x} \rangle \tag{8}$$

$$\langle \mathbf{v} \mid \alpha\, \mathbf{u} \rangle = \overline{\alpha} \langle \mathbf{v} \mid \mathbf{u} \rangle \tag{9}$$

This means that the scalar product is linear in the first variable and *semi-linear* in the second variable. (For a real vector space the scalar product is linear in both variables).

In the mathematical literature it is common to include a scalar product as part of the character of a particular vector space. This means that if two vector spaces are equal in all other respects but have distinct scalar products, they are regarded as two different vector spaces. In this presentation, however, we take the liberty of regarding the scalar product as an additional feature of a vector space. The above stated properties $1 - 5$ of a scalar product simply define it as a special type of mapping from two vectors to a scalar, and for a general vector space there will be many, typically infinitely many, ways of choosing such a mapping. In a particular signal processing application it may be the task at hand that dictates or suggests which scalar product that is appropriate for a particular signal space, and it may even be the case that the scalar product varies for different parts of the signal.

The message that should get across is that for many signal processing applications we can choose the scalar product rather arbitrary as long as it satisfies properties $1-5$. In a similar way, there is no particular basis that we must use for our signal space. Both the basis and the scalar product can be chosen arbitrary within very general limits and it is for this general context that the issue of how to compute the coordinates of a particular signal $\mathbf{v} \in V$ should be seen. Clearly, there are particular choices of both basis and scalar product that could simplify this computation significantly, but if both the scalar product and the basis are stipulated by the problem we are trying to solve, we need a general approach to the computation of coordinates.

2.2 Orthogonality and norm

Two vectors $\mathbf{v}$ and $\mathbf{u}$ are orthogonal when

$$\langle\, \mathbf{v} \mid \mathbf{u} \,\rangle = 0. \tag{10}$$

The scalar product induces a norm in the vector space, defined as

$$\|\mathbf{v}\|^2 = \langle\, \mathbf{v} \mid \mathbf{v} \,\rangle. \tag{11}$$

2.3 Dual basis

Given a basis $\{\mathbf{e}_k\}$ we can construct a basis $\{\tilde{\mathbf{e}}_k\}$ that satisfies

$$\langle\, \mathbf{e}_k \mid \tilde{\mathbf{e}}_l \,\rangle = \delta_{kl} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \tag{12}$$

This basis is unique and is called the *dual basis* of $\{\mathbf{e}_k\}$.

> The dual basis depends on which scalar product that is used for the vector space V. This means that formally, we should say that $\{\tilde{\mathbf{e}}_k\}$ is the dual basis of $\{\mathbf{e}_k\}$ relative to the scalar product $\langle \cdot \mid \cdot \rangle$. However, as long as we can assume that the scalar product is known it is sufficient to just say that $\{\tilde{\mathbf{e}}_k\}$ is the dual basis of $\{\mathbf{e}_k\}$.

Notice that the definition of dual basis given above does not say much about how to obtain it. For now, we will leave that issue to the side and instead look at the main motivation for having a dual basis.

2.4 Computing coordinates

By means of the dual basis, the following result can be obtained

$$\begin{aligned}\langle\, \mathbf{v} \mid \tilde{\mathbf{e}}_l \,\rangle &= \langle\, c_1\, \mathbf{e}_1 + c_2\, \mathbf{e}_2 + \ldots + c_m\, \mathbf{e}_m \mid \tilde{\mathbf{e}}_l \,\rangle = \\ &= c_1 \langle\, \mathbf{e}_1 \mid \tilde{\mathbf{e}}_l \,\rangle + c_2 \langle\, \mathbf{e}_2 \mid \tilde{\mathbf{e}}_l \,\rangle + \ldots\ldots + c_l \langle\, \mathbf{e}_l \mid \tilde{\mathbf{e}}_l \,\rangle + \ldots + c_m \langle\, \mathbf{e}_m \mid \tilde{\mathbf{e}}_l \,\rangle = \\ &= c_1 \cdot 0 + c_2 \cdot 0 + \ldots\ldots + c_l \cdot 1 + \ldots + c_m \cdot 0 = c_l\end{aligned} \tag{13}$$

This means that the coordinates of a vector $\mathbf{v} \in V$ can be obtained as the scalar products between $\mathbf{v}$ and *the corresponding dual basis vectors.*

> Maybe you were taught in basic linear algebra that the way to compute coordinates of a vector is to form the scalar product between $\mathbf{v}$ and the basis vectors. This approach assumes that the basis vectors are orthonormal and does not work in the general case. In the general case you should instead use the dual vectors!

2.5 Change of basis

Let us assume that we have two bases of V, $\{\mathbf{e}_k\}$ and $\{\mathbf{e}'_k\}$. An arbitrary vector $\mathbf{v}$ can then be expressed as the linear combination of the elements in either of the two bases:

$$\mathbf{v} = c_1\,\mathbf{e}_1 + c_2\,\mathbf{e}_2 + \ldots + c_m\,\mathbf{e}_m \tag{14}$$
$$\mathbf{v} = c'_1\,\mathbf{e}'_1 + c'_2\,\mathbf{e}'_2 + \ldots + c'_m\,\mathbf{e}'_m \tag{15}$$

According to the result above, the coordinates $\{c_k\}$ and $\{c'_k\}$ can then be calculated by forming the scalar products between $\mathbf{v}$ and the dual basis vectors $\{\tilde{\mathbf{e}}_k\}$ and $\{\tilde{\mathbf{e}}'_k\}$, respectively.

Alternatively, we can write

$$c'_l = \langle\, \mathbf{v} \mid \tilde{\mathbf{e}}'_l \,\rangle = \langle \sum_k c_k\,\mathbf{e}_k \mid \tilde{\mathbf{e}}'_l \,\rangle = \sum_k c_k \,\langle\, \mathbf{e}_k \mid \tilde{\mathbf{e}}'_l \,\rangle \tag{16}$$

which shows that the "new" coordinates can be written as linear combinations of the "old" coordinates. The coefficients of these linear combinations are given as the scalar products between the "old" basis vectors and the "new" dual basis vectors.

In the finite dimensional case, the above equation can be written as a matrix product

$$\mathbf{c}' = \mathbf{M}\,\mathbf{c}, \tag{17}$$

where $\mathbf{M}$ is an $n \times n$ matrix with elements given as

$$M_{kl} = \langle\, \mathbf{e}_l \mid \tilde{\mathbf{e}}'_k \,\rangle \tag{18}$$

and the vectors $\mathbf{c}$ and $\mathbf{c}'$ consists of the old and new coordinates.

2.6 Infinite dimensional vector spaces

So far, the results have been derived with the assumption that the vector space is of finite dimension n. This assumption can be removed if it is done with some care, but for certain infinite dimensional vector spaces such an extension can be done in a rather straight-forward way. For example, for l^2 and L^2 the following scalar products can be defined

$$V = l^2: \qquad \langle\, f \mid g \,\rangle = \sum_k f_k \cdot \overline{g_k} \tag{19}$$

$$V = \mathcal{L}^2: \qquad \langle\, f \mid g \,\rangle = \int_{-\infty}^{\infty} f(x)\,\overline{g(x)}\,dx \tag{20}$$

and from these, the dual basis can be derived for any choice of basis.

3 A simple example

Let us now illustrate the results so far with a practical example. In this example, the following notation will be used:

- Bold face, e.g., $\mathbf{b}$, is used to denote a vector.
- Italic face, e.g. f_k, is used to denote scalars.

Consider a discrete signal $\mathbf{f}$ consisting of n samples

$$\mathbf{f} = \{f_m, \qquad m = 0, 1, \ldots, n-1\} \tag{21}$$

Together these samples form a finite sequence of numbers, but we can also see $\mathbf{f}$ as a vector in an n-dimensional vector space V, see figure 8. As a basis for V we can use the so-called *trivial basis* $\{\boldsymbol{\delta}_k\}$, where each basis vector $\boldsymbol{\delta}_k$ is given as

$$\boldsymbol{\delta}_k = \{\delta_{km}, \qquad m = 0, 1, \ldots, n-1\} \tag{22}$$

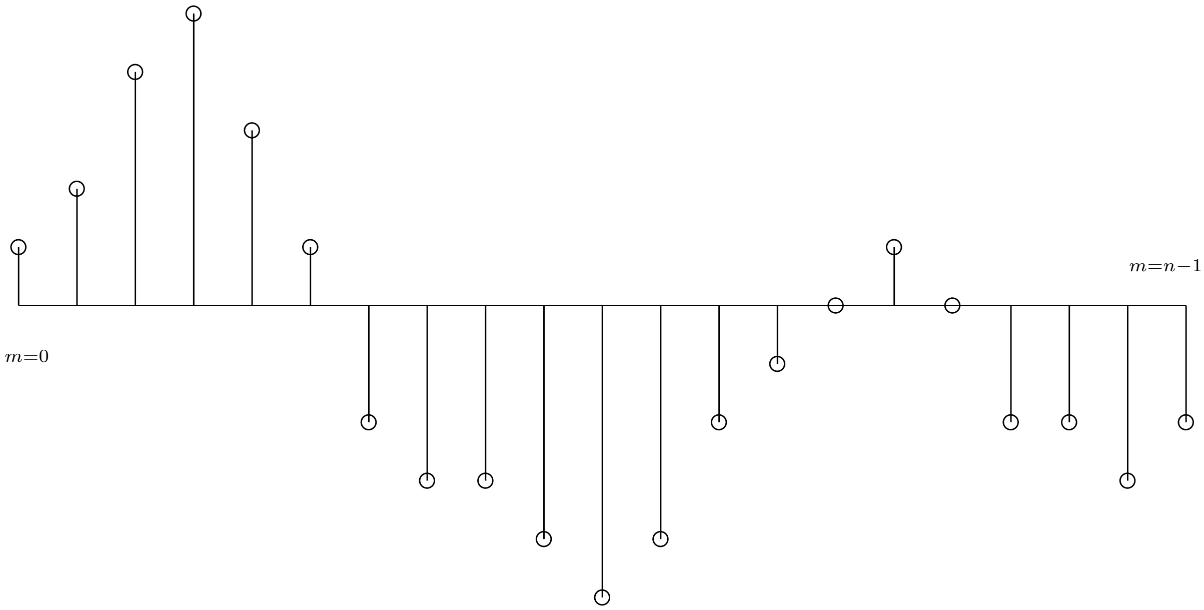

Figure 8: An example of the signal $\mathbf{f}$.

See figure 9. This basis is trivial in the sense that the coordinates of the signal $\mathbf{f}$ are trivially given as the sample values:

$$\mathbf{f} = \sum_{m=0}^{n-1} f_m \,\boldsymbol{\delta}_m \tag{23}$$

As a scalar product in this space we choose the sum of products of the samples values. However, soon we will consider complex valued functions, and to allow the scalar product to be consistent with the previously stated properties 1) - 5), we must take the complex conjugate of the sample values in the second variable:

$$\langle\, \mathbf{f} \mid \mathbf{g} \,\rangle = \sum_k f_k \cdot \overline{g_k} \tag{24}$$

Notice that the trivial basis is its own dual basis, since

$$\langle\, \boldsymbol{\delta}_k \mid \boldsymbol{\delta}_l \,\rangle = \sum_{m=0}^{n-1} \delta_{km} \,\overline{\delta_{lm}} = \delta_{kl} \tag{25}$$

Let us now introduce a second basis for V, defined by the following sequences

$$\mathbf{b}_k = \{ e^{\frac{2\pi i k m}{n}}, \qquad m = 0, 1, \ldots, n-1 \} \tag{26}$$

that give us a unique sequence for each k. We note that

$$\langle\, \mathbf{b}_k \mid \mathbf{b}_l \,\rangle = \sum_{m=0}^{n-1} e^{\frac{2\pi i k m}{n}} \, e^{-\frac{2\pi i l m}{n}} = \sum_{m=0}^{n-1} e^{\frac{2\pi i (k-l) m}{n}} = n\,\delta_{kl} \tag{27}$$

This means that the dual basis of the basis $\{\mathbf{b}_k\}$ is given by

$$\tilde{\mathbf{b}}_k = \frac{1}{n}\,\mathbf{b}_k \tag{28}$$

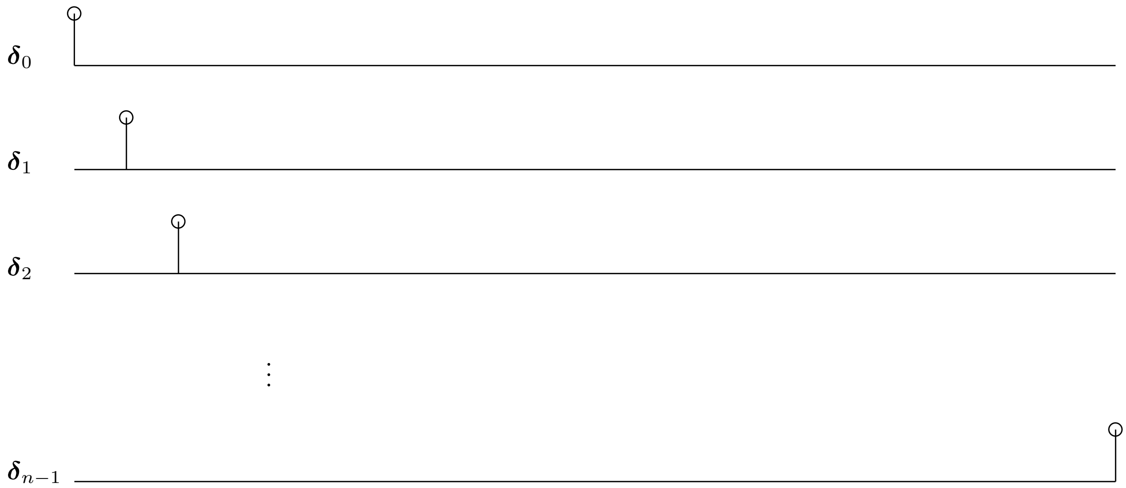

Figure 9: The trivial basis.

We know that the coordinates of $\mathbf{f}$ relative to the trivial basis are given as f_m. What are the coordinates of $\mathbf{f}$'s relative to the new basis? Let these coordinates be denoted as f'_m. Coordinates of a vector are given as the scalar products between the vector and the dual basis vectors:

$$f'_k = \langle\, \mathbf{f} \mid \tilde{\mathbf{b}}_k \,\rangle = \frac{1}{n} \sum_{m=0}^{n-1} f_m \, e^{-\frac{2\pi i k m}{n}}. \tag{29}$$

We recognise this expression as the discrete Fourier transform (DFT) of the discrete signal $\mathbf{f}$.

This means that we can see a DFT of $\mathbf{f}$ as a calculation of the coordinates of $\mathbf{f}$ relative to a basis of complex exponential functions!

Alternatively, we can see a DFT as a change of coordinates, from the trivial basis to the basis $\mathbf{b}_k$. This follows from

$$f'_k = \sum_l f_l \,\langle \text{ old basis} \mid \text{new dual basis } \rangle = \sum_l f_l \,\langle\, \boldsymbol{\delta}_l \mid \tilde{\mathbf{b}}_k \,\rangle \tag{30}$$

But

$$\langle\, \boldsymbol{\delta}_l \mid \tilde{\mathbf{b}}_k \,\rangle = \sum_m \delta_{lm} \, \frac{1}{n} \, e^{-\frac{2\pi i k m}{n}} = \frac{1}{n} \, e^{-\frac{2\pi i k l}{n}} \tag{31}$$

which leads to

$$f'_k = \frac{1}{n} \sum_l f_l \, e^{-\frac{2\pi i k l}{n}} \tag{32}$$

This is the same expression as in (29).

3.1 Summary

We can describe the DFT operation, mapping a sequence of n signal samples to n frequency components either as

- A computation of coordinates for the signal vector. In this case, the samples constitute the signal vector, and the DFT produce the coordinates of this vector relative to the "DFT basis" $\mathbf{b}_k$.

- A change of coordinates. In this case, the signal samples are coordinate of the signal vector relative to the trivial basis and the DFT transforms these coordinates to coordinates relative to the "DFT basis" $\mathbf{b}_k$.

3.2 Two views

Most of the common linear transforms that exist in the literature, e.g.,

- the Fourier transform
- the Laplace transform
- the z-transform
- the wavelet transform

can be described in similar ways, i.e., in terms of coordinate computations and change of basis. Which view we take in any particular case is relatively arbitrary. The important thing is that in both cases, the result can be interpreted as coordinates of the signal relative to some specific basis of the signal space. The latter view, however, allows us to consider the actual signal as an abstract entity that we do not really have to know explicitly, as long as we know its coordinates relative to some basis.

In the following, we will see that the important issue is not exactly which basis we use, but rather which character the basis has. As was mentioned earlier, the set of functions that represent a signal does not necessarily have to be a basis, it may instead be a subspace basis or a frame. In all three cases, the set of functions may or may not be orthogonal (the extension of this idea to a frame is called a thigh frame). Furthermore, the set of representing functions may or may not be normalised. These characterisations are often more important than exactly which set is being used.

4 More about bases and dual bases

We can summarise the results so far about coordinates and dual bases as

$$\mathbf{v} = \langle\, \mathbf{v} \mid \tilde{\mathbf{e}}_1 \,\rangle\, \mathbf{e}_1 + \langle\, \mathbf{v} \mid \tilde{\mathbf{e}}_2 \,\rangle\, \mathbf{e}_2 + \ldots + \langle\, \mathbf{v} \mid \tilde{\mathbf{e}}_n \,\rangle\, \mathbf{e}_n \tag{33}$$

where $\{\mathbf{e}_k\}$ is a basis for the signal space V and $\{\tilde{\mathbf{e}}_k\}$ is the corresponding dual basis (relative to the scalar product used in V).

It is straight-forward to show (at least in the finite dimensional case) that the dual basis of the dual basis is the original basis:

$$\tilde{\tilde{\mathbf{e}}}_k = \mathbf{e}_k. \tag{34}$$

As a consequence of this result we can write

$$\mathbf{v} = \langle\, \mathbf{v} \mid \mathbf{e}_1 \,\rangle\, \tilde{\mathbf{e}}_1 + \langle\, \mathbf{v} \mid \mathbf{e}_2 \,\rangle\, \tilde{\mathbf{e}}_2 + \ldots + \langle\, \mathbf{v} \mid \mathbf{e}_n \,\rangle\, \tilde{\mathbf{e}}_n \tag{35}$$

4.1 Dual coordinates

Since $\langle\, \mathbf{v} \mid \mathbf{e}_k \,\rangle$ are the coordinates of $\mathbf{v}$ relative to the dual basis, we will refer to them as the *dual coordinates* of $\mathbf{v}$. This concept assumes that both a basis and a scalar product, and therefore also a dual basis, are well-defined.

Given a basis and a corresponding dual basis we can now write either

$$\mathbf{v} = c_1\, \mathbf{e}_1 + c_2\, \mathbf{e}_2 + \ldots + c_n\, \mathbf{e}_n \tag{36}$$

or

$$\mathbf{v} = \tilde{c}_1\, \tilde{\mathbf{e}}_1 + \tilde{c}_2\, \tilde{\mathbf{e}}_2 + \ldots + \tilde{c}_n\, \tilde{\mathbf{e}}_n \tag{37}$$

where $\tilde{c}_k$ are the dual coordinates of $\mathbf{v}$ relative to the basis $\{\mathbf{e}_k\}$. Let us now use the expression for how coordinates transform when changing from one basis to another. Transforming coordinates from the basis $\{\mathbf{e}_k\}$ to $\{\tilde{\mathbf{e}}_k\}$ gives us

- the old basis = $\{\mathbf{e}_k\}$
- the new basis = $\{\tilde{\mathbf{e}}_k\}$
- the dual of the new basis = $\{\mathbf{e}_k\}$

This gives

$$\tilde{c}_k = \sum_m c_m \, \langle \, \mathbf{e}_m \mid \mathbf{e}_k \, \rangle. \tag{38}$$

We can also transform coordinates from $\{\tilde{\mathbf{e}}_k\}$ to $\{\mathbf{e}_k\}$:

- the old basis = $\{\tilde{\mathbf{e}}_k\}$
- the new basis = $\{\mathbf{e}_k\}$
- the dual of the new basis = $\{\tilde{\mathbf{e}}_k\}$

This time, we get

$$c_l = \sum_k \tilde{c}_k \, \langle \, \tilde{\mathbf{e}}_k \mid \tilde{\mathbf{e}}_l \, \rangle. \tag{39}$$

4.2 The dual basis revisited

Let us now try to derive an expression for how to calculate the dual basis $\{\tilde{\mathbf{e}}_k\}$ given a certain basis $\{\mathbf{e}_k\}$ and scalar product. Using (35) with $\mathbf{v} = \mathbf{e}_m$, a given basis vector $\mathbf{e}_m$ can be expressed as

$$\sum_k \tilde{\mathbf{e}}_k \, \langle \, \mathbf{e}_m \mid \mathbf{e}_k \, \rangle = \mathbf{e}_m, \tag{40}$$

i.e., it is a linear combination of the dual basis vectors, where the dual coordinates are given as

$$\langle \, \mathbf{e}_m \mid \mathbf{e}_k \, \rangle. \tag{41}$$

Equation (40) can be represented in matrix form as

$$\sum_k \tilde{\mathbf{e}}_k \, G_{km} = \mathbf{e}_m, \tag{42}$$

where

$$G_{km} = \langle \, \mathbf{e}_m \mid \mathbf{e}_k \, \rangle. \tag{43}$$

Notice the order of the indices in the left and right hand side of (43).

Solving for $\tilde{\mathbf{e}}_k$ in (42) gives

$$\tilde{\mathbf{e}}_k = \sum_l \mathbf{e}_l \, (G^{-1})_{lk}, \tag{44}$$

where $\mathbf{G}^{-1}$ is the matrix inverse of $\mathbf{G}$. (44) provides us with a recipe for how to compute the dual basis vectors: compute the elements of the matrix $\mathbf{G}$, invert it, and use the elements of $\mathbf{G}^{-1}$ in a linear combination with the basis vectors.

4.3 The operator G

In the last derivations, $\mathbf{G}$ is referred to as a matrix, and this is correct in the case that V is of finite dimension. $\mathbf{G}$ contains the scalar products between all possible combinations of the basis vectors with other basis vectors. If V is of finite dimension n there are n basis vectors, and $\mathbf{G}$ can be represented as an $n \times n$ matrix that holds all possible scalar products between pairs of the n basis vectors. In the infinite dimensional case, however, we can think of $\mathbf{G}$ as an infinite dimensional array of scalars which is indexed by an infinite range of integers. In general, therefore, we will refer to $\mathbf{G}$ and its inverse $\mathbf{G}^{-1}$ as operators rather than as matrices, since in both the finite and infinite dimensional case they can be used as a linear operator onto the set of basis vectors to produce the dual basis vectors, or vice versa, depending on whether we consider $\mathbf{G}$ or $\mathbf{G}^{-1}$.

From (44) follows immediately

$$\langle\, \tilde{\mathbf{e}}_k \mid \tilde{\mathbf{e}}_m \,\rangle = \sum_l \langle\, \mathbf{e}_l \mid \tilde{\mathbf{e}}_m \,\rangle\, (G^{-1})_{lk} = \sum_l \delta_{lm}\, (G^{-1})_{lk} = (G^{-1})_{mk} \tag{45}$$

This means that the operator $\mathbf{G}^{-1}$ is obtained by forming all possible scalar products between the dual basis vectors.

Notice that $G_{kl} = \overline{G}_{lk}$, i.e. $\mathbf{G}$ is *Hermitian*, which applies also to $\mathbf{G}^{-1}$.

We will now investigate further what use we can have of $\mathbf{G}$ as an operator. Let $\mathbf{v}, \mathbf{u} \in V$ which means that

$$\mathbf{v} = \sum_k c_k\, \mathbf{e}_k \tag{46}$$

$$\mathbf{u} = \sum_k d_k\, \mathbf{e}_k \tag{47}$$

where $\{c_k\}$ and $\{d_k\}$ are the coordinates of $\mathbf{v}$ and $\mathbf{u}$, respectively, relative to the basis $\{\mathbf{e}_k\}$. The scalar product between $\mathbf{u}$ and $\mathbf{v}$ then becomes

$$\langle\, \mathbf{v} \mid \mathbf{u} \,\rangle = \langle\, \sum_k c_k\, \mathbf{e}_k \mid \sum_l d_l\, \mathbf{e}_l \,\rangle = \sum_k \sum_l c_k\, \overline{d}_l\, \langle\, \mathbf{e}_k \mid \mathbf{e}_l \,\rangle = \sum_k \sum_l c_k\, \overline{d}_l\, G_{lk}. \tag{48}$$

This means that $\mathbf{G}$ allows us to express the scalar product between $\mathbf{u}$ and $\mathbf{v}$ in terms of their coordinates in the basis.

Previously we derived an expression for the dual coordinates as a linear combination of the original coordinates, (38). Now that $\mathbf{G}$ is defined, this expression can be written as

$$\tilde{c}_k = \sum_m c_m\, G_{km} \tag{49}$$

This means that $\mathbf{G}$ transforms original coordinates (relative to the original basis) to dual coordinates relative to the dual basis. Consequently, $\mathbf{G}^{-1}$ transforms dual coordinates back to original coordinates.

In the literature, $\mathbf{G}$ is sometimes referred to as the *metric tensor*. However, since $\mathbf{G}$ is dependent on a specific choice of basis, it is more correct to say that it contains the coordinates of the metric tensor relative to this basis.

4.4 Orthonormal basis

Some bases are their own dual basis. For example, we have seen that the trivial basis is its own dual basis. Consequently, such a basis is characterised by

$$G_{kl} = \langle\, \mathbf{e}_l \mid \mathbf{e}_k \,\rangle = \delta_{kl} \quad \Rightarrow \quad \mathbf{G} = \mathbf{I}. \tag{50}$$

Since this implies that all the basis vectors have norm = 1 and are mutually orthogonal, we refer to such a basis as being orthonormal. In short, an orthonormal basis is its own dual basis.

If you have an orthonormal basis, the coordinates of a vector $\mathbf{v}$ are indeed obtained by taking the scalar product between $\mathbf{v}$ and the basis vectors since they are the same as the dual basis vectors. If you are lucky, the problem at hand may allow you to use an orthonormal basis in the solution. In the general case, however, this is not what you should expect and to determine the dual basis is then the way forward.

Given a particular basis $\{\mathbf{e}_k\}$ we may in some cases be free to choose the scalar product such that this basis becomes orthonormal. In this case the metric tensor $\mathbf{G}$ is equal to the identity operator.

4.5 Orthogonal bases

A basis may also be orthogonal, i.e., all the basis vectors are orthogonal but not necessary normalised. An example of an orthogonal basis is the basis of complex exponential functions in the previous example (DFT). In that case we have

$$\langle\, \mathbf{e}_k \mid \mathbf{e}_l \,\rangle = n\, \delta_{kl} \tag{51}$$

where $n =$ is the linear dimension of the signal. In the case that $n > 1$, the basis vectors are not normalised but they are orthogonal.

4.6 Analysis and reconstruction

From the presentation so far it should be clear that given a certain basis $\{\mathbf{e}_k\}$ of our signal space V we can determine the coordinates of a signal $\mathbf{v} \in V$ by computing the scalar products $\langle\, \mathbf{v} \mid \tilde{\mathbf{e}}_k \,\rangle$. In this case we refer to the dual basis vectors as *analysing functions*. Vice versa, if we analyse the signal by means of the original basis vectors, the result are the dual coordinates of the signal. Furthermore, the dual coordinates that results from an analysis with the basis vectors can be transformed to the original coordinates by means of $\mathbf{G}^{-1}$.

Furthermore, given the coordinates $\{c_k\}$ of $\mathbf{v}$ relative to a basis $\{\mathbf{e}_k\}$, $\mathbf{v}$ can be *reconstructed* as

$$\mathbf{v} = c_1\, \mathbf{e}_1 + c_2\, \mathbf{e}_2 + \ldots + c_n\, \mathbf{e}_n. \tag{52}$$

In this case we have a *reconstructing basis* or *synthesising basis* since it is used to rebuild $\mathbf{v}$ from its representation in terms of coordinates. Similarly, the dual basis can be used to reconstruct $\mathbf{v}$ from its dual coordinates.

The analysing and reconstructing bases are each others dual.

5 Finite dimensional signal spaces

Many of the cases discussed in this presentation are derived from a discrete and truncated signal, i.e., it is an element of a signal space with a finite dimension n. In these cases the signal vector $\mathbf{f}$ can be represented as an element of $\mathbb{R}^n$, and in this presentation this will typically be in terms of an n-dimensional column vector:

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \tag{53}$$

where $\{f_k\}$ are the various samples of the signal, e.g., intensity values in a digital image of a neighbourhood around some point x_0. This representation can also be used for the basis vectors and if they are complex valued we have to extend the signal space to $\mathbb{C}^n$ rather than $\mathbb{R}^n$. In fact, to make the following results more useful, we already now state that the signal space is $\mathbb{C}^n$. This still allows particular observations of the signal vector $\mathbf{f} \in V = \mathbb{C}^n$ to have only real elements.

The signal samples f_k can, in general, be seen as coordinates of the signal relative to the trivial basis $\{\boldsymbol{\delta}_k\}$, where

$$\boldsymbol{\delta}_k = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{position } k \tag{54}$$

5.1 Scalar product

Let $\mathbf{f}$ and $\mathbf{g}$ be two vectors in the signal space $V = \mathbb{C}^n$. In order to compute the scalar product between $\mathbf{f}$ and $\mathbf{g}$ we need the matrix $\mathbf{G}_0$ with elements $G_{0,km}$ given as

$$G_{0,km} = \langle\, \boldsymbol{\delta}_m \mid \boldsymbol{\delta}_k \,\rangle. \tag{55}$$

This is the metric tensor related to the trivial basis.

Notice that $\mathbf{G}_0$ does not have to be a unit matrix, but rather a general Hermitian matrix with some additional properties derived from the scalar product. We will use the index "0" to indicate that this metric tensor is derived from the trivial basis, where instead $\mathbf{G}$ without an index is used to denote the metric tensor related to a general basis.

In accordance to (48), the scalar product between $\mathbf{f}$ and $\mathbf{g}$ can be expressed as

$$\langle\, \mathbf{f} \mid \mathbf{g} \,\rangle = \sum_k \sum_l \bar{g}_l \, G_{0,lk} \, f_k. \tag{56}$$

This can also be written as the following matrix product:

$$\langle\, \mathbf{f} \mid \mathbf{g} \,\rangle = \mathbf{g}^\star \, \mathbf{G}_0 \, \mathbf{f} \tag{57}$$

where $\mathbf{g}^\star$ denotes transpose of the vector and complex conjugation of its elements.

Notice the order between $\mathbf{f}$ and $\mathbf{g}$ in the left and right hand sides of the last expression.

5.2 The dual coordinates

We have introduced the trivial basis in V, but so far only to express the signal values (the samples) as coordinates in terms of this basis. In most practical cases, there is another basis of V that we use for the primary purpose of the signal analysis.

Let us assume that $\{\mathbf{e}_k\}$ is such a basis of V. We could then be interested in the dual coordinates of some $\mathbf{v} \in V$ relative to this basis. These dual coordinates can be collected as the elements of a vector $\tilde{\mathbf{c}} \in \mathbb{C}^n$ according to

$$\tilde{\mathbf{c}} = \begin{pmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \vdots \\ \tilde{c}_n \end{pmatrix} = \begin{pmatrix} \langle\, \mathbf{f} \mid \mathbf{e}_1 \,\rangle \\ \langle\, \mathbf{f} \mid \mathbf{e}_2 \,\rangle \\ \vdots \\ \langle\, \mathbf{f} \mid \mathbf{e}_n \,\rangle \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1^\star \, \mathbf{G}_0 \, \mathbf{f} \\ \mathbf{e}_2^\star \, \mathbf{G}_0 \, \mathbf{f} \\ \vdots \\ \mathbf{e}_n^\star \, \mathbf{G}_0 \, \mathbf{f} \end{pmatrix} \tag{58}$$

This can also be expressed as

$$\tilde{\mathbf{c}} = \mathbf{E}^\star \, \mathbf{G}_0 \, \mathbf{f} \tag{59}$$

The matrix $\mathbf{E}$ holds the basis vectors $\{\mathbf{e}_k\}$ in its columns:

$$\mathbf{E} = \begin{pmatrix} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ | & | & & | \end{pmatrix} \tag{60}$$

and $\mathbf{E}^\star$ is the conjugate transpose of this matrix. Since $\{\mathbf{e}_k\}$ is a basis, $\mathbf{E}$ has size $n \times n$ and is of full rank (not singular).

5.3 Analysing operator

We will now define what is called an *analysing operator* relative to a set of basis vectors. Let the set of basis vectors be represented by the matrix $\mathbf{E}$ and let the scalar product in V be defined by the matrix $\mathbf{G}_0$. The corresponding analysing opeator, here denoted $\mathbf{A}$, is then given as

$$\mathbf{A} = \mathbf{E}^\star\, \mathbf{G}_0, \tag{61}$$

This means $\mathbf{A}$ applied to some signal $\mathbf{f}$ corresponds to computing the dual coordinates of $\mathbf{f}$, denoted $\tilde{\mathbf{c}}$, relative to the basis in $\mathbf{E}$. This follows from

$$\tilde{\mathbf{c}} = \mathbf{A}\,\mathbf{f} = \mathbf{E}^\star\, \mathbf{G}_0\, \mathbf{f}. \tag{62}$$

The analysing operator simply computes the scalar products between a signal (or in general an element of the signal space V) and the basis vectors in $\mathbf{E}$. The result comes out as the dual coordinates of the signal relative to the basis.

5.4 Reconstruction and reconstructing operator

From the previous results it follows that if we have the dual coordinates of $\mathbf{f}$, the vector $\mathbf{f}$ can be reconstructed in terms of a linear combination with the dual basis vectors $\{\tilde{\mathbf{e}}_k\}$:

$$\mathbf{f} = \sum_k \tilde{\mathbf{e}}_k\, \tilde{c}_k. \tag{63}$$

This can also be expressed as the matrix product

$$\mathbf{f} = \tilde{\mathbf{E}}\, \tilde{\mathbf{c}}, \tag{64}$$

where the matrix $\tilde{\mathbf{E}}$ contains the dual basis vectors in its columns:

$$\tilde{\mathbf{E}} = \begin{pmatrix} | & | & & | \\ \tilde{\mathbf{e}}_1 & \tilde{\mathbf{e}}_2 & \dots & \tilde{\mathbf{e}}_n \\ | & | & & | \end{pmatrix} \tag{65}$$

We refer to $\tilde{\mathbf{E}}$ as a *reconstructing operator*, in this context also denoted $\mathbf{R}$. This means that the signal vector $\mathbf{f} \in V$ can be expressed as

$$\mathbf{f} = \mathbf{R}\, \tilde{\mathbf{c}} = \mathbf{R}\,\mathbf{A}\,\mathbf{f}, \tag{66}$$

By first applying an analysing operator $\mathbf{A}$ onto the signal we compute its dual coordinates relative to some basis. From these dual coordinates $\mathbf{f}$ can then be reconstructed by means of the reconstructing operator $\mathbf{R}$. Apparently, it must be the case that

$$\mathbf{R}\,\mathbf{A} = \mathbf{I} = \text{identity matrix.} \tag{67}$$

Notice that $\mathbf{A}$ and $\mathbf{R}$ have different character. $\mathbf{A}$ forms scalar products while $\mathbf{R}$ forms linear combinations.

5.5 Computing the dual basis

From the last relations follows that

$$\text{Equation (67)} \quad \Rightarrow \quad \tilde{\mathbf{E}}\,\mathbf{E}^\star\,\mathbf{G}_0 = \mathbf{I} \quad \Rightarrow \quad \tilde{\mathbf{E}}\,\mathbf{E}^\star\,\mathbf{G}_0\,\mathbf{E} = \mathbf{E} \tag{68}$$

$$\text{Equation (12)} \quad \Rightarrow \quad \mathbf{E}^\star\,\mathbf{G}_0\tilde{\mathbf{E}} = \mathbf{I} \quad \Rightarrow \quad \mathbf{E}\,\mathbf{E}^\star\,\mathbf{G}_0\tilde{\mathbf{E}} = \mathbf{E} \tag{69}$$

Notice that the matrix $\mathbf{E}^\star\,\mathbf{G}_0\,\mathbf{E}$ represents the scalar products between all possible pairs of vectors in the basis $\{\mathbf{e}_k\}$, i.e., $\mathbf{E}^\star\,\mathbf{G}_0\,\mathbf{E} = \mathbf{G}$. This allows us to rewrite (68) as

$$\tilde{\mathbf{E}} = \mathbf{E}\,(\mathbf{E}^\star\,\mathbf{G}_0\,\mathbf{E})^{-1} = \mathbf{E}\,\mathbf{G}^{-1}, \tag{70}$$

which is the matrix form of the previously derived expression (44) for how to compute the dual basis vectors.

Alternatively, from (69) we can also express the dual basis vectors as

$$\tilde{\mathbf{E}} = (\mathbf{E}\,\mathbf{E}^{\star}\,\mathbf{G}_0)^{-1}\,\mathbf{E}, \tag{71}$$

This means that each dual basis vector is the result of a fixed operator $(\mathbf{E}\,\mathbf{E}^{\star}\,\mathbf{G}_0)^{-1}$ applied to the corresponding basis vector:

$$\tilde{\mathbf{e}}_k = (\mathbf{E}\,\mathbf{E}^{\star}\,\mathbf{G}_0)^{-1}\,\mathbf{e}_k. \tag{72}$$

This expression will be used later on when we discuss frames of a vector space.