Beat Extraction and Analysis: Determining the Tempo from an Audio Signal

Matt Aldrich
May 11, 2002
Silliman College, 2004
Electrical Engineering
EENG 236b
Advisor: Peter J. Kindlmann

Abstract

Every piece of music has an underlying pulse that unifies the distinct sounds, textures, and phrases found within each measure. This is what is known as a tempo. This paper describes the process in which a "beat detector" circuit provides a visual interpretation of the tempo. Included in this discussion are the motivations for having such a circuit, the design and implementation of the circuit, and the design's advantages as well as drawbacks.

Section 1
Introduction

Objective
To design and produce a beat tracking unit that processes the quarter note beats of music and provides a visual representation.

Background
Within any given genre of music, there is an underlying pulse that unifies the individual timbres, intonations, and beats. It has long been the ability of humans to successfully locate and identify this recurring pulse regardless of changing tempos or meters. This process, commonly referred to as "foot-tapping," comes to humans without much thought or computation. The tempo at which the foot tapping occurs is called the pulse or beat. In a broader sense, this "foot tapping" can be analyzed and processed. This information is commonly referred to as the tempo.

Applications
Uses for a visual tempo of music are quite varied. The tempo is used to classify and label music. The tempo may be further manipulated to trigger specific events strictly related to the beat of the song. An intelligent algorithm along with the correct acquisition of the quarter notes will yield the Beats Per Minute (BPM) of a song.

Intended Application
The concept of tempo and tracking the BPM is bound tightly to various genres of music. Such a genre is Electronica[1]. The strong "bassy" pulse within this type of music makes it a very good candidate in beat tracking because the central bass pulse can easily be filtered and separated from the music. From a disc jockey's standpoint, the ability to be aware of and control the BPM is important. By using two turntables (with +/- 10% pitch adjust) and a flashing LED to mark the pulse, the disc jockey is able to effectively and accurately mix together two records of different BPMs and produce the illusion that one song simply "fades" into the other. Having an LED flash on the pulse provides a visual interpretation of the beat, allowing the DJ to effectively mix two records without hearing the music, or more realistically, if the music is difficult to hear.

Motivation
Currently, beat triggering devices are implemented into audio mixers and some DJ equipment. However, they are isolated, fixed devices. This projects aims to create a very robust detection scheme that leaves a large margin of adjustment open to the user. Most commercial designs have a wide range of triggering problems. The result is a flashing LED that is not very accurate. In this project, the detection scheme requires some input from the user about the detection threshold. This leads to an accurate LED flashing. The remaining sections of this report include the methods I used to develop the hardware, the results of my work, a discussion of the data, and some concluding statements. The final pages constitute the appendix that contains the circuit schematic as well collected data.

---

[1] Electronica is genre of music that combines both digital sounds and acoustic ones to deliver a type of music that is highly energetic, simply organized, and very dance oriented. Within the genre of Electronica are smaller genres such as House, Trance, and Drum n' Bass. These styles all have similar traits and place an emphasis on tempo and rhythm.

Section 2
Methods & Results

It is easiest to describe the processes used to extract the beat by breaking down the system into blocks. First, an acceptable form of music is selected. As mentioned previously, this project detects the quarter note pulse. For this reason, electronic music with a distinct beat is needed. The system consists of four primary blocks. The block diagram below describes the process.

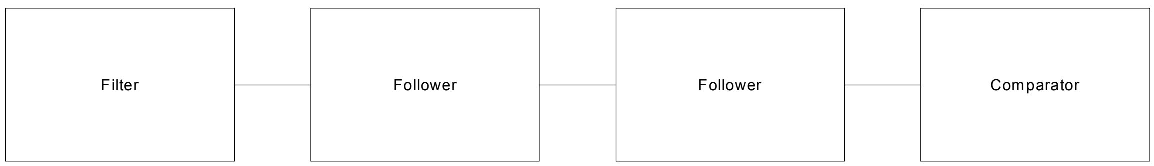


Figure 1 Design Outline

A more detailed chart of the beat extraction process is as follows. These blocks represent each individual circuit of the project.

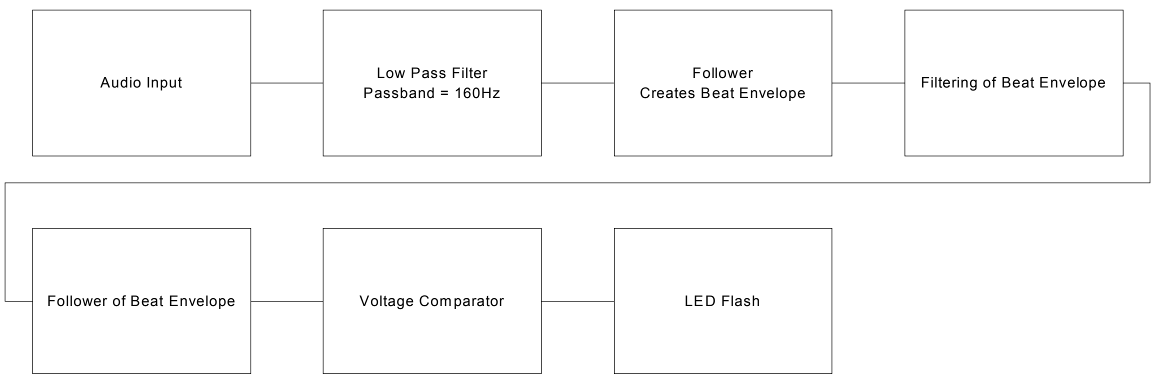


Figure 2 Actual Blocks

It is easiest to describe how each element was conceived, tested, implement by discussion each particular block, beginning with the filter.

<u>Low Pass Filter</u>
A robust filter design is necessary to ensure the stability of the system. In fact, many hours of prototyping simply went into the design of the filter. The first decision was active versus passive. An active filter network is preferential to a passive network for several reasons. The low input impedance of the op-amp prevents excessive loading of the source and the low output impedance prevents the filter from being affected by the load it is driving. Quite simply, an active filter is more practical. In this case, a second order active filter was selected. The filter in this configuration is a Sallen-Key low pass. The two RC networks provide the filter with a quick -40 dB/decade roll off above the critical frequency.

<u>Low Pass Filter</u> - *Determining the proper pass band*

This is an integral decision in relation to the entire project. Fortunately, the music has a very distinct sound. Much of this music is synthesized and as a result, each instrument generally has its own distinct frequency range. The quarter note pulse is produced by a synthesized bass drum. Although different songs have different sounding bass drums, they encapsulate a similar frequency range. At frequencies 200 Hz and below, much of the other noise if filtered out and the "bassy" elements of the song are left. However, it is possible to select too small a pass band. In the case of a very narrow pass band, too much of the bass is attenuated, leaving very little to analyze. A frequency cutoff at 160Hz attenuates many frequencies that produce false triggering. Also this frequency allows for distinguishable recognition of the beat in comparison to the other sounds. (Please refer to ***Appendix A***: Filter Data for a complete graphical analysis of filter design).

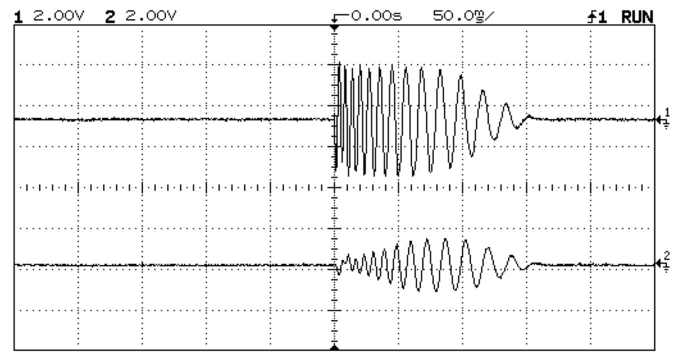


Figure 3 – 160Hz Filter
The top display is the actual bass drum waveform.
The bottom display is the filtered waveform.
Notice the high frequency attack of the bass drum is filtered out.

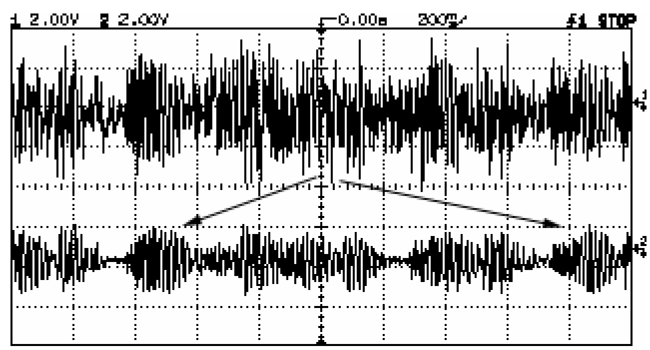


Figure 4 – Filtering of Music
Arrows are pointing to bass drum impulse.

Voltage Follower 1

The second stage is essentially a peak detector. The peak detector takes input from the low pass filter and creates a waveform from it. The general idea behind this filter is as follows: The first operational amplifier with the diodes detects the peak of the circuit. The second operational amplifier is a simple follower. The detector portion "remembers" the peak value of a signal. When a positive voltage is fed to the non-inverting input the output voltage of the op-amp forward biases the diode and charges up the capacitor. This charging last until the inverting and non-inverting inputs are at the same voltage, which is equal to the input voltage. When the non-inverting input voltage exceeds the voltage at the inverting input, which is also the voltage across the capacitor, the capacitor will charge up to the new peak value. Consequently, the capacitor voltage will always be equal to the greatest positive voltage applied to the noninverting input. Once charged, the time that the peak detector "remembers" this peak value is based upon an RC time constant. However, the capacitor cannot remember forever and thus the capacitor will slowly discharge towards zero. To minimize this rate of discharge, a voltage follower can be used to buffer the detector's output from any external load.

Voltage Follower 1 - *Determining the RC Constant*
This is another crucial step in an efficient design. There are several design considerations that must be taken into account. First, since the music is very linear and has a distinguishable quarter note pulse, the RC constant must quick enough to capture the entire pulse. The second consideration is that the bass drum impulse consists from both high frequencies and low frequencies. The filter attenuates the high frequencies, however the goal of the first follower is to create an accurate waveform, therefore, the RC constant must be fast enough to get all the peaks of the bass drum impulse. A time constant of
5.6 x 10^(-7) s was used.

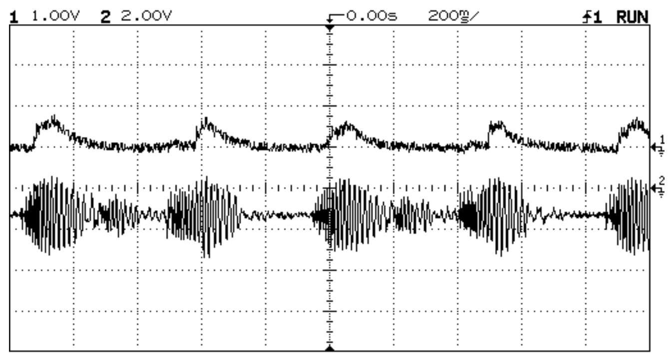


Figure 5 – Peak Detection & Follower

Voltage Follower 1 - *Smoothing the Output* [labeled in block as '**filtering beat envelope"**]
As Figure 5 shows, the design is quite robust and detects all of the peaks of the bass drum impulse. However, the output consists of many tiny saw tooth waves. To circumvent this phenomenon, a simple RC filter is attached to the output. The attached RC network has a time constant of 3.9 x 10^(-8) s. The resulting output is uniform and streamlined. Another benefit of a smooth output is a reduction of false triggering (see Comparator discussion).

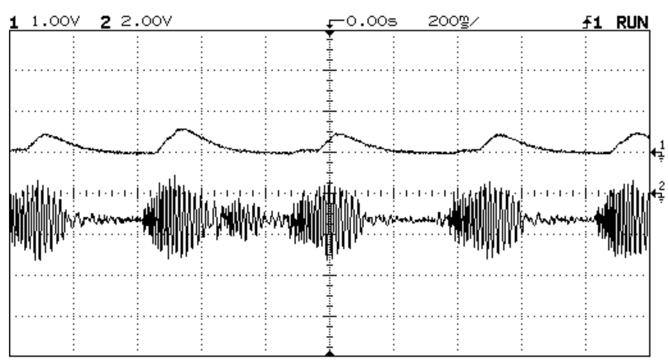


Figure 6 –Same Input with RC network

Voltage Follower 2
The idea behind the second follower is to create a threshold necessary for comparison. The circuit theory and configuration is exactly the same as the previous follower. However, the RC constant is larger. In this case it is 1.12 ms. The design idea behind this follower is a somewhat different then the first. The first follower simply "rides" the bass drum impulse, whereas the second follower "rides" beat to beat. The second follower must have a slow decay because it

needs to remain accurate for comparison, however, it must not have such a slow time constant that it will miss other bass drum impulses.

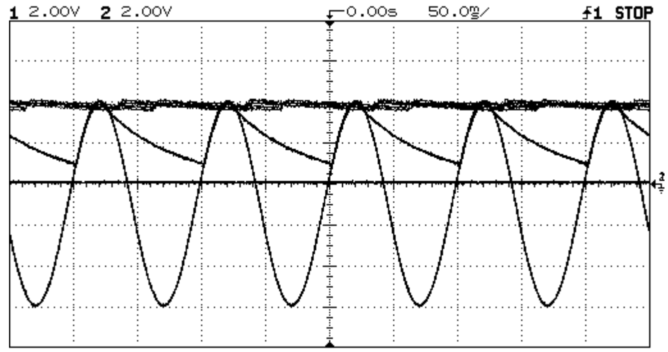


Figure 7 – Input Signal (sine Wave 5VPP@ 10Hz) & Followers 1, 2
Follower 1: Fast decay
Follower 2: Slow decay

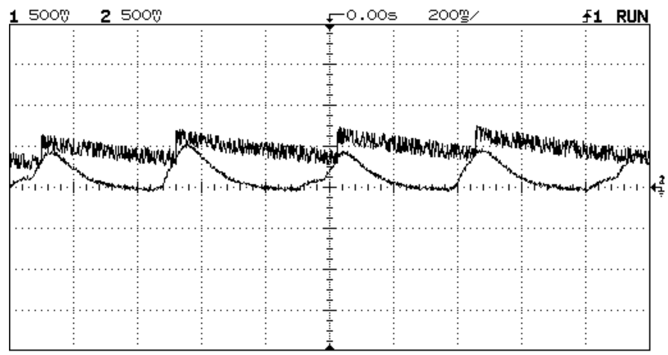


Figure 8 – Audio Imput
The second follower "rides" on top of a smooth quarter note impulse
Notice that the bass drum impulse has a decay ≈ 200 ms

<u>Comparator</u>

The goal of the comparator is simply to output a square wave when there is a difference between two input signals. A 311 operational amplifier provides this function. The output from the second follower is the threshold. That means that data that is greater than the "window" created by the threshold will trigger the op-amp to output a square wave. In this design, approximately 2/3 of the second follower is taken. It is then connected with a 10K potentiometer that provides a "variable threshold." This helps prevent false triggering. Finally, this output can used to trigger almost anything. In this case, it is an LED.

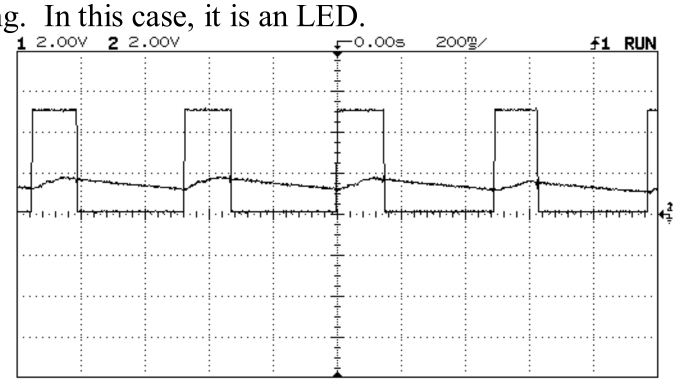


Figure 9 – Triggering of the Comparator
Step Function indicates a peak is detected.

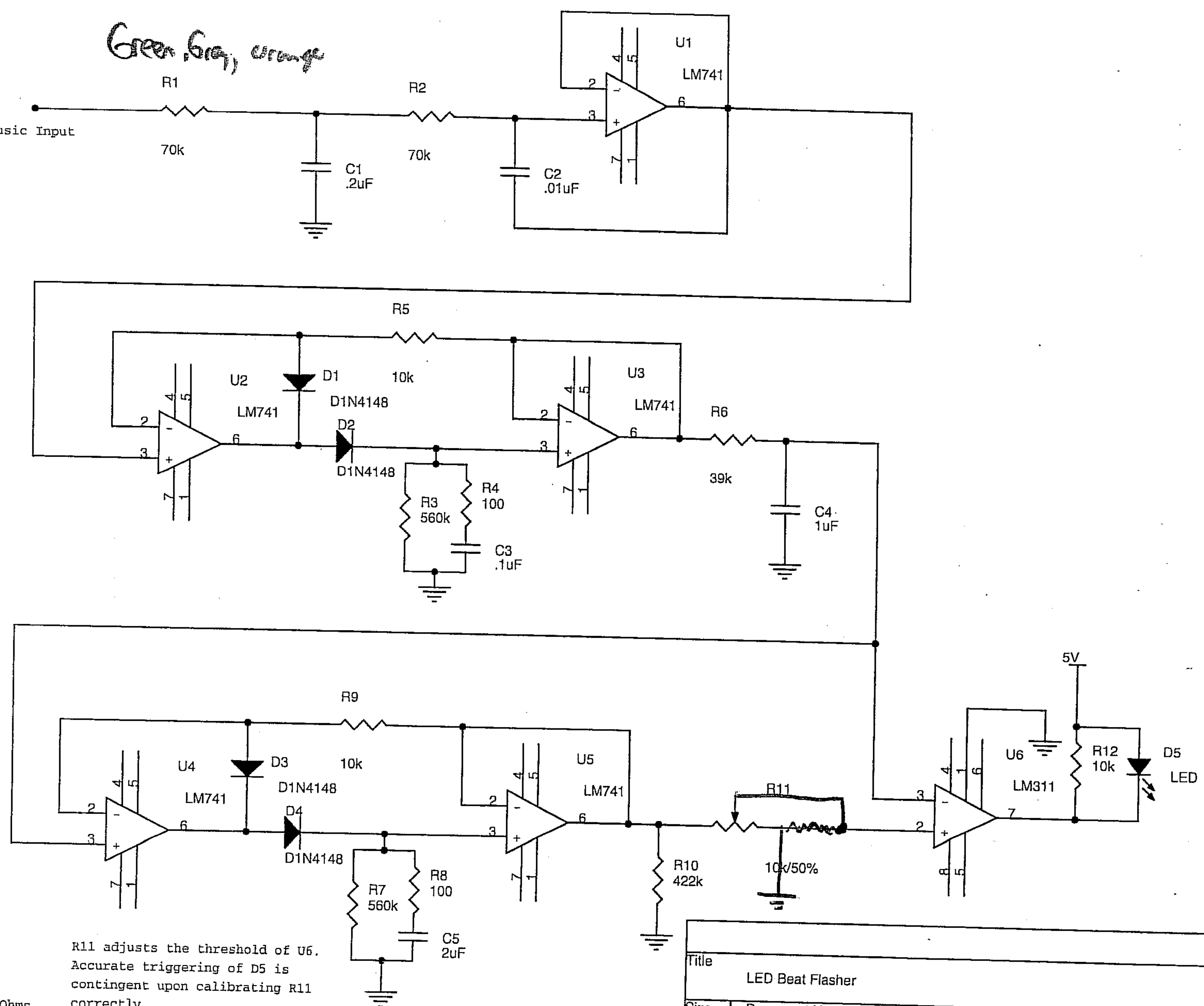
Section 3
Discussion

In general, the system's response is quite accurate. Once the input is connected, it is sometimes necessary to adjust the threshold. This is required because the audio signal can be of varying amplitude. The system is able to detect the quarter note beat more than 95% of the time and appears more stable than commercially available detectors. However, there are some cases where the LED will flash with two or more quick bursts. This is the system's response to two large beats of similar amplitude that occur quickly in a row. This occurrence is easily preventable however. By connecting the output to a one shot timer, the double triggering effect can be eliminated. This is one way of improving the design. Another novel idea is to re-design the low pass filter. It is possible that a lower cut off frequency eliminates this double triggering effect. In some cases where no solid beat exits, the circuit will find peaks regardless. This also results in false triggering. However, the intended application of the circuit is in conjunction with an audible signal, therefore the user knows that there is beat. Overall, the circuit provides the user with a distinguishable result.

What makes this circuit exciting is that there is room for many customizations. It terms of flexibility, the user may want access to a variable pass band filter. This is in addition to the variable threshold control the user currently has. Now that the initial filtering design is implemented, this project has the potential to branch out into many different directions. Besides a visual flash of the beat, it is possible to use a programmable micro controller to display the actual beats per minute of the music. In this case, the user has a visual cue of the beat and an explicit numerical calculation. This design could also be employed in a large-scale visual circuit, where preprogrammed patterns could be synced to the tempo and displayed visually via lights and strobes.

Section 4
Conclusion

In any robust design, the first implementation is probably not the most stable. It is an ever evolving process where expanding, rethinking, and redesigning the system is crucial. In this case, the circuit performs very well. However, there are ways to improve its stability and overall design. What makes this circuit unique is that it has a wide range of applications. In terms of completion, this design is done; however the addition of a BPM monitor is a novel and proprietary idea worth of investigation.



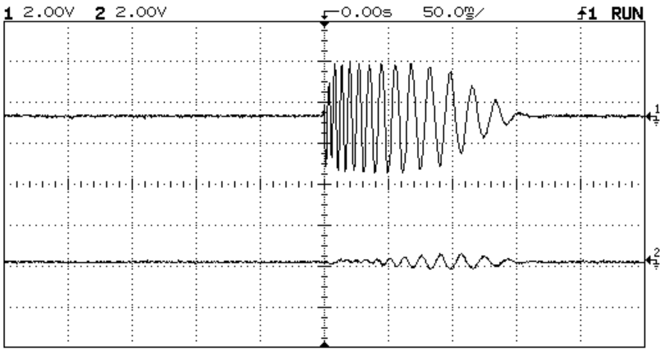
Op-amp PWR +/- 15V
U1-U5 pin 7=V+ pin 4=V-
U6 pin 8=V+ pin4=V-
U6 pin 1 MUST be GND.
All resistor values are in Ohms.
All capacitor values are in Farads.

R11 adjusts the threshold of U6.
Accurate triggering of D5 is
contingent upon calibrating R11
correctly.

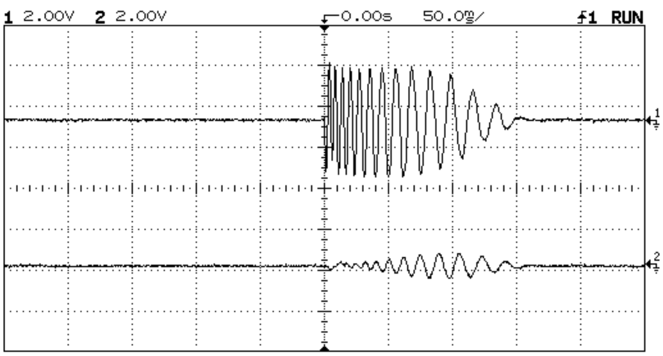
Title	LED Beat Flasher	
Size A	Document Number: MHA236B Matthew Aldrich, Yale University, Silliman College matthew.aldrich@yale.edu	Rev B
Date:	Tuesday, June 11, 2002	Sheet 1 of 1

Appendix A

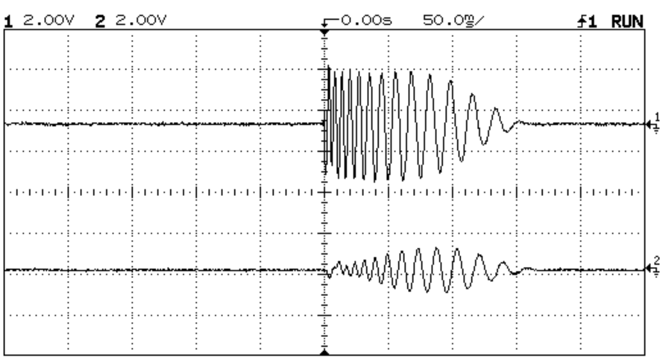
<u>Filter Data</u>



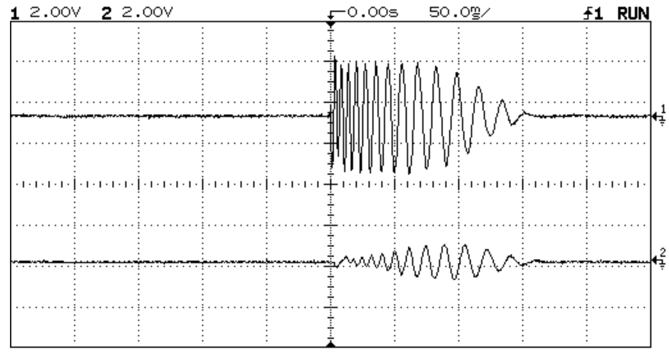
Preliminary data using Krohn-Hite Filter 60Hz pass band. High Frequency origins of bass drum impulse are too filtered.



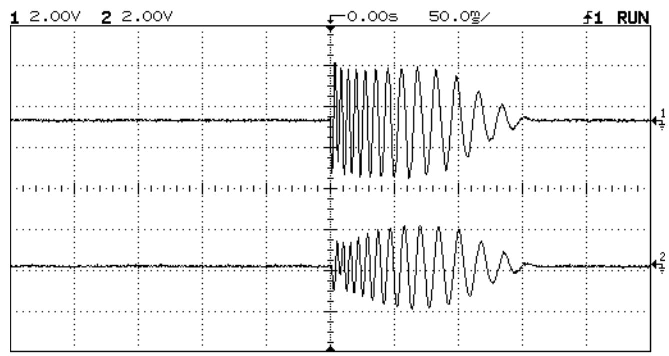
Krohn-Hite Filter 60Hz Pass band.



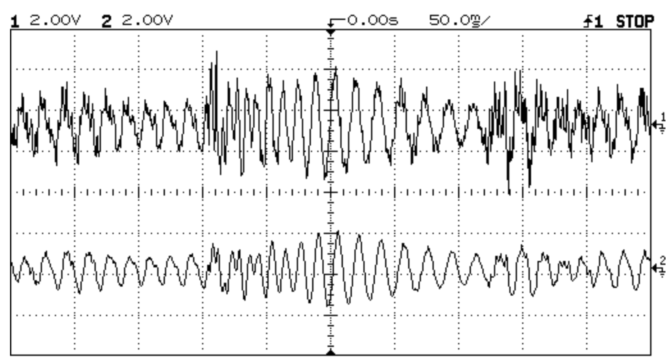
Krohn-Hite Filter 100Hz Pass band.



Krohn-Hite Filter 140Hz Pass band.



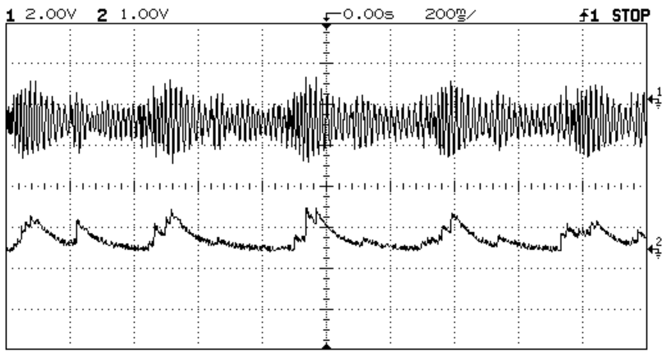
Krohn-Hite Filter 300Hz Pass band.



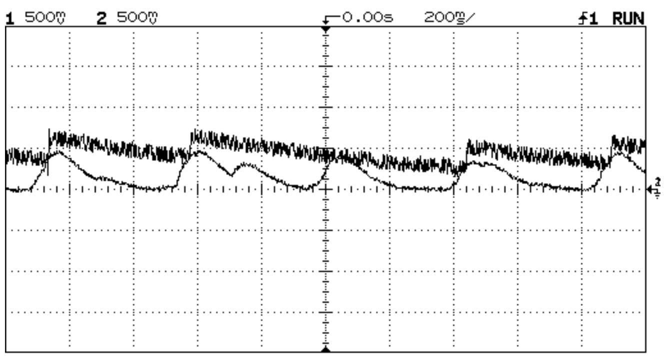
Music Input. The bass drum impulse spans [-100:100]

Appendix B

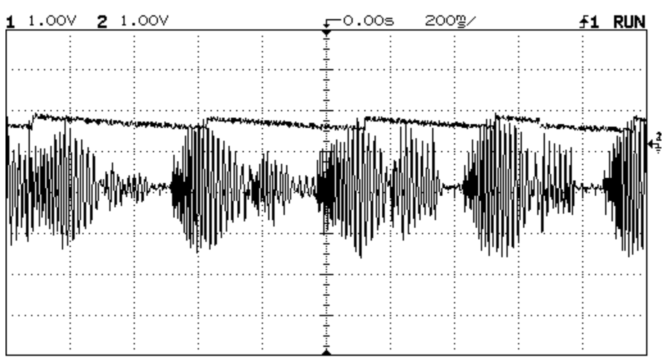
Follower, Peak Detector Data, & Output



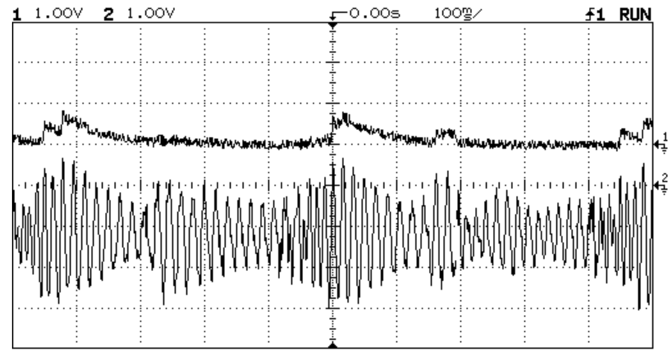
Follower 1 detecting beats of music input, note jagged edges on beats.



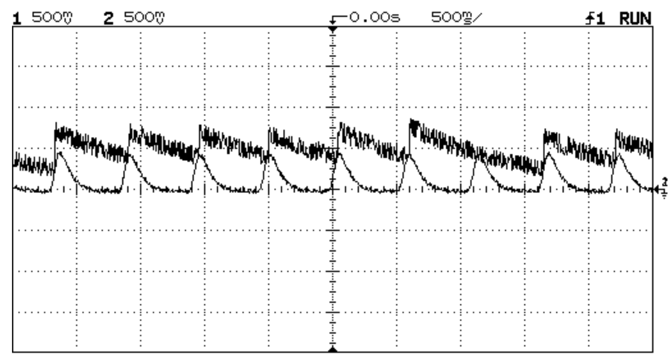
Second stage follower "riding" peaks of first. Note the double peak at ≈ -250ms.



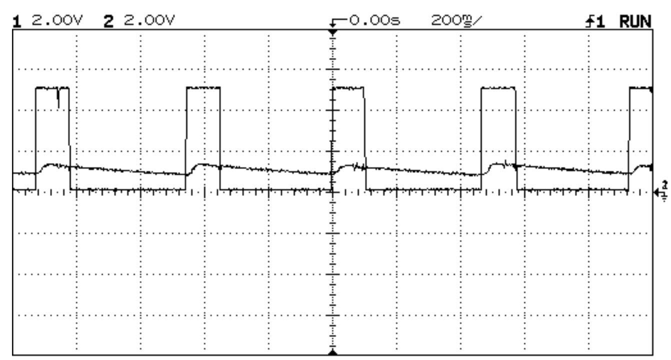
Output of low pass filter & second follower.



Another example of filter & second stage follower.



Both followers showing "missed" peak. This may or may not cause a triggering problem.



LM311 Output during beat detection.