# Annotated bibliography – Opus audio codec

MUMT 621 Presentation 2. February 09, 2021. Sevag Hanssian, 260398537

- Jean-Marc Valin et al. *High-Quality, Low-Delay Music Coding in the Opus Codec*. 2016. arXiv: `1602.04845` [`cs.MM`]. URL: `https://arxiv.org/abs/1602.04845`
  This is the published paper for the music (CELT) details of the Opus codec (and some minor information on the speech/SILK mode).

- *Xiph.org: About*. URL: `https://xiph.org/about/`
  Xiph.org's "About" page on their website, where a full list of their projects, mission statement, and origins (the 1998 MP3 situation) are outlined.

- Jean-Marc Valin. *Opus 1.3 Released*. URL: `https://jmvalin.ca/opus/opus-1.3/`
  Jean-Marc Valin (one of the principal authors of Opus and members of Xiph.org) talks about the Opus 1.3 release and describes the ambisonic features and the music-speech detection neural network.

- Shahram Shirani. "Speech Compression". In: *ELEC 728, McMaster University Department of Electrical Engineering* (2010). URL: `https://www.ece.mcmaster.ca/~shirani/multi10/speech%20compression.pdf`
  Lectures slides from a McMaster's Electrical Engineering course describing speech compression and linear prediction coding in terms of the larynx, throat, glottis, excitation signal and tube coefficients.

- *MPEG – The Moving Picture Experts Group*. URL: `https://www.mpegstandards.org/`
  MPEG's website from where I copied their mission statement.

- *What's New – Oct. 31, 1998*. URL: `https://web.ncf.ca/aa571/wn103198.htm`
  Computall, a defunct tech company, describes how theyhad to cease work on their MP3 software, and published the emails they received from Fraunhofer and Thomson re: MP3 royalties in 1998.

- *What We Do – RIAA*. URL: `https://www.riaa.com/what-we-do/`
  RIAA's website where I copied their mission and anti-piracy statements.

- Jean-François Fortin Tam. *Understanding codecs and containers*. URL: `http://www.pitivi.org/manual/codecscontainers.html`

Small website by Pitivi.org (open source media player) describing the difference between containers and codecs which I used to help word my explanation.

- Stephen W. Webb. "RIAA v. Diamond Multimedia Systems: The Recording Industry Attempts to Slow the MP3 Revolution, Taking Aim at the Jogger Friendly Diamond Rio". In: *7 RICH. J.L. & TECH. 5.* 2000. URL: `https://core.ac.uk/download/pdf/232774502.pdf`
  A report on the court case of RIAA versus Diamond, the MP3 player manufacturer.

- Steven S. Skiena. *The Algorithm Design Manual.* London: Springer, 2008. DOI: `10.1007/978-1-84800-070-4`
  Skiena's famous Algorithm Design manual, from which I looked up lossy and lossless compression.

- Péter Rucz. *Examination of lossy audio compression methods.* 2018. URL: `https://last.hit.bme.hu/sites/default/files/documents/audio_labor_en.pdf`
  An examination and description of lossy psychacoustic audio compression methods, which I used to describe the idea that lossy audio codecs go hand-in-hand with psychoacoustics to ensure the loss is minimally percetible.

- Frank Kurth. "An Audio Codec For Multiple Generations Compression Without Loss Of Perceptual Quality". In: (Aug. 2002)
  A paper describing what generation loss is, to add context to Xiph.org's recommendation that Opus should not be used for archival due to generation loss.

- Julian Spittka, Henrik Astrom, and Koen Vos. *RTP Payload Format and File Storage Format for SILK Speech and Audio Codec.* 2009. URL: `%7Bhttps://tools.ietf.org/html/draft-spittka-silk-payload-format-00%7D`
  The IETF draft RFC for the SILK audio codec, containing many technical details.

- Julien Epps and W.H. Holmes. "A new technique for wideband enhancement of coded narrowband speech". In: Feb. 1999, pp. 174–176. ISBN: 0-7803-5651-9. DOI: `10.1109/SCFT.1999.781522`
  A report on different speech and audio bands (narrowband, wideband) in my speech vs. music slide.

- Brian Moore. "Effects of Sound-Induced Hearing Loss and Hearing Aids on the Perception of Music". In: *Journal of the Audio Engineering Society* 64

(Mar. 2016), pp. 112–123. DOI: `10.17743/jaes.2015.0081`
A study by famous psychoacoustic researcher Brian C. J. Moore, describing that hearing aids designed for speech frequencies are not so good at music (which helps me provide rationale for why Opus contains different coding strategies for speech and music).

- K. Vos et al. "Voice coding with opus". In: *135th Audio Engineering Society Convention 2013* (Jan. 2013), pp. 722–731. URL: `https://jmvalin.ca/papers/aes135_opus_silk.pdf`
A second paper on Opus by the codec authors, this time focused on the speech/SILK part of the codec.

- N. Ahmed, T. Natarajan, and K. R. Rao. "Discrete Cosine Transform". In: *IEEE Transactions on Computers* (1974). DOI: `10.1109/T-C.1974.223784`. URL: `https://www.ic.tu-berlin.de/fileadmin/fg121/Source-Coding_WS12/selected-readings/Ahmed_et_al.__1974.pdf`
The original paper presenting the DCT, comparing it to the DFT, and describing its special property of "spectral compactness."