A Critical Analysis of Design Flaws in the Death Star

Luke Skywalker

99652154

Report submitted in partial fulfilment of the requirements of the module Project (E) 448 for the degree Baccalaureus in Engineering in the Department of Electrical and Electronic Engineering at Stellenbosch University.

Supervisor: Dr O. W. Kenobi

October 2099

Acknowledgements

I would like to thank my dog, Muffin. I also would like to thank the inventor of the incubator; without him/her, I would not be here. Finally, I would like to thank Dr Herman Kamper for this amazing report template.

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Plagiaatverklaring / *Plagiarism Declaration*

1. Plagiaat is die oorneem en gebruik van die idees, materiaal en ander intellektuele eiendom van ander persone asof dit jou eie werk is.
   Plagiarism is the use of ideas, material and other intellectual property of another's work and to present is as my own.

2. Ek erken dat die pleeg van plagiaat 'n strafbare oortreding is aangesien dit 'n vorm van diefstal is.
   I agree that plagiarism is a punishable offence because it constitutes theft.

3. Ek verstaan ook dat direkte vertalings plagiaat is.
   I also understand that direct translations are plagiarism.

4. Dienooreenkomstig is alle aanhalings en bydraes vanuit enige bron (ingesluit die internet) volledig verwys (erken). Ek erken dat die woordelikse aanhaal van teks sonder aanhalingstekens (selfs al word die bron volledig erken) plagiaat is.
   Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism

5. Ek verklaar dat die werk in hierdie skryfstuk vervat, behalwe waar anders aangedui, my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik ingehandig het vir bepunting in hierdie module/werkstuk of 'n ander module/werkstuk nie.
   I declare that the work contained in this assignment, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

Studentenommer / *Student number*	Handtekening / *Signature*
Voorletters en van / *Initials and surname*	Datum / *Date*

Abstract

English

The English abstract.

Afrikaans

Die Afrikaanse uittreksel.

Contents

Declaration ii

Abstract iii

List of Figures v

List of Tables vi

Nomenclature vii

1. Introduction 1

1.1. A brief history of musical synthesisers . . . 1
1.2. Basic modular synthesiser building blocks . . . 3
1.2.1. The VCO . . . 3
1.2.2. The VCF . . . 4
1.2.3. The VCA . . . 4
1.2.4. The ADSR envelope . . . 4
1.3. A basic monophonic modular setup . . . 6
1.4. An overview of synthesis techniques . . . 7
1.4.1. Additive synthesis (Fourier synthesis) . . . 7
1.4.2. Subtractive synthesis . . . 7
1.4.3. FM synthesis . . . 8
1.4.4. Physical modelling . . . 8
1.4.5. Sampling . . . 8
1.4.6. Wavetable synthesis . . . 9
1.5. Wavetable frequency conversion . . . 10

2. Summary and Conclusion 13

Bibliography 14

A. Project Planning Schedule 15

B. Outcomes Compliance 16

List of Figures

1.1. Examples of fundamental modules. 3
1.2. ADSR behaviour. 5
1.3. Block diagram of a very basic monophonic modular setup. 6
1.4. Frequency scaling using upsampling and downsampling 10
1.5. The effects of linear interpolation in the frequency domain. 11

List of Tables

Nomenclature

Variables and functions

$p(x)$	Probability density function with respect to variable x.
$P(A)$	Probability of event A occurring.
ε	The Bayes error.
ε_u	The Bhattacharyya bound.
B	The Bhattacharyya distance.
s	An HMM state. A subscript is used to refer to a particular state, e.g. s_i refers to the i^{th} state of an HMM.
$\mathbf{S}$	A set of HMM states.
$\mathbf{F}$	A set of frames.
$\mathbf{o}_f$	Observation (feature) vector associated with frame f.
$\gamma_s(\mathbf{o}_f)$	A posteriori probability of the observation vector $\mathbf{o}_f$ being generated by HMM state s.
μ	Statistical mean vector.
Σ	Statistical covariance matrix.
$L(\mathbf{S})$	Log likelihood of the set of HMM states $\mathbf{S}$ generating the training set observation vectors assigned to the states in that set.
$\mathcal{N}(\mathbf{x}\|\mu, \Sigma)$	Multivariate Gaussian PDF with mean μ and covariance matrix Σ.
a_{ij}	The probability of a transition from HMM state s_i to state s_j.
N	Total number of frames or number of tokens, depending on the context.
D	Number of deletion errors.
I	Number of insertion errors.
S	Number of substitution errors.

Acronyms and abbreviations

VCA	Voltage-controlled amplifier
ADSR	Attack, Decay, Sustain, Release
VCF	Voltage-controller filter
VST	Virtual studio technology
DAW	Digital audio workstation
FM	Frequency modulation
LUT	Lookup table
LFO	Low frequency oscillator
IR	Impulse response
VCO	Voltage-controlled oscillator
CV	Control voltage
LPF	Low-pass filter
HPF	High-pass filter
BPF	Band-pass filter
FIR	Finite impulse response
IIR	Infinite impulse response

Chapter 1

Introduction

1.1. A brief history of musical synthesisers

An instrument usually described as a “synthesiser” or “synth” is any electronic device or software that can generate audio signals. This can be done through a variety of techniques that have been developed and have evolved during the 20th and 21st centuries.

One of the most well-known earlier synthesisers is the theremin, invented by Leon Theremin and patented in 1928. The theremin consists of an oscillator whose frequency and amplitude are controlled by two “antennas” (they look like antennas but are not, strictly speaking).

The device operates through capacitive coupling with the “antennas” and the performer’s hands. The capacitance changes as the performer moves his/her hands, which subsequently controls the amplitude and frequency of the oscillator. Due to the environmental sensitivity, the analogue theremin is notoriously difficult to keep in tune as room conditions change (temperature, positioning etc.).

The vacuum tube (or thermionic valve) triode was used before the silicon transistor to control currently flow with a separate control voltage. This formed the basis of many keyboard synthesisers which became popularised in the late 1930s. The Novachord built by the Hammond Organ Company is a prime example, which used 146 vacuum tubes for 72 VCAs.

A non-keyboard style of synthesis - modular synthesis - was popularised by the companies Moog and Buchla in the 1960s. Analogue electronics were used for synthesis where multiple modules generate control voltages in tandem to modulate parameters. Common building blocks are VCOs, envelope generators (ADSR), VCFs, VCAs, sequencers, wave-shapers, and noise generators. These are still the fundamental aspects of most synthesisers to date. Emulation or cloning of original Moog or Buchla hardware such as the 4-pole ladder VCF is still sought after in the current commercial market.

In contrast, modern Eurorack modules (a standard for modular hardware specification) offer a wide variety of complex options, which are often analogue or digital-analogue hybrid. An example is the Make Noise MATHS module which is very popular in the modular synth community. It provides features such as amplification, integration, summation, function generation, and is considered an essential module by many influencers.

Due to the complexity of analogue electronics, most synths were still monophonic in the 1960s and 1970s, or had very limited polyphony. Each extra note required duplication of electronics, which, in turn, requires more fine-tuning. The introduction of digital technology in the 1980s allowed for more flexible polyphony at affordable prices. The Yamaha DX7 is an incredibly well-known early digital synthesiser released in 1983, which used FM-synthesis (see section 1.4). The DX7 was used on records by U2, Toto, Queen, Elton John, and jazz virtuoso Chick Corea. Sampling synthesis, which is very similar to wavetable synthesis, was utilised in the late 1980s by other digital keyboard products such as the Roland D-50, the Fairlight CMI, and drum machines used in the conception of the hip-hop genre.

The introduction of more powerful computation led to the development of software synths and VSTs for DAWs which use a variety of synthesis techniques such as FM, additive synthesis, subtractive synthesis, physical modelling, and wavetable synthesis. Wavetable synthesis is very popular in all music genres and sound design for film. Some of the most popular instrument VSTs used are Serum by Xfer Records, Massive by Native Instruments, and PIGMENTS by Arturia. The aforementioned VSTs focus on wavetable synthesis with sampling, filtering, parameter modulation and FM capabilities. Serum is extensively used in EDM genres by artists such as DeadMau5 and Virtual Riot.

1.2. Basic modular synthesiser building blocks

This overview focuses on modular synthesiser building blocks directly but is relevant to most forms of synthesis (especially in this thesis) since most standard modern synthesis products is based on the building blocks popularised by Moog and Buchla. Example modules will be shown, discussed, and compared to features present in commercial wavetable synthesisers and features to be considered for design in this thesis.

It should be noted that modular synthesis is a niche market, with many commercial and boutique manufacturers along with hobbyists wanting to contribute. Hence, there are many modules with a wide variety of features that make them unique. Many modules fulfil multiple purposes instead of one. The examples chosen in this section are simplistic popular modules that serve only a single purpose.

(a) Doepfer A-111-3 **(b)** IntelliJel UVCF **(c)** MFB VCA **(d)** Doepfer A-140

Figure 1.1: Examples of fundamental modules.

1.2.1. The VCO

VCO modules commonly include 1V/octave CV inputs for frequency and provide the basic waveforms as outputs either separately via a switch or simultaneously. They can be analogue or digital in nature and can use a variety of synthesis techniques to generate their waveforms. They often come with the ability set the offset tuning voltage and can be used to create FM signals through control voltages. Extra features such as wave folding are sometimes also present.

The Doepfer A-111-3 Micro Precision VCO/LFO is an analogue VCO that can also operate in LFO configuration, either with a linear or exponential voltage control. Sync (for phase/frequency syncing) and PWM CV inputs are also available. All the basic waveforms are present, except for the sinusoid which is notoriously difficult to generate with analogue electronics, which is commonly implemented by a high-Q unstable filter.

1.2.2. The VCF

The VCF is an incredibly important module that forms the basis of subtractive synthesis techniques. Most synths also offer filtering capabilities, such as the widely used Nord Stage 3.

Filters can come in many types, often designed with unique characteristics. This can include special control voltage behaviour, feedback path saturation to limit resonance while adding additional harmonics, or the ability to achieve exceptionally high Q values that cause purposeful instability that allow filters to also function as a sinusoidal oscillator (which many VCOs do not generate).

Thus, filters for musical applications are usually not designed to be as "clean" and stable as possible. Instead, they focus on usability and uniqueness. Filter types can include a switchable LPF, BPF or HPF mode, a ladder filter, 12dB/octave or 24dB/octave varieties and a state variable filter configuration.

The IntelliJel UVCF is popular state variable filter that simultaneously outputs a 2-pole low-passed, 2-pole high-passed and 1-pole band-passed signal which has a cut-off that can be modulated by 2 separate 1V/octave control voltages. It can also be set to have a high Q-value so that it can act as a sinusoidal VCO due to filter instability.

1.2.3. The VCA

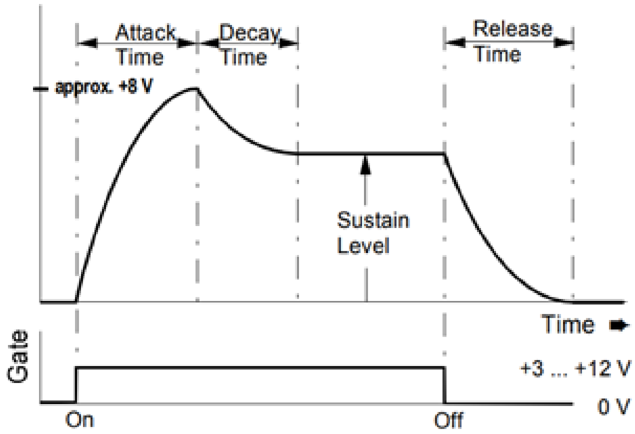
The VCA has the primary purpose of performing the multiplication of signals for uses in AM and otherwise. It acts as an amplifier with a voltage controllable gain. It is often used in conjunction with an LFO to create a tremolo effect or with an ADSR envelope to shape the transient of signal to emulate bowing or plucking and removing clicks and pops that can occur with the immediate triggering of signal. Many VCOs only output a continuous signal. Hence, a VCA is required to mute any oscillators that are not triggered.

The ring modulation effect can also be achieved by multiplying 2 signals in the audible frequency range together.

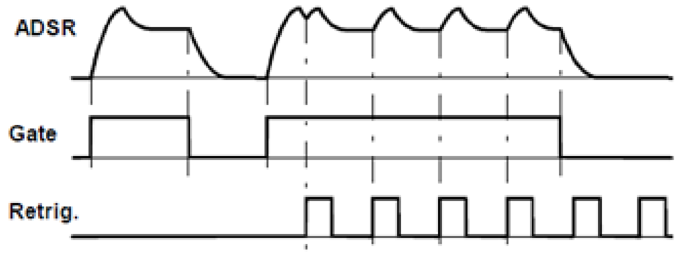
The MFB VCA is module that has 3 different inputs and 2 CV inputs that modulate the gain. The operation of the various inputs is specific to this module and out of the scope of this thesis.

1.2.4. The ADSR envelope

The ADSR envelope is a critical component in synthesis used to achieve realistic sounds. It is often used to modulate filter cut-off to allow for dynamic subtractive synthesis. It is also used for AM to emulate the natural attack and decay characteristics of real instruments. It can emulate plucking, strumming, bowing, and blowing techniques found in real instruments. It can also be used in FM to recreate the typical pitch modulation



(a) The typical ADSR curve with gate signal.



(b) Typical re-triggering operation.

Figure 1.2: ADSR behaviour.

found when striking percussive instruments *REF?*.

ADSR envelopes are available in most wavetable synthesisers for parameter modulation, such as Serum, Massive and Ableton's stock Wavetable VST instruments. Many keyboards also include this feature, such as the Nord Stage 3.

The ADSR envelope consists of 4 phases. The envelope curve is initiated with a gate "on" trigger after which a rising function is started. Once a threshold is reached, determined by the attack time, the decay state is activated. The decay is specified by a decay time parameter. The function decreases until a sustain level is reached, which is a parameter set by the performer. The sustain level is usually less than the peak achieved by the attack phase. The sustain phase remains constant until the gate signal changes state to indicate an "off" trigger, initiating the release phase. The release phase is a decreasing function that decreases until zero is reached (or close to zero in the case of an RC circuit), determined by a release time parameter.

There are thus 4 parameters that can be set by the performer: attack time, decay time, sustain level and release time. The A, D and R phases are usually exponential functions implemented by an RC circuit. This is well suited for AM and FM, since octaves are exponential in nature (doubling in frequency) and human hearing is logarithmic in nature *REF?* – an exponential volume change is perceived as linear.

ADSR state changes are often triggered by a comparator and other switching ciruitry when the exponential function reaches a fraction of its final charge value *REF?*. Reaching 0.67 of the final value is a common threshold *REF?*. The decay and sustain states are often a single RC circuit *REF?*. *Doepfer schem?*

The Doepfer A-140 ADSR Envelope Generator is a classic envelope generator with a gate CV input, an envelope output, and a negated envelope output. It also has a retrigger input that allows the "on" trigger to occur again, reinitiating the attack phase, independent of the current phase of the envelope.

Digitally, an ADSR envelope generator can be implemented by a 5-state state machine (off, A, D, S, R). It can be designed to provide retrigger functionality. Retriggers can often be required in monophonic synths where a note can be played before a previous note is released, thus requiring a retrigger without reaching the release phase. For computational efficiency, an exponential LUT can be used.

1.3. A basic monophonic modular setup

[describe figure 1.3]

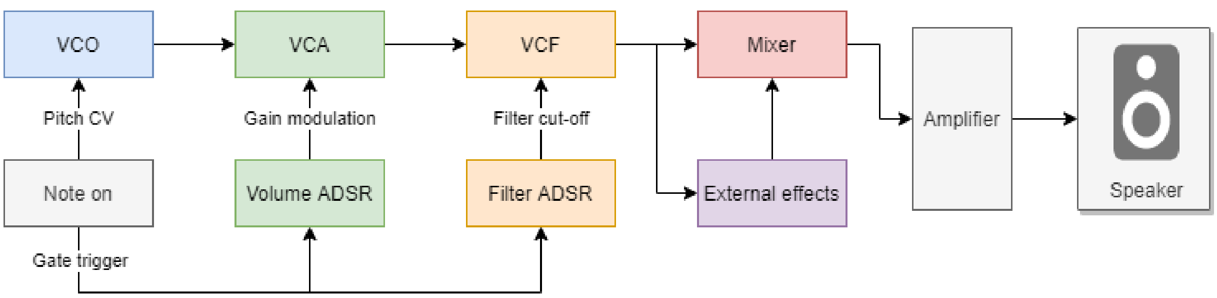


Figure 1.3: Block diagram of a very basic monophonic modular setup.

1.4. An overview of synthesis techniques

Most techniques are often combined in commercial products and operate in a similar way. The similarities, differences and operation principles will be explored in this section. This is will also explore why wavetable synthesis can be considered the most flexible and computationally robust technique.

1.4.1. Additive synthesis (Fourier synthesis)

The principle of operation is based on the harmonic series of a time signal:

$$y(t) = \sum_{n=0}^{\infty} a_n sin(\omega nt + \phi_n)$$

Various sinusoids are added together with different amplitudes and phases to produce a signal. The amplitudes and phases and frequencies may be time varying as well, which ties into physical modelling techniques that accounts for the time dependent timbre of most instruments. Modulating frequency directly ties in with FM synthesis.

Computing and adding many sinusoids in real-time can be computationally expensive – which can be reduced with a sinusoid LUT, as is directly done in wavetable synthesis. If time-varying amplitudes and phases are not present, additive synthesis can be completely replaced by wavetables (LUTs). If enough memory is present, caching the resulting waveform in a LUT would require less computation and thus have more robust performance. Serum has some functionality for creating wavetables using the principles of additive synthesis.

1.4.2. Subtractive synthesis

This technique is very simple and is possible in most synthesisers. It requires a harmonically rich source signal (generated by any means, such as direct computation, LUTs or analogue electronics) like a square wave:

$$y(t) = A \cdot sgn(sin(\omega nt + \phi_n))$$

It consists of odd harmonics with amplitudes:

$$a_n \propto \frac{1}{n}$$

The source signal is then passed through a filter to further shape the harmonics. Any filter can be used. Time varying filters with modulated parameters are usually prevalent.

This option is almost always present and/or possible to achieve in most synthesisers that offer filtering capabilities. Many products usually offer a selection of base waveforms which often includes most of or all the basic waveforms (sinusoid, triangle, sawtooth and square).

1.4.3. FM synthesis

This technique uses the same principles as FM for data communication, except in the audible frequency range. The resulting waveform is of the form:

$$y(t) = g(\omega(t))$$

where $g(\omega(t))$ is a periodic function with time-dependent frequency ω. This technique can produce unique and interesting results depending on the functions chosen for g and ω. Emulation of drum-like sounds such as toms and growling sounds often occurring in EDM genres are easily possible with this technique.

The choices offered for chosen for g and ω are product dependent but can often include the basic waveforms for g and ADSR envelopes and LFOs for ω. Multiple oscillators modulating each other's frequency, often in a coupled or recursive manner, is common, as in the stock Ableton VST plugin Operator, which is a FM-centric VST.

Many non-FM-centric synths also offer a vibrato feature, which requires the use of dedicated vibrato LFO that slightly modulates the source signal's frequency. This is present in VSTs such as Omnisphere 2 by Spectrasonics, which is wavetable and sample-based.

FM synthesis is often combined with wavetable synths such as the Serum and Massive VSTs. It is also easily achievable in modular synth setups since most oscillator modules allow for controlling their frequency with a voltage signal.

1.4.4. Physical modelling

This method involves simulating the sound source of interest. It is usually separated in continuous models for bowed or blown instrument or impulsive models such a struck or picked instruments.

A variety of methods can be used, such as IR modelling, analytical simulation (differential equations), frequency domain modelling as mentioned under additive synthesis, and waveguide synthesis such as the Karplus-Strong plucked string algorithm.

This type of synthesis is not relevant to the topic of this thesis.

1.4.5. Sampling

Sampling synthesis is the technique of using pre-recorded audio samples to reproduce sounds. An example would be to record every key of a piano at different volumes and then assigning a sample to trigger when conditions are met. The Kontakt player by Native Instruments is a popular sample player plugin into which third-party sample libraries can be loaded into to reproduce high-quality and realistic audio. High quality samples often take enormous amounts of effort to make, which results in a high commercial price point as can be seen in the Omnisphere 2 VST and the Spitfire Audio Kontakt libraries.

Recorded samples can also be manipulated to increase or decrease their pitch, allowing for a wide variety of options to the performer.

It is very similar to wavetable synthesis, where a predefined buffer (LUT) is used to generate sound. However, sampling often uses large buffers that are not necessarily intended to reproduce a periodic waveform (but sometimes do for continuous sound produced by instruments such as flutes), but instead a one-shot or partially looped ([dan worral?]) triggered signal, ideal for percussive instruments. Samples and wavetables can be manipulated and modulated in the same way.

This technique is computationally efficient but may require a large amount of memory to store the samples – often in the order of gigabytes, as for Kontakt libraries.

Another related synthesis technique is granular synthesis, where sampled audio is divided into "grains" which are treated like wavetables. They grains are looped, layered, and randomised to produce a soundscape often referred to as a "synth pad".

1.4.6. Wavetable synthesis

[Some basic WT synth descriptions]

1.5. Wavetable frequency conversion

[should be in a different chapter]

This section focuses on modelling wavetable sampling and determining the effects of linear interpolation and frequency scaling in the frequency domain.

Suppose we want to store a periodic signal $y[n]$ in a buffer consisting of 2^b samples where $b \in \mathbb{N}$. The buffer size choice being a power of 2 is arbitrary.

The fundamental frequency of the buffer as normalised digital frequency (cycles per sample) is thus 2^{-b}.

If we require the buffer to store H harmonics including the fundamental, we use the Nyquist frequency to determine the maximum number of harmonics the buffer can store.

$$H_{max}2^{-b} = 0.5 \Rightarrow H_{max} = 2^{b-1} \tag{1.1}$$

Therefore, a bigger buffer size results in the ability to store more harmonics.

Now we consider the follow process shown in figure 1.4 to scale a sampled signal's frequency ($y[n]$) by a factor of $\eta = \frac{M}{L}$.

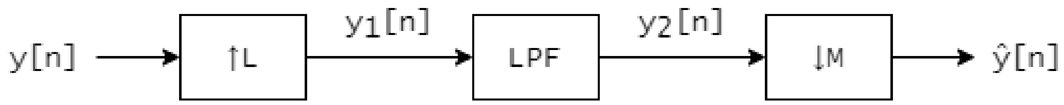


Figure 1.4: Frequency scaling using upsampling and downsampling

With reference to figure 1.4, the signals in the frequency conversion process are related as follows:

$$y_1[n] = y[n]_{\uparrow L}$$

$$y_2[n] = y_1[n] * h[n]$$

$$\hat{y}[n] = y_2[n]_{\downarrow M}$$

As per the scaling theorem, the frequency axis of $DTFT\{y_1[n]\}$ contracts by a factor L. A LPF is used to remove unwanted copies of the spectrum, which would result in aliasing if the frequency axis is expanded by a factor M through downsampling.

There are many choices for the LPF, but for arbitrary frequency scaling, M and L will become large to achieve close approximation for any real number. Thus, the obvious filter choice type is FIR, which can be easily computed per sample by simply looking at next and previous samples stored in the wavetable LUT. Upsampling and downsampling happens implicitly for frequency scaling using wavetables (see subsection **??**), and is only used to predict the effects of interpolation.

The simplest FIR choice would be the linear interpolator, which can characterised by its impulse response:

$$h[n] = tri(\frac{n}{L}) = \frac{1}{L} rect(\frac{n}{L}) * rect(\frac{n}{L})$$

Taking the DTFT,

$$H(f) = DTFT\{rect(\frac{n}{L})\}^2 = \frac{sin^2(L\pi f)}{L sin^2(\pi f)}$$

Note that $H(f) = 1$ for $L = 1$ and has zeroes at $f = \frac{p}{L}, p \in \mathbb{N} \setminus \{0\}$. Figure 1.5 shows the effect of the linear interpolation process in the frequency domain. The bandwidth of $y[n]$ is represented as a triangular pulse.

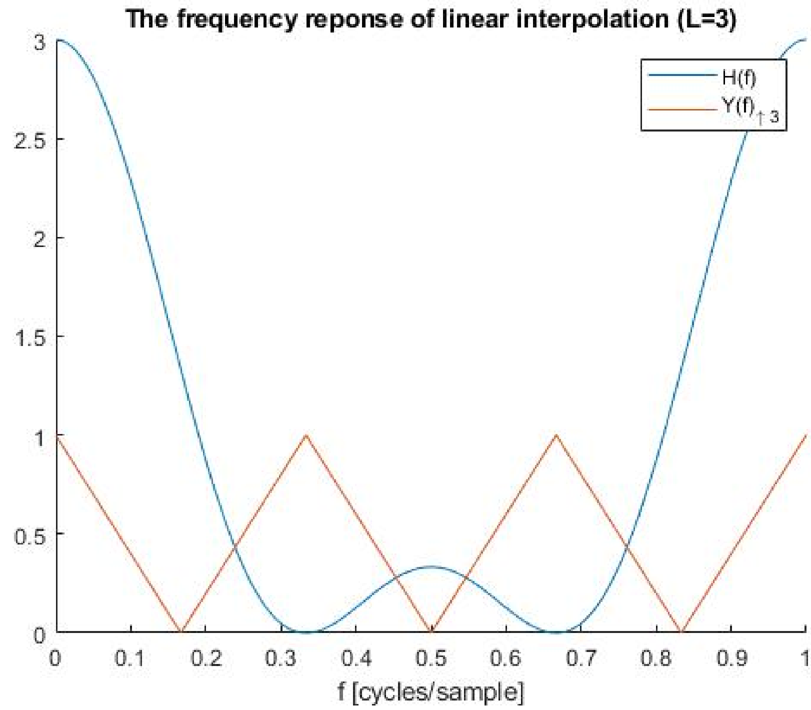


Figure 1.5: The effects of linear interpolation in the frequency domain.

Better, but more computationally intensive versions of this filter, can be obtained by convolving the $rect(\frac{n}{L})$ function an arbitrary number of times with itself, which will correspond to increasing the exponent in the frequency domain, thus reducing the amplitude of the side-lobes of $H(f)$. However, after convolution, the width of $h[n]$ increases, requiring more samples to be taken into consideration when interpolating. This is a naive approach, which also attenuates higher frequencies, and is out of the scope of this thesis. The linear interpolator only requires 2 samples in the wavetable LUT. This will be discussed later in subsection **??**.

The linear interpolation filter has a fixed cutoff which is a function of L. When the frequency is scaled up, i.e. $M > L$, the original spectrum with a full bandwidth will alias since the linear interpolation filter does not account for this. To combat this issue, we can use a variety of buffers for a single waveform, each containing a different number of harmonics, thus bandlimiting the signal to combat aliasing for when $M > L$. Assuming that the waveform is bandlimited to B cycles/sample, we need to ensure that

$$\frac{M}{L}B \leq 0.5$$

as per the Nyquist criterion. Furthermore, B is determined by the number of harmonics h, i.e. $B = h2^{-b}$.

$$\Rightarrow h2^{-b} \leq \frac{L}{M}0.5 \Leftrightarrow h \leq 2^{b-1}\frac{L}{M} = 2^{b-1}\frac{1}{\eta} \tag{1.2}$$

We can find the maximum number of harmonics (including the fundamental) h_{max} for a given digital frequency f_0. First, we find the required frequency scaling factor:

$$2^{-b}\eta = f_0 \Rightarrow \eta = f_0 2^b \tag{1.3}$$

Substituting equation 1.3 into equation 1.2, we find

$$h \leq \frac{0.5}{f} \Rightarrow h_{max} = \frac{0.5}{f_0}$$

If we use K buffers to store a varying number of harmonics, with the buffers indexed with $i = 0, 1, ..., K$, and store 2^{i+1} harmonics in each buffer, we are restricted to $K_{max} = b-1$ as per equation 1.1. The exponential number of harmonics accounts for the exponential nature of octaves (a doubling in frequency to reach the next octave).

We can find the closes index of a given digital frequency f_0 which will contain the maximum number of harmonics (h_{max}), without aliasing, stored in the buffer i as follows:

$$i = min(\lfloor log_2(h_{max})\rfloor, K) = min(\left\lfloor log_2(\frac{0.5}{f_0})\right\rfloor, K)$$

Chapter 2

Summary and Conclusion

Bibliography

Appendix A

Project Planning Schedule

This is an appendix.

Appendix B

Outcomes Compliance

This is another appendix.