Why two pressings of the "same" CD will inevitably sound (at least a little bit) different

Brian Cantwell Smith
Information, Philosophy,
Cognitive Science
University of Toronto

Digitally is often characterized ostensively—a digital system is *like:*

Or, as *like the integers:*

1, 2, 3, 4, 5, 6, 7, …

According to John Haugeland, a digital system requires:

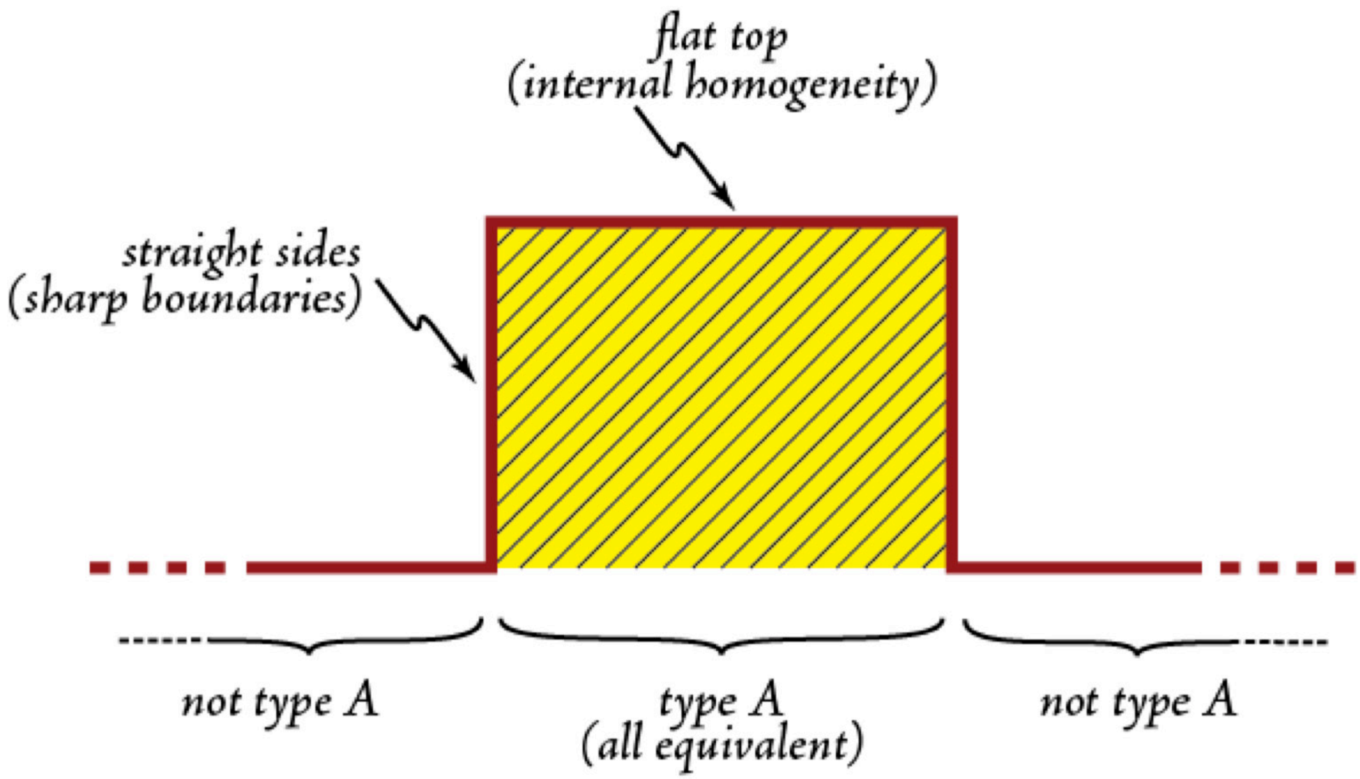
1. A set of distinct **types**
2. Each type must have a set of *absolutely identical, indistinguishable* (for purposes of the system) **tokens**
   a) Cf. checkers, chess, tic-tac-toe
   b) Cf. 0's and 1's in a "computer"

Haugeland's "token manipulation"

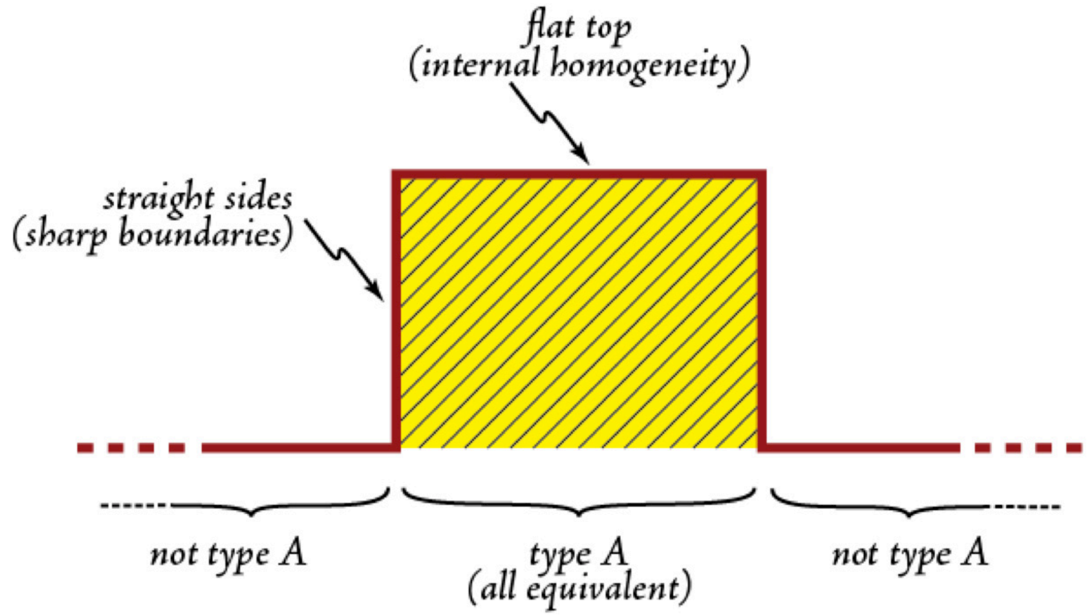
3. Questions must have absolute, definite, yes/no answers:
   a) "Is α a token of type β?"
   b) "What type is α a token of?"
4. No *ambiguity*, no *vagueness*, no *matters of degree*
5. I.e.: *perfect copyability*, *perfect reproducibility*, *absolute determination of types*, etc.
6. In other words: a **perfect** system of **utterly reliable interchangeable parts**

Eli Whitney to the max!

An iconic representation of this characterization of digitality



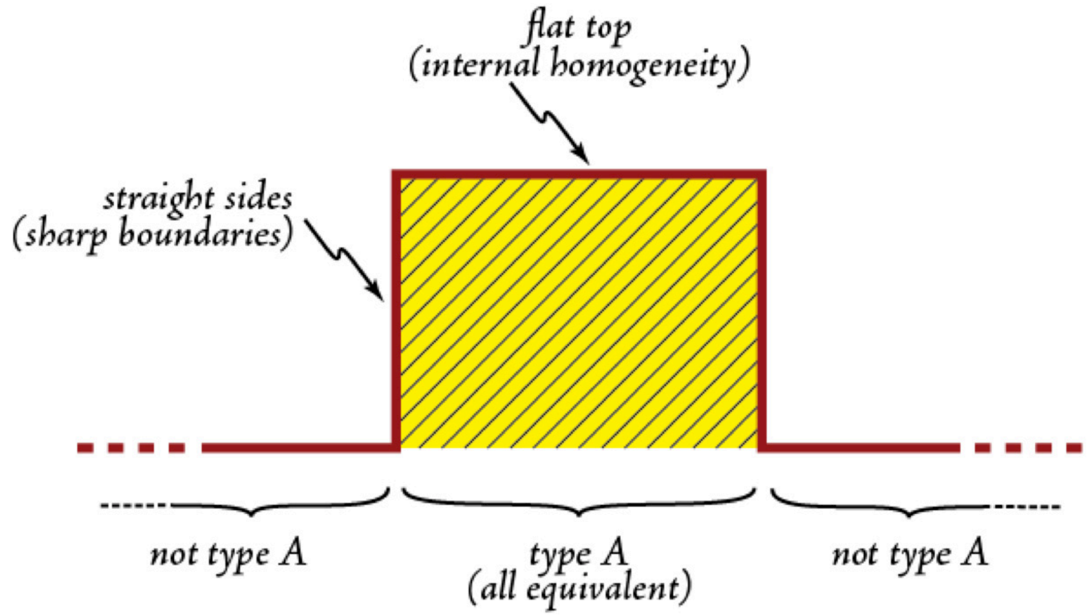
Problem



1. So far, this sounds like abstract mathematics
2. Discrete, perfect, types & tokens
3. What does this have to do with computing, and with the digital revolution?
4. And how is anything like this—anything with this sort of "perfect or perfected clarity"—**possible** in the messy, disheveled world we live in—a world of friction, decay, sloppiness, etc.?

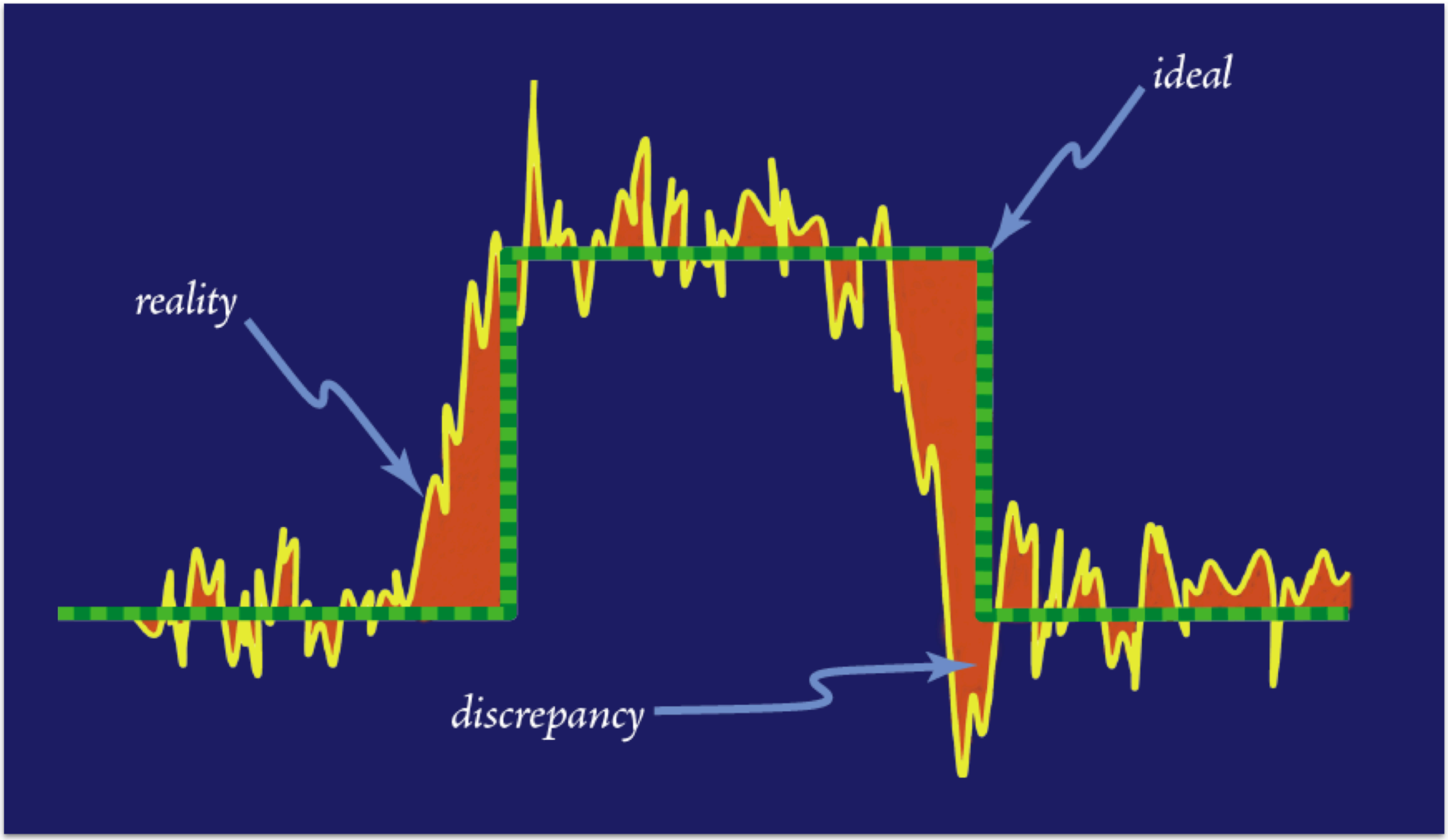
The $1,000,000 question!

Problem (cont'd)



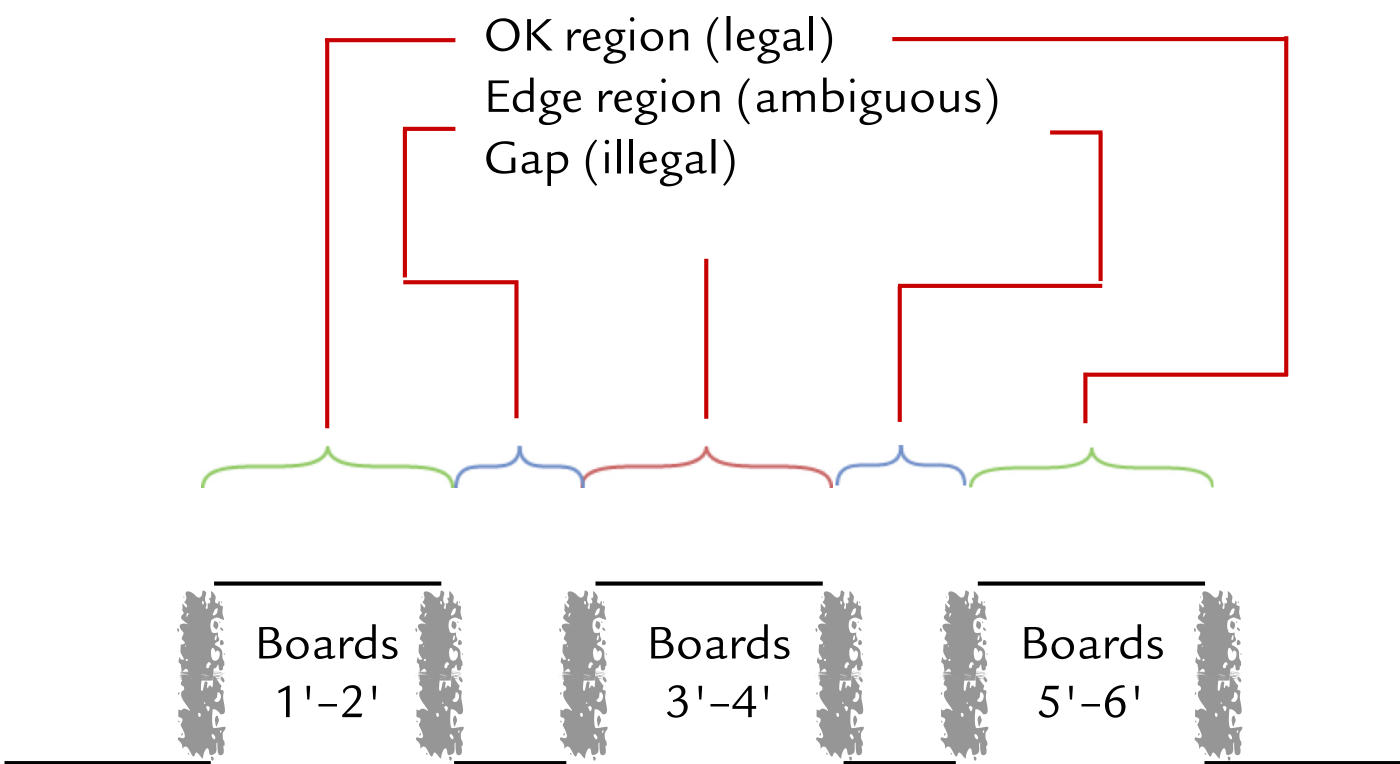
5. Haugeland doesn't tell you! (how to have digitality).
6. His account of digitality is a good account of **what digitality gives you** — of **what digitality is *for*.**
7. It is **not** a good account of **what digitality is** — especially of *how digital systems can be constructed in the physical world.*

So how can we actually *have* digital systems—systems of such perfection?



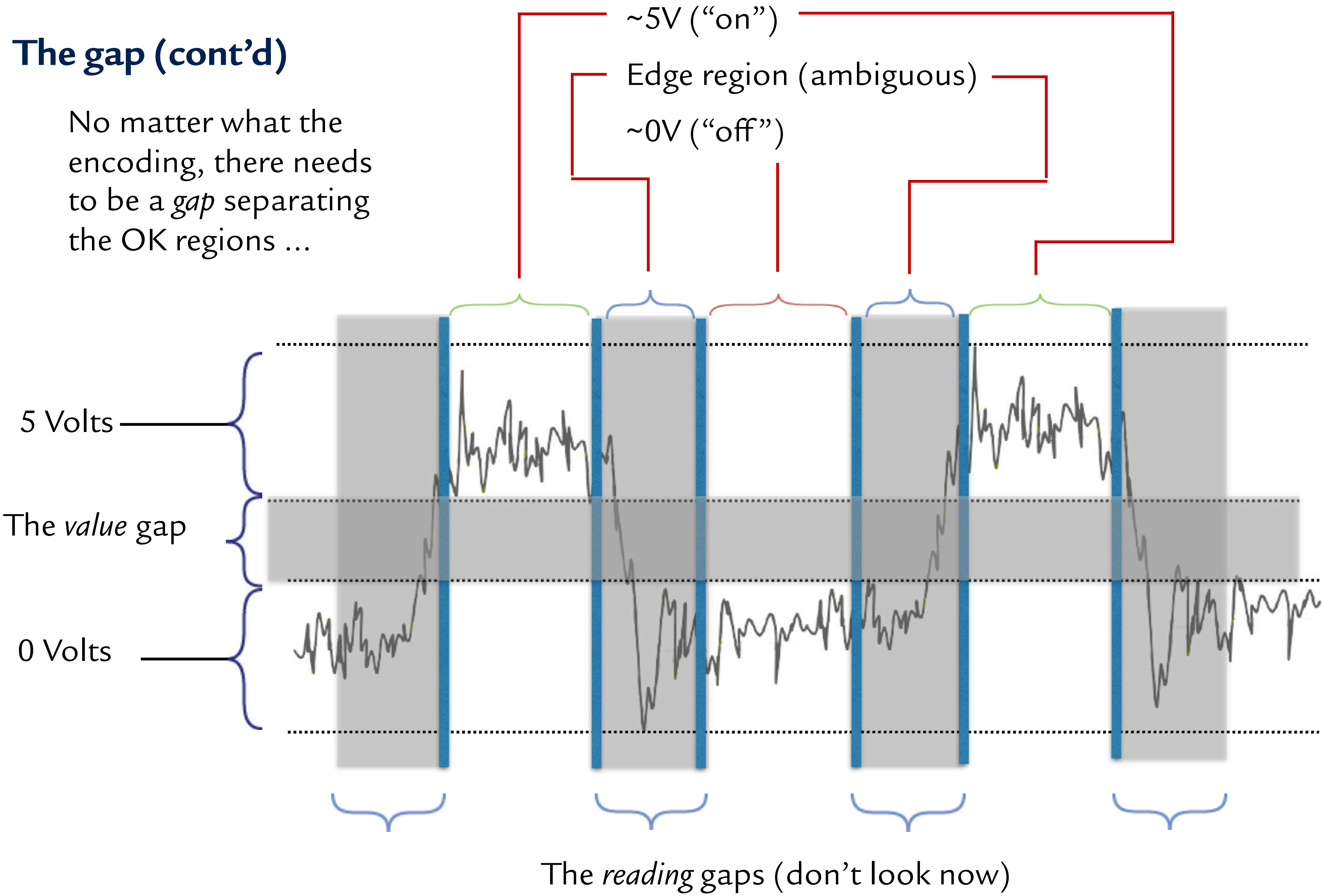
The critical gap

One builds in a "gap" between every state that is OK, so that no two legal states abut. Between all the OK ones is an ***illegal region***.

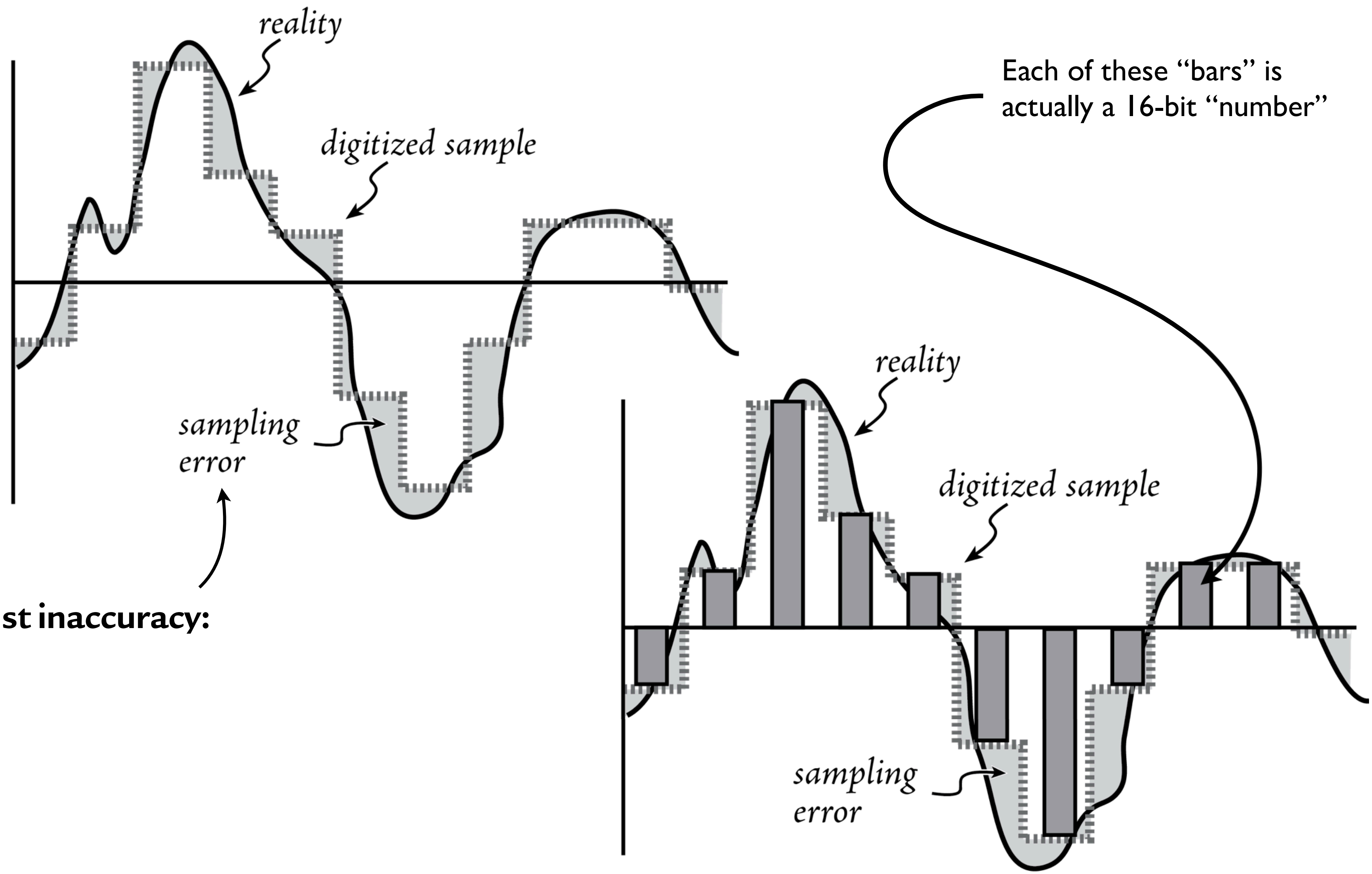


The gap (cont'd)

No matter what the encoding, there needs to be a *gap* separating the OK regions …

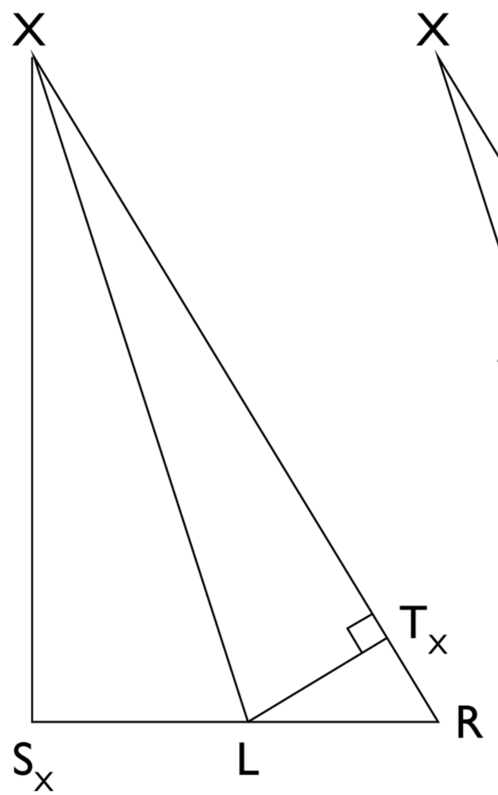


CDs: The First step is an Analog to Digital (A/D) conversion of the music

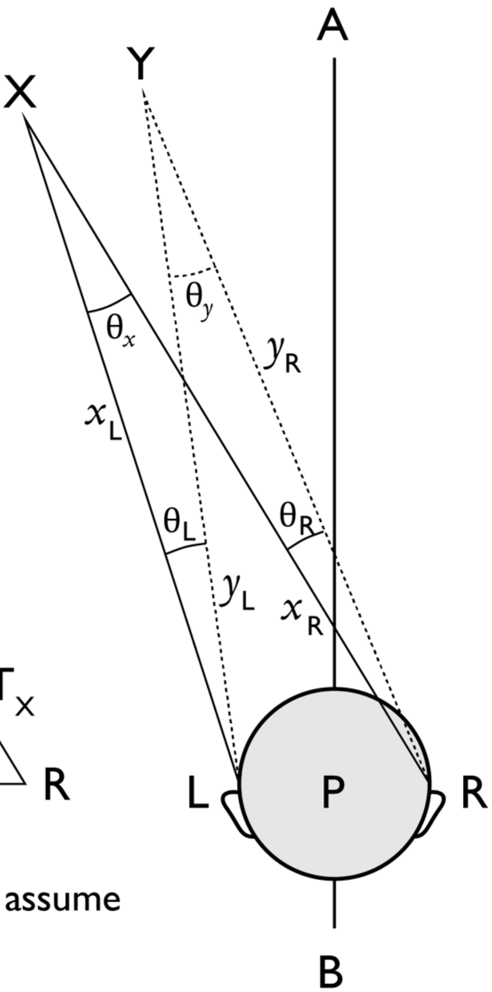


1st inaccuracy:

Temporal requirements for directional discrimination (e.g., a twig breaking in the woods)



(Working diagram for X; assume an analogous one for Y)



Assumptions:

1) XL = YL ≈ 10 m (discriminate a sound ocurring at X from a sound occurring at Y from 10 m away)
2) LR ≈ 0.2032 m (=8distance between two ears)
3) $\theta_L \approx \theta_R \approx 2°$ (assume we can distinguish two breaking twigs 2° apart from this distance)
4) RLX ≈ 110° (just as an example; 20° off straight forward)
5) 343 m/sec = speed of sound in air

Conclusion:

To determine, from Δs in time of arrival alone, the direction from which a sound is coming, to a resolution of 3° (at 20° off normal), requires being able to distinguish a time delay of 0.2045 msec and a time delay of 0.1758 msec—i.e., requires a **21.1 µsec acuity** in discerning Δs in arrival time at R vs. L, which is the period of of a 47.4KHz tone, or the Nyquist sampling rate of 23.7KHz, which—coincidentally?—almost perfectly aligns with the upper frequency limit of human hearing.

Calculations for X:

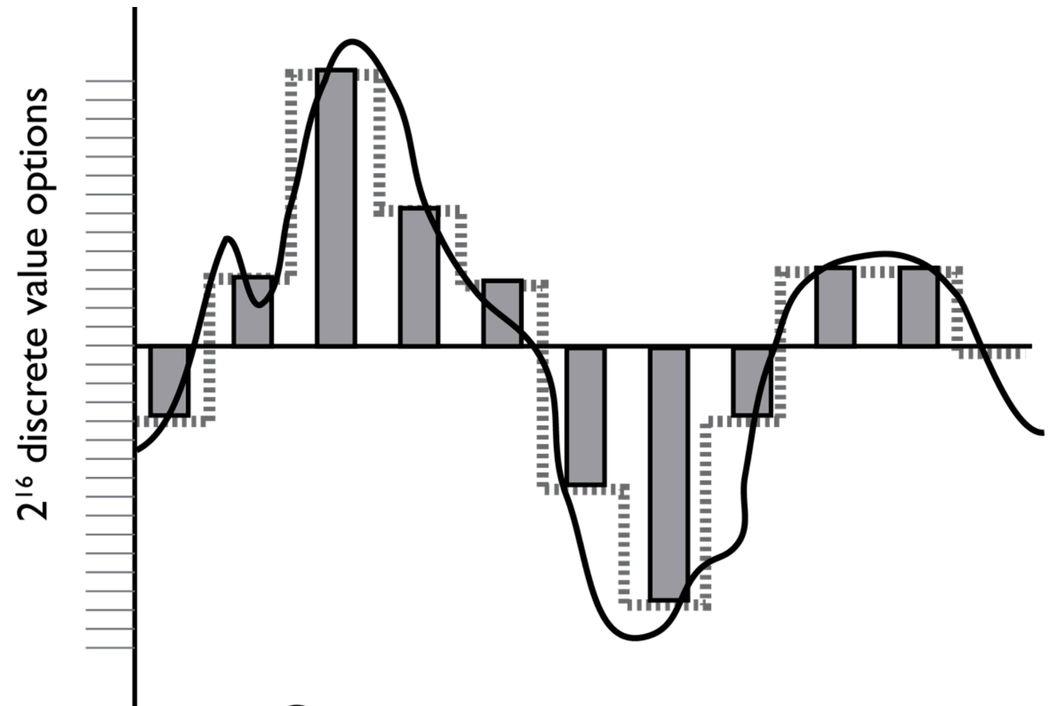
1) $LS_X = XL \cdot \sin(20°) = 3.4202$ m
2) $S_XX = XL \cdot \cos(20°) = 9.3969$ m
3) $S_XXR = \tan^{-1}((3.4202+0.2032)/9.3969) = 21.0694°$
4) $LXR\ (\theta_L) = 1.0694°$
5) $RLT_X = 21.0694°\ (= 180°–(70°+(90°–1.0694°)))$
5) $T_XR = LR \cdot \sin\ (21.0694°) = 0.0719$ m
6) $XT_X = XL \cdot \cos\ (1.0694°) = 9.9983$ m
7) $XR = XT_X+T_XR = 10.0702$ m
8) $XR–XL = 0.0702$ m $= 7.02$ cm
9) 0.2045 msec = Δ in arrival time at L and at R for a sound originating at X (= the period of ~5KHz tone, or the Nyquist sampling rate of ~2.5KHz)

Calculations for Y:

1) $LS_Y = YL \cdot \sin(18°) = 3.0902$ m
2) $S_YY = YL \cdot \cos(18°) = 9.5106$ m
3) $S_YYR = \tan^{-1}((3.0902+0.2032)/9.5630) = 19.1003°$
4) $LYR\ (\theta_R) = 1.1003°$
5) $RLT_Y = 19.1003°\ (= 180°–(72°+(90°–1.1003°)))$
5) $T_YR = LR \cdot \sin\ (19.1003°) = 0.0649$ m
6) $YT_Y = YL \cdot \cos\ (1.1003°) = 9.9982$ m
7) $YR = YT_Y + T_YR = 10.0631$ m
8) $YR–YL = 0.0631$ m $= 6.31$ cm
9) 0.1834msec = Δ in arrival time at L and at R for a sound originating at Y (= the period of ~5.5KHz tone, or the Nyquist sampling rate of ~2.7KHz)

These calculations, and the fact that the highest frequency that people can hear (c.22KHz), suggest that a temporal accuracy of ~25 *microseconds* should suffice—i.e., should be good enough for all purposes of human perception.

However …



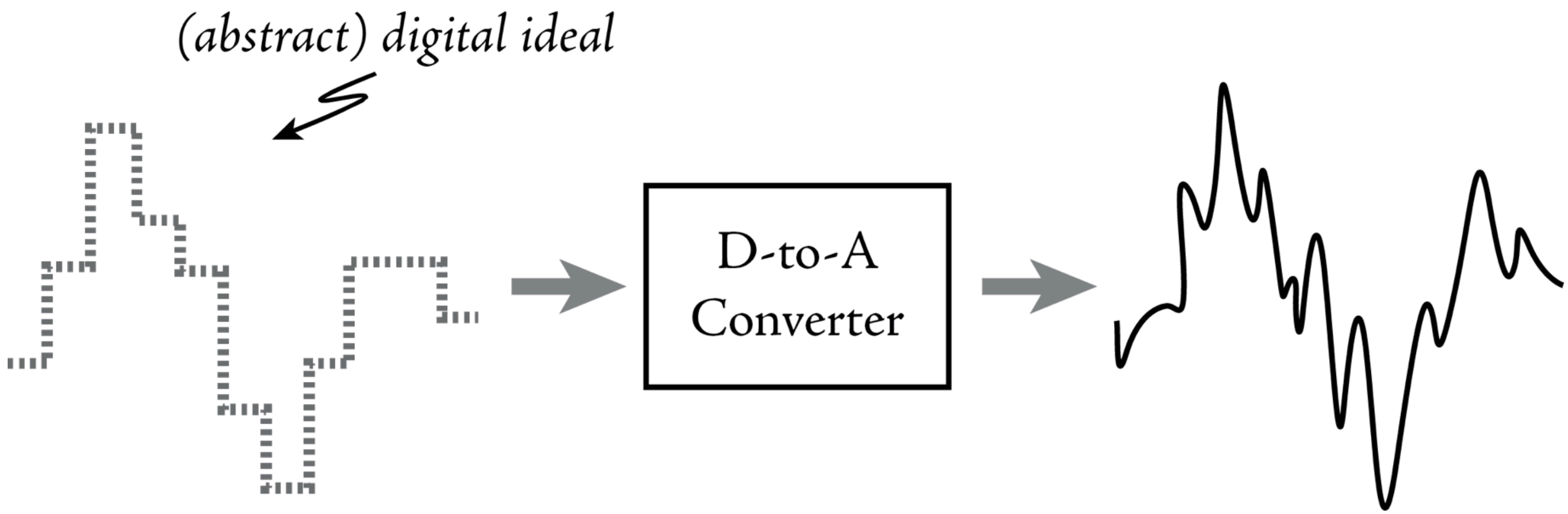
Q: What *temporal resolution* (accuracy) is required so that it is equivalent to the *amplitude resolution* (accuracy)?

A:
1) Sampling rate is 44.1KHz, implying that pulses are 22.7 miscroseconds apart.
2) 1 in 2^{16} accuracy (1 part in 65,536) means that the "clock" must be accurate to ±346 *picoseconds!*

100,000 times more temporally demanding!

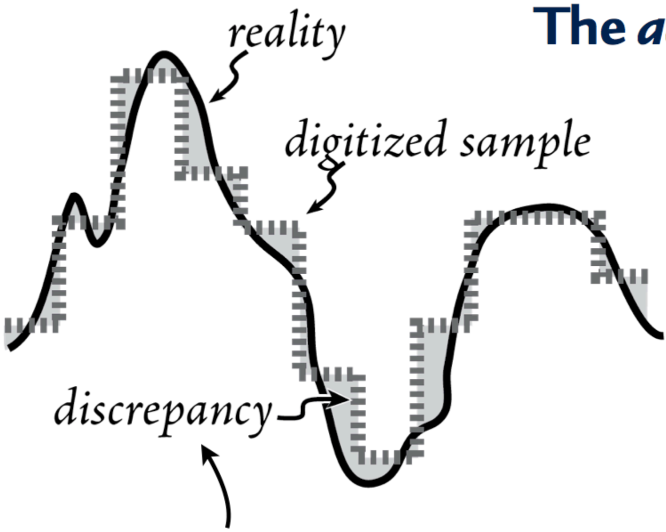
CDs — The Output step: Digital to Analog (D/A) conversion

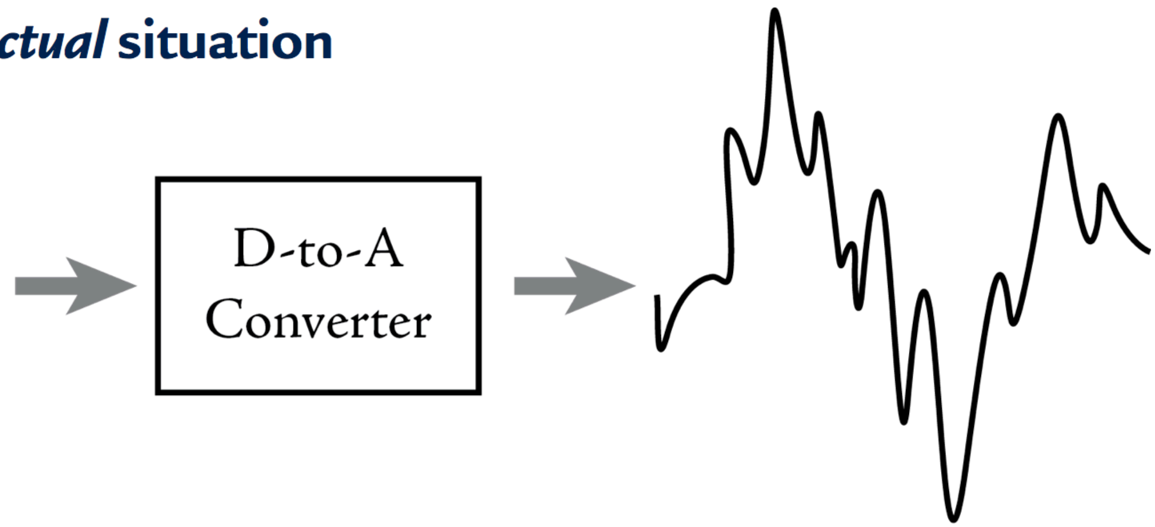
The *imagined* situation



CDs — The Output step: Digital to Analog (D/A) conversion (cont'd)

The *actual* situation





2nd inaccuracy:

The smoking gun!

1. The chips in the D/A convertor *cannot* respond to the digital ideal (which after all does not really "exist" physically)
2. Rather, they have to respond to the analogue signal that encodes the digital ideal, and must therefore respond to every "vagary and vicissitude" in that encoding

Moral

1. The chips in the D/A convertor
   a) Are *supposed* to respond to the digital ideal that the analogue system encodes
      – With everyone recognizing the 1st inaccuracy: the sampling error
   b) *Actually* respond to the *analog signal* that "implements" the digital ideal
      – I.e, also including the 2nd inaccuracy: the discrepancy
2. So from a digital point of view, the two pressings may be (bitwise) "the same"
3. But from an analogue point of view, there is effectively zero change of their being *analog-identical*
   a) There will inevitably be a certain amount of dust in the pits
   b) There will be inevitably be differences in the quality/density of the plastic
   c) Power variations in the power to the lasers burning the pits may vary, so the pits may be burned to slightly different depths
   d) Etc.!