Paul Lunn | Researcher, University of Huddersfield, UK
p.lunn@hud.ac.uk

Andy Hunt, PhD | Researcher, University of York, UK
adh@ohm.york.ac.uk

Listening to the invisible
Sonification as a tool for astronomical discovery

Abstract
Sound has been used for scientific investigation for many years; the stethoscope and the Geiger counter are just two examples. Sonification is a method of transforming data into sound. The listener can then explore the data sonically, which can reveal hidden structures and relationships not apparent through visualisation. This paper discusses the advantages of sonification and introduces the reader to techniques such as audification, parameter mapping and model-based sonification. It provides case studies of astronomy-based sonification and concludes with a brief discussion of current work on the sonification of radio astronomy data as part of the Search for Extra-Terrestrial Intelligence (SETI).

Keywords
astronomy, audification, interdisciplinary, parameter mapping, science, sonification

Introduction
Just for one moment, close your eyes and listen.

What cues about your environment can you establish just by listening? Are you sat in large reverberant hall or a small-enclosed room? Can you hear a clock ticking? In what direction is the clock? We are often unaware of the richness of information sound can portray to us. For example, if we pour water into a vessel it will resonate with increasing pitch, providing an acoustic indication of its fullness. When driving a car we often use the sound of the engine to anticipate changing gear.

Sonification describes a process where raw data is analysed or explored through the medium of sound. In Greg Kramer's 'Sonification report' this is defined as 'The transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation' (Kramer et al., 1999).

Sonification is a series of techniques that take data as an input and generate sound as an output. Generally the input data is a numeric series such as a column of a spreadsheet, and it can be multidimensional. The sonification algorithm is a computer program that transforms data into sound. This algorithm can be implemented in a traditional programming language, such as C++/Java, or a graphical based environment such as MAX/MSP. The output is a synthesised sound, an audio file or a MIDI file that is later played through a synthesiser.

The phrase 'sonification' has parallels with the term 'visualisation'. Whereas visualisation is a process of representing data in a graphical format, sonification is a process where data is represented sonically. Listening to data can reveal patterns and structures that may not be apparent through visual methods. Hermann (2010) identifies several areas where sonification techniques could be utilized within a scientific context. Process monitoring, rapid summary of large datasets, searching for patterns in data, and exploratory data analysis. Sonification techniques have been applied in many scientific disciplines, including astronomy, particle physics, chemistry, mechanical engineering, medicine, seismology and meteorology.

Many scientific investigations involve the analysis and exploration of large multidimensional datasets that have traditionally been perceived by visualisation techniques. There are limitations to visual perception, such as temporal resolution and difficulties in representing multiple dimensions of data. Sonification can provide an alternative to visualisation techniques or used in conjunction to enhance analysis.

Multidisciplinary aspects of sonification

Sonification combines skills found in many areas such as music technology, computer science, sound design, composition and performance, data mining, acoustics and physics. An excellent example of multidisciplinary sonification is the *sonEnvir* project (Campo et al., 2006), which investigated implementing sonification techniques in a range of scientific disciplines including neurology, theoretical physics, signal processing and sociology. The project produced a series of tools based around the Supercollider programming language (Cycling74, n.d.).

Why use sound?

There are several characteristics of auditory perception that sonification exploits (Kramer, 1994: 7). The ear is excellent at perceiving time based information, such as rhythm and pitch. The highest pitch we can theoretically perceive is 20Khz (in practice it is often much lower) – that is we can detect acoustic vibrations that oscillate over 20,000 times a second. If such a frequency rate was to be presented visually – much of the information would be blurred or lost completely. The ear is better at detecting rapid or transient changes than the eye.

We are capable of perceiving several sounds simultaneously. Whilst listening to a classical music performance, you will be able to distinguish the individual components of the orchestra concurrently. This ability means that we can listen to several sonified streams in parallel, which is advantageous when dealing with multidimensional information.

Backgrounding is the ability of our auditory system to relegate sounds to a lower priority. Although we are constantly surrounded by sound, we are not aware of most sounds until our attention is drawn to them. For example, you may not be aware of road traffic outside until a passing motorist sounds their horn.

Our hearing is multidimensional; we have the ability to localise sound. If a tiger snaps a twig behind us, we are alerted to this danger that we didn't see because it was not in our field of vision. Unlike a visualisation system, we do not have to be orientated in the direction of the sonification. An example of this is the Geiger counter. As it provides its information in clicks we do not have to constantly observer the meter, so we can safely walk around taking measurements.

A sonification can be 'eyes free' – enabling the user to listen to the data while occupied with another task. So the addition of sound can increase the amount of data presented without increasing visual overload. Our hearing is constant; we do not have the ability to stop listening. This can be useful when monitoring data because if the data was visual we could miss an event if we blinked or looked away momentarily.

There are of course some disadvantages of using sound. Auditory perception is relative. When comparing two stimuli we can only state if they are the same or if one is larger than the other; we cannot give an absolute value. Many characteristics of hearing are co-independent upon others, so a change in one will change how we perceive another – for example our perception of loudness changes with pitch.

Some individuals may find the use of sound to be an irritation. Consideration should be given to environmental issues of sound; someone quietly working in the next office may not appreciate hearing your sonification. The hearing capabilities of the listener may also be a consideration. Some may have noise-induced hearing loss or even amusia (tone deafness).

Sonification techniques

There are several algorithmic techniques used to transform data to sound, in this section we discuss these processes and provide examples within a science.

Auditory graphs

Most readers will be familiar with the concept of a graph, having an x and y axis and a series of data points plotted between. Auditory graphs are a sonic representation of graphical data. Each data point on the graph is mapped to a note on a synthesizer. The pitch of the note depends upon where the point falls upon the Y axis. Starting at the XY origin each point is played from left to right – so that the output is a series of notes that corresponds to the input sequence. Multidimensional graphs can be represented by converting each dimension into a different instrument, for example piano, guitar and violin. This form of sonification can be beneficial for presenting graphical data to both visually impaired and sighted users.

Audification

Audification is where data is directly converted to sound. Techniques are based around converting the data to digital audio samples, which can then be saved in an audio file. This practice is beneficial for working with large datasets. One second of CD quality audio requires 44,100 data points. If the playback length is too short the listener may not be able to distinguish any sonic features or recognize patterns.

Audification has been applied to seismology with relative success. Speeth (1961) found that it was difficult to differentiate atomic bomb explosions from earthquakes just by visual observation of seismograph data. By audification, listeners were able to distinguish between the two types of events.

In an engineering context, audification was used by Pauletto and Hunt (2004) in collaboration with Westland Helicopters to analyse flight data. There are numerous flight sensors on a helicopter, which the Westland engineers would print out in graph form and place on the floor so that they could examine all the data visually at once. Sonifying the data was found to accelerate the diagnostic process.

Audification was utilised by Pereverzev et al. in their research on weakly coupled super fluids. Presenting data visually demonstrated no useful information, but when the data was directly audified Pereverzev was able to hear frequency components that led to a discovery.

> If the electrical output of the displacement transducer is amplified and connected to audio headphones, the listener makes a most remarkable observation. As the pressure across the array relaxes to zero there is a clearly distinguishable tone smoothly drifting from high to low frequency during the transient, which lasts for several seconds. This simple observation marks the discovery of coherent quantum oscillations between weakly coupled superfluids. (Pereverzev et al., 1979)

Parameter mapping

Parameter mapping is a sonification technique where input data is used to control a characteristic of a synthesised sound such as pitch, loudness, timbre, rhythm and melody. This is perhaps the most common form of sonification. In ‘sinification’ the data is mapped to the frequency of a sine wave oscillator. Midification describes a process where the data is converted to MIDI note values that can then be ‘played’ on a MIDI synthesizer. As there are numerous acoustic characteristics of sound, parameter mapping can be used for sonifying multidimensional datasets.

An early example of parameter mapping was demonstrated by Sara Bly (1982) who showed that sound could be used to distinguish between three species of Iris flowers. She mapped characteristics such as sepal length and width to pitch, volume and timbre. When this was played back each variety of iris had its own characteristic ‘sound’. Bly found that most listeners had the ability to accurately identify each species through sound alone.

An audio-visual browser was developed by Grond et al. (2010) to explore the structure of RNA. The software developed provides a visualisation of RNA structure that the user can interact with, that is, focus upon a particular section or expand out. The software maps RNA shape information to parameters of sound such as loudness, phase and timbre. This is a good example of how sonification can enhance visualisation techniques.

Model-based sonification

Developed by Thomas Hermann (2002), this moderately new technique data is used to form points in a multidimensional space. Each point of data has physical laws imposed upon it, which dictate how the data points relate to each other and the modelled space. The user then excites the data points by stimulating the system and the reaction of the points is sonified. In model-based sonification the data forms an instrument that is played by applying a stimulus. The prepared piano is a good analogy for model-based sonification. This instrument has physical objects inserted between its strings and dampers with the purpose of altering timbre. For example, a composer may attach ping pong balls, paper clips or spoons. In this analogy the objects are data, and the piano is the data space. The presence of objects on the strings changes the sound of the piano when the instrument is played, very much how the data points interact with the model space when excited by an input.

Interactive sonification

Interactive sonification is an area that focuses on the importance of interaction between the listener and the sonification (Hermann and Hunt, 2005). Rather than passive listening, the user interacts with the data. They may have a controller that allows them to navigate the sound, replaying a section, and adjusting playback speed and position.

Sonification toolkits

Several authors have developed open source sonification toolkits. A toolkit is a set of off-the-peg software that provides sonification functionality. The toolkit can be an entire application or a library of functions that can be incorporated into other tools. These applications make sonification more accessible to those new to the area, providing a quick and easy introduction to sonification techniques.

The *sonification sandbox* (Walker and Cothran, 2003) is a standalone Java application developed to facilitate the creation of auditory graphs. The user can upload vector data in a comma separated values (CSV) format and control many parameters to the graph output such as voice, note length, etc.

MAX/MSP is a visual programming language for sound, which is popular with music composers. Instead of text the user manipulates graphical elements to build a program. *aeSon Toolkit* is a framework that adds sonification functionality to MAX/MSP's impressive set of audio objects. This toolkit includes objects to extract data from files, to transform the data and to map to synthesisers.

Some toolkits have been developed to extend the functionality of existing programming environments. The use of these tools enables users to rapidly incorporate sonification into their existing research. *SKDtools* (Miele, 2003) is a library that extends the MATLAB (MathWorks, n.d.) numerical computing environment, and *Sonipy* (Sonipy, n.d.) is a set of modules that add sonification modules to the Python computer programming language.

Case study of sonification within astronomy

Exploring the vacuum of deep space with sound may seem to be counter-intuitive, but there are some interesting examples of sonification within an astronomical context.

The Jet Propulsion Laboratory (JPL) has a multimedia presentation called *Spooky Sounds* intended to introduce concepts of sonification of space-based data (Jet Propulsion Laboratory, n.d.). This resource includes a variety of sonifications including Jupiter's radio emissions, Ganymede's magnetosphere, Cassini spacecraft flybys of Enceladus and Saturn, and Voyager 1 recordings of Jupiter's bow shock. This site is an interesting introduction to space-based sonification.

Sonification techniques were used to detect the impact of micrometeoroids upon the Voyager 2 spaceship while investigating Saturn's rings (Scarf et al., 1982). There was a problem with the spacecraft that its controllers were trying to pinpoint using visual methods – but this only resulted in visual noise. By sonifying the data, a 'machine gun' sound was perceived. This machine gun sound was caused by micro-meteoroid impacts on the spacecraft.

Selene (Selenological and Engineering Explorer) was a lunar orbiter spacecraft launched by JAXA (Japan Aerospace Exploration Agency) in 2007. Sobue et al. (2010) developed a Geographical Information System based upon laser altimeter data obtained from the mission. The altimeter provides a topographical contour of the surface of the moon. The system developed was a web-based Java application called *Moonbell* (Higashiizumi et al., 2009). This application sonifies the altitude along the route by converting the measured reading into musical notes. The interface is highly configurable, allowing the user to adjust many parameters, such as speed, note range, instrument and volume.

Cosmic Microwave Background Radiation (CMBR) is a nearly uniform space-based radiation that is received in all directions with the majority of its power spectrum contained within the microwave bandwidth. CMBR is a background 'noise' signal that is found all over the universe. It is now considered that this is the faint afterglow of radiation generated by the big bang. The Planck Visualization Project is a NASA education and public outreach initiative to increase public understanding of the CMBR. This project plans to create multimedia displays using CMBR data sourced from the Planck space observatory. Sonification of this data is also being implemented (Van der Veen, n.d.). This team has also sonified the power spectrum of the CMBR (Van der Veen, 2009, 2010).

xSonify is a Java-based sonification toolkit developed to assist in the development of astronomy-based data (Candey et al., 2006). It is designed to implement sonification functionality of NASA's Space Physics Data Facility (SPDF). The SPDF is an online collection of data collated during space-based missions from 1963 onwards. Data is organised by spacecraft and then individual instruments. This resource includes 'heliospheric, magnetospheric, ionospheric and upper-atmospheric data from all NASA and some non-NASA space physics missions' (NASA, 2009).

Finally, audification functionality has been incorporated in relatively new software used to detect the presence of exoplanets (planets orbiting a star other than our own sun) (Systemic, 2011). This approach is experimental but further investigation may be beneficial.

Summary and further work

There are several examples where sonification techniques have proven advantageous in scientific research. The work of Pereverzev, Scarf and Speeth demonstrates that transforming data into sound can provide insights not revealed by visualisation alone. There is a history of sonification within astronomy, from the detection of micro meteors at Saturn, to contemporary research on cosmic microwave background radiation and exoplanet detection. Nevertheless, are there spheres of astronomical research that would warrant investigation using sonification techniques?

Wall (2010) describes the deluge of information that is facing modern astronomers, who are increasingly trying to find a needle in a haystack of data. This data overload has motivated the formation of several citizen science projects such as Galaxy Zoo (2010) and Stardust@Home (n.d.). These projects distribute images to the public so that they can classify them. There is essentially more astronomical data available than there are astronomers to process it. Sonification is a tool that could be utilised for the exploration of large amounts of data.

This paper's authors are currently researching into sonifying radio waves used in the Search for Extra-Terrestrial Intelligence (SETI). SETI is a project initiated in 1961 to search for evidence of intelligent life by the detection of electromagnetic radiation emitted by extra-terrestrial technology (SETI Institute, 2010). If there is intelligent alien life, it *may* use radio waves for communication, much like earth-based television transmission. Radio waves will propagate through the universe, and might be detected as they reach earth. SETI's approach is to selectively position the Allen Telescope Array at a co-ordinate and then take measurements. This process produces a large amount of data that is freely available to researchers through the SetiQuest project (SetiQuest, 2011). We will be investigating how sonification techniques can be implemented in conjunction with a high performance computer cluster to assist in rapid data exploration.

References

Bly S (1982) Sound and computer information presentation. Unpublished doctoral thesis (UCRL-53282), Lawrence Livermore National Laboratory and University of California, Davis, CA, cited in Gaver WW, Auditory interfaces. In: Helander M, Landauer TK and Prabhu PV (eds) (1997) *Handbook of Human–Computer Interaction*, Vol. 1, 1003–1041.

Candey RM, Schertenleib AM and Diaz Merced WL (2006) xSonify: Sonification tool for space physics. In: *Proceedings of the ICAD, 12th International Conference on Auditory Display*, London, UK.

Campo A de, Dayé C, Frauenberger C, Vogt K, Wallisch A and Eckel G (2006) Sonification as an interdisciplinary working process. In: *Proceedings of the 12th International Conference on Auditory Display*, London, UK,

Cycling74 (n.d.) Available at: http://cycling74.com/products/maxmspjitter/ (accessed 4 February 2011).

Galaxy Zoo (2010) Galaxy Zoo Hubble. Available at: http://www.galaxyzoo.org/ (accessed 9 February 2011).

Grond F, Janssen S, Schirmer S and Hermann T (2010) Browsing RNA structures by interactive sonification proceedings of ISon 2010. In: *Proceedings of the 3rd Interactive Sonification Workshop*, KTH, Stockholm, Sweden, 7 April.

Hermann T (2002) Sonification for exploratory data analysis. PhD thesis, Faculty of Technology, Bielefeld University, Germany.

Hermann T (2010) Improving interaction with sonification: Applications from science to smart environments, keynote presentation – DAFx 2010. Available at: http://dafx10.iem.at/proceedings/papers/Hermann_DAFx10_K1_slides.pdf (accessed 26 January 2011).

Hermann T and Hunt A (2005) An introduction to interactive sonification. *IEEE Multimedia* 12(2): 20–24.

Higashiizumi I, Shimada T, Yamaguchi T, Higa S, Vincent T, Hirokawa J, Kobayashi K and Koike H (2009) Moonbell. Available at: http://wms.selene.jaxa.jp/selene_sok/about_en.html (accessed 15 December 2010).

Jet Propulsion Laboratory (n.d.) Spooky Sounds. Available at: http://www.jpl.nasa.gov/multimedia/sounds/?msource=enews102805&tr=y&auid=1197511%E2%80%9D (accessed 9 December 2010).

Kramer G (1994) An introduction to auditory display. In: Kramer G (ed.) *Auditory Display: Sonification, Audification, and Auditory Interface*. Reading, MA: Addison-Wesley.

Kramer G, Walker B, Bonebright T, Cook P, Flowers J, Miner N and Neuhoff J (1999) Sonification report: Status of the field and research agenda. International Community for Auditory Display. Available at: http://sonify.psych.gatech.edu/publications/pdfs/1999-nsf-report.pdf

MathWorks (n.d.) MATLAB and Simulink for technical computing. Available at: http://www.mathworks.co.uk/ (accessed 4 February 2011).

Miele JA (2003) Smith-Kettlewell display tools: A sonification toolkit for Matlab. In: *Proceedings of the 2003 International Conference on Auditory Display*, Boston, MA, USA, 6–9 July.

NASA (2009) SPDF – information for new users. Available at: http://spdf.gsfc.nasa.gov/new_users.html (accessed 3 December 2010).

Pauletto S and Hunt A (2004) Interactive sonification in two domains: Helicopter flight analysis and physiotherapy movement analysis. In: *Proceedings of the Interactive Sonification Workshop*, Bielefeld, Germany.

Pereverzev SV, Loshak A, Backhaus S, Davies JC and Packard RE (1997) Quantum oscillations between two weakly coupled reservoirs of superfluid 3He. *Nature* 388: 449–451.

Python.org (n.d.) Python programming language – official website. Available at: http://www.python.org/ (accessed 4 February 2011).

Scarf FL, Gurnett DA, Kurth WS and Poynter PL (1982) Voyager 2 plasma wave observations at Saturn. *Science* 215(4532): 587–594.

SETI Institute (2010) Homepage. Available at: http://www.seti.org/ (accessed 4 February 2011).

SetiQuest (2011) SetiQuest homepage. Available at: http://setiquest.org// (accessed 9 February 2011).

Sobue S, Araki H, Tazawa S, Noda H, Kamiya I, Yamamoto A, Fujita T, Higashiizumi I and Okumura H (2010) Proposal of lunar GIS with visualized and auditory data of Japan's lunar explorer 'KAGUYA (SELENE)'. *Journal of Communication and Computer* 7(1) (Serial No. 62).

Sonipy (n.d.) An extensive open-source python framework for data sonification research and auditory display. Available at: http://www.sonification.com.au/sonipy/index.html (accessed 4 February 2011).

Speeth S (1961) Seismometer sounds. *Journal of Acoustical Society of America* 33(7): 909–916.

Stardust@Home (n.d.) Stardust@Home homepage. Available at: http://stardustathome.ssl.berkeley.edu/ (accessed 9 February 2011).

Systemic (2011) Systemic: Characterizing planetary systems. Available at: http://oklo.org/ (accessed 7 February 2011).

Van der Veen J (n.d.) Planck Visualization Project: Seeing and hearing the CMB. Powerpoint presentation. Available at: www.physics.ucsb.edu/~jatila/papers/Planck-for-AAPT-2010.pdf (accessed 16 December 2010).

Van der Veen J (2009) From sound waves to microwaves: 'Listening' to the oldest light of the universe with the Planck Mission. Available at: http://learningcenter.nsta.org/products/symposia_seminars/fall09/NASA/webseminar3.aspx (accessed 21 December 2010).

Van der Veen J (2010) Planck Visualization Project: Seeing and hearing the CMB. In: *Science Education and Outreach: Forging a Path to the Future ASP*. Astronomical Society of the Pacific Conference Series, Vol. 431.

Wall M (2010) Astronomy overload: Scientists shifting from stargazing to data mining. Available at: http://www.space.com/9308-astronomy-overload-scientists-shifting-stargazing-data-mining.html (accessed 2 November 2010).

Walker BN and Cothran JT (2003) Sonification sandbox: A graphical toolkit for auditory graphs. In: *Proceedings of the 2003 International Conference on Auditory Display*, Boston, MA, USA, 6–9 July.