3. The Audio System in the Linux Sound Room

(This section last updated August 2006)

3.1. Acoustic, analog, digital and magnetic audio signals

A *signal* is an energy source. An *audio* signal (sound) can exist in one of four physical states: *acoustic*, *analog* (electronic), *magnetic* or *digital*. We can only hear sounds when they are in an acoustic state — when something is vibrating at a rate somewhere between about 20 hertz and 16 kHz. These vibrations produce fluctuations in air pressure, which are sensed by the cilia (tiny hairs) within the cochlea of our inner ears and converted into electrical impulses that are fired to the brain.

☞ **Analog signals:** A *microphone* is an electroacoustic *transducer* (converter) that converts air pressure variations into a corresponding (*analog*ous) AC electrical signal. Any difference between the "shape" of the acoustical vibrations and the up-and-down (or back-and-forth) "shape" of the alternating current (which we can display on an *oscilloscope*) is heard as *distortion*. If the analog signal is not identical to the original acoustic signal, it will not sound exactly the same.

An *amplifier* controls the voltage level (the power, or intensity) of an analog signal. Although distortion can be introduced by any electrical circuit within an audio signal path, amplifiers are the most common source. A *pre-amplifier* boosts low intensity analog signals, most often from a microphone, up to a usable operating *line* level required for good audio quality by most electronic circuits. A *power amplifier* provides a substantial increase in signal power, measured in watts, necessary to drive loudspeakers, which must move a lot of air around. A *loudspeaker* is an electroacoustic transducer that functions in a manner opposite to that of a microphone, converting an (analog) alternating current into physical vibration of one or more speaker cones. A *headphone* is a miniature loudspeaker (or, more accurately for stereo, two miniature loudspeakers). Like all speakers, headphones require power amplification of the input signal. Because of their small size, headphones cannot produce frequencies below about 200 hz. accurately, or at all, and should not be used for critical monitoring.

☞ **Digital signals:** Digital audio signals consists of a stream of numbers, called *samples*. These samples represent the rapid up-and-down amplitude variations of a sound at evenly spaced points of time, called the *sampling rate*. Another important factor in digital signal representation, called the *word size* or *bit depth*, is the number of on/off *bits* (two-position switches) in each sample.

Computers can only process audio signals when these signals are in a digital state. Computers can only deal with numbers, not with variations in air pressure or in electrical voltage. Acoustic signals (air pressure variations) cannot be converted directly into digital samples. Only analog (electrical) signals can be digitized for processing or recording by computer-based audio equipment. An *Analog-to-digital* converter (**ADC**) performs this sampling process, taking a reading, or "snapshot," of an input voltage level thousands of times a second and spitting out a number for each reading. Similarly, digital samples cannot be used to make a loudspeaker cone vibrate back and forth. A digital audio signal must first be converted into corresponding analog voltage fluctuations by means of a *digital-to-analog converter (***DAC**).

In addition to performing conversion, *ADC*s and *DAC*s also send the analog signal through a *low pass filter*, a circuit that passes low frequency components within a sound but eliminates high frequencies. The low-pass output filter on a *DAC* smooths out the discrete, staircase-like "steps" between adjacent samples into continuous (smoothly curved) voltage changes. The *anti-aliasing* low-pass input filter on an *ADC* eliminates all frequencies greater than 1/2 of the sampling rate, called the *Nyquist frequency*, within the analog input signal. To produce total attenuation at the Nyquist frequency, *ADC* filters apply progressively greater amounts of attenuation to frequencies between 40 % and 50 % of the sampling rate.

Digital representation of sound requires at least two samples to represent a frequency. Any frequency greater than 1/2 of the sampling rate (the *Nyquist frequency*) will not be represented correctly, but rather will *fold over*, or *alias*, according to the formula

Sampling Rate — original frequency

Example: If our sampling rate is 20000, a frequency of 12000 hertz within an input sound would actually sound at 8000 hertz (20000-12000), creating a spurious new frequency not present within the original sound. With the same 20000 sampling rate, an input frequency of 14,400 hertz would "fold over" to 5600 hertz, producing even more noticeable timbral distortion. An *ADC* with a sampling rate of 20000 should include an anti-aliasing filter that attenuates frequencies between 8 kHz and 10 kHz and eliminates all input frequencies above 10 kHz, thus eliminating the aliasing illustrated in these examples.

Today, the following sampling rates are used in professional digital audio systems:

- 44100 (44.1 k, employed on compact discs).
- 48000 (48 k), the default rate used on most video DVD discs. 48k also was the default on most 2 and 8 channel audio DAT decks, such as the Alesis *ADAT* and Tascam *DA-88*, which are now obsolete, and also is used in some standalone hard disk recording systems
- 96000 (96 k) ; gradually becoming the new standard for professional audio production
- 86200 (88.2 k) : less common than 96 k; also available on some recording and playback systems and on DVD-Audio format
- 192000 (192 k) : available on some higher end systems; not widely used because of its enormous data storage requirements, and because most people can hear no qualitative difference between 96k and 192k

44.1k and 48k systems enable us to capture frequencies up to about 18 kHz (above the range of hearing for us humans) This might seem more than adequate, but higher sampling rates can provide smoother low pass filtering, and thus, subjectively, more "air" or "spaciousness" in recordings.

To obtain the highest possible audio quality, especially in capturing very rapid attack and decay transients, and complex musical textures, with greater clarity and crispness, however, increased *bit depth* is generally more important than high *sampling rates*. Between the mid 1970s and the late 1990s 16 bit resolution was considered "professional quality." Each bit provides 6 dB of signal resolution and dynamic range. 16 bits theoretically provide about 90 dB of dynamic range (although the actual dynamic range, and signal-to-noise ratio of any audio system always will be less than the theoretical maximum). 20 bit systems afford a theoretical 114 dB dynamic range, 24 bit systems a theoretical 138 dB range, and 32 bits a theoretical 186 dB range.

64 bit word size is available today as an internal signal processing option with some audio applications, such as *Csound*. At first glance this might appear to be overkill, since our auditory systems have a dynamic range somewhere between 120 and 130 dB. However, 32 or 64 bit resolution can make a quite audible difference in signal quality and clarity when we process reverberant and other types of complex signals, where very slight but repetitive, incremental round off errors produced by smaller word sizes can result in quantization error, grunge and even artifacts.

Due to limitations in analog circuitry very few hardware DAC and ADC converters today exceed 20 or, at most, 22 bit resolution. 24 bit converters can read and write 24 bits, but the last 2, 3 or 4 bits often do not actually produce any change in the waveform. To illustrate the precision of 24 bit resolution, if 24 bits with all bits on represents the distance between New York and Los Angeles, the least significant bit represents a distance of about one meter.

96k 24 bit systems require hefty computing power, and, more importantly, large storage capacity (at least by audio standards). Each second of a 44100 16 bit stereo recording consumes 176400 bytes, and one megabyte of computer, disk or tape memory holds only about 5.6 seconds of music. 96k 24 bit recordings can require as much as four times this storage capacity.

One of the first decisions you will need to make when beginning a composition or making a recording in the studios is whether to employ 44.1k 16 bit resolution (44/16) or 96k 24 bit (9624) resolution. 9624 resolution provides higher audio quality, but also some limitations and drawbacks. Audio compact discs — still the most widely available playback medium available today — requires samples in 44/16 format. Most DAT decks, and many other types of portable digital audio recording and playback systems, can only record and play back in 48k or 44.1 k 16 bit resolution. Many of the audio applications in both ECMC studios do not support 9624 resolution. The majority of our *sflib* soundfiles are 44/16 samples. 9624 requires much greater disk space and more processing power and I/O. When mixing in real time, fewer tracks may be available; synthesis and sound modification algorithms run more slowly.

There are significant differences in audio quality between a $50 DAC/ADC chip and converters costing thousand of dollars. Cheap converters, for example, can suffer from *jitter*, which occurs when the bits for successive samples do not change at exactly the same instant, leading to incorrect conversion and the addition of noise and "harshness" to the signal. The highest quality converters available today are made by companies such as Apogee, but are very expensive, often costing five thousand dollars or more for two channel A/D/A (anal-to-digital and digital-to-analog conversion). That is the reason we do not have banks of Apogees in our studios.

☞ **Magnetic signals:** Both analog and digital signals can be converted into corresponding magnetic signals, which then can be written to a magnetic tape coated with zillions of tiny iron oxide particles (tiny magnets). Magnetic recording to tape (performed on 2-channel *DAT* decks the legacy Otari *DTR-8S* decks in the ECMC studios, and, in the past, on the eight channel multitrack digital recorders by Alesis and Tascam and on analog cassette recorders) has several disadvantages. Cueing (accessing) particular spots on a tape is comparatively slow, and tapes are less robust than discs and other newer media formats. Tapes have a usable life of about 15 years for analog signals, perhaps double this longevity for digital signals. After this period, a tape begins to degrade physically, and the iron oxide particles begin to lose their magnetic polarity, resulting in distortion, noise and digital data errors.[1]

3.2. Audio equipment in the Linux sound room

The Linux sound room (room 52) contains two separate audio systems:

• a primary system for playing mono, stereo and quad soundfiles from either *madking* or *gesualdo* over the four *Genelec 1031* loudspeakers and the Genelec subwoofer, and for recording monophonic microphone or analog signals into soundfiles on *madking*; and

• a dedicated 8-speaker system that is used for playing 8 channel soundfiles from the *madking* sound disk over eight *Mackie 824* speakers that are arranged in the 8 corners of a cube surrounding the listening position

Most or all of your work in the studio likely will be done using the primary, 1/2/4 channel audio system. The playback signal path for this audio system is shown in this table:

madking ->	*madking Delta 1010* ->	*Coleman A/B switcher* ->	*Blue Sky BMC controller* ->	*Genelec speakers*
gesualdo ->	*gesualdo Delta 1010* ->			
computer	audio interface	audio source switcher	bass management & volume control	playback

An *audio interface* routes audio signals into and out of a computer system. *madking* and *gesualdo* each have their own *Delta 1010* audio interface, located in the audio rack near the rear of room 52. (The third Delta 1010 interface in the rack is for the 8-channel system.) The audio outputs of both the *madking* and the *gesulado Delta 1010* units are connected to inputs on a *Coleman* model *A/B 5.1* switcher, which switches the audio signal source between the *madking* and gesualdo. The mono, stereo or quad audio signal from *madking* or from *gesualdo* then is routed by the Coleman switcher to a Blue Sky *BMC bass management controller*, which filters out low frequencies from all of the input channels and routes them to the subwoofer. The *BMC* unit also allows us to control the loudness of all playback channels with a single black knob on the *BMC* remote. From the *BMC*, the audio signals are routed to the Genelec speakers for playback.

When playing back mono soundfiles, some audio applications send all of the the signal out channel 1, to the left front Genelec speaker, while other applications send this mono signal out both channels 1 and 2, to both of the front Genelec speakers. Stereo soundfiles are routed to the two front speakers.

On *madking*, compositions can be created in either 2 (stereo), 4 (quad) or 8 audio channels.[2] Eight channel playback requires use of the separate 8-channel audio system. The software and hardware on Windows system *gesualdo*, by contrast, is configured primarily for conventional stereo reproduction.

The "four speakers in a square surrounding the listener" setup of the Genelec speakers in the primary audio system works particularly well for horizontal *ambisonic* sound reproduction, allowing us to localize

[1] A digital system with good *error correction* capabilities, however, can compensate for all or most of these errors.

[2] While it also is possible to create a 3,5,6 or 7 channel work, no ECMC user has yet found a good reason to do this.

sounds at any virtual (perceived) two dimensional point on a horizontal plane between front and rear, left and right. The audio channels and speakers are numbered according to the most common American multichannel numbering scheme, which generally treats channels in stereo pairs:

1 (left front)		2 (right front)
	⇑ Listener(s) facing forward ⇑	
3 (left rear)		4 (right rear)

By contrast, Europeans usually number channels in a clockwise circle beginning from the left front, so that the four channels of a quad soundfile normally would be routed to these speaker positions:

1 (left front)		2 (right front)
	⇑ Listener(s) facing forward ⇑	
4 (left rear)		**3** (right rear)

Whenever you create a multichannel composition, be certain to indicate the intended routing of each audio channel to the desired speaker location. If you do not do this, some of your audio channels (like rear channels 3 and 4 in this example) may be played back reversed — something particularly detrimental to ambisonic works.

However, ambisonic processing, discussed in section 7 of this *Users' Guide*, is by no means the only way to employ the four speaker setup in room 52 effectively. Traditionally intensity panning quad techniques, using the rear speakers for reverberant ambience or other types of "effects," and several other possibilities also are available.

By contrast, the MIDI studio includes 5 speakers plus a subwoofer, and is designed to facilitate conventional cinema-style *5.1 surround* processing techniques, available in several leading commercial audio applications and plugins, as well as stereo reproduction. If you want to create a *5.1* piece, you should use the MIDI studio for this purpose.

Full three dimensional *ambisonic* processing, which additionally includes a low-to-high sound localization height dimension for all sound sources, requires a minimum of 6 loudspeakers and preferably 8 or more speakers, which must be positioned at two height levels. The secondary, 8-channel audio system in room 52, employing a separate sound server (computer) to route soundfile samples from the *madking* sound disk through a Delta *1010* audio interface to eight Mackie model *824* speakers, is designed for playback of 8 channel ambisonic works. Usage of this 8-channel system is discussed in more detail at the end of section 7 of this *Users' Guide*.

The following table lists the audio equipment in use in the Linux sound room as of August 2006. **Normaled** input and output connections between devices, indicated in the two right hand columns of this table, are permanently wired, and require no patching. All of the common recording and playback signal paths in the studio have been normaled. The only situation in which you will need to make a patch chord connection is if you bring some external auxiliary (*AUX*) device, such as a hardware synthesizer or a laptop, into the studio for use.

Audio equipment in the Linux sound room (room 52) (Components of the 8-channel system are shown in square brackets [and])			
Unit:	Manufacturer and model	Normaled input(s) from:	Normaled output(s) to:
(1) ANALOG SIGNAL SOURCES:			
condenser **microphone**	*Audio Technica 3035*		*Voicemaster* pre-amp
analog signal **pre-amplifier**	*Focusrite* **Voicemaster**	microphone	*madking Delta 1010*
(2) MONITOR DEVICES & CONTROLS:			
audio interface (*madking*) audio interface (*gesualdo*) [8 chan. audio interface]	*Delta* **1010** *Delta* **1010** *Delta* **1010**	*Voicemaster*, *madking* *gesualdo* 8 chan. server	*Coleman* switcher *Coleman* switcher 8 Mackie speakers
Audio source switcher	*Coleman A/B 5.1*	*madking* & *gesualdo Delta 1010*s	*BMC*
Bass management controller	Blue Sky *BMC*	*madking Delta 1010*	Genelec speakers
5 active **loudspeakers** [8 active **loudspeakers**]	*Genelec:* 4 model **1031**s and a **1092** subwoofer *Mackie 824*	*BMC* 8 chan. *Delta 1010*	
(3) RECORDING & PLAYBACK DEVICES:			
Otari audio **DAT** deck (rarely used today)	*Otari* **DTR-8S**	Analog: *Voicemaster* AES/EBU: from *madking Delta 1010* Coax: from *gesualdo Delta 1010*	Analog: to *O3D* AES/EBU: to *madking Delta 101* Coax: to *gesualdo Delta 1010*
Computers: Linux PC (*madking*) Windows PC (*gesualdo*) [8-channel server]		madking *Delta 1010* gesualdo *Delta 1010* *madking snd* disk	madking *Delta 1010* gesualdo *Delta 1010* 8 channel *Delta 1010*
(4) MIDI Signals :			
MIDI controller MIDI interface	MIDI keyboard M-Audio **Midisport 8x8**	MIDI signals from MIDI keyboard & *madking*	*Midisport 8x8* MIDI signals to *madking*

Audio signal sources:

A signal source to be recorded can be acoustic, electronic or digital in origin.

An *acoustic* signal is transduced by a microphone, then pre-amplified (and perhaps processed for control of amplitude and/or timbre) by a pre-amplifier, then routed to the *analog-to-digital converters (ADC)* on the recording device.

Occasionally users bring synthesizers, samplers, electric guitars or other instruments with *analog* outputs into the studio for digitizing. These *auxiliary* signal sources must be patched into the *Line In* jack on the Focusrite *Voicemaster* pre-amp, and then are treated much like a microphone signal input. If you wish to capture the *digital* audio output of some external device for recording, consult a staff member to determine the best procedure to use.

Recording and playback devices:

Recording devices:

All recording and playback devices in the ECMC studios are digital. The ECMC computer systems that can be used for digital recording include:

- Linux PC *madking* in room 52 (monophonic recording only)
- Windows PC *gesualdo* in room 52 (monophonic recording only)
- MIDI studio *Windows* PC system *igor*, which includes an RME *Multiface* audio interface and supports stereo recording
- *Mac Pro* computer *wozzeck* in the MIDI studio, which employs an RME Hammerfall audio interface and supports stereo recording

• Two portable laptop systems, available with some restrictions to advanced ECMC users on a sign-out basis, for remote recording and concert playback:

☞ A Macintosh Powerbook with a *Mark of the Unicorn Traveler* audio interface that includes eight pre-amp inputs and eight outputs. This is generally the preferred system for remote recording, and currently is the only ECMC system capable of performing high quality analog-to-digital conversion for more than two input channels.

☞ A Dell model *8100* laptop with an RME *Multiface* audio interface. This machine can be booted either into Linux or into *Windows XP*. We generally use this system for real time synthesis, signal processing and playback in concert performances.

As of this writing the Linux studio (room 52) employs a single Audio Technica microphone and a single Focusrite *Voicemaster* pre-amp for microphone signals, so only monophonic recording is possible. The MIDI studio (room 54) includes two *Voicemaster* pre-amps and two Audio Technica microphones, making stereo recording possible; The Dell 8100 Windows/Linux laptop system includes a *Voicemaster Pro* pre-amp, affording monophonic recording, while the *Powerbook* laptop and its *Traveler* audio interface allow recording up to eight channels simultaneously. Thus ECMC users have several options as to where and how they make audio recordings.

Some signal quality and practical considerations:
The hardware quality and specs of a microphone pre-amplifier and analog-to-digital converters, the sampling rate and word size of the recording, and the environment in which a recording is made all will affect the resulting audio quality.

☞ 96k 24 bit recording is available on all four of the computers in the Linux and MIDI studio sound rooms, as well as on our two portable laptop systems. However, if your composition will wind up on a compact disc, it will ultimately be heard in 44.1k 16 bit resolution. If you make source recordings at 96/24 resolution, but do not apply a good *dithering* processing algorithm when converting these source recordings to 44.1/16, the audio quality often will sound <u>worse</u> than if you had made good 44.1/16 source recordings in the first place. Most available dithering programs are plug-ins to complex audio mixing applications such as *Cubase* and *Nuendo*, and many can be somewhat confusing to use, especially the first few times. The Windows application *resample*, currently available on *gesualdo*, is easier to use when converting sampling rates and word sizes, and provides good dithering algorithms when downsampling from 9624 to 44/16. The question you must always ask about 96/24 recording is: Is it worth it?

☞ Ambient room *noise level* and modal resonances are vital considerations that affects signal quality. None of our studios is fully optimized for recording, for practical reasons (they also must serve as teaching studios, and they are small) as well as financial reasons. We have moved all computers out of rooms 52 and 54 in order to eliminate computer fan noise and disk chatter from these rooms. However, there is not a whole lot more we can do to reduce noise from the ventilation systems, from room 120 above and the hallway outside. This problem of modal resonances (uneven responses to different frequencies within a room) is worse in the Linux studio than in the MIDI studio because the room is smaller.

☞ Since the Linux sound room currently has only one *Voicemaster* pre-amp, stereo recordings must be made in the MIDI studio, or with the portable *Powerbook*. Recordings made in room 54 or with one of our the laptops must be transferred quickly to *madking* or gesualdo and then deleted from the room 54 system or laptop.

I really cannot recommend a single, "best" method or platform for making <u>all</u> of your recordings. Your working preferences may not be the same as mine. Additionally, some instruments, such as the piano, marimba and drum kit, generally are more successfully recorded with two or more microphones, (even if this recording ultimately is bounced down to mono), or in a particular type of room environment.[3] After you have made some test recordings of various types of sound sources on two or more computer systems, employing various software recording applications, you probably will find the platform, software and procedures best suited to most of your recording needs, and you also should know when and how to employ alternative recording systems and procedures as the occasion demands.

---

[3] The reference texts *The Microphone Book* by John Eargle and *Professional Microphone Techniques* by David Huber and Phillip Williams, available in room 53, can provide you with valuable information on ways to mic various types of instruments and sound sources.

3.3. Monitoring devices and controls

We can *monitor* audio signals both aurally (by listening to them over the studio loudspeakers or headphones) and also visually (by means of the hardware or software signal level meters or oscilloscope). The principal monitoring devices in the Linux studio include

- the Genelec and Mackie loudspeakers;
- headphones (used mostly when recording);
- signal routing devices including the Coleman switcher, the M-Audio *Delta 1010* interfaces and the Blue Sky *BMC* controller; and
- software that provides visual monitoring of waveform displays and signal levels.

The **Coleman A/B** switcher

Operation of the *Coleman A/B* switcher, a one unit blue panel located in the audio rack at the rear of the Linux sound room, is extremely simple. Push the red button IN to select *madking* as the audio signal source, and push this button OUT to select *gesulado* as the sound source.

The **Delta 1010** audio interfaces

The three *Delta 1010* interfaces for *madking*, *gesualdo* and the 8-channel server each provide 10 inputs and 10 outputs —8 channels of analog inputs and outputs, and a stereo SP/DIF signal for straight digital transfers (no conversion).[4] The three *Delta 1010* units are mounted on the audio rack in room 52. Two position switches for all analog inputs and outputs on the rear panel of the *1010* toggle between either +4 dB (for connecting "professional grade" equipment) and -10 dB (for "semi-pro" equipment level). These switches should always be set to the +4 position, with the buttons OUT. In addition to its audio converters and connections, the *1010*s include both *word clock* I/O that can sync to any other word clock device for sample accurate transfer of digital signals, and also MIDI signal input and output connections.

The companion *1010* PCI cards installed inside the computer backplanes include a digital mixer that can be controlled by an *envy24* software panel on *madking* and by the application *Totalmix* on *gesualdo*. These virtual mixer handles signal channel routings, panning, soloing, muting and so on. However, in all likelihood you always will want to use the settings we have programmed and will not need to change these defaults. For more information on the *Delta 1010*s, consult the hard copy manual in room 52.

The **Genelec** and **Mackie** loudspeakers

The **Genelec** model *1031* speakers and model *1092* subwoofer speaker are extremely high quality *active* monitors. "Active" means that each speaker includes its own built-in power amplifier. The *1031*s are *bi-amped* speakers, with two physical speakers cones: a metal dome *tweeter* that transduces higher frequencies within audio signals, and a larger 8 inch polymer composite cone that transduces lower frequencies. An internal electronic *crossover* network filters high and lower frequencies respectively to the tweeter and woofer cones, and also includes protection circuitry to shut down the speaker when excessive signal level is detected. (But do not put this protection circuitry to the test —just in case.) The *1031*s are capable of producing frequencies between 47 Hz and 22 kHz (+/- 3 dB) without audible distortion at up to 120 dB SPL (peak response) at a distance of one meter.

The companion 1092 subwoofer contains two 8 inch cone drivers with large magnets and a 180 watt RMS power amplifier. The frequency response of this sub is 33 to 80 Hz (+/- 2.5 dB) and it can put out a 115 dB SPL level at a distance of one meter. Note that both the *1031*s and the *1092* sub can produce frequencies between 47 and 80 hertz, but the sub can generate these frequencies at higher levels without distortion. Note also that much music does not require a subwoofer. For example, playback of a string quartet, where the lowest note —the open C string on the cello —has a frequency of 65.4 Hz, could be handled very well by the Genelec *1031*s without the subwoofer.

The eight **Mackie** model **824** loudspeakers, used only for playback of 8 channel soundfiles, are high quality professional monitors, although not quite as precise as the Genelec *1031*s. The *824*s can produce frequencies as low as 47 hertz with almost no measurable distortion (or down to 37 hertz with slight distortion when no subwoofer is available). Although the *824*s, like the Genelec *1031*s, are designed to function

[4] The SP/DIF connection on *madking*'s *1010* provides digital transfer of signals between *madking* and the Otari DAT deck, which we now use only in rare cases, generally to make "protection" copies of critical recordings.

as *nearfield* monitors (placed at a distance of between one and three meters from the listening position), they are powerful enough that we also use a companion set of eight *824*s with our remote laptop systems for concert playback of 8 channel works. (Eight channel output usually does not require as much power to drive any individual speaker as stereo output.)

Both the Genelec and Mackie speakers include several switches and pots on the back panel to optimize playback for various types of speaker placement and room environments. We have calibrated these controls for optimum performance in the Linux sound room. Never attempt to change any of these speaker settings for any reason.

The **Blue Sky BMC** controller

The *BMC* ("bass management controller") manufactured by the Blue Sky company links the *madking Delta 1010* audio interface to the Genelec speakers and performs two very important audio monitoring functions in the studio:

1) It provides a single rotary volume control pot that controls the gain of all five Genelec loudspeakers in the studio. Without this master control, if we were using all five speakers and wanted to increase or to reduce the volume of all channels we would need to raise or lower five faders.

2) It provides *bass management* for all audio channels —something that may require a little explanation.

   As noted above, the four Genelec *1031* speakers have a frequency response that extends down to 47 Hz, while the *1092* subwoofer can produce undistorted frequencies, at higher gains, down to 33 Hz. Electronic crossover circuits within the **BMC** filter out all frequencies below 80 HZ from all four of the four principal audio channels and send these bass frequencies to the subwoofer.

If we play back a recording of the lowest note (*a0*) on the piano, which has a frequency of 27.5 hertz, in room 52 or room 54, the fundamental frequency and the second harmonic (55 Hz) will be produced by the sub, while all higher harmonics will emanate from one or more of the *1031* speakers. This might seem odd, in theory, since a single sound is reaching our ears from two different source locations. In fact, some acousticians and audiophiles argue that this <u>will</u> "color" (or "distort") the piano tone timbre in subtle ways, that subwoofers should not be used, and that only "full range" speaker enclosures (typically triamped, with three speakers capable of producing the full audible range of frequencies) should be employed. "Full range" speakers are expensive and often rather unwieldy, however, and thus are not commonly found in recording studios. In addition, the conventional acoustical wisdom is that the ear is largely (or even completely) insensitive to phase, and thus to source location, for very low frequencies because of their very long wavelength. (The low pitched piano tone in our example has a wavelength of 12.5 meters, or 41.09 feet.)

BMC controllers are often used with 5.1 surround systems, as in the ECMC MIDI studio. The main panel of the *BMC* includes inputs and outputs for five audio channels and a subwoofer. (We do not use the center channel input or output in room 52.) The blue remote control unit, located near the computer monitors in rooms 52 and 54, includes the following controls:

- a large black potentiometer in the lower left corner, which serves as the **master gain control** for our studios to raise or lower the volume of all audio channels by the same amount
- directly above this pot, two *GAIN* buttons; the $\boxed{\text{MUTE}}$ button mutes all channels, while the $\boxed{\text{REF}}$ button restores the gain of all channels to a "normal listening level" that we have calibrated
- in the *MENU* portion of the remote, 6 navigation buttons that normally are used only when calibrating the system. You should have no reason to do this.

Should you need more detailed information, consult the hardcopy *BMC* manual in room 52 or room 54. A *BMC* unit also is available for use with our two remote laptop systems.

3.4. Signal Levels

Among the most fundamental concerns in audio applications is obtaining proper signal amplitude levels, especially when passing a signal from one device to another. The softest sound we can hear and the loudest sound (one at a listener's threshold of pain, or potential damage to the hearing mechanism) have an amplitude ratio of greater than a million to one. A logarithmic **decibel** scale is used to express electrical voltage ratios between two analog signal levels, such as between the input and output of an amplifier. A *gain* (increase) of 6 dB equals a doubling in amplitude. An attenuation of 6 dB results when we reduce the amplitude of a signal by half. Note that the decibel scale expresses a ratio between two signal levels, rather than absolute physical levels. *0 dB* is a reference level, which may be a very "soft" signal, a very "loud" one, or anywhere in between.[5]

The following chart indicates logarithmic voltage ratios in decibels, corresponding linear amplitude ratios, and digital integer representation of these ratios over the 90 dB range possible with 16 bit systems:

dB ratios:	linear ratios	digital integer representation (numbers ranging between +32768 and -32767)	
0 dB	1:1 (1.) (*reference level*)	+/- 32767	(requires 16 bits)
-6 dB	1/2 (.5)	+/- 16384	(requires 15 bits)
-12 dB	1/4 (.25)	+/- 8192	(requires 14 bits)
-18 dB	1/8 (.125)	+/- 4096	(requires 13 bits)
-24 dB	1/16 (.0625)	+/- 2048	(requires 12 bits)
-30 dB	1/32 (.03125)	+/- 1024	(requires 11 bits)
-36 dB	1/64 (.015625)	+/- 512	(requires 10 bits)
-42 dB	1/128 (.0078125)	+/- 256	(requires 9 bits)
-48 dB	1/256 (.0039 +)	+/- 128	(requires 8 bits)
-54 dB	1/512 (.00195 +)	+/- 64	(requires 7 bits)
-60 dB	1/1024 (.00097 +)	+/- 32	(requires 6 bits)
-66 dB	1/2048 (.000488+)	+/- 16	(requires 5 bits)
-72 dB	1/4096 (.000244+)	+/- 8	(requires 4 bits)
-78 dB	1/8192 (.000122+)	+/- 4	(requires 3 bits)
-84 dB	1/16384 (.000061+)	+/- 2	(requires 2 bits)
-90 dB	1/32768 (.000030+)	+/- 1	(requires 1 bit)

On a 20 bit system (with sample values ranging between +/- 524,272) a 114 dB dynamic range is theoretically possible. On a 24 bit system (with sample values ranging between +/ 8,388,352) a 138 dB dynamic range is theoretically possible. As note earlier, however, these theoretical maximums are rarely if ever achieved with hardware circuitry or transducers.

20 dB represents a linear ratio of approximately 10 to 1. The human auditory system is sensitive to a dynamic range somewhere between 120 dB (1,000,000 to 1) and 130 dB.

0 dB	1 : 1	(threshold of hearing)
20 dB	10 : 1	
40 dB	100 : 1	
60 dB	1000 : 1	(normal conversation or telephone ring level)
80 dB	10,000 : 1	
100 dB	100,000 : 1	(chain saw)
120 dB	1,000,000 : 1	(threshold of pain)
130 dB		(nearby jet take off)

A signal ratio of 1 dB is about the smallest difference between two signal levels that we can perceive.

The *dynamic range* of an audio signal (or "program") is the dB ratio between the highest and lowest amplitudes encountered within the program. Elevator music may have a dynamic range of only 10 dB — maybe even less for the really sappy stuff. The dynamic range of symphonic works by Mahler, Strauss or Tchaikovsky may approach 120 dB, and rock music peaks can reach 150 dB.

Over time, prolonged exposure (more than a few minutes at a time) to sustained amplitude levels above 90 dB or so — a potential hazard for rock musicians, in particular —can lead to two types of permanent hearing loss: (1) inability to hear higher frequencies, and (2) the more serious condition of *tinnitus*, characterized by the perception of a persistent loud buzzing sound. Prolonged listening over headphones at very high levels is especially dangerous.[6]

[5] However, there are a few specific measuring scales, such as *SPL* ("sound pressure level") that employ dB ratios and do express absolute physical intensity levels with respect to a fixed physical reference level.

[6] For more information on listening levels and hearing loss see *http://www.headwize.com/articles/hear-*

Physical amplitude levels often correspond only very roughly with psychoacoustically perceived "loudness" levels, especially at lower and medium dynamic levels. Differences in timbre, particularly in high frequency energy levels, are often more important to our perception of loudness than amplitude differences. Very low frequencies —below 60 hertz or so —will never sound "loud," even when cranked up to very high amplitudes. In addition, percussive sounds often have very sharp attack spikes, lasting only a few milliseconds, which may barely register on meters or register so quickly that we cannot visually monitor them adequately. However, memory features on *peak reading meters*, including the meters on some DAW software displays, display and hold signal peaks, including even most transient spikes, for several seconds or until the meter is reset, a very useful feature that does enable us to detect spikes that may clip.

The dynamic range of an electronic circuit, such as an amplifier or equalizer, is the difference between the noise floor (the amount of noise introduced by the circuit) and the hottest signal that the device can process or record without distortion. When recording signals with a very wide dynamic range, it sometimes is necessary or desirable to reduce the dynamic range of the input signal in order to avoid both distortion, in hot passages, and, during playback, unacceptably poor signal-to-noise ratios (often heard as amplifier thermal noise hiss) during very soft passages. In digital recording one must be extremely careful that signal peaks do not clip, since the results can be disastrous —either a click, if we are lucky, or a burst of noise —and these "*overs*" can be very difficult or tedious to correct. Signal *limiters*, such as the *compressor/limiter* circuit within the *Voicemaster* preamplifier, can prevent most digital clips. Compression also can be applied during playback, if the playback system is unable to reproduce the input signal cleanly.

When working with digital audio workstation (DAW) systems visual monitoring of signal levels usually is more helpful and important while recording, mixing or processing than during playback. Software mixing applications and soundfile utilities can detect and alert us to digital *overs* so that we can make the necessary adjustments. During playback, signal level displays provide information only on the intensity of signals being sent to the power amplifiers and loudspeakers. The more important question, at this stage, is whether the power amps and speakers can handle reproduce this signal level cleanly, without any distortion.

3.5. Using the Voicemaster pre-amplifier

The Focusrite *Voicemaster* pre-amplifier is used to boost and control the signal level, and in some cases also to adjust the timbre, of monophonic microphone signals. As its name implies, the internal circuitry and external control knobs of the *Voicemaster* are optimized for ease of use when amplifying vocal microphone sources. However, excellent amplification of instrumental and other types of acoustic sound sources, and of electronic sources, also can be achieved with this unit. The studio *Audio Technica* mic is permanently connected to the *Mic In* jack on the back of the *Voicemaster*. (Monophonic signals from an electronic instrument, such as a synthesizer, must be patched into the appropriate input jack on the back panel of the *Voicemaster* —most often the +4dBu *LINE INPUT* jack, for 1/4 inch patch cords, or else the *LINE OP +4 dBu* jack for female XLR patch cords.) This section summarizes the basic operation of the *Voicemaster*. For additional information on this unit, consult the *Voicemaster* manual, which is available in the studio.

Adjusting the signal level for recording

There are two gain pots on the *Voicemaster* that determine signal level:

(1) a switchable MIC/LINE *GAIN* pot in the *DISCRETE TRANSISTOR INPUT* section near the far left of the unit. This is the actual pre-amplifier, which provides most of the amplification, and it is the most critical control on the unit. For microphone inputs this pot can provide between 0 and 60 dB of gain. For *Line level* inputs (e.g. from a hardware synthesizer) the pot provides *gain* levels between -10 dB and + 10 dB.

(2) a *MASTER FADER* at the far right of the *Voicemaster*, normally centered at 0 dB, which can provide up to 10 dB of additional gain. This pot is used to make fine adjustments in signal level, and also to boost the gain by a few dB if necessary without over-driving the pre-amplifier gain pot into distortion.

To set the gain of a microphone signal:

• When setting an initial signal level, always begin by *bypassing* the circuits for the *Noise Reducing*

ing_art.htm.

Expander, the *EQ* circuits, the *Vocal Saturator*, the *De-Esser* and the *Opto-Compressor*. Make certain that the *In/Out* switch for each of these circuits is *Out*.

• The +48V switch on the *DISCRETE TRANSFORMER INPUT* section should be pressed IN, so that the mic receives phantom power from the *Voicemaster*.

• Set the *Master Fader* pot initially to 0 dB gain (so that this circuit is passive, neither increasing nor decreasing the signal level)

• Now adjust the pre-amplifier *GAIN* pot to bring the mic signal up to a good line level. However, do not overdrive this amplifier. You almost never would want to set this pot wide open (at +60 dB, or 5 o'clock). Watch the *O/L* ("overload") LED near the GAIN pot. If it lights up briefly you are approaching overload (analog distortion) at the pre-amp. If the *O/L* LED lights continuously, you are overdriving the pre-amp.

• Check the output level of the *Voicemaster* on the *PEAK OUTPUT LEVEL (VU)* LED meters under the *MASTER FADER*. Unfortunately, these LEDs are in 5 dB increments, which is less detailed resolution than we would like. Remember that the output of the *Voicemaster* is still an analog signal, so that it well may be possible to exceed *0 dB* on the *Voicemaster* LED meters without creating audible distortion.

• If you already have the pre-amp gain pot cranked up above +42 dB, and if the pre-amp *O/L* LED lights occasionally, but you still need more gain, raise the *MASTER FADER* pot a few dB.

• Once you have obtained a good initial operating signal level at the *Voicemaster*, begin monitoring this signal (aurally and visually) at the input to the the recording device (the DAT deck or a software recording application running on one of our DAW systems), and compare it with the signal level at the *Voicemaster*. Remember that the *Voicemaster*'s *PEAK OUTPUT LEVEL* LEDs tell you only the signal level at the *Voicemaster*. The hardware or software meters of the recording device indicate the actual signal levels that will be recorded. You may need to make additional gain adjustments at the *Voicemaster* to obtain an optimum recording level. While setting this level, USE YOUR EARS, and use visual monitoring only to confirm or clarify what you are hearing.

Using the compressor/limiter, expander, saturator, de-esser and EQ circuits on the *Voicemaster*

In general, it is best to keep recording paths as simple as possible, never routing a signal through any circuit that is not needed. However, applying *compression* can help to avoid clipping of signal peaks, or smooth out an uneven performance; an *expander* can be used reduce background noise; a *de-esser* circuit can attenuate objectionably prominent mouth or breathing noises from vocal signals; and *EQ* (*equalization*) circuits can be used to brighten a dull signal, or to take the edge off an overly shrill signal source.

Before introducing any of these circuits into the signal path to try to correct one of these problems, however, first try to fix the problem by adjusting the microphone placement or the distance and angle between the sound source and the mic. If signal processing still is indicated, introduce the processing circuits one at a time, so that if new problems appear you will know which circuit is responsible for any new signal degradation. I recommend working first with amplitude, then with timbral considerations, and introducing the five *Voicemaster* processing circuits in the following order:

Amplitude adjustments:

(1) the *Opto-compressor* : guard against clipping and tame hot signals

(2) the *"Noise reducing" expander*: deal with room noise

Timbral adjustments:

(3) the "*vocal saturator*" : I don't recommend articially trying to "warm up" signals, although this is much prized in pop music recordings. If you want to try adding some simulated tube distortion, for added "warmth," read the *Voicemaster* manual description of this circuit for more information.

(4) the *de-esser* : take care of objectionable sibilants, fricatives and plosives

(5) the 3 band *EQ* circuits : finally, adjust the overall timbre of the sound source

While experimenting with any of these five circuits, it is a good idea to perform A/B comparison tests, listening to the signal both with the circuit engaged and with the circuit bypassed. Also, these 5 circuits often will alter the overall amplitude level of a signal, generally reducing the amplitude but in certain cases increasing the signal level. When you have all of the pots for one of these five circuits set as you wish, re-check the overall amplitude level, and make any necessary adjustments at the pre-amp or master fader *GAIN* pots.

Compressors and *expanders* both reduce the dynamic range of a signal through attenuation, but at opposite ends of the dynamic range. Compressors reduce the amplitude of high intensity signals. *Expanders*, by contrast, reduce the amplitude of low intensity signals:

High amplitude	0 dB	|	(digital clipping point)
("loud") signals	-10 dB	|	<- **compressor** attenuates signals in this range
	-20 dB	|	-------- compressor threshold
	-30 dB		
Low amplitude	-40 dB	|	--------- expander threshold
("soft") signals	-50 dB	|	<- **expandor** attenuates signals in this range
	-60 dB	|	

Compression and limiting

A **compressor** is an electronic circuit that progressively reduces the amplitude of an input signal, by ever greater amounts, whenever this signal level exceeds a preset threshold:

---- *indicates input to* compressor, - - - *indicates output*

+10 dB			
0 dB	*(desired peak level)*		*0 dB*
-10 dB	*(threshold set to -20 dB)*		*-12 dB*
-20 dB			
-30 dB		*progressively greater*	
-40 dB		*amplitude attenuation*	

In the example above, the threshold is set to 20 dB below the desired peak level. When the signal is below this amplitude level, the output of the compressor is unity gain (no attenuation). Above this threshold, the compressor applies increasing amounts of attenuation. The compression ratio in this example is 3:1. An input signal at 0 dB is reduced to -12 dB. An input signal at +12 dB would be reduced to 0 dB (barely avoiding clipping). An input signal greater than +12 dB would still distort.

A **limiter** is simply a compressor with a very high compression ratio, producing very sharp attenuation above the threshold level:

---- *indicates input to* limiter, - - - *indicates output*

+20 dB			*+20 dB*
+10 dB			
0 dB	*(desired peak level)*		*0 dB*
-10 dB	*(threshold set to -20 dB)*		*-10 dB*
-20 dB			
-30 dB		*very sharp attenuation*	

In the example above a very high compression ratio, such as 20 : 1, is used to sharply limit high level signals (possibly transient peaks), thus avoiding distortion.

To use the *Opto-Compressor* circuit on the *Voicemaster*:

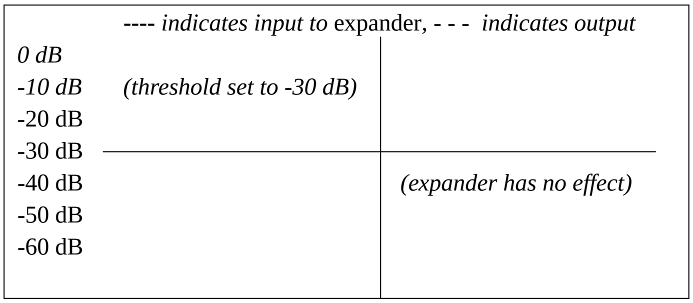
- Push the IN button for the *Opto-compressor* circuit so that it is illuminated and the circuit is introduced into the signal path.
- For *compression*, set the *THRESHOLD* pot to somewhere between -6 dB and 20 dB. For *limiting*, set the *THRESHOLD* pot to about -3 dB, and perhaps engage the *FAST* and *HARD RATIO* buttons. The *Voicemaster* controls are somewhat better suited to general purpose compression than for hard limiting.
- Adjust the *RELEASE* pot to taste. The quicker the release time, the louder the signal appears to be.
- Aurally and, on the LED meters within the *Opto-Compressor* panel, visually monitor the amount of

compression. Too much compression can drain the life out of a signal. On the other hand, if the LEDs never light, the compressor is having no effect.

- If the compression results in reduced amplitude, raise the *OUTPUT* pot to bring the signal back up to a good, hot level.
- The *TREBLE* pot can be raised to brighten the timbre of the signal if desired.

Using the *Noise Reducing Expander*

The term **expander** can be confusing. There are several types of expanders, which work in different fashions. The type available on the *Voicemaster* is called a **downward expander,** which operates only on low amplitude signals. One might imagine that an "expander" would increase the amplitude of low intensity signals, but actually the reverse is true. A *downward expander* expands the ratio between the threshold level and the low intensity input signal, reducing this intensity still further:



Expanders generally are used to try to minimize background noise, either from the performers (e.g. breathing noises, lip pops, foot scrapes, etc.) or from the room environment (e.g. ventilation noise). In pop music *expanders* also are often used as *noise gates*, "shutting down" the circuit amplifier completely (producing silence) when the program falls below the expander threshold, producing abrupt, "staccato" decays in sounds. Use *downward expansion* carefully, or you may get "pumping" or decays that sound unnaturally abrupt.

To use the *Noise Reducing Expander* :

- Push the IN button for the *Noise Reducing Expander* circuit so that it is illuminated.
- Set the *THRESHOLD* pot, generally to between -30 and -40 dB.
- Set the *DEPTH* pot, generally to about half way between *0* (no effect) and *FULL* (maximum attenuation).
- While performing audio tests, watch the LED meters in the *Noise Reducing Expander* panel of the *Voicemaster* and see how much attenuation is being applied to low level signals.
- Do not use the *GATE* (noise gate) button except for special effects. And you know my loathing of most special effects.

Note that the *expander* can be used simultaneously with *compression* to attenuate both high intensity and low intensity signals (leaving only "medium" intensity signals unaltered). The pre-amp or master fader *GAIN* pot then can be raised to compensate for reduced amplitude.

The Opto De-esser circuit

A *de-esser* circuit, used primarily when recording vocal sources, senses and limits high frequencies, generally in the range of 3 kHz through 8 kHz, whenever these frequencies exceed a pre-defined amplitude threshold. Frequencies within this range most often are associated with vocal sibilant and fricative consonants, but they also result from bow scrapes and other noise-like elements within instrumental tones. *De-essing* can be especially helpful when one employs very close micing of vocal signals.

To use the Opto De-esser *circuit on the Voicemaster*:

- Push the IN button for the *Opto De-esser* circuit so that it is illuminated and the circuit is added to the signal path.
- Set the *CUT FREQ* ("cutoff frequency," or the center frequency of the band of frequencies to be attenuated) to about 3 kHz for male voices, somewhat higher (perhaps 4 or 5 kHz) for female voices.
- Adjust the *THRESHOLD* pot to somewhere between 0 and -5 dB. If you set the *THRESHOLD* level too low, you will filter the signal too much, and it may sound dull and lifeless.

Equalization

An **equalizer** (*EQ*) is an amplifier that operates only on a selected band (range) of frequencies within a signal. The frequency response of many equalizers is a bell-shaped curve. In Example 1 below an EQ applies up to 10 dB of gain to frequencies between 750 Hz and about 1250 Hz (.5 octave centered at 1 kHz). In Example 2 the EQ attenuates a band of frequencies between 950 and about 1050 Hz. (.1 octave centered at 1 kHz):

	Example 1: **EQ** *boosts a frequency band:*			*Example 2:* **EQ** *attenuates a frequency band:*	
+10 dB					
0 dB					
-10 dB					
		FC =		**F**C =	
	no effect	*1 kHz*	*no effect*	*no effect* *1 kHz*	*no effect*
		+3dB --- +3dB		*- 3dB --- - 3dB*	
		EQ bandwidth		*EQ bandwidth*	
		= .5 octave		*= .1 octave*	

With **parametric** equalizers three "parameters" can be adjusted by the user:

- center frequency, or *Fc* : the frequency at the center of the bell-shaped curve;
- bandwidth : the range of frequencies above and below the *center frequency* that will be attenuated or boosted; and
- the *gain* (boost or attenuation) applied to the band of frequencies

With **graphic** equalizers, the *center frequency* and the *bandwidth* are fixed, and only the *gain* can be adjusted.

Care must always be exercised when applying gain to a band of frequencies so that you do not add distortion within this band. Meters may indicate that the overall signal level is acceptable, but an EQ circuit is a (frequency dependent) amplifier, and may be distorting frequencies within the selected frequency band. Attenuation with EQ circuits generally is safer than boosting frequencies bands. As always, trust your ears, and do not rely too heavily on visual monitoring, which is only an aid.

In general it is not a good idea to EQ a signal while recording, unless you know exactly what you are doing. If you do apply EQ while recording, be conservative, since you are hearing the sound in isolation rather than within the context of a mix, and there is no *undo*. Remember that EQ can be applied later to a soundfile with many soundfile editing and some mixing applications.

The *Voice Optimized EQ* circuits of the *Voicemaster* are designed for ease-of-use when recording vocal signals. As such, they give us somewhat less control over the EQ parameters than I would like. The EQ unit consists of one partially parametric and two graphic EQ circuits in series, set to the three frequency bands that most frequently required adjustment in vocal signals:

(1) a mid-low *WARMTH* band, adjustable between -12 dB attenuation and +8dB gain, with a companion adjustable center frequency *TUNING* pot adjustable between 120 (bass voice) and 600 hertz (high soprano).
(2) a mid range *PRESENCE* band used to adjust the gain of frequencies centered at 1.5 kHz
(3) a shelving high frequency *BREATH* band that controls the gain of frequencies above 10 kHz

To use the *Voice Optimized EQ* circuits on the *Voicemaster*:

- Push the EQ_IN button on the *EQ* panel so that it is illuminated and the EQ circuits are introduced into the signal path.
- Begin with "flat" (no effect) settings, with the *WARMTH, PRESENCE* and *BREATH*T pots set to 0 dB.
- Mid-lows ("*WARMTH*" band) : If the vocal source has a thin, harsh or distant quality, try raising the *WARMTH* pot about 3 dB. If, instead, the vocal source sounds "dull," "tubby" or otherwise bottom-heavy, lower the *WARMTH* pot about 4 dB. Very close micing can produce the "proximity effect," boosting low frequencies, and reducing the *WARMTH* pot by a few dB can eliminate this effect.
  Play with the *TUNING* pot, which typically should be set somewhere between 200 and 300 hertz for male voices or lower pitched instruments, or between 300 and 500 hertz for female voices and higher

pitched instruments (e.g. flute).

• Mid range ("*PRESENCE*" band) : Raising the pot on this graphic EQ circuit about 3 dB or so will tend to make the signal seem "closer," and will help it to "cut through" complex textures. Trimming this band by a few dB will give the sound a more "mellow" or "distant" quality.

• High frequencies : The *BREATH* control pot, which controls the gain of frequencies centered around 10 kHz and above, allows us to control the "airiness" of a signal. (On many signals it will have little or no effect.)

• A fourth band of EQ, which can be switched in and out by tapping the curiously named *Absence* button on the EQ panel, is a notch filter that reduces the gain of mid high frequencies. I don't recommend this circuit because the user has no control over its parameters.