Wwise

HDR User's Guide

2013.1

Wwise

HDR User's Guide

Wwise 2013.1 Build 4668

Table of Contents

Chapter 1. HDR User's Guide

Understanding HDR

High dynamic range audio (HDR audio) is a technique to design a mix using level values spanning across a very high dynamic range as occurs in nature. HDR is also a run-time system that dynamically maps this wide range of levels to a range that is more suited to your sound system's digital output.

In the real world, the audible dynamic range that spans from the threshold of human hearing to the loudest possible sound in air is several times wider than the dynamic range offered by speakers at game play levels. The role of an HDR system is to collapse or "compress" the whole real life dynamic range, approximately 190 dB, into 96 dB (the dynamic range available for a digital device), and even less in practice due to floor noise levels.

In HDR photography, local tone mapping is applied independently to various regions of an image to enhance the contrast within each region. HDR audio works in the same way; it performs sound level mapping instead of tone mapping, and does so locally in time. Thus, at any given moment, the system automatically adapts the mapping based on the levels of sounds that constitute the audio scene.

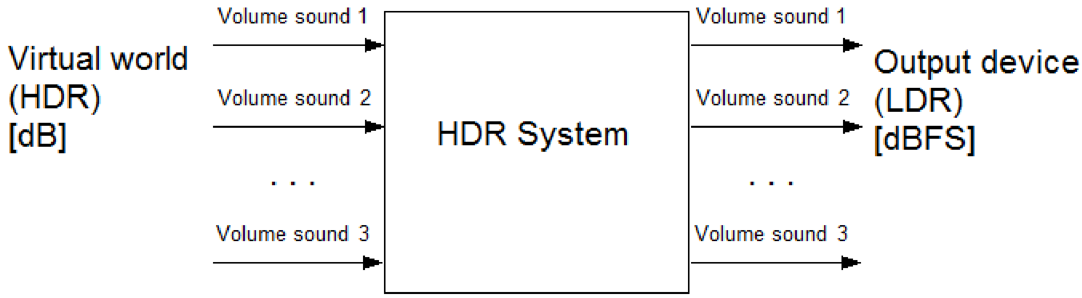
HDR Glossary:

Term	Definition
Decibel (dB)	A logarithmic measure of the level of a sound compared to the level of another sound or an arbitrary reference value. One decibel equals 20 * log10(A/AR) when dealing with amplitude. A difference of +20 dB means that a sound's amplitude is 10 times greater than the reference.
Decibel full scale (dBFS)	A logarithmic measure of the amplitude of a signal compared with the maximum that a device can handle before clipping occurs. A value of 0 dBFS is thus the loudest sound that can be generated by the digital audio output. 16-bit digital audio output devices range from 0 dBFS down to -96 dBFS. The level of the audio signal coming out of the Master Audio Bus in Wwise should therefore lie between these values.

HDR examples

In HDR audio, you can assign volume values to sounds of the game's virtual world that span over a much larger dynamic range than the standard 96 dB of 16-bit output devices, much like they would in the real world. It is the task of the HDR system to translate these values into dBFS, as illustrated in the figure below.

Figure 1.1. HDR system inputs and outputs



The inputs are the sound levels of the virtual world, expressed in decibels (dB) relative to an arbitrary reference. The values can be chosen arbitrarily high or low, and thus have a high dynamic range. The outputs are the levels of the corresponding sounds in dBFS. The range of these values depends on the output device, which typically has lower dynamic range than the input.

In its simplest form, the HDR system operates as follows: at each time slice, the system selects the sound assigned the highest volume in the virtual world, automatically maps it to the output value of 0 dBFS, and then maps all other sounds proportionally.

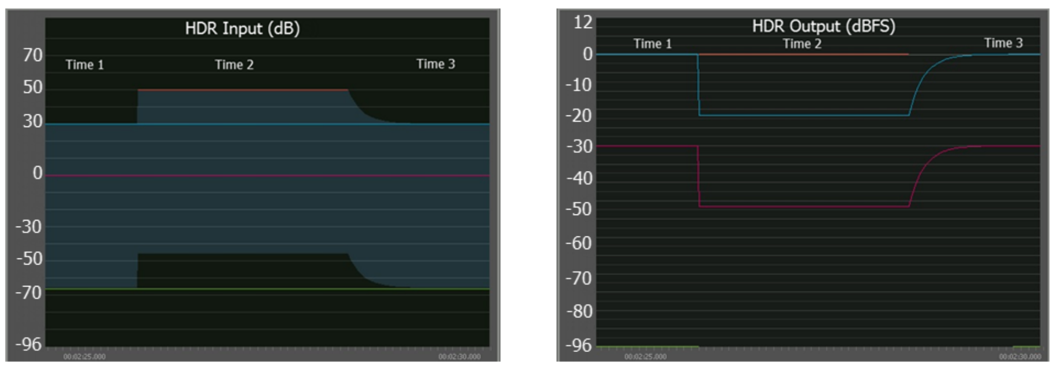
Let's use an example to illustrate this. Suppose that at a given moment ("time 1"), sound "blue" plays at +30 dB in the virtual world, as illustrated on the left side of Figure 1.2, "HDR window". The reference (0 dB) is arbitrary. Because "blue" is the loudest sound at time 1, it plays at 0 dBFS at the output of the HDR system. Another sound, "purple", plays at 0 dB in the virtual world, that is, 30 dB below sound "blue". Thus, it comes out at -30 dBFS at the output of the HDR system. A third sound, "green", plays at -66 dB in the virtual world, which results in -96 dBFS at the output of the HDR system. Since the output of the system is constrained to a dynamic range of 96 dB, the level of "green" corresponds to the lower bound of all audible sounds. At time 1, any sound that is softer than "green" is inaudible.

In the previous example, the range [-66, +30] dB in the virtual world (input side) is referred to as the **HDR window**. It is represented by the blue region on the left side of Figure 1.2, "HDR window". The HDR window has a fixed width, determined by the dynamic range of the output. For a 16-bit device, it is equal to 96 dB at most, but in practice it is usually smaller. At time 1, the sound at +30 dB is the loudest in the virtual world, and any sound below -66 dB is inaudible because it is below the HDR window.

Suppose that later, at time 2, another sound, "orange", starts playing at +50 dB in the virtual world. To accommodate this new louder sound, the HDR system slides the window up by 20 dB, so that its bounds are now [-44, +50] dB on the input side. All sounds are then mapped to the new values. The sound at +50 dB plays at 0 dBFS, the sound at +30 dB now plays at -20 dBFS, and the sound at -66 dB is now below the window and is therefore completely inaudible. When the orange sound stops playing at time 3, the window gently slides back to where it was, and other sounds take their former volume.

Figure 1.2. HDR window

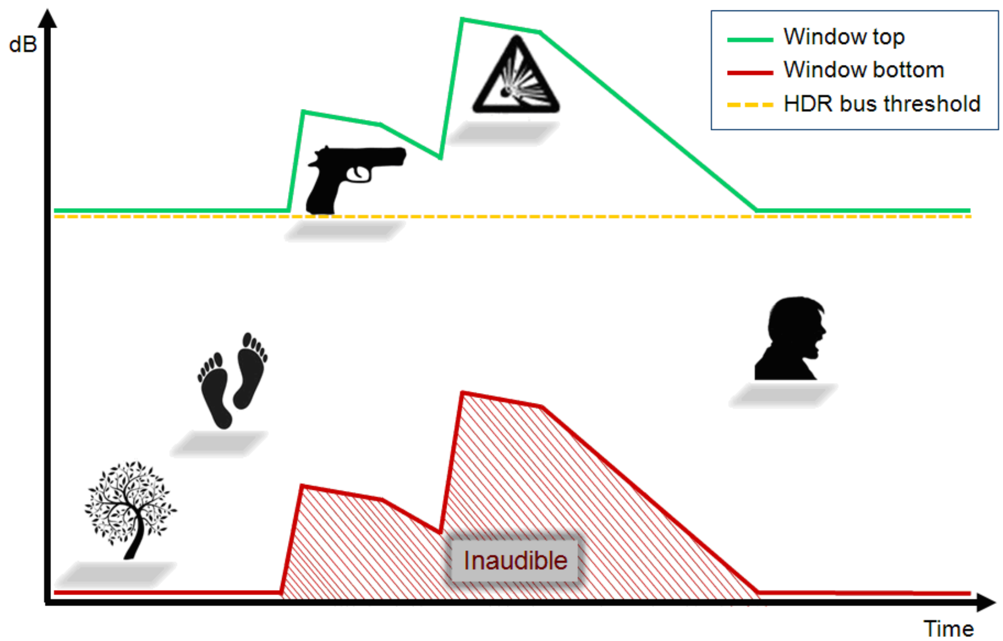
	Time 1		Time 2		Time 3	
	Input	Output	Input	Output	Input	Output
Orange	-	-	50	0	-	-
Blue	30	0	30	-20	30	0
Purple	0	-30	0	-50	0	-30
Green	-66	-96	-66	-∞	-66	-96



On the left, sound levels are represented at the input of the system, in decibels (dB) with arbitrary reference. The HDR window is represented by the blue region. At time 1, the top of the window is aligned with the loudest playing sound (blue). At time 2, another sound (orange) starts playing at +50 dB, and the window slides up immediately to accommodate it. The sound at -66 dB (green) is then clearly below the window, and thus inaudible. The resulting levels in dBFS at the output of the system are shown on the right. When the HDR window slides up by +20 dB because of the orange sound, the volume of other sounds drops by -20 dB. During this time, the green sound is completely excluded from the output. Notice that the orange sound plays at the same output level as the blue sound during time 1. When the orange sound stops playing, the window gently slides down to its former level, and the volume of other sounds increases accordingly.

The HDR system works like a dynamic range limiter/compressor. It affects your mix by making soft sounds inaudible when loud sounds play, and making them audible again when playing alone. The relative levels of sounds between one another in the HDR world are preserved, creating the illusion of a greater dynamic range, while in fact they are compressed within the output device's lower dynamic range. Furthermore, thanks to the system's automatic volume ducking when louder sounds play, your mix will be cleaner and have better focus. The next figure illustrates this principle.

Figure 1.3. HDR overview, as seen on the input side of the system



The window slides up only when louder sounds play. When the window slides up on the input side, the sound volume drops on the output side. What was formerly audible, such as the sound of leaves in a tree, can become completely inaudible when a gunshot is played. The actual volume of sounds at the output of the system depends on the distance between them and the top of the window, at any given moment. Here, the gunshot and explosion would come out of the system at the same level if played alone, but because the explosion is considered louder than the gunshot and effectively ducks its volume, listeners are left with the impression that it is indeed louder.

Using HDR

In Wwise, users need to select a bus to act as a converter between HDR levels and full (device) scale. The input levels that this HDR bus consumes are the logical levels that you set in Wwise. Thus, sounds routed to an HDR bus can have their volume set far beyond 0 dB. The only thing that counts is their position within the HDR window, which is placed dynamically by the HDR bus according to what is playing. The HDR bus therefore acts as a logical limiter/ compressor by exposing controls that are similar to those of an audio limiter/ compressor. It has ballistics (infinitesimal attack, user-defined release) to control how the HDR window slides in time. It also has a threshold, which can be seen as the lowest possible position of the HDR window.

Usage of dB SPL

In prior literature, HDR audio systems often express input side volume levels in terms of Sound Pressure Levels units (dB SPL). dB SPL are a measure in decibels whose reference (0 dB SPL) corresponds to the threshold of human hearing. The notion of SPL does not exist in Wwise because it adds unnecessary complexity, pollutes the interface, and does not make the system more usable. Instead, the input side reference is left arbitrary, and it is up to you to define it. If you wish to work in dB SPL, you can set the volume of sound structures to positive dB SPL values directly. On the other hand, the default range of volume sliders in Wwise goes up to +12 dB only, so it might be more practical to choose another reference, and make the necessary subtraction to find the corresponding relative dB level from the desired SPL value. For example, you may decide that 100 dB SPL is your reference at 0 dB. Then a sound at 80 dB SPL needs to have its volume slider set to -20 dB, a sound at 130 dB SPL needs to have its volume slider set to +30 dB, and so on. You also need to set the HDR bus threshold accordingly.

- Enabling HDR in your Project
- Routing Sounds to an HDR Bus
- Monitoring the Window
- Setting the Dynamics of HDR

Enabling HDR in your Project

To enable HDR in your project, you simply need to define which bus or buses will act as HDR subsystems.

To enable HDR on a bus:

1. Inspect a bus object.
2. Click the **HDR** tab of the **Property Editor.**
3. Click **Enable HDR.**

The volume of all sound structures routed to the HDR bus, or to one of its descendants, is affected by the HDR bus. Sounds that are not routed to the HDR bus are not, but can coexist with sounds routed to the HDR bus within your sound design. You can also have more than one HDR bus in a project, as long as they are completely independent: an HDR bus may not have another HDR bus as an ascendant or descendant.

As mentioned previously, sounds at the top of the HDR window will come out at 0 dBFS. Use the bus volume slider of the HDR bus to scale down the output of the HDR system before mixing it with other non-HDR sections of your project.

Routing Sounds to an HDR Bus

To have sound objects in the actor-mixer hierarchy or music hierarchy use the HDR system, you need to route these objects to the HDR bus.

To route a sound to an HDR bus:

1. Inspect the sound object to show its content.
2. In the **Property Editor**, go to the **General Settings** tab.
3. Select the HDR bus as the **Output bus**.

Monitoring the Window

The Voice Monitor view displays the volume of voices and their envelope (if available).

To understand and debug the HDR system:

1. Open the **Voice Monitor** view.
2. Drag your HDR bus into the view's context.
3. Set the Mode to either Bus input or Bus output. Bus input mode displays voice levels in decibels as seen by the HDR bus, before the HDR compression, while Bus output mode displays them at the output of the HDR bus, after HDR compression and the bus output gain. Most figures in this document are screenshots of the "two sides" of the Voice Monitor view.

Setting the Dynamics of HDR

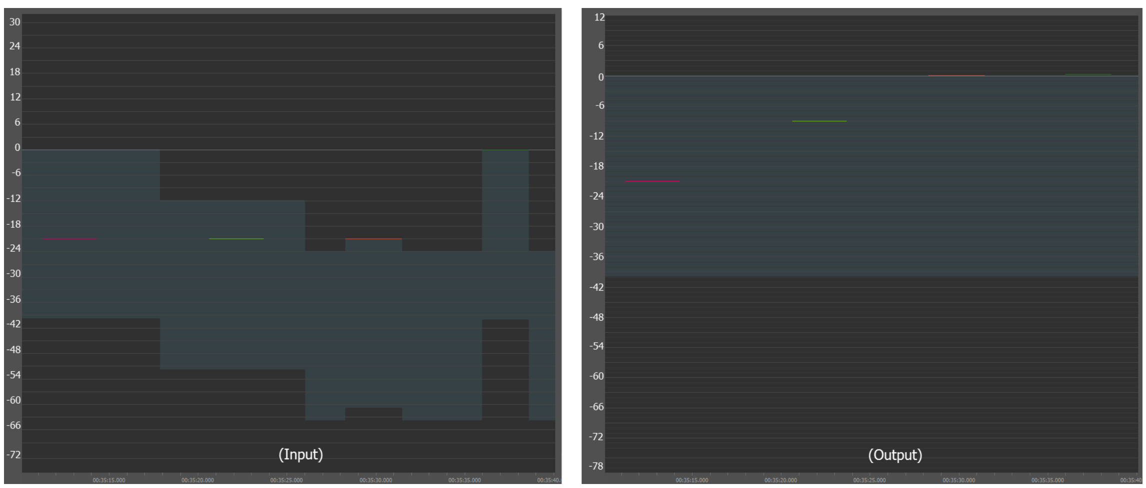
Each HDR bus maintains an HDR window whose position and width is defined by the level of the loudest sound at any moment, and the project's volume threshold respectively. The behavior of the HDR window can be edited in the HDR tab of each HDR bus. Since the HDR bus acts like an audio limiter/compressor, its controls are similar.

- Working with the HDR Threshold
- Working with HDR Ballistics
- Working with the HDR Ratio

Working with the HDR Threshold

Like the threshold of a typical audio limiter/compressor, the HDR threshold defines the minimum input level above which the HDR window may slide (refer to the yellow line in Figure 1.3, “HDR overview, as seen on the input side of the system”), or in other words, the minimum level above which the compressor "kicks in". When only soft sounds are playing, it directly impacts their level at the output: the farther away they are from the window top, the lower their output level is. When sounds play above this threshold, then softer sounds are ducked down automatically. It is important to understand that with an infinite compression ratio (we will discuss about the ratio later), any two sounds whose input level is above the HDR threshold will play back at the same output level when played alone, regardless of their input level as shown in Figure 1.4, “Effect of HDR threshold”.

Figure 1.4. Effect of HDR threshold

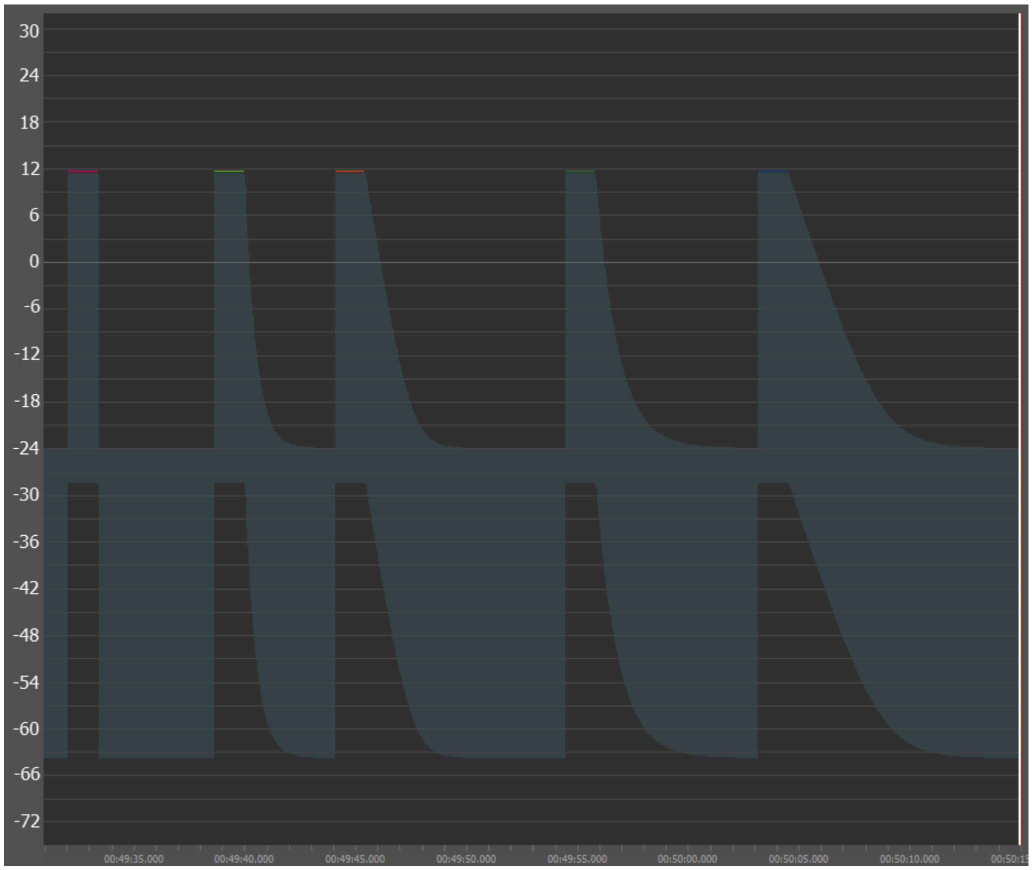


On the figure above, the left side shows input; the right side shows output. The first time, the sound is set at -21 dB with a threshold of 0 dB, and comes out at -21 dBFS. The second time, the threshold is at -12 dB and the sound comes out at -9 dBFS. The third time, the sound is above the threshold of -24 dB, and thus drives the window and comes out at 0 dBFS. The sound is replayed at an input level of 0 dB, even higher above threshold, and still comes out at 0 dBFS

Working with HDR Ballistics

Standard audio compressors implement ballistics (attack/release) that define how compression behaves in time. Likewise, the HDR bus has user controls for the release time and shape, which let you define how the HDR window releases to a lower value. Figure 1.5, “HDR release: 0, 0.5 exponential, 0.5 linear, 1 exponential, 1 linear” shows various release times and modes. Adjust it to minimize artifacts such as pumping. Note however that there is no control for attack time, because the attack of an HDR system has to be instantaneous, like an audio limiter.

Figure 1.5. HDR release: 0, 0.5 exponential, 0.5 linear, 1 exponential, 1 linear

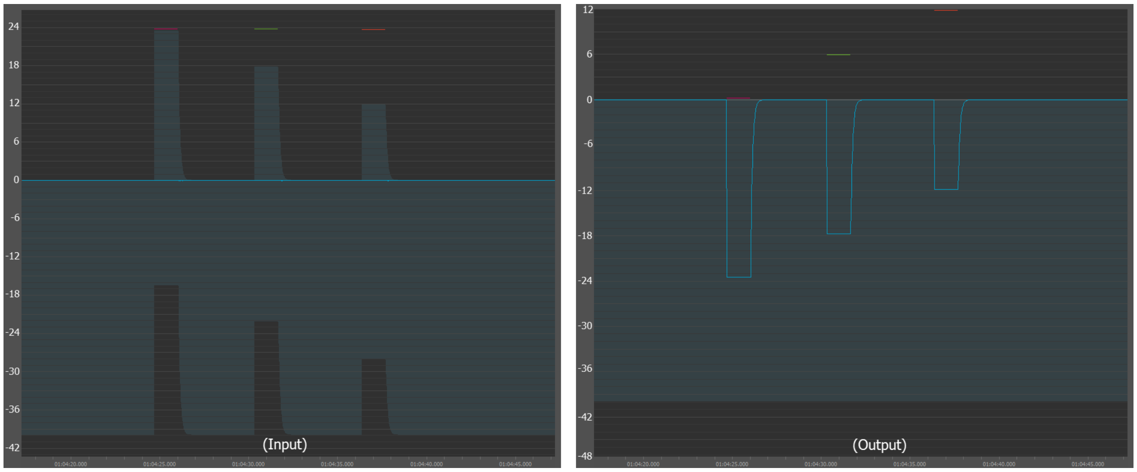


Working with the HDR Ratio

So far, the HDR system has been presented like an audio limiter. With the HDR window top following exactly the level of the loudest sound, every sound above the HDR threshold is played back at the same output level when played alone, as was displayed in Figure 1.4, “Effect of HDR threshold”. As previously described, while HDR increases the perceived dynamic range between sounds relative to each other, it makes the "absolute" dynamic range smaller. It is possible to gain some of this "absolute" dynamic range back by transforming the HDR limiter into an HDR compressor. You do this by using a finite, smaller compression ratio (it is infinite by default). The compression ratio defines how closely the HDR window follows the loudest sound, and is a function of the difference between the sound's level and the HDR threshold. Figure 1.6, “Effect of HDR ratio” illustrates the effect of the ratio. Note that with non-infinite compression ratios, loud sounds may overshoot the HDR window, and thus come out of the

HDR system above 0 dBFS. Give yourself enough headroom by lowering the volume of the HDR bus and/or other buses downstream in the signal path.

Figure 1.6. Effect of HDR ratio



The sound has an input level of +24 dB (left panel), and played in HDR bus with ratio 100:1, 4:1, 2:1. The corresponding output levels of 0 dBFS, + 6 dBFS and +12 dBFS are displayed on the right side. Another sound has an input level of 0 dB and is ducked down by -24 dB, -18 dB and -12 dB. At 2:1, only half of the energy above threshold is used to drive the HDR window up, hence ducking the other sounds by -12 dB. Notice that the level difference of 24 dB is preserved all along. Also, when using smaller ratios, sounds may stand above the window top, hence above 0 dBFS, so you should reserve sufficient headroom by decreasing the volume of the bus or its parents.

Working with Amplitude Envelopes

Since the HDR system works with logical volume values set in Wwise, it is unaware of the actual amplitude of the input sounds, and therefore sees them as black boxes with constant volume throughout the duration of the sound. Most of the time, the amplitude of sounds vary. Imagine an impact sound with a transient and a decaying part. If this sound was loud enough to fix the position of the HDR window, this position will remain unnaturally constant during the whole duration of the sound. See the resulting effect on the window in Figure 1.7, “HDR window with a decaying impact sound without envelope”.

In Wwise, it is possible to let the HDR system peek into the black box by enabling envelope tracking. In the HDR tab of the desired sound's properties, select the **Enable Envelope** check box: the amplitude envelope of the audio file is then analyzed and is attached to the sound's metadata. At run-time, the HDR system uses this data to move the window appropriately, as can be seen in

Figure 1.8, “HDR window with a decaying impact sound with envelope, playing above a softer, steady background sound”.

For the following figure, the input is on the left, the output on the right and the corresponding output wave is below.

Figure 1.7. HDR window with a decaying impact sound without envelope

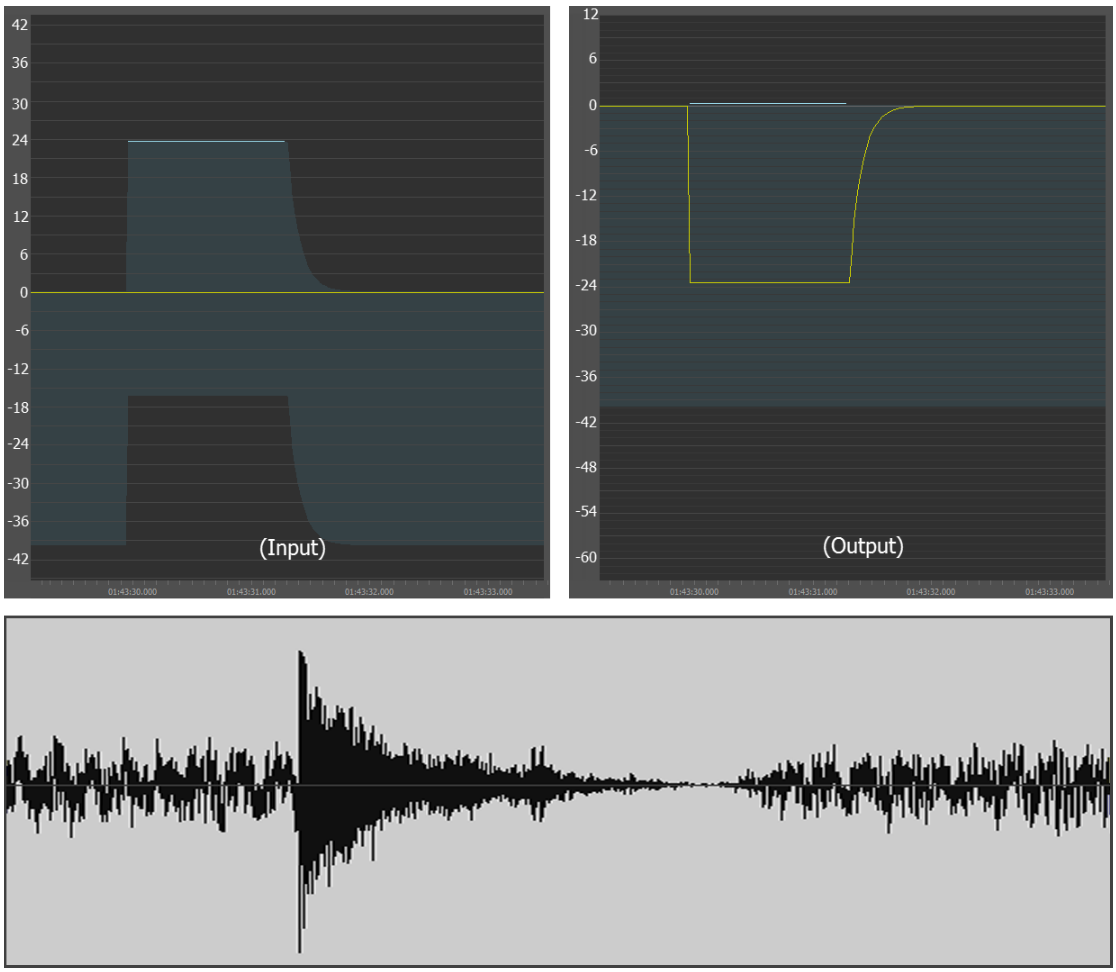
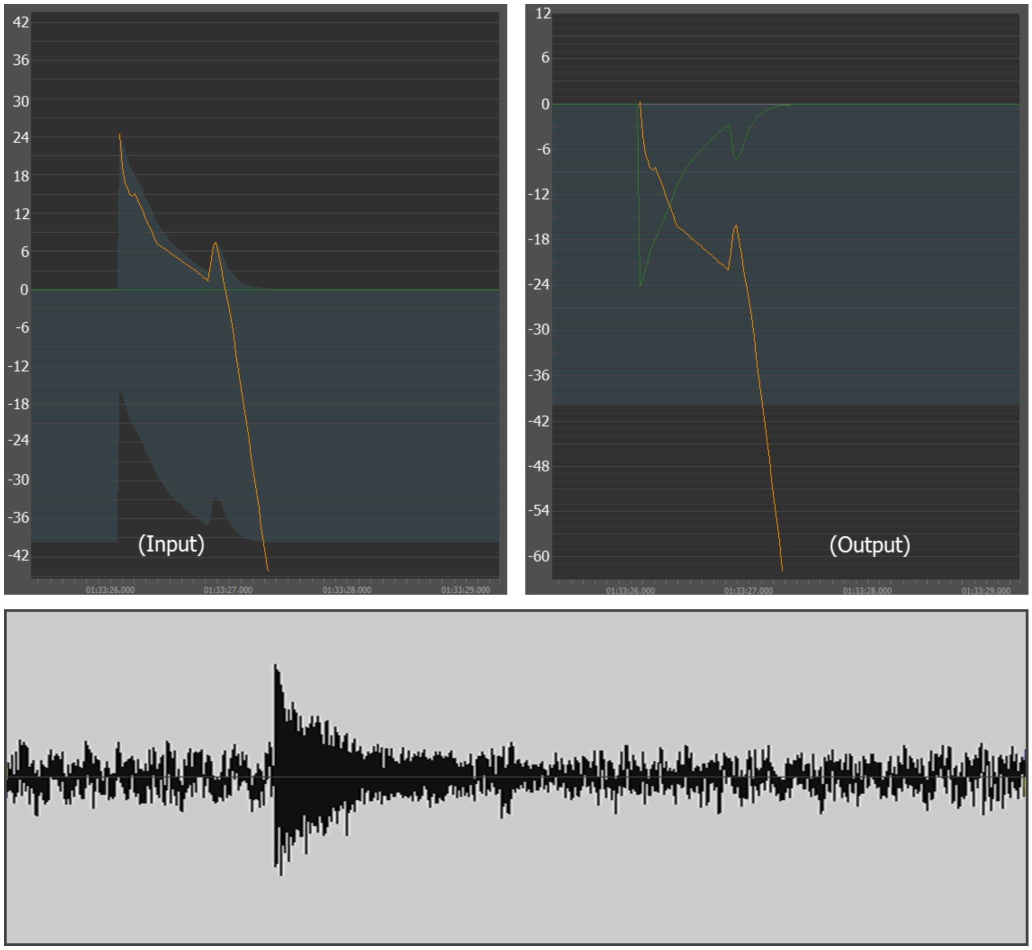


Figure 1.8. HDR window with a decaying impact sound with envelope, playing above a softer, steady background sound



Notice the lull after the impact sound in the first figure above, caused by the HDR system interpreting it as being constant.

Envelope Sensitivity and Manual Editing

Once you enable the envelope, you can preview the result by opening the source editor (next figure). If you are unhappy with the precision and/or number of points, you can change the value of the sensitivity slider in the envelope tracking group box (see Figure 1.10, "HDR tab in Property Editor"). You can work with a group of sounds by enabling the envelope and setting the sensitivity on a higher-level sound structure, an actor-mixer for example. Note that the effect of the sensitivity is dependent on the original audio data, and may therefore be different for different variations of similar sounds. You can also edit the envelope manually for each sound; set the sensitivity control so that you obtain a curve that is close to the desired result, and move, remove or add new points.

Figure 1.9. Source editor in RMS mode, with envelope display

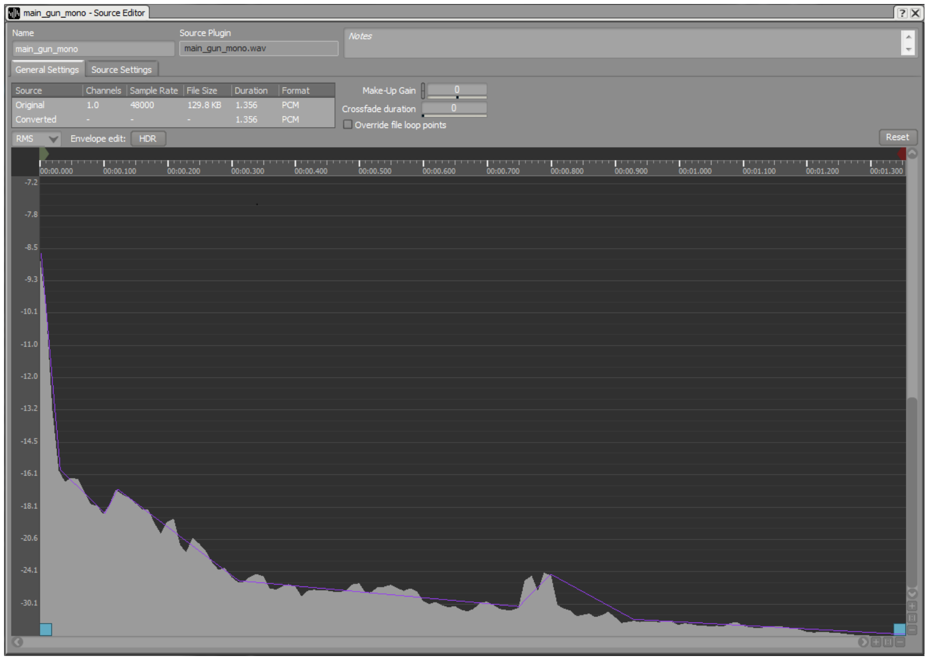
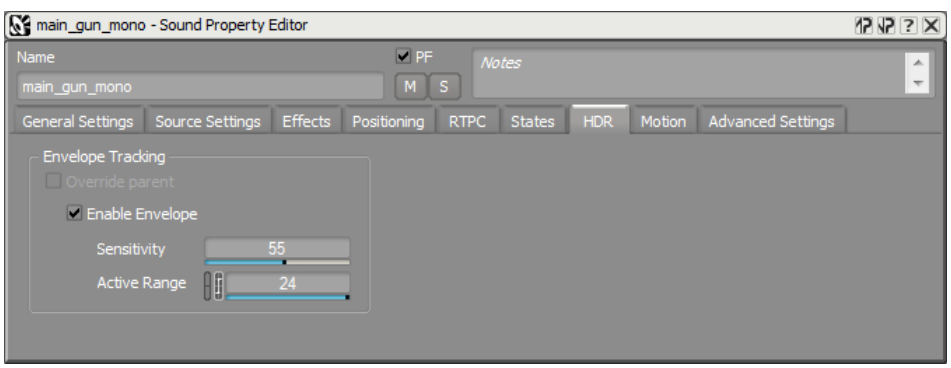


Figure 1.10. HDR tab in Property Editor

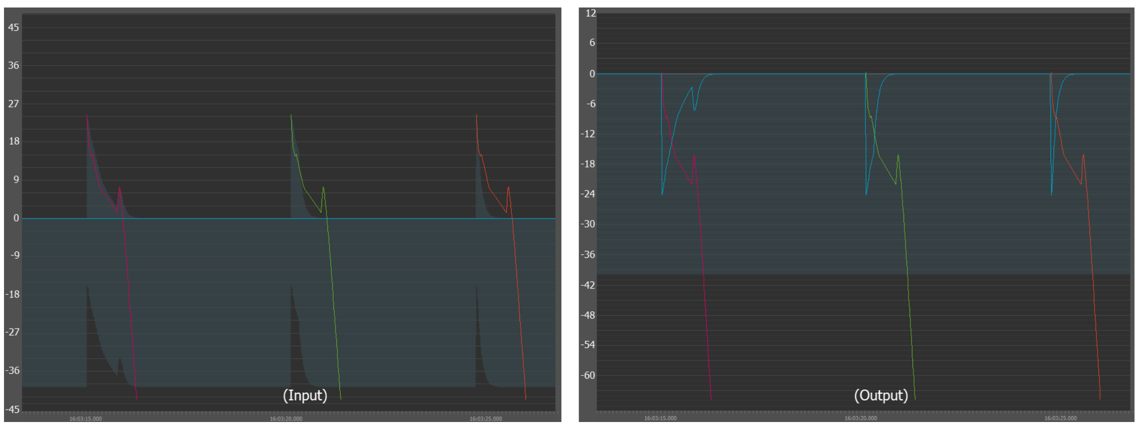


Region of Interest: Active Range

It has been shown that in typical cases, sounds that have large level differences such as an impact sound that ranges from 0 dB to -96 dB should not drive the HDR window for their whole duration, but should instead do so only during a limited period of time. The rationale for this comes from the following paradox: HDR acts as a priority system based on loudness, but the softer parts of a loud sound should not be given as much importance. For example, the decaying tail of a grenade sound has very low importance compared to the transient, and thus should not cover up the impact of a shotgun, even though the former is technically louder, on an absolute scale. In this case, the region of interest of the grenade sound should be limited to its first impact. It moves the HDR window up and therefore ducks everything else. On the other hand, once in the tail and playing say, 12 dB below the peak value, it is usually not desired to continue ducking the other sounds, even if the grenade's volume minus 12 dB is still above the level of all the other sounds in the audio scene. Restricting the HDR window control to a region of interest can be regarded as a way of gracefully blending together various "tonal regions" to continue with the HDR imaging analogy.

In Wwise, the way to define the region of interest is to declare a range in decibels from the peak of the sound, the active range. When the sound's envelope drops below the active range, the sound is considered outside of its region of interest, and is not considered by the HDR system as being able to drive the HDR window. Figure 1.11, “Active range” illustrates this with three similar sounds having different active ranges.

Figure 1.11. Active range



The same decaying sound is played three times with active range set to 96 dB, 12 dB and 6 dB respectively, above a steady background sound. When the sound drops by that amount of decibels below its peak, the window stops following it and instead releases back to idle. The movement of the window distinctly affects the level of the background sound. On the other hand, the window or active range has no impact on the decaying sound itself.

More about HDR

It is important to realize that the envelope of a sound only has an effect on other sounds, and never on itself. This last characteristic is what mostly distinguishes the HDR system from an audio compressor. At any given time, the loudest sound is scaled by its peak value, but is unaffected by its own envelope, which contributes to making the HDR system sound transparent. In consequence, level differences between input sounds are not strictly preserved when envelopes are used. This has the benefit of blending sounds of varying "local loudness" together in the audio scene, as was discussed in the previous section. But you may find that the audio scene loses a bit of its realism when the HDR window moves wildly. In HDR imaging, the same situation arises when using a large amount of smaller tonal regions. Lots of examples of unrealistic HDR photos can be found on the web. We thus recommend that you proceed carefully in setting levels above the HDR threshold: do not just pluck in values blindly, but instead carefully set the level and design the envelope of your loudest sounds in order to make space for them in the mix.

Avoid enabling envelopes for nothing, as they require extra memory. Envelopes are useless for soft sounds that never go above the HDR threshold, so make sure to disable them.

Beware of envelopes that wiggle above and below the active range, as this might cause unwanted behavior. In the following figure, a long explosion sound goes in and out of the active range, provoking rapid changes of the HDR window. In this case, you might want to coarsen the approximation of the envelope and/or edit it manually. You may even want to proceed artistically and design an envelope that does not exactly correspond to the reality, by ducking the audio scene only during the initial impact of the sound, and recovering sooner during the rumbling part.

Simpler envelopes with fewer points result in less erratic amplitude changes of the audio scene, and this is often preferable.

Figure 1.12. Editing envelopes (a)

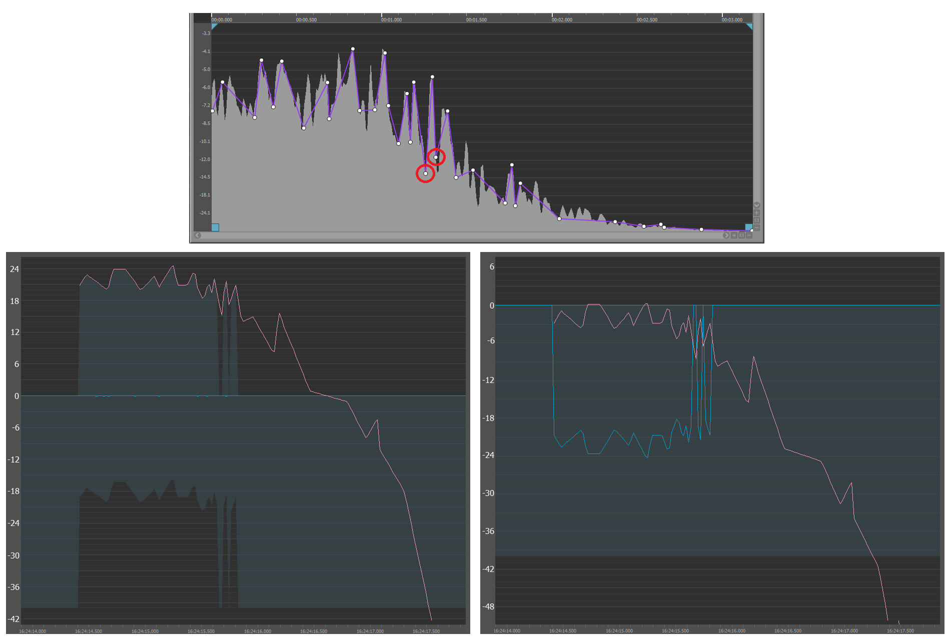


Figure 1.13. Editing envelopes (b)

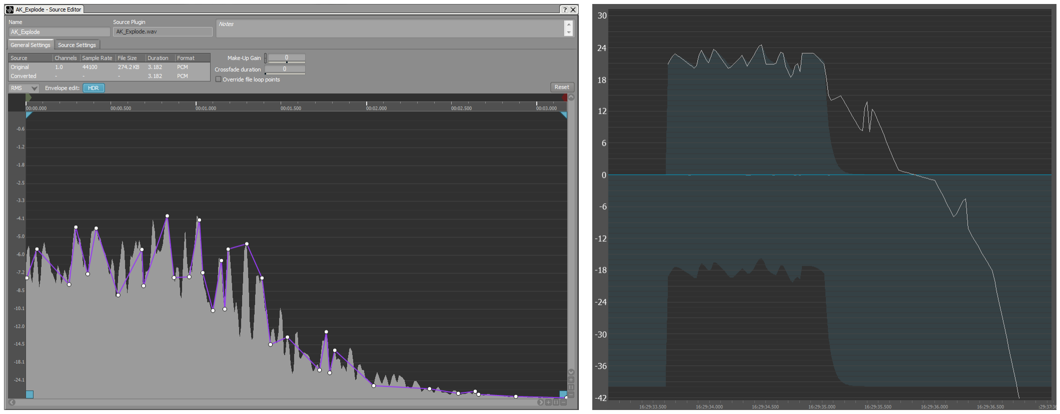
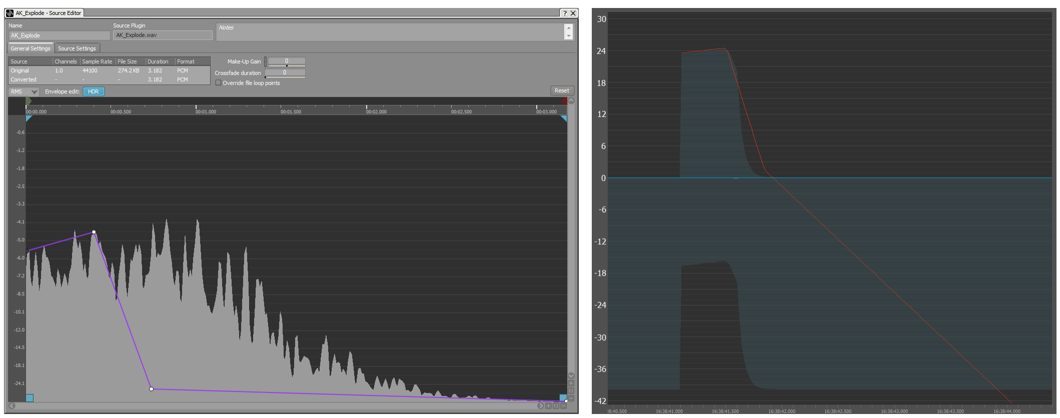


Figure 1.14. Editing envelopes (c)



In (a), the envelope of an explosion sound is displayed in the source editor. The points circled in red are below the active range. During playback, the envelope goes in and out of the region of interest, resulting in rapid movements of the HDR window. The drastic effect on the volume of the background sound is very explicitly illustrated in the resulting output. To fix this problem, the envelope was edited manually in (b) by removing the two offending points. Recall that editing the explosion envelope only affects the HDR window, that is, other sounds, but not the explosion itself. For aesthetic reasons, the designer may also decide to duck other sounds only during the initial impact of the explosion, but let them take their full volume earlier, during the following rumble by editing the envelope somewhat like in (c). Interestingly, in all three examples the explosion is played back identically.

HDR and the Wwise Voice Pipeline

As previously mentioned, the HDR system works with logical volumes set in the Wwise authoring tool, and ignores the actual amplitude of audio data. The levels considered by the HDR system's logic correspond to those at the input of the HDR bus, as can be seen in the Voice Monitor when inspecting the HDR bus in Input bus mode. These levels depend on the voice volumes, which are the sum of contributions from the actor-mixer hierarchy, control buses of the master-mixer hierarchy, actions, RTPC and distance attenuation, as well as on gains of each mixing bus located upstream of the HDR bus in the signal flow. When multiple signal paths lead to the HDR bus, for example when using auxiliary sends, the path with maximum gain is used by the HDR logic.

At each audio frame, the sound engine begins by computing the volume of all voices as it is seen by the HDR bus. It then executes the HDR system's logic, which returns the global HDR gain/attenuation according to the position of the HDR window, and reapplies this gain into each voice. Once this is done, it processes voices as usual: voices are evaluated against the volume threshold to decide if they should be virtual, data is produced, and volumes are applied.

Make-Up Gain and Source Normalization

The HDR system's logic ignores two volume properties supplied by Wwise: source normalization and make-up gain(s). These volume controls are primarily used to normalize the audio assets independently of their logical volume, as you could have done in your wave editor prior to importing them into Wwise. For example, you can use the localization make-up gain to even out volume differences between two languages, but the HDR system should have the same behavior across all languages. Thus, it is important to have volume controls that are ignored by the HDR logic, and source normalization and make-up gains play this role.

Note that the virtual voices system, as well as the Voice Monitor calculations, does not take into account the make-up gain and the source normalization. However, virtual voices do take the HDR attenuation into account. Thus, voices can become virtual when they are below the HDR window.

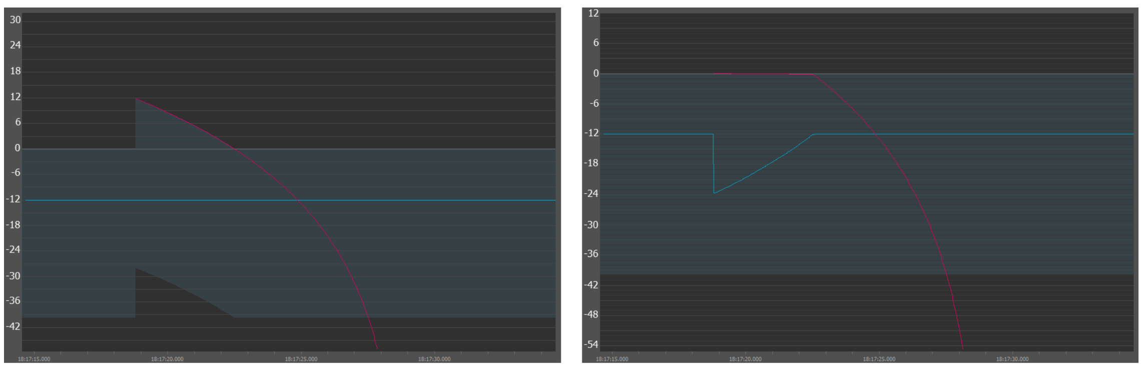
The make-up gain slider in the Source Settings tab of the Property Editor can also be used with HDR for aesthetic purposes. Since it is transparent from the HDR system's point of view, it can be seen as a post-HDR volume (in terms of signal flow), allowing you to change the volume of sounds independently of the HDR window. Normally, any sound playing alone above the HDR threshold comes out of the system at the same level. With the make-up gain, you can effectively make it louder. For example, you might want the player's guns of your first-person shooter to be much louder than the rest, while benefiting from the HDR system's dynamic mixing capabilities. In this case, give them a boost using the make-up gain.

HDR and Distance Attenuation

Distance attenuation curves can be designed as usual. However, be aware that there is a somewhat unexpected behavior caused by the HDR system.

Recall that when you play a sound that is above the HDR threshold and is the loudest sound of the audio scene, it comes out of the system at 0 dBFS. If that sound is played back 50 meters away, but remains above threshold and remains the loudest one in your scene, it will still come out of the system at 0 dBFS. Consequently, you will be left with the impression that the attenuation curve does not work at closer distances. It is not the case. What happens is that the increasing input volume at closer distances is actually used to duck other sounds instead of increasing the output volume. If nothing else is playing, you will not notice this, but you need to keep it in mind when designing attenuation curves for louder sounds.

Figure 1.15. Effect of distance on loud sounds



Here, the listener is steadily stepping away from a loud sound (in red). While the volume seems correct on the input side (left), the output (right) exhibits a plateau when the loud sound remains above the HDR threshold in spite of the increasing distance, giving the false impression that distance attenuation is not working. In fact, at closer distances, the volume offset exceeding the threshold affects other sounds (in blue) instead of the sound in question.

Using the HDR Window as an Input Variable

The HDR window position can be monitored by looking at the window meter in the HDR tab of a bus. You can also assign a game parameter to its value. From there, you can map the game parameter to any property of any object via RTPC. For example, you could use the window position to tweak the maximum number of sound instances.

To attach a game parameter on the HDR window position:

1. Inspect the HDR bus to see its content in the **Property Editor**.
2. Click the **HDR** tab.
3. In section **Window Top Output Game Parameter**, click [**...**] to browse for a game parameter.

4. Set the range for which the game parameter will be set, in decibel on the HDR bus input.