Full name: _____   MIT ID: _____

21M.380 Music and Technology
Sound Design

First quiz (qz1)
Physics & perception of sound, digital audio, Pd
basics

Monday, March 14, 2016
30 minutes, 10% of total grade

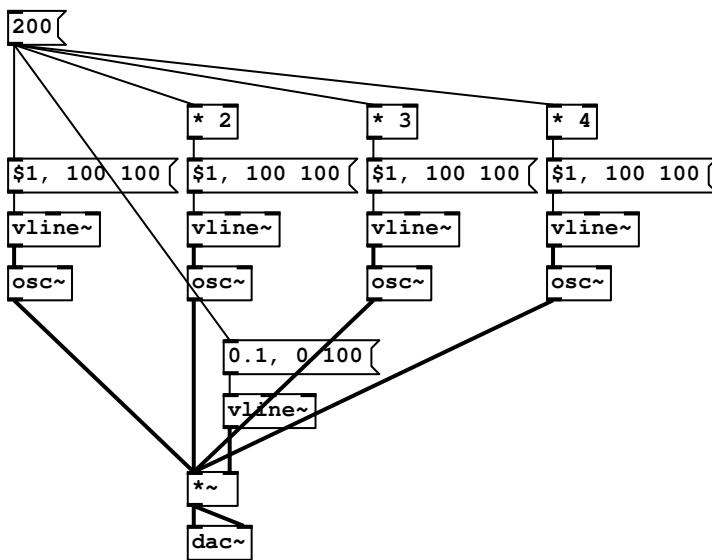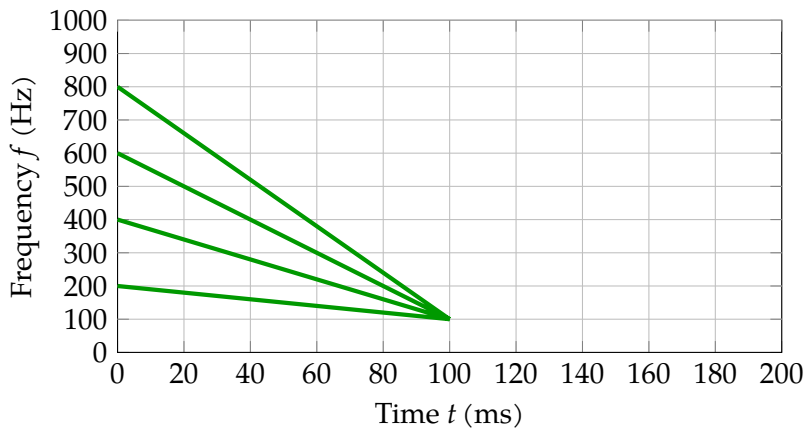## Pd patch for questions 1.1 (p. 2) and 2.2 (p. 4)

Diagram for question 1.1 (p. 2)

1 Physics of sound

1.1 Spectrogram (15%)

In the diagram provided for this purpose on page 1, draw a spectrogram of the sound that the Pd patch from page 1 generates when the user hits the [200(message in the top left corner at $t = 0$ (assuming the DSP is on).[1]

1.2 Resonating pipe (15%)

The length l of a pipe closed on one end coincides with $\frac{\lambda_1}{4}$, a quarter of the wavelength that corresponds to its fundamental frequency f_1. Such a pipe will produce only odd harmonics $(f_1, f_3, f_5,$ etc.), with even harmonics $(f_2, f_4,$ etc.) absent.
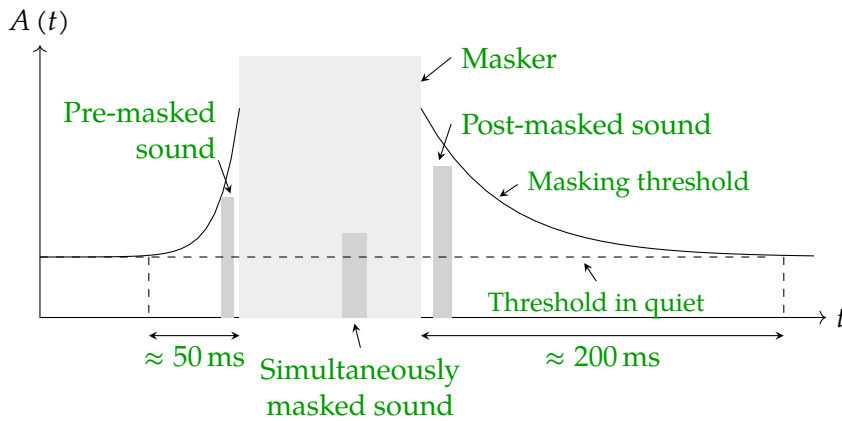
- What is the frequency f_5 of the fifth harmonic created by a pipe closed at one end with a length of $l = 3.4\,\text{m}$? Clearly show how you derive your result and quantify the result in Hz; don't leave fractions. This problem requires knowledge of the speed of sound c in air.

$$l = \frac{\lambda_1}{4} \rightarrow \lambda_1 = 4l$$

$$c = \lambda_n \cdot f_n \rightarrow f_1 = \frac{c}{\lambda_1}$$

$$f_n = n \cdot f_1 \rightarrow f_5 = 5 \cdot f_1$$

$$f_5 = \frac{5 \cdot c}{4 \cdot l} = \frac{5 \cdot 340\,\text{m}\,\text{s}^{-1}}{4 \cdot 3.4\,\text{m}} = 125\,\text{Hz}$$

- Describe the two aspects of the pipe's spectrum that would change if its closed end was opened.

  1. <u>The entire spectrum would shift up an octave.</u>

  2. <u>The even harmonics $f_2, f_4, f_6,$ etc. would appear.</u>

[1] Do not worry about a precise representation of amplitude, it is sufficient to indicate the correct frequencies, onset times, and stop times of all partials.

2   Perception of sound

2.1   Psychoacoustics (15%)

- Which psychoacoustic phenomenon does the figure above illustrate? State its precise name. If there are multiple flavors of this phenomenon, indicate which of these the above figure refers to.

  Masking in the time domain

- Label all arrows in the above figure to support your description, using professional terminology. Provide rough estimates for the length of the time periods indicated by the two horziontal arrows.

- Describe the effect that this phenomenon has on the perception of the three sounds that are represented in the figure by the three dark gray blocks. How does this relate to the other elements in the figure?

  All three sounds are masked by the sound represented by the light gray block, which temporarily raises the threshold of hearing. Even though all three sounds lie above the threshold in quiet (and are thus audible if the masker is absent), they are all below this raised threshold, so the presence of the masker makes them inaudible. This phenomenon applies to sounds that occur simultaneously with, shortly after, or – somehwhat counterintuitively – just before the masking sound.

2.2 Describing sound (15%)

Describe to a listener who has never experienced the Pd patch from page 1, and who has no technical understanding of (electro)acoustics whatsoever, what the patch sounds like. Embrace any possibilities that written text offers you in this regard. Anything goes, as long as your description employs text in some capacity (it should not be purely pictographic) and captures the sound's essential perceptual characteristics.

This is an open-ended question, so I can at best provide an exemplary answer. In technical terms (provided here only as background information rather than as part of the exemplary answer), we are listening to a rapidly fading glissando over 100 ms from a harmonic sound with 4 harmonics on a fundamental of 200 Hz to a single frequency at 100 Hz. This sounds not dissimilar to your average laser gun and could perhaps best be encapsulated through a mixture of onomatopoeia and typesetting, i.e., *Pew!*

3 Digital audio

3.1 Digital audio file formats (5%)

- What is the difference between lossless and lossy compression in the context of digital audio formats?

  Lossless compression is reversible, whereas lossy compression is not.

- For both lossless and lossy compression, name one example of a container format (file name extension) *and* a corresponding audio codec.

  – Lossless compressed:

  Container: .flac . Codec: FLAC

  – Lossy compressed:

  Container: .ogg . Codec: Vorbis

3.2   Sampling (5%)

- You are trying to optimize a digital audio system that plays music from before 1970, which – as we all know – only people over 40 years of age listen to. Which sample rate would you choose for this system, such that it just covers its audience's entire range of hearing? Show how you derived your result.

  - Upper range of human hearing: $20\,\text{kHz} - \frac{1\,\text{kHz}}{\text{life decade}}$
  - So $f_{max} = 16\,\text{kHz}$ is roughly the highest frequency that 40-year olds can perceive.
  - To comply with the sampling theorem, choose the sample rate $f_S \overset{!}{>} 2 \cdot f_{max} = 32\,\text{kHz}$

3.3   Digital artefacts (5%)

- You will hear a sound example, first without, then with a digital artefact. Which artefact is this an example of?

  ○ Clipping
  ⊗ Aliasing
  ○ Click

- What has likely caused this artefact?

  ○ A bit rate that was chosen too low
  ○ A signal level that was chosen too high
  ⊗ A sample rate that was chosen too low

- Describe the artefact, this time in purely technical terms (ideally by a single phrase). How does the original spectrum compare with the spectrum that contains the artefact?

  <u>Inharmonic distortion</u>

3.4   Quantization and sampling (10%)

Check any statements that are correct (false positives are penalized).

☒ 96 kHz is a sample rate widely used in professional audio production.

☐ A bit depth of 16 bit is sufficient to prevent aliasing in most situations. (Aliasing relates to sample rate, not bit depth.)

❑ Quantization creates periodically repeating copies of the original spectrum (sidebands). (No, but sampling does.)

☒ With 9 bit, one can express 512 different amplitude values.

❑ Lossless compressed audio file formats exploit the phenomenon of the missing fundamental in order to reduce data.

❑ The dynamic range of a digital signal quantized at 16 bit is roughly 144 dB. (No, $\Delta L \approx 6 \cdot N = 6 \cdot 16 = 96$ dB)

❑ 256 kbit s⁻¹ is a popular bit depth for uncompressed audio.

☒ One can truthfully restore signals which have been sampled at a sample rate that complies with the sampling theorem, as long as no quantization has occured.

❑ The highest frequency that can be faithfully reproduced with a sample rate of 8 kHz is 16 kHz. (No, $f_{max} = f_s/2 = 4$ kHz)

☒ A bit depth of 24 bit provides a dynamic range that exceeds that of the human auditory system. ($\Delta L \approx 6 \cdot 24 = 144$ dB > 130 dB)
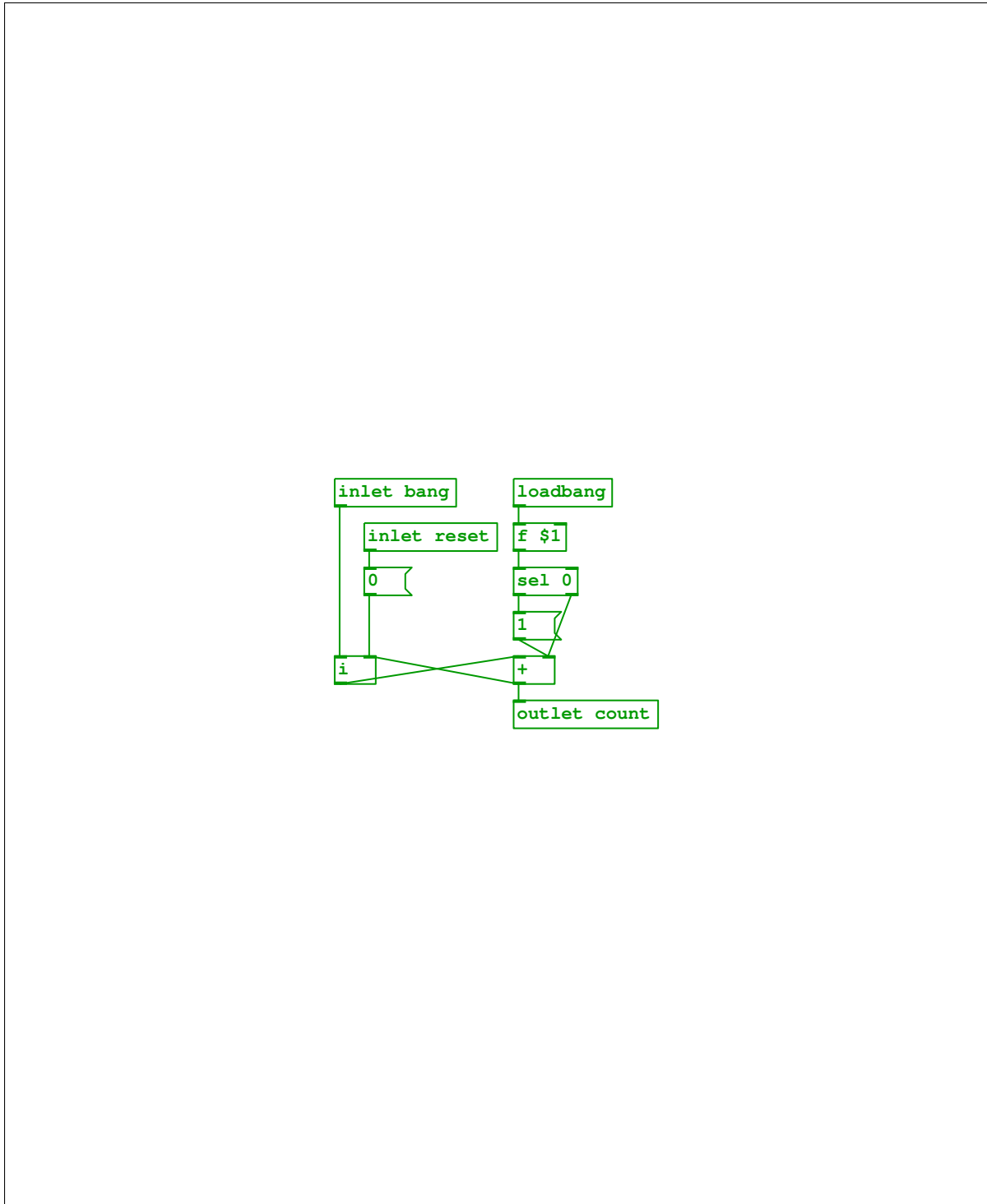
4 Pure Data basics

4.1 Happy patching! (15%)

In the provided space on page 7, implement a Pd abstraction `counter.pd` according to the following specifications:[2]

- The abstractions first and only creation argument $1 should allow the user to set the increment, i.e., the amount by which the counter will increase upon each [bang(that it receives on its left inlet (see below). If $1 is not provided (in which case Pd assumes it to be zero), the increment should default to 1.

- When the abstraction's left inlet receives a [bang(message, it should increment an integer counter and immediately output the new count at the abstraction's only outlet. Counting should start at the value of the increment (not zero), so the first three [bang(mesages sent to the left inlet of a [counter 2] should yield 2, 4, 6 (not 0, 2, 4).

- The abstraction's right inlet should accept a [bang(message that resets the counter, such that it starts the counting sequence again once it receives the next [bang(on its left inlet (but not earlier). The counting sequence should start at the value of the increment (not zero) also after a reset.

[2] You can use the margins on this page for sketches (which will not be graded).

Answer to question 4.1

```
┌────────────┐      ┌──────────┐
│ inlet bang │      │ loadbang │
└────────────┘      └──────────┘
    ┌──────────────┐ ┌────────┐
    │ inlet reset  │ │ f $1   │
    └──────────────┘ └────────┘
       ┌───┐           ┌───────┐
       │ 0 │           │ sel 0 │
       └───┘           └───────┘
                         ┌───┐
                         │ 1 │
                         └───┘
  ┌───┐                ┌───┐
  │ i │                │ + │
  └───┘                └───┘
                     ┌───────────────┐
                     │ outlet count  │
                     └───────────────┘
```

21M.380 Music and Technology: Sound Design
Spring 2016