SAMPLING THEOREM ASSOCIATED WITH THE DISCRETE COSINE TRANSFORM

Jelena Kovačević[2,1] and Markus Püschel[1]

[1]Electrical and Computer Engineering and [2]Biomedical Engineering
Carnegie Mellon University
Email: jelenak@cmu.edu, pueschel@ece.cmu.edu

ABSTRACT

One way of deriving the discrete Fourier transform (DFT) is by equispaced sampling of periodic signals or signals on a circle. In this paper, we show that an analogous derivation can be used to obtain the DCT (type 2). To achieve this goal, we replace the circle by a line graph with symmetric boundary conditions, and define signal space, filter space, and filtering operation appropriately. Further, we derive the corresponding sampling theorem including the proper notions of "bandlimited" and "sinc function." The results show that, in a rigorous sense, the DCT is closely related to the DFT, and can be introduced without concepts from statistical signal processing as is the current practice.

1. INTRODUCTION

For years, researchers have been using the DCT in image processing, with the theoretical justification that it is an input-independent transform best approximating (under suitable assumptions) the signal-dependent Karhunen-Loève transform [1, 2, 3]. However, recent research [4, 5] offers a possibly more satisfying explanation. Namely, the DCT is a Fourier transform (in a strict mathematical sense), if the space of signals, the space of filters, and the filtering operation are chosen appropriately. The authors call such a choice a signal model and show that models associated with standard time signal processing are directed, whereas models associated with the DCTs are undirected and are consequently called space models (since space, in contrast to time, has no inherent direction).

As within this theory it is possible to have notions of filtering and Fourier transform different from the usual ones associated with time, an immediate task is to extend other concepts fundamental to signal processing. Two such concepts are sampling and the associated sampling theorem—the focus of this paper. To give a more concrete idea of what we want to do, consider Fig. 1. On the top left we start with periodic signals, or, equivalently, signals given on a circle (which explains the term "compact"). The spectrum is discrete (top right) and the associated Fourier transform is the (continuous time) Fourier series, or CTFS. Sampling (left column) leads to discrete periodic signals, or, signals on a discrete circle. The spectrum becomes periodic (bottom right) and the associated Fourier transform is now the DFT. In fact, the DFT can be derived this way from the Fourier series. An associated (known) sampling theorem states which sampled signals can be perfectly reconstructed from their samples.

The first contribution of this paper is to identify the same diagram for the DCT (focusing on the best-known DCT type 2 [3], denoted as DCT-2); the result is shown in Fig. 2 and explained later.

This work was supported by NSF through award 0310941, and by the PA State Tobacco Settlement, Kamlet-Smith Bioinformatics Grant.

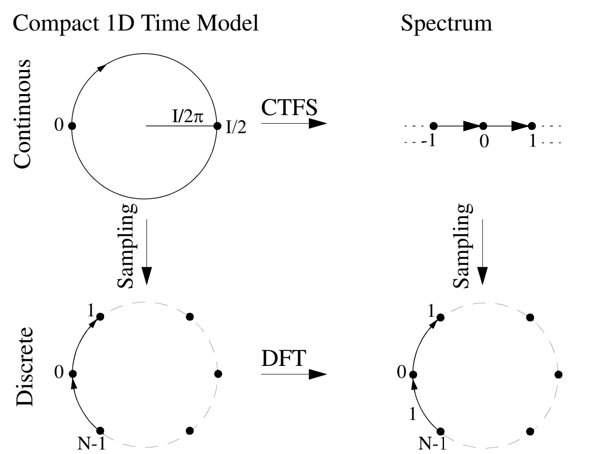


Fig. 1. Compact continuous and discrete time models (directed), connected by sampling. This is one way of deriving the DFT.

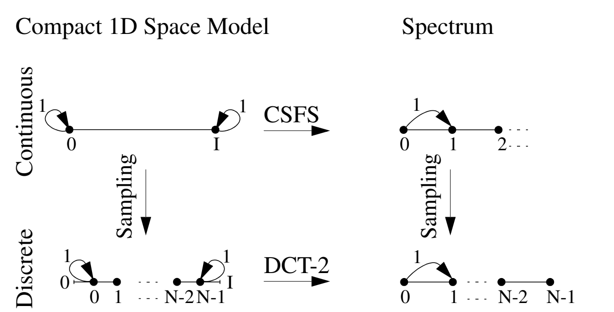


Fig. 2. Compact continuous and discrete "space" models (undirected), connected by sampling. These can be used to derive the DCTs and DSTs (only DCT-2 is shown). The details, and the associated sampling theorem are provided in this paper.

The second, and main contribution, is the derivation of the associated sampling theorem including the proper notions of bandlimited subspace and the equivalent of the sinc function.

There are two additional benefits to this exercise. First, we obtain the continuous structure underlying the discrete DCT. Second, we derive the DCT not based on the KLT, and without any concepts from statistical signal processing.

Organization of the paper. In Section 2, we derive the sampling theorem associated with Fig. 1. The corresponding derivation for Fig. 2 in Section 3 will then be completely analogous. Section 3

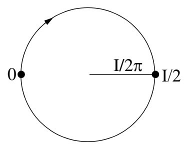
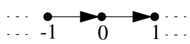
Compact Continuous 1D Time Model	
Signal model	
Signal symmetry	$s(kI+t) = s(t),\ k \in \mathbb{Z}$
Filter symmetry	$h(kI+t) = h(t),\ k \in \mathbb{Z}$
Invariant subspaces	$\mathcal{S}_k = \{ae^{j\frac{2\pi k}{I}t} \mid a \in \mathbb{C}\}, k \in \mathbb{Z}$
FT (=CTFS)	$\hat{s}_k = c\int_0^I s(t)e^{-j\frac{2\pi k}{I}t}dt$
Symmetry of FT	None
Spectrum visualized	
Inverse FT	$s(t) = \frac{1}{cI}\sum_{k\in\mathbb{Z}} \hat{s}_k e^{j\frac{2\pi k}{I}t}$
Frequency response	$\hat{h}_k = c\int_0^I h(t)e^{-j\frac{2\pi k}{I}t}dt$
Inverse frequency response	$h(t) = \frac{1}{cI}\sum_{k\in\mathbb{Z}} \hat{h}_k e^{j\frac{2\pi k}{I}t}$

Table 1. Essential concepts for the compact continuous 1D time model. Signals and filters both belong to $\mathcal{L}^1([0,I))$. The constant $c \neq 0$ in the Fourier transform definition can be chosen arbitrarily.

also identifies the proper definitions of signals, filters, and filtering (convolution) yielding the structures shown in Fig. 2 and thus underlie the DCT.

2. SAMPLING THEOREM FOR THE COMPACT 1D TIME MODEL

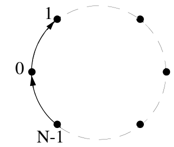
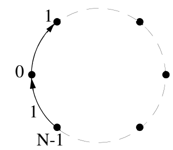
To derive the sampling theorem for the 1D compact *space* model (Fig. 2), we start with the compact *time* model (Fig. 1), which, although less well known than the usual infinite 1D time model, is still intuitive and will thus show how to proceed in the space case. First, we define the signal space, filter space, and the notion of filtering; this is what we call a signal model (following [4]).

Continuous compact time model. To precisely explain what we mean by compact continuous time we define the following (the most important concepts are summarized in Table 1):

Signal model: We consider continuous periodic $\mathcal{L}^1$-signals with fundamental period in the interval $[0,I)$. Equivalently, the signal space is $\mathcal{S} = \mathcal{L}^1([0,I))$, where the interval parameterizes a circle with circumference I or diameter $I/(2\pi)$.[1] The space of filters $\mathcal{H}$ is also $\mathcal{L}^1([0,I))$. Filtering is defined as usual on the circle (continuous circular convolution). The signal space is closed under filtering with these definitions.

Fourier transform: To find the Fourier transform, one first has to identify the eigenspaces under filtering. It is well known that these are spanned by complex exponentials: each $\mathcal{S}_k = \{ae^{j\frac{2\pi k}{I}t} \mid a \in \mathbb{C}\} \leq \mathcal{S}$, $k \in \mathbb{Z}$, is a simultaneous eigenspace for all filters in $\mathcal{H}$. The Fourier transform expands a signal as a series in these exponentials;

[1] We could also choose $\mathcal{L}^2$ signals but in the compact case, $\mathcal{L}^1 \supset \mathcal{L}^2$.

Compact Discrete 1D Time Model	
Signal model	
Sampling period	$T = \frac{I}{N}$
Sampled signal	$s_T(t) = \sum_{n=0}^{N-1} s(nT)\delta(t-nT)$
FT (=DFT)	$\hat{s}_{T,k} = c\sum_{n=0}^{N-1} s(nT)e^{-j\frac{kn2\pi}{N}}$
Symmetry of FT	$\hat{s}_{T,mN+k} = \hat{s}_{T,k}$
Spectrum visualized	
FT of sinc filter	$\hat{l}_k = \begin{cases} cT & 0 \leq k \leq N-1, \\ 0 & \text{otherwise.} \end{cases}$
Sinc filter	$l(t) = \frac{1}{N}e^{j\pi t\frac{N-1}{NT}}\frac{\sin(\frac{\pi}{T}t)}{\sin(\frac{\pi}{NT}t)}$
Bandlim. subspace	$\mathcal{S}_{BL} = \{s \in \mathcal{S} \mid \hat{s}_k = 0, k < 0, k \geq N\}$
Basis for $\mathcal{S}_{BL}$	$b = \{l(t-nT) \mid 0 \leq n < N\}$

Sampling Theorem: For $s(t) \in \mathcal{S}_{BL}$,

$$s(t) \ = \ \sum_{n=0}^{N-1} s(nT)l(t-nT)$$

Table 2. Essential concepts for the compact discrete 1D time model obtained by sampling the model in Table 1.

the coefficients of the series are projections onto the $\mathcal{S}_k$. The Fourier transform for this case is well known and is called continuous-time Fourier series (CTFS, see Table 1).

Sampling. The sampling process and the derivation of the sampling theorem can be described using the following steps, which we will apply later for the model associated with the DCT (the main concepts are summarized in Table 2):

Sample the signal: We first define the sampling period, $T = I/N$, which implies N samples. To place the samples at equispaced locations on the circle, we can start at any location in the interval $[0,T)$; we choose 0. (Different starting points lead to slightly different versions of the DFT below.) Sampling can be described as a multiplication by a train of Dirac pulses T apart:

$$s_T(t) \ = \ s(t)\sum_{n=0}^{N-1}\delta(t-nT) \ = \ \sum_{n=0}^{N-1} s(nT)\delta(t-nT).$$

We then find the Fourier transform of the sampled signal by applying the CTFS to get

$$\hat{s}_{T,k} = c\int_0^I s_T(t)e^{-j\frac{2\pi k}{I}t}dt = c\sum_{n=0}^{N-1} s(nT)e^{-j\frac{kn2\pi}{N}}, \qquad (1)$$

where $c \neq 0$ is a normalizing constant that can be chosen arbitrarily. With $c = 1$, and denoting $W_N = e^{j\frac{2\pi}{N}}$, we recognize the above as the DFT of a sequence of length N (see Fig. 1).

The above process gives rise to the compact discrete time model shown in the bottom row of Fig. 1. The signal lives on a circle with N points (it is discrete periodic with a fundamental period of length N) and the shift moves the signal by one sample clockwise. Applying the shift N times yields the original signal. In the spectrum, the points denote the Nth roots of unity and depict the periodic nature of the DFT. Note that all four graphs are directed.

Extract the original spectrum using an ideal lowpass filter: From (1), we see that after sampling the spectrum becomes periodic, that is, $\hat{s}_{T,mN+k} = \hat{s}_{T,k}$. Thus, for reconstruction, we need to extract only one period of it by applying an ideal lowpass filter. In the time domain, this is equivalent to the convolution of the signal and the inverse Fourier transform of this lowpass filter. This is the filter with the cut-off "Nyquist frequency." We will in general term it as *sinc*, as it has to be exactly 1 at $t = 0$ and 0 at all other sampling points $t = nT$, just as the sinc in the infinite continuous case. The exact form of the sinc is given in Table 2.

Find the space of bandlimited signals: We can then define the space $\mathcal{S}_{BL}$ of those signals bandlimited exactly to the bandwidth of the sinc (see Table 2). We also define the basis b for the space as the set of sincs translated by multiples of T (same table).

Sampling theorem. Based on our discussion, the sampling theorem can be seen as the expansion of signals belonging to $\mathcal{S}_{BL}$ using the translated sincs $l(t-nT)$, and it is simply expressed; For a signal $s(t)$ belonging to $\mathcal{S}_{BL}$:

$$s(t) \;=\; \sum_{n=0}^{N-1} s(nT)l(t-nT).$$

3. SAMPLING THEOREM FOR THE COMPACT 1D SPACE MODEL

The question now is: How do we repeat the previous sequence of steps to get the well-known DCT-2 instead of the DFT? The main problem is in identifying the underlying continuous model, that is, the space of signals, space of filters, and the filtering operation. A summary of the following is in Table 3.

Continuous compact space model. We start by identifying signal and filter spaces.

Signal model: We consider signals on the interval $[0, I]$, which are symmetrically extended to the left and to the right:

$$\begin{aligned} s(-t) &= s(t), \\ s(I+t) &= s(I-t). \end{aligned} \qquad (2)$$

This implies that the signals are $2I$-periodic, since $s(2I + t) = s(I + (I + t)) = s(-t) = s(t)$. The structure produced by these extensions is not a circle but the one shown in Table 3, under the heading "Signal model". It graphically depicts the domain of the signal. The extensions at the left and right boundaries as in (2), become loops with weights 1 at $t = 0$ and $t = I$, respectively. The structure can be parameterized by $[0, I]$, yielding the signal space $\mathcal{S} = \mathcal{L}^1([0, I])$.

As the filter space, we also choose $\mathcal{H} = \mathcal{L}^1([0, I])$. We define filtering (convolution) of $s \in \mathcal{S}$ with $h \in \mathcal{H}$ as

$$(h * s)(\tau) \;=\; \frac{1}{I}\int_0^I h(t)\frac{1}{2}(s(\tau - t) + s(\tau + t))dt. \qquad (3)$$

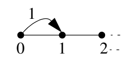
Compact Continuous 1D Space Model	
Signal model	
Signal symmetry	$s(kI + t) = s(kI - t),\ k \in \mathbb{Z}$
Filter symmetry	$h(kI + t) = h(kI - t),\ k \in \mathbb{Z}$
Invariant subspaces	$\mathcal{S}_k = \{a\cos(\frac{k\pi}{I}t) \mid a \in \mathbb{C}\}, k \geq 0$
FT (=CSFS)	$\hat{s}_k = c\int_0^I s(t)\cos(\frac{k\pi}{I}t)dt$
Symmetry of FT	$\hat{s}_{-k} = \hat{s}_k$
Spectrum visualized	
Inverse FT	$s(t) = \frac{1}{cI}\hat{s}_0 + \frac{2}{cI}\sum_{k\geq 1}\hat{s}_k\cos(\frac{k\pi}{I}t)$
Frequency response	$\hat{h}_k = c\int_0^I h(t)\cos(\frac{k\pi}{I}t)dt$
Inverse frequency response	$h(t) = \frac{1}{cI}\hat{h}_0 + \frac{2}{cI}\sum_{k\geq 1}\hat{h}_k\cos(\frac{k\pi}{I}t)$

Table 3. Essential concepts for one of four possible compact continuous 1D space models. Signals and filters both belong to $\mathcal{L}^1([0, I])$. The constant $c \neq 0$ in the definition of the Fourier transform can be chosen arbitrarily.

Direct computation shows that, with this definition, the signal space is closed under filtering. Since this form of filtering operates symmetrically on $s(t)$, we call the model undirected or a space model; pictorially, the line in the graph does not contain an arrow.

The definition of filtering in (3) is equivalent to taking a signal and a filter, both symmetrically extended outside $[0, I]$, viewing them as $2I$-periodic, applying the filtering from the compact time model (circular convolution), and reducing the (symmetric on $[0, 2I)$) result to $[0, I]$.

We can define three other compact space models similarly, by considering all combinations of symmetric and antisymmetric extensions to the left and to the right. To obtain the DCT-2, the above compact model is the right starting point.

Now we proceed as in Section 2; we find the appropriate notion of Fourier transform.

Fourier transform: The subspaces invariant under filtering are given by $\mathcal{S}_k = \{a\cos(\frac{k\pi}{I}t) \mid a \in \mathbb{C}\} \leq \mathcal{S}$ for $k \geq 0$. Note that $\mathcal{S}_{-1} = \mathcal{S}_1$; thus the structure of the spectrum shown in Table 3. This symmetry is pictorially shown as a transition between $k = 0$ and $k = 1$ with weight 1 and can be explained as follows: going to the left from $k = 0$ would lead us to $\hat{s}_{T,-1}$, which, since it does not exist, is redirected to $\hat{s}_{T,1}$; in other words, $\hat{s}_{T,-1} = \hat{s}_{T,1}$. The Fourier transform expands a signal $s(t) \in \mathcal{S}$ in a series in the above cosine functions. Analogously to the CTFS, we call it continuous space Fourier series (CSFS); it is given in Table 3.

Sampling. By virtue of sampling, we produce a discrete model; the main concepts we need are summarized in Table 4. We proceed as in the time case.

Sample the signal: As opposed to the time case, where we could have started equispaced sampling at any point $t \in [0, T)$, here, the situation is different; Only the starting points $t = 0$ and $t = T/2$ allow equispaced sampling. Any other starting point has a distance to itself (via the left signal extension) that is not a multiple of T.

Similarly, we have two choices where to end sampling at the right boundary, namely at I or $I - T/2$, for a total of four choices. These lead to four different versions of the DCT. To get the DCT-2, we have to sample from $T/2$ to $I - T/2$ as we do next. With this choice, and to obtain N samples, we set again $T = I/N$. The sampled signal is then

$$s_T(t) \;=\; \sum_{n=0}^{N-1} s(nT + \frac{T}{2})\delta(t - nT - \frac{T}{2}),$$

while its CSFS is

$$\hat{s}_{T,k} \;=\; c\sum_{n=0}^{N-1} s(nT + \frac{T}{2})\cos(\frac{k(n+\frac{1}{2})\pi}{N}), \qquad (4)$$

which, with $c = 1$, is nothing else but the DCT-2. In words, the DCT-2 is the Fourier transform for a continuous signal given on $[0, I]$, symmetrically extended to the right and left, and sampled at the N equispaced points $T/2, 3T/2, \dots, (2N-1)T/2$ (with the notion of filtering as in (3)).

This way of sampling gives rise to a compact discrete space model in which the signal lives on the discrete structure shown in Table 4, under the heading “Signal model”. This graph, as opposed to the signal model graph in Table 2, is undirected as only symmetric filters are available (the discrete counterparts of the continuous filters in (3)).

Extract the original spectrum using an ideal lowpass filter: The spectrum of the sampled signal in (4) has the following symmetry properties: $\hat{s}_{T,-k} = \hat{s}_{T,k}$ on the left, and $\hat{s}_{T,N} = 0$, $\hat{s}_{T,N+k} = -\hat{s}_{T,N-k}$ on the right; these yield the spectrum structure shown in Table 4. For reconstruction, we again need to extract only one “period,” that is, the spectrum from $0, \dots, N-1$. The inverse Fourier transform (that is, the inverse CSFS) of the corresponding ideal lowpass filter is the sinc for this case:

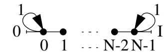
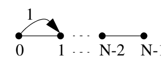
$$l(t) \;=\; \frac{1}{2N}\frac{\sin(\frac{\pi}{T}t)}{\tan(\frac{\pi}{2NT}t)}. \qquad (5)$$

Find the space of bandlimited signals: Similarly to what was done for the time case, we can define the space $\mathcal{S}_{BL}$ of those signals bandlimited exactly to the bandwidth of the sinc (see Table 4). We also define the basis b for $\mathcal{S}_{BL}$ as the set of those sincs translated by multiples of T starting at $T/2$ (same table).

Sampling theorem. The sampling theorem now states that for a signal $s(t)$ belonging to the bandlimited space spanned by the translated sincs $l(t - nT - \frac{T}{2})$, the signal can be reconstructed from its samples as:

$$s(t) \;=\; \sum_{n=0}^{N-1} s(nT + \frac{T}{2})l(t - nT - \frac{T}{2}).$$

Sampling theorems for all 16 DCTs and DSTs. As mentioned above, there are four choices of continuous compact space models and, for each, there are four choices of how to sample. This gives a total of 16 cases, which yield the 16 known DCTs and DSTs. We can follow the same “recipe” we presented for the DCT-2. However, pitfalls abound. For example, for the continuous models which are antisymmetric at the left boundary, signals are necessarily 0 at $t = 0$. Thus, when sampling at points $t = nT$, one needs to start at $t = T$. Also, for the same models, signal and filter space are different, and thus the Fourier transform of a signal and the frequency response of a filter are computed differently. This again shows that one has to be very careful in distinguishing between signals and filters. In those cases, the sincs we found (which are filters) are not signals as well. However, their translated versions are and can be used as basis functions for the bandlimited space. We derived the sampling theorems for all 16 DCTs and DSTs; however, due to the lack of space, these will be given in a future paper.

Compact Discrete 1D Space Model	
Signal model	
Sampling period	$T = \frac{I}{N}$
Sampled signal	$s_T(t) = \sum_{n=0}^{N-1} s(nT + \frac{T}{2})\delta(t - nT - \frac{T}{2})$
FT (=DCT-2)	$\hat{s}_{T,k} = c\sum_{n=0}^{N-1} s(nT + \frac{T}{2})\cos(\frac{k(n+\frac{1}{2})\pi}{N})$
Symmetry of FT	$\hat{s}_{T,mN+k} = (-1)^m \hat{s}_{T,mN-k}$
Spectrum visualized	
FT of sinc filter	$\hat{l}_k = \begin{cases} cT/2 & 0 \le k \le N-1, \\ cT/4 & k = N, \\ 0 & \text{otherwise.} \end{cases}$
Sinc filter	$l(t) = \frac{1}{2N}\frac{\sin(\frac{\pi}{T}t)}{\tan(\frac{\pi}{2NT}t)}$
Bandlim. subspace	$\mathcal{S}_{BL} = \{s \in \mathcal{S} \mid \hat{s}_k = 0, k < 0, k \ge N\}$
Basis for $\mathcal{S}_{BL}$	$b = \{l(t - nT - \frac{T}{2}) \mid 0 \le n < N\}$

Sampling Theorem: For $s(t) \in \mathcal{S}_{BL}$,

$$s(t) \;=\; \sum_{n=0}^{N-1} s(nT + \frac{T}{2})l(t - nT - \frac{T}{2})$$

Table 4. Essential concepts for the compact discrete 1D space model obtained by sampling the continuous model in Table 3 as explained in the text.

4. REFERENCES

[1] N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosine transform,” *IEEE Trans. on Computers*, vol. C-23, pp. 90–93, 1974.

[2] Y. Yemini and J. Pearl, “Asymptotic properties of discrete unitary transforms,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 4, pp. 366–371, 1979.

[3] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, 1990.

[4] M. Püschel and J. M. F. Moura, “Algebraic theory of signal processing: Foundation and 1-D time,” submitted for publication.

[5] M. Püschel and J. M. F. Moura, “Algebraic theory of signal processing: 1-D space,” submitted for publication.