

MacPOD / WinPOD User Manual

Version 2.0.4

Vancouver, November 10, 2009.

© 2009, C. Rolfe

info@thirdmonk.com

Quick Start	3
Example: Time-stretching.....	3
Example: Pitch-shifting	3
Auto-stretch	4
Granular Synthesis.....	5
Granulation of Sampled Sound - A little history.....	5
Synchronous vs. Asynchronous Granulation.....	6
Grain Parameters.....	7
Example: Controlling Grain Density	8
Reson / EQ and Spectrogram	8
Reson.....	8
Example: Flanger	8
EQ.....	9
Sonogram	9
Audio Driver	10
Recording-to-disk.....	10
Presets	11
MIDI.....	11
Supported file formats.....	11
Appendix A: MIDI Implementation.....	12

The following document explains MacPOD/WinPOD's features and operation.

Quick Start

To start using MacPOD, click **OPEN FILE** to open a WAVE (.wav) or AIFF (.aiff) sound-file. Click **play** to start sound playback.



Figure 1: Transport controls

The **main volume faders** control left and right output levels.

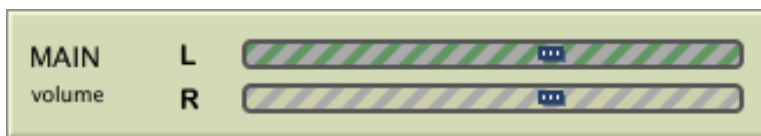


Figure 2: Main volume faders

Example: Time-stretching

The **tempo slider** controls the evolution of the sound without affecting its pitch. A tempo of 50%, for example, plays the sound at half-tempo while preserving the original pitch. A rate of 0% freezes playback at the current location. A rate of -100% plays the file in reverse. Note, you can also click on the tempo number box and type in a new tempo value directly *including values outside the slider's range*.

Adjust the tempo to begin exploring granular time-stretching (granulation). Longer time stretches (0%-20%) of speech sounds or other richly varying sound events can be especially mesmerizing. Time-compression can be achieved by typing in a value higher than 100%.

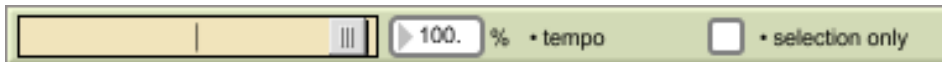


Figure 3: Tempo slider

Example: Pitch-shifting

The **pitch slider** changes the pitch independently of tempo. A pitch value of 100% plays back at the original pitch. A pitch value of 50% plays the sound at half-pitch (one octave lower); a pitch value of 25% plays the sound at quarter-pitch (two octaves lower). Pitch-shifting is limited to between 25% and 400%.

The **pitch range** number box varies the pitch of individual grains randomly within the specified percentage. A pitch value of 100% plus a pitch range of 6%, for example, results in grains ranging from the original pitch to about one semitone higher.



Figure 4: Pitch slider

Pressing the **stop** button ceases all audio output. Alternatively, the **pause** (the double-bar ||) pauses playback at its current position while continuing to output sound. When playback is paused, you can freely scrub to a new position using the **position slider** (blue bar overlaid on top of the waveform display).

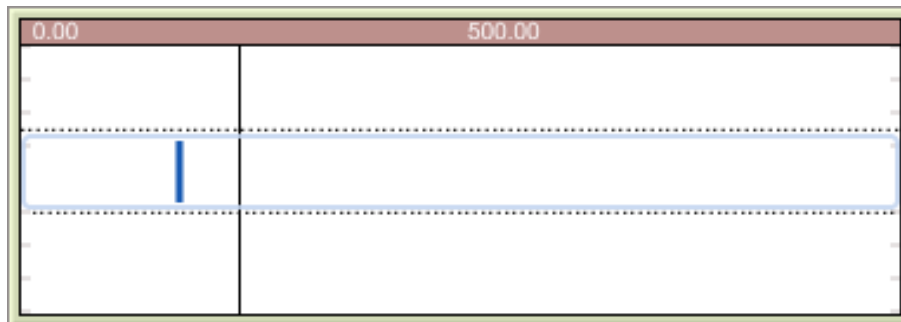


Figure 5: Waveform display and position slider

Clicking and dragging anywhere in the waveform display outside the position slider lets you select part of the sound-file. Toggle the **selection only** option to play just the selected portion.

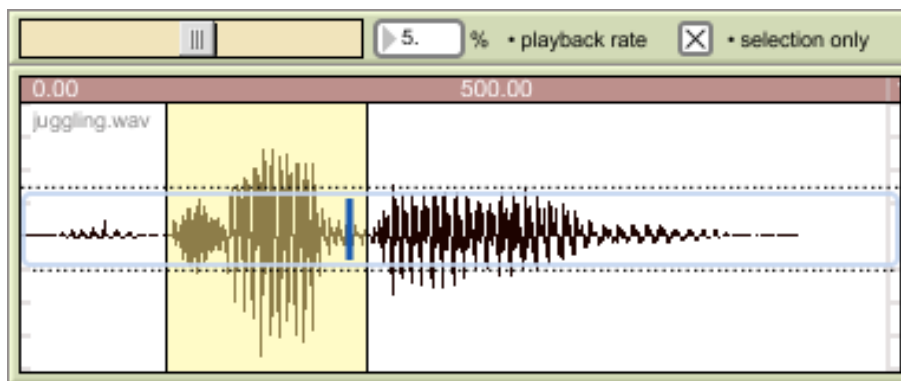


Figure 6: Selection only

Auto-stretch

The **auto-stretch option** automatically adjusts the tempo slider so that transients are stretched less than stable sounds [a *transient* is a short period of higher energy such as consonants within speech or the attack portion of a sound]. The **stretch slider** controls the degree of time-stretching. The **sensitivity slider** determines how responsive the detection circuit is to transient energy – generally, more sensitivity means less stretching of transients. The **lookahead number box** sets how far ahead (in milliseconds) of the playback head the detection circuit scans. The default of 100 milliseconds works well in most cases.

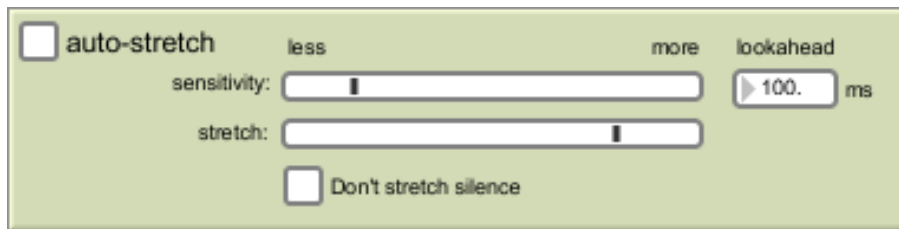


Figure 7: Auto-stretch

Granular Synthesis

First, a little history of granular synthesis and time-stretching.

Granular synthesis was first suggested as a computer music technique by Iannis Xenakis (1971) and Curtis Roads (1978)¹. Barry Truax implemented real-time granular synthesis for the first time in 1986 on the POD system developed at Simon Fraser University.

The basic unit of granular synthesis is the **grain**, a short (30-120 ms) particle of sound created by windowing a fixed waveform. A simple grain generator can be constructed using a tone generator and an amplitude envelope generator. Typical control parameters include:

- pitch
- duration
- delay (between grains)
- spatialization

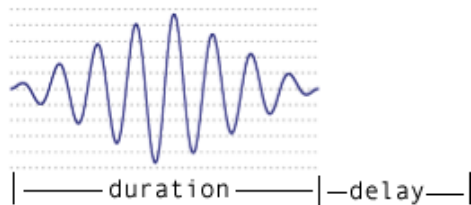


Figure 8: Single grain

MacPOD/WinPOD supports all of these parameters, although left and right stereo assignment of grain streams is automatic. The program does, however, preserve stereo left and right when granulating stereo files.

Granulation of Sampled Sound - A little history

In the early 1950's, Dennis Gabor (working with Pierre Schaeffer, the originator of many Musique Concrete and tape studio techniques) developed the Tempophone, a spinning playback head that, when clamped onto a reel-to-reel tape machine, allowed independent control over pitch and duration. The perceived playback rate was determined by the tape transport speed *plus* the rate of rotation allowing the user to scrub slowly through a sound event without causing a change in pitch.

Playback of sampled sound on a computer is achieved using grain generators, or, **grain streams**.

¹ - <http://www.sfu.ca/~truax/gran.html> (accessed Nov. 1, 2009).

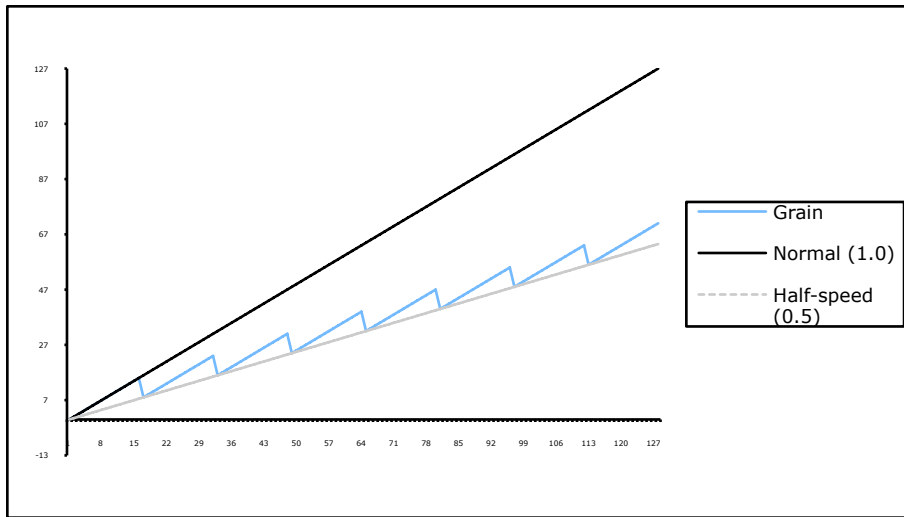


Figure 9: Granulation (half-speed; original pitch)

Fig. 10 shows the spinning playhead technique from the perspective of playing a sound file. The grain generator begins at the start of the file, plays a short segment, then jumps to an earlier position, and plays forward, and so on, repeating portions of the sound.

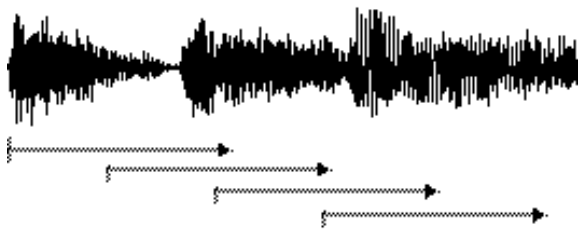


Figure 10: Granulation (half-speed; original pitch)

Each repositioning, however, creates discontinuities that might cause audible clicks, so grain generators also have windowing (fadein/out) functions.

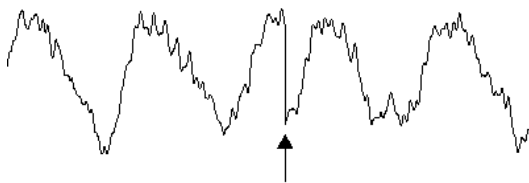


Figure 11: Discontinuity at grain edges - needs windowing

Synchronous vs. Asynchronous Granulation

Multiple grain generators (**streams**) can be used to smooth the granulation effect. When stream envelopes operate in-phase the streams are said to be **synchronous**, otherwise they are **quasi-synchronous**, or **asynchronous**.

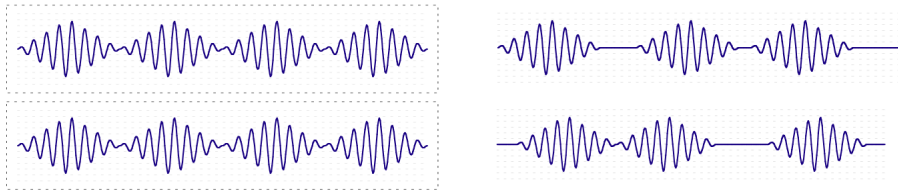


Figure 12: Synchronous vs. asynchronous grain streams

When grain streams are synchronous, amplitude modulation and intermodulation effects become quite pronounced (which sounds like a mechanical resonator). By adding a certain amount of randomness to the grain durations or delays, the overlapping effects becomes less mechanical and less apparent.

MacPOD/WinPOD uses a trapezoidal window, as did the original POD GSAMX system.

Grain Parameters

When the tempo slider is set to 100% (normal speed), you hear the sound-file at its original speed. As you move the tempo slider toward 50% or lower, time-stretching occurs and you begin to hear the grain streams coming into effect. Now MacPOD/WinPOD's grain parameters become important.

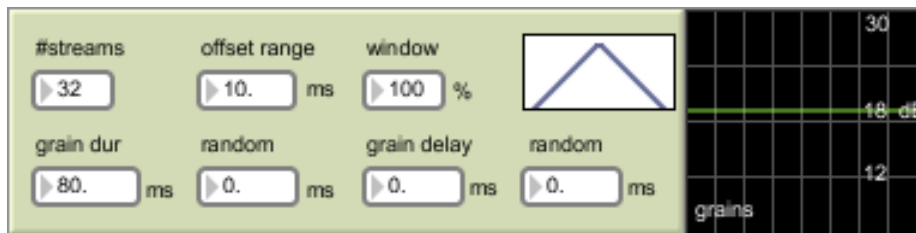


Figure 13: Grain parameters

The first parameter, **#streams**, simply turns grain streams on and off. Typically, you'll have all the grain streams on but for lower density effects you may want to use fewer streams.

Grain dur (short for **grain duration**) controls the size of each grain in milliseconds². Grain durations below 50 ms. tend to muddy lower frequencies – 50 ms. (1/20th of a second) is a very short time, not enough for the ear to hear and recognize lower frequencies, which have longer wavelengths than high frequencies. Grain durations above 120 ms. or so begin to result in audible echo or delay effects.

Grain delay controls the time (in milliseconds) between grains in each stream.

Both **grain dur** and **grain delay** can be varied by a **random** amount to avoid mechanical-sounding modulation effects caused by synchronous grain streams. Because each stream is varying randomly, they rarely sync up. The **grain amplitude** display to the right of the grain parameters will indicate whether the streams are synchronous or asynchronous. When the green line is straight, the streams are synchronous. When the green line is jumping around, the streams are asynchronous.

Offset range also adds randomness by varying each grain stream's read position by a random amount. Even when grain streams are operating synchronously, a small amount of offset range can be used to lessen the mechanical-sounding resonance; larger offset ranges (several 1000 ms.'s) can be used to jump around in the sound-file, a sometimes interesting effect.

² A millisecond is 1/1000th of a second.

Finally, the **window** parameter controls the shape of the trapezoidal window used. Be cautioned, however, that a 0% window may result in clicks at grain edges.

Example: Controlling Grain Density

Another useful granular effect (in addition to time-stretching and pitch-shifting) is achieved by actively varying grain density.

Begin by opening a WAVE (.wav) or AIFF (.aiff) file. Speech works well for this example, but any file of a few seconds duration will do.

Set the **tempo slider** to 0% (freeze).

Set the **#streams** parameter to **1** and listen to the single grain stream in the right channel. If you don't hear anything, try clicking and dragging on the **position slider** until you find a patch of sound to work with.

Now try varying the **grain delay** to hear the effect. Longer delays between grains means fewer grains, in other words, lower grain density.

Set the **#streams** parameter to **2** and listen to a second grain stream in the left channel.

Try varying the **grain delay random** parameter to hear the grain streams become asynchronous.

Lastly, increase the **offset range** parameter to listen to grains from different positions in the sound-file.

After playing around with 2 streams for a while, slowly increase the **#streams** to 32. More streams means more grains, in other words, higher grain densities.

As a general tip: randomness, whether from grain dur random, grain delay random, offset range or pitch range softens or defeats amplitude and intermodulation effects.

Reson / EQ and Spectrogram

MacPOD/WinPOD's equalizer section contains two sections: a resonator filter (**reson**) and a parametric **EQ**, containing a low / mid / and high filter. Each is explained below.

Reson

The **reson** filter allows you to emphasize certain frequencies using a variable **Q** bandpass filter. At high **Q**'s, especially with **feedback** and/or **feedforward** added, the resonator begins to *ring*. **Rate** and **mod** automatically vary the **reson freq** to achieve dramatic flanging and chorusing effects.

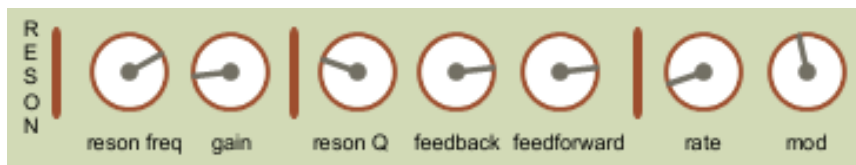


Figure 14: Reson filter

Example: Flanger

To demonstrate the **reson** filter, open a WAVE (.wav) or AIFF (.aif) file. Set the parameters as follows (or, experiment on your own):

reson freq = 1154 Hz

gain = .286

Q = 23%

feedback = 81%

feedforward = 81%

rate = 0.24 Hz

mod = 20.5%

The effect is a familiar flanger/chorus.

EQ

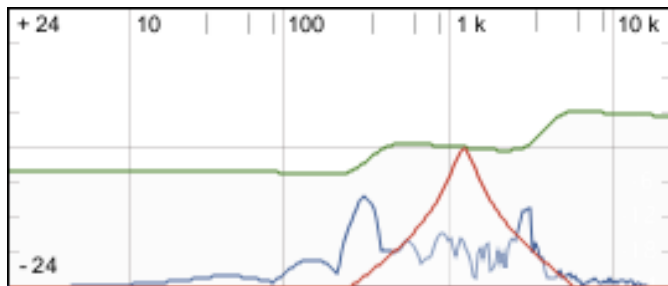


Figure 15: EQ display and spectrogram

The amplitude/frequency response of both the **reson** filter and the parametric **EQ**, as well as a spectrogram of the resulting (post-EQ) signal are displayed above the controls. It is a relatively quick matter to set the EQ to boost high frequencies, for example, whether using the reson filter, parametric EQ or both. Note that the parametric MID freq also provides a Q control to change the bandwidth of the middle parametric EQ.



Figure 16: EQ section

Additionally, the reson/EQ section contains an option to bypass both the reson filter and EQ. There's also a **clear filter** button to reset all filters to neutral values.

Sonogram

In addition to the spectrogram display in the reson/EQ section, MacPOD/WinPOD offers a colorful sonogram view, which is useful in observing the evolution of the sound over time. Frequency, from 20 Hz to 20 kHz, is displayed along the y-axis.

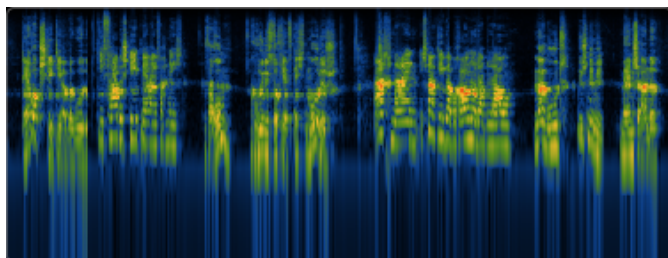


Figure 17: Sonogram view

Audio Driver

To set hardware options such as sampling rate, device options and output channels, click on the **audio driver...** button. This opens the default audio driver page from the Max³ runtime application on top of which MacPOD/WinPOD is built.

Two important points to note:

The audio driver sampling rate should agree with the sampling rate of the audio files you are using, otherwise sample rate conversion will occur and sound files will not sound at their original pitch.

And, I/O and signal vector sizes should be set to 1024. The built-in output limiter requires that they be less than 2048, and MacPOD/WinPOD operates most efficiently at a setting of 1024.

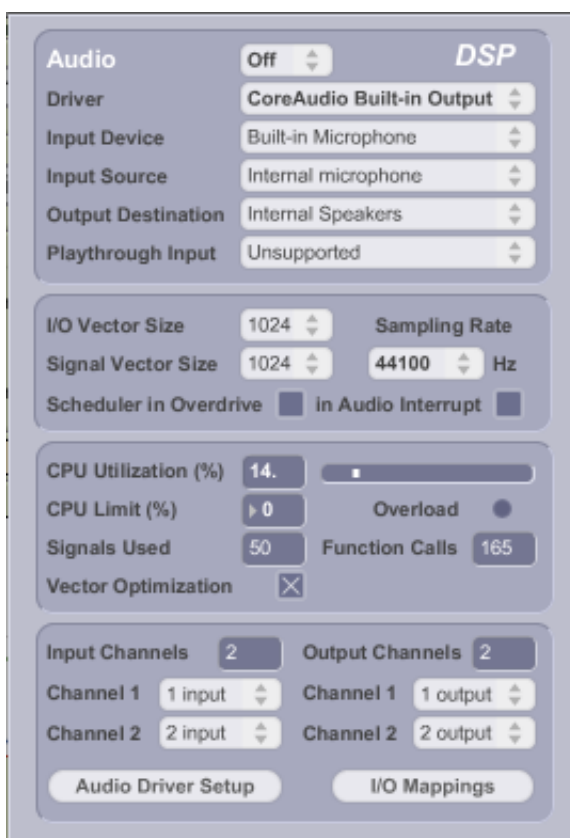


Figure 18: Audio driver

Recording-to-disk

The **record-to-disk** button is used to create an output sound-file. Assuming enough disk space, all subsequent audio will be recorded to a stereo file at the current sampling rate until you hit the **stop** button. The default word size is 16 bits, but you can change the setting for the next file to be recorded in the **record options**....

³ see <http://www.cycling74.com> for information on MaxMSP.

Typically, you'll **stop** playback first and select **record-to-disk**. Name the output file and then set the parameters, position slider, etc. to what you want. Nothing is recorded until you hit the **play** button. If MacPOD/WinPOD is already playing, sound recording begins as soon as you name and save the output file.

The amount of time recorded so far is displayed in hours : minutes : seconds : milliseconds to the right of the record-to-disk button.



Figure 19: Record-to-disk

Presets

Presets offer a quick way to store and recall all parameter settings (except Main Volume, which is not stored).

To store settings as a preset, mouse or type the preset slot number in the number box to the right of the **store** button, then click the store button.

To **recall** a preset, type the preset number into the leftmost number box (or use the arrow buttons). Presets are recalled using a fade time (in milliseconds) to interpolate smoothly between settings.



Figure 20: Presets

To **save** a set of presets (up to 127) as an .xml file, click on the write button, then name and save the file. The **read** button can be used to recall previously saved files.

MIDI

All of MacPOD/WinPOD's parameters, including transport, grain parameters, reson/EQ, tempo and pitch are mapped onto MIDI continuous controllers. This can very useful in controlling the program using a fader pack, MIDI controller or from another application. Additionally, MacPOD/WinPOD responds to MIDI program change messages by selecting the appropriate preset. For specific controller assignments refer to the *MIDI Implementation* section in *Appendix A*.

Supported file formats

MacPOD can read WAVE (.wav) or AIFF (.aif) sound files at standard sample rates (32 kHz, 44.1 kHz, 48 kHz and 96 kHz) and word sizes (8, 16, or 24 bit). Playback is determined by the capabilities of your hardware device.

To open and playback MP3 files, use the **File** menu's **Import...** option.

Appendix A: MIDI Implementation

Controller ID	Parameter	Range
1	tempo	-100% ... 100%
7	volume	-INF ... 0.0 dB
20	transport play/stop	0 ... 63 = stop; >= 64 = play
21	transport pause	0 ... 63 = pause; >= 64 = resume
22	position	0 ... max
30	auto-stretch on	0 ... 63 = off; >= 64 = on
31	auto-stretch lookahead	0 ... 1000 ms.
32	auto-stretch sensitivity	min ... max
33	auto-stretch stretch amt	min ... max
34	auto-stretch silence	0 ... 63 = off; >= 64 = on
40	grain #streams	0 ... 32
41	grain offset range	0 ... 1000 ms.
42	grain window	0 ... 100%
43	grain duration	2 ... 1000 ms.
44	grain duration random	0 ... 200 ms.
45	grain delay	0 ... 4000 ms.
46	grain delay random	0 ... 4000 ms.
47	pitch	25% ... 400%
48	pitch range	0% ... 100%
50	filter bypass	0 ... 63 = off; >= 64 = on
60	reson freq	min ... max
61	reson gain	min ... max
62	reson Q	min ... max
63	reson feedback	min ... max
64	reson feedforward	min ... max
65	reson rate	min ... max
66	reson mod	min ... max
70	EQ low freq	min ... max
71	EQ low gain	min ... max
72	EQ mid freq	min ... max
73	EQ mid gain	min ... max
74	EQ mid Q	min ... max
75	EQ high freq	min ... max
76	EQ high gain	min ... max
MIDI Program Change	Preset	1...127