Einführung in die digitale Signalverarbeitung - Übung (3135 L 372) - WS10/11
FG Audiokommunikation, Institut für Sprache und Kommunikation, TU Berlin
Prof. S. Weinzierl, Frank Schultz, Thanassis Lykartsis

Übung 9, 13.01.2011: DFT - Einführung

Musterlösung, Fehler/Ungereimtheiten bitte melden[1]

Aufgabe

In dieser Übung soll es vorrangig, darum gehen, die DFT mit Matlab 'händisch' zu programmieren und anhand dieses Vorgangs mit der Analyse- und Syntheseformel vertraut zu werden. Außerdem sollen prosaisch noch ein paar Anmerkungen zum Wesen dieser fundamental wichtigen Gleichungen gegeben werden, die hoffentlich zum allgemeineren Verständnis beitragen.
Lösung:

Einführung

Das DFT (Analyse) / IDFT (Synthese) - Transformationspaar war nach [Opp04] für eine Transformationslänge N mit dem Twiddle Faktor

$$W_N = \mathrm{e}^{-\mathrm{j}\,\frac{2\,\pi}{N}} \tag{1}$$

gegeben durch

$$X[k] = \sum_{n=0}^{N-1} x[n]\,W_N^{+k\,n} \qquad x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]\,W_N^{-k\,n} \tag{2}$$

Es ist nützlich sich klarzumachen - da in anderer Literatur manchmal so benutzt - dass genauso die Transformationspaare mit unterschiedlicher Normierungskonvention

$$X[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]\,W_N^{+k\,n} \qquad x[n] = \sum_{k=0}^{N-1} X[k]\,W_N^{-k\,n} \tag{3}$$

und

$$X[k] = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} x[n]\,W_N^{+k\,n} \qquad x[n] = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} X[k]\,W_N^{-k\,n} \tag{4}$$

gelten. Es muss offensichtlich die Normierung $K = K_{\mathrm{IDFT}} \cdot K_{\mathrm{DFT}} = \frac{1}{N}$ erfüllt sein, um ein geschlossenes Transformationspaar

$$x[n] = \mathrm{IDFT}\,(\,\mathrm{DFT}\,(\,x[n]\,)\,) \tag{5}$$

[1] frank.schultz@tu-berlin.de, 20. Januar 2011, 04:52

zu erhalten. Man kann sich also entweder K_{IDFT} oder K_{DFT} aussuchen, optimiert nach Anwendungsfall, der jeweilig andere Faktor ist dann bestimmt. Zudem kann genauso das Vorzeichen des Twiddle-Faktors invertiert sein (x als zu analysierendes Signal und X als Spektrum jenes zu betrachten, erfolgte nämlich mit einer gewissen Wilkür, die zwar erheblichen Praxisbezug bei Audiosignalen hat, aber eben Wilkür bleibt)

$$W'_N = \mathrm{e}^{+\mathrm{j}\,\frac{2\,\pi}{N}} \tag{6}$$

, dies wird manchmal bei einigen physikalischen Anwendungen, insbesondere in der Fourierakustik verwendet. Für das allgemein gültige DFT/IDFT-Paar kann dann also geschrieben werden

$$\boxed{X[k] = K_{\mathrm{DFT}} \sum_{n=0}^{N-1} x[n]\, W_N^{\pm k\, n} \qquad x[n] = K_{\mathrm{IDFT}} \sum_{k=0}^{N-1} X[k]\, W_N^{\mp k\, n}} \tag{7}$$

mit

$$\boxed{W_N = \mathrm{e}^{-\mathrm{j}\,\frac{2\,\pi}{N}} \qquad K_{\mathrm{IDFT}} \cdot K_{\mathrm{DFT}} = \frac{1}{N}} \tag{8}$$

Kehren wir zurück zur in [Opp04] und den meisten anderen DSP-Büchern verwendeten Konvention und schreiben den Twiddle-Faktor aus. Dieser ist zwar nützlich wenn man platzsparend Formelapparate z.B. zur FFT herleiten will, ist aber evtl. hinderlich, wenn man sich das erste Mal in die Thematik reindenkt. Es wird also

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n]\, \mathrm{e}^{-\mathrm{j}\,2\,\pi\,\frac{k}{N}\,n} \\ x[n] &= \frac{1}{N} \sum_{k=0}^{N-1} X[k]\, \mathrm{e}^{+\mathrm{j}\,2\,\pi\,\frac{n}{N}\,k} \end{aligned} \tag{9}$$

Mit der Euleridentität geschrieben

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n] \left(\cos(2\,\pi\,\frac{k}{N}\,n) - \mathrm{j}\,\sin(2\,\pi\,\frac{k}{N}\,n)\right) \\ x[n] &= \frac{1}{N} \sum_{k=0}^{N-1} X[k] \left(\cos(2\,\pi\,\frac{k}{N}\,n) + \mathrm{j}\,\sin(2\,\pi\,\frac{k}{N}\,n)\right) \end{aligned} \tag{10}$$

Es wird in beiden Gleichungen eine Folge x oder X mit einer cos-Folge und einer sin-Folge multipliziert. Dies erinnert stark an die Vorgehensweise der Korrelation, die für zeitkontinuierliche Signale $x_1(t)$ und $x_2(t)$

$$y(\tau) = \int_{t=-\infty}^{+\infty} x_1(t) \cdot x_2(t+\tau)\mathrm{d}t \tag{11}$$

oder für eine endliche Integration

$$y(\tau) = \frac{1}{T} \int_{t=t_1}^{t_1+T} x_1(t) \cdot x_2(t+\tau)\mathrm{d}t \tag{12}$$

das Korrelationsergebnis $y(\tau)$ in Abhängigkeit einer Verschiebungszeit τ liefert. Die Korrelation kann natürlich auch wieder zeitdiskret formuliert werden

$$y(m) = \frac{1}{N} \sum_{n=n_1}^{n_1+N-1} x_1(n) \cdot x_2(n+m) \tag{13}$$

Verlaufen also $x_1(t)$ und das um τ verschobene $x_2(t+\tau)$ mit gleicher oder aber invertierter Polarität exakt gleichförmig, wird das Ergebnis $y(\tau)$ den größten Absolutwert erreichen. Nimmt man jetzt ein harmonisches cos/sin-Signal als x_1 oder x_2, kann man über verschiedene τ feststellen, wann die Korrelation des zu analysierenden Signals und der einzelnen cos/sin-Schwingung am größten ist. Das bedeutet aber, das man für Spektralanalyse immer dieses bestimmte τ suchen muß, was nicht gerade elegant ist. Die DFT nutzt nun die Eigenschaft orthogonaler Signale

$$\frac{1}{N} \sum_{n=0}^{N-1} x[n] \cdot y[n] = 0 \qquad \frac{1}{T} \int_{t=t_1}^{t_1+T} x(t) \cdot y(t) = 0 \tag{14}$$

aus, (wobei in N ganzzahlige Vielfache von Periodendauern reinpassen müssen, ein Beweis für Glg. 14 entfällt, kann aber analytisch oder mit Matlab numerisch für cos und sin leicht nachvollzogen werden) und fällt damit in die Kategorie der Orthogonaltransformationen. Hier macht man sich zu Nutze, dass das zu analysierende Signal einmal mit cos und einmal mit sin mit $\tau = 0$ korreliert wird und dann über den Phasenunterschied ermittelt werden kann, ob es 'mehr' sin oder 'mehr' cos enthält. Die DFT kann also als Spezialfall der Korrelation mit harmonischen Schwingungen angesehen werden, denn Audiosignale entstehen laut Fourier aus einer Überlagerung von cos/sin-Schwingungen (Synthese, IDFT) und können demzufolge auch wieder in diese Einzelkomponenten (Analyse, DFT) zerlegt werden. Es gibt andere Anwendungen wo es aufgrund der Bizarrheit des zu analysierenden Signals nützlich ist nicht mit harmonischen, sondern anderen Orthogonalsignalen (z.B. rechteckförmige Signale in der Bildverarbeitung, sogenannte Wavelets) zu arbeiten, weil diese Orthogonalsignale 'besser' mit dem zu analysierenden Signal korrelieren. Die Fouriersynthese tut sich z.B. schwer zu einem absolut sauberen Rechtecksignal zu konvergieren.
In der Analysegleichung wird die Korrelation offensichtlich für N verschiedene cos und N verschiedene sin Signale vorgenommen. Wählt man z.B. $k = 1$ wird Glg. 9

$$X[1] \;=\; \sum_{n=0}^{N-1} x[n] \left(\cos(2\,\pi\,\frac{1}{N}\,n) - \mathrm{j}\,\sin(2\,\pi\,\frac{1}{N}\,n) \right) \tag{15}$$

Das Signal $x[n]$ wird also exakt mit einer Periode eines cos/sin Signals für $0 \leq n \leq N-1$ multipliziert und summiert und liefert $X[1]$. Für $k = 2$ passen exakt 2 Schwingungen in die Fensterlänge N usw. Zu einem späteren Zeitpunkt muss noch festgestellt werden, welche Ergebnisse $X[k]$ in der Analyse tatsächlich informationsredundant sind, wenn man, wie bei Audiosignalen $x \in \mathbb{R}$ wählt. Hilfreich, ohne zu weit vorzugreifen, ist allerdings noch eine Überlegung für $k = 0$. Dann nämlich wird

$$X[0] \;=\; \sum_{n=0}^{N-1} x[n] \left(\underbrace{\cos(2\,\pi\,\frac{0}{N}\,n)}_{1} - \underbrace{\mathrm{j}\,\sin(2\,\pi\,\frac{0}{N}\,n)}_{0} \right) \tag{16}$$

zu

$$X[0] = \sum_{n=0}^{N-1} x[n] \tag{17}$$

Benutzt man jetzt noch die anschaulichere Konvention $K_{\mathrm{DFT}} = \frac{1}{N}$

$$X[0] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \tag{18}$$

folgt daraus offensichtlich der Mittelwert oder Gleichanteil (den man in Audiosignal gerne vermeidet) der Folge $x[n]$. Die DFT/IDFT hat keine spezielle Anforderungen an den gewählten Zahlenbereich und kann $x \in \mathbb{C}$ und $X \in \mathbb{C}$ darstellen. Der tatsächlich resultierende Zahlenbereich hängt von der speziellen Anwendung ab und wird in der folgenden Übung untersucht werden. Was man sich aber als Vorgeschmack auf Folgendes anschauen kann ist das Verhalten des Twiddle-Faktors W_N, der ja für die verschiedenen n und k ausgerechnet werden muss. In den Grafiken 1 und 2 ist der Winkel des Twiddle-Faktors in Grad für alle möglichen Kombinationen

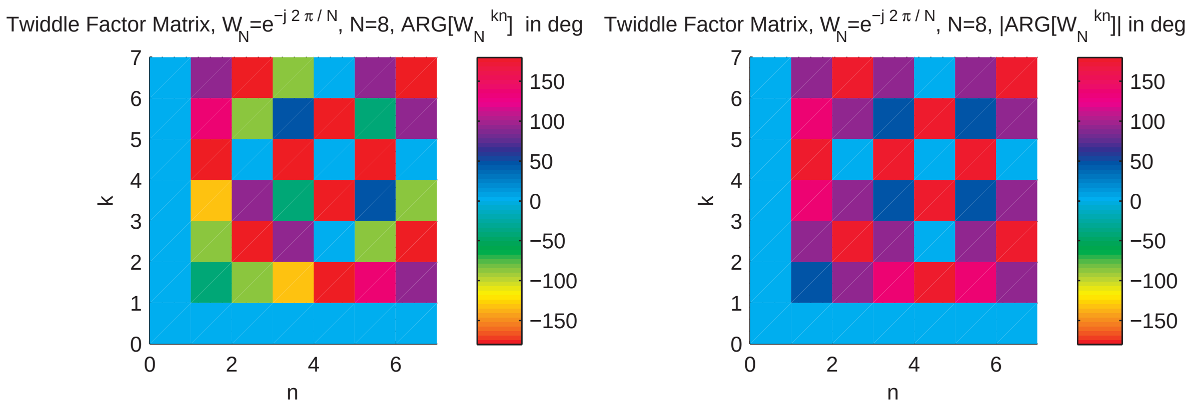

Abbildung 1: farbig codierte Winkel der Twiddle Faktor Matrix nxk zu Veranschaulichung von Symmetrien, N=8

$n \times k$ mittels eines sogenannten 'surf'-Plots grafisch aufgetragen. Man erkennt jeweils im linken Bild, dass offensichtlich gleiche Winkel für verschiedene n,k-Kombinationen im Twiddle-Faktor auftreten. Diese Symmetrie wird noch deutlicher, wenn man im rechten Bild nur noch den Winkelbetrag (also ohne Vorzeichen) darstellt. Wenn man also weiß, welche n,k-Kombinationen den gleichen Winkel oder Winkelbetrag haben, kann man in der Berechnung der DFT erheblich einsparen, indem man nur einmal den entsprechenden Winkel errechnet und dann für die jeweiligen n,k-Kombinationen benutzt, also redundante Rechenschritte vermeidet. Dies ist mit eine der Grundideen bei der Programmierung effizienter Algorithmen für die DFT-Formel, die in der Literatur unter Fast Fourier Transformation (FFT) geführt werden.

Matlab

Glg. 10 soll nun in Matlab programmiert werden, siehe 'DFT_Intro.m'. Es ist zudem lohnenswert sich auch Gedanken zu machen, wie ein Rechner mit komplexen Zahlen umgeht, wenn

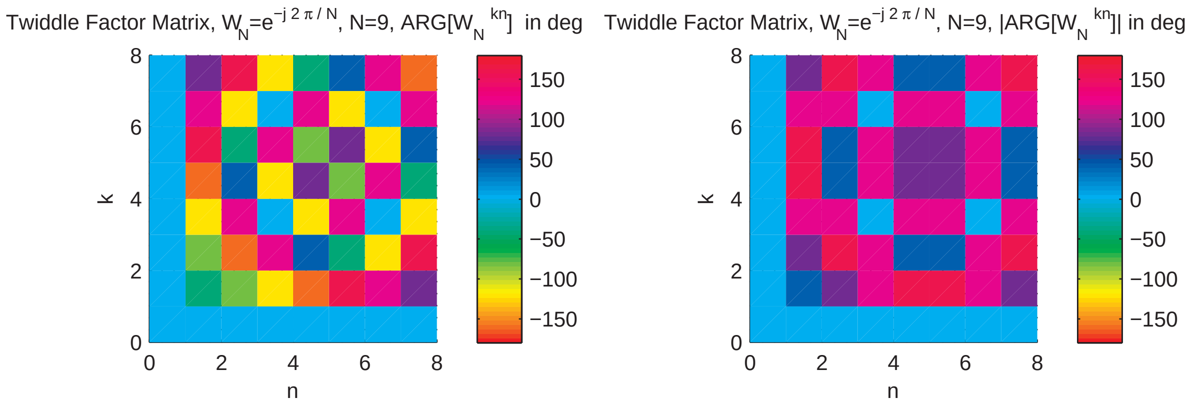

Abbildung 2: farbig codierte Winkel der Twiddle Faktor Matrix nxk zu Veranschaulichung von Symmetrien, N=9

man keine so hochentwickelte Sprache wie Matlab, sondern C oder sogar noch maschinennäher Assembler benutzt; Spektralanalyzer benutzen in der Regel nämlich kein Matlab :-)...

Literatur

[Opp04] Oppenheim A.V., Schafer R.W., Buck J.R. (2004): Zeitdiskrete Signalverarbeitung. München: Pearson Studium, 2. Auflage.