Amendments

In this document, requested changes are shown *in italics like this*. Responses follow immediately, and use a regular font.

In the document comparing the original and amended theses (`diff.pdf`), unchanged text is black, while deletions are shown in red with a strike-through, and additions are shown in blue with a wavy underline. This comparative document was produced using an automated tool (`latexdiff`). This program shows the changes accurately, but occasionally splits paragraphs in unexpected ways (for example between pp. 87-90). It is recommended to view this document in colour, and to read it alongside the final version of the thesis, which has more legible formatting.

Section numbers have been added in the amended thesis, and can be used to locate the following changes.

Internal Examiner

Add title "Abstract" to the abstract.

Done

Page 7 line 13 "is also an problem" should be "is also a problem"

Introduction 2nd bullet point: Done

Figure 2.1 is not referred to in the text

3.1.1 2nd paragraph: Added reference, now numbered figure 3.1 due to extra Theory chapter.

Page 19 last sentence of first paragraph – sentence needs qualifying with "when…"

3.1.3 2nd paragraph: Gives a concrete example of the effects of exponential complexity to qualify the preceding claim.

Page 20 "number to checks" should be "number of checks"

3.2 2nd paragraph: Done

Page 24 second paragraph line 5 – shouldn't "statistical" be "stochastic"

4.1.4 1st paragraph: Done

Page 25 last paragraph – "25s" should be "25μs". Also KHz should be kHz.

4.1.4 3rd paragraph: Done

Page 27 fifth paragraph – "which a common descriptor" should be "which is a common descriptor"

4.2.1 last paragraph: Done

Page 27 can you give a reference for equation 9 (or are you implying this equation is obvious from your preceding description?)

4.2.1 penultimate paragraph: The equation should follow from the description, yes. Added the sentence beginning "It follows…" to clarify. Also added intermediate form to eq. 4.4.

Page 28 KHz should be kHz

Fixed throughout

Page 30 "be this same" should be "be the same"

4.2.3.3 2nd paragraph: Done

Page 34 following equation 23 please clarify what c, T and X are.

5.1.1 following eq. 5.7: Done

Figure 4.2 "Black lines" should be "Blue lines". Is there any way you can redraw the tetrahedral one so it's easier to see (maybe put it in twice for the two orientations? Or use different colours for the two orientations?)

5.2.1: Maybe the print quality was poor, but the lines are black. Added diagram 5.3 showing tetrahedral structure in more detail. Added explanatory text in the second paragraph of 5.2.1.

Page 40 "like and obvious" should be "like an obvious". Also KHz should be kHz.

5.2.2.2 4th paragraph: Done

Page 40 bottom paragraph says following figure but it is referring to the figure after the following one.

5.2.2.2 6th paragraph: Text refers directly to figure 5.5

Page 41 KHz should be kHz.

Fixed throughout

Page 42 Figure 4.4 on each graph a line goes from (0,0) to the top left corner and bottom right corner of the excitation signal – this appears to be a presentational error and needs correcting. Also it is unclear to me why the amplitudes on Dirac, Ricker and PCS are "filled in" – is this fluctuation in amplitude that is lost in the resolution or a presentational problem? If the former just clarify if the latter please fix.

5.2.2.2 fig. 5.5: Figure adjusted to remove the diagonal lines. The graphs are filled-in due to high-frequency fluctuations, which is clarified in the figure annotation.

Figure 4.6 not referred to in text.

5.3.3 final paragraph: This paragraph added to explain the need for figure 5.7, which is referenced in the text.

Page 49 last line of first paragraph – "exceptional" is it possible to be a little more precise or use more restrained language…

6.1 1st paragraph: changed "without requiring exceptional computing power" to "but for a much lower computational cost than would be possible with any individual method".

Page 49 3rd paragraph suggest delete word "maybe"

6.2.1 1st paragraph: Done

Page 50 "no concrete rule to place" should be something like "no concrete rule about where to place"

6.2.2 1st paragraph: Now reads "no concrete rule governing where to place"

Page 50 Last sentence – is it possible in any way to indicatively quantify "very large", "reasonable" and "excessive" or reword to be more precise?

6.2.2 2nd paragraph: Clarified derivation of overall simulation complexity.

Page 51 first paragraph. Suggest adding a sentence to complete the paragraph saying what Wayverb does (you've said what it doesn't do!)

6.2.2 3rd paragraph: Added explanation of Wayverb's behaviour, along with justification.

Page 53 "the to end" should be "the top end"

6.3.2 2nd paragraph: Done

Page 55 KHz should be kHz

Fixed throughout

Page 57 fifth paragraph – needs a little clarification – AB pair add "spaced" and I presume you are referring to "omnis". Similarly for "XY" add "co-incident" and I presume cardiods or figure-of-8.

7.2 2nd paragraph: Done

Page 58 first line – "four coincident directional" should be 3 plus an omnidirectional

7.2 4th paragraph: Done

Page 58 "front the microphone" should be "front of the microphone"

7.3.1 2nd paragraph: Done

Page 59 bottom sentence. Looks like you are taking advantage of left right symmetry in arriving at 288 – if so please claify this.

7.3.2 last paragraph: Not sure exactly how to interpret this. Wayverb assumes that both ears share the same location, at least for the purposes of histogram creation, and therefore the same raw data is used to compute the left and right ear responses. Added an explanation of the number of horizontal and vertical divisions.

Page 65 KHz should be kHz

Fixed throughout

Page 67 KHz should be kHz

Fixed throughout

Page 75 KHz should be kHz

Fixed throughout

Equations can't be told from figures – don't like style of referencing figures

Fixed throughout

Page 78, 84 and 87 have unlabelled tables. Please label and then refer to in text.

All are tables now labelled and referenced in the text.

Page 85 Figure 8.2 title at top should say "T30" not "RT30"

9.2.1: Done

Page 87 fifth paragraph first sentence change "several" to "three"

9.2.2 1st paragraph: Done

Figure 8.4 and 8.5 are not referred to in the text – please correct

9.2.2 last paragraph: Fig. 9.5 referenced

9.2.3 last paragraph: Fig. 9.6 referenced in final paragraph of "Direct Response Time"

Page 100 "acoustics arbitrary" should be "acoustics of arbitrary"

10.2 1st paragraph: Done

Conclusions should be a section in its own right

Moved the Analysis section of the Evaluation chapter onwards to a new Conclusion chapter.

References – despite the explanatory sentence at the start I feel that references 78 onwards - which are not cited in the text - should be deleted or placed in a Bibliography (or similarly named) section

Removed unused references.

I'm not convinced about the style used for referring to figures. I've not seen it done this way before. Unless you can convince me with an example that this is a valid approach I'd like you to change "figure (x.x)" to "Figure 2.3". The same applies to referring to equations. In particular page 74 second last paragraph needs to make it clear it is referring to equation 41 and equation 44

Changed to a more normal "fig. x.y" scheme for figures, "eq. x.y" for equations, and "table x.y" for tables.

External Examiner

Bring together background theory into a single chapter that precedes the more focused technical work and discussion. This should for instance consider basic room acoustic theory and metrics for success used – at the moment these aspects are scattered throughout the other chapters and in some cases repeated. This proposed chapter is essentially a restructure of work already presented rather than asking for new literature work to be completed. However, it should also include a proper consideration of this research as it is applied in more creative contexts. Although the nature of the thesis and the field is highly multidisciplinary, given its arts/humanities, science/engineering backgrounds, an MA-Research thesis should have its foundation and justification in the creative arts and relevant humanities focused scholarship.

Added Chapter 2, which covers equations governing wave behaviour, boundary models, and reverb time estimates and measurements. The majority of the information on boundary modelling has been moved from the beginning of the Boundary Modelling chapter. The "creative context" is quite a different topic to the background theory, so this was added to the existing Context chapter, rather than to the new Theory chapter. The creative context can be found in section 1.4, under the heading "Acoustic Simulation and the Creative Arts".

Each chapter needs a proper summary. At the moment the chapters move on from one to the other tackling various technical issues, but with no reflection on what has been achieved in each case, or considering the implication of any interim results (if given).

All chapters now have a reflective summary which attempts to fulfil this requirement.

A proper Conclusion chapter is needed, rather than just a section in another chapter. This should be separated off from the current existing final chapter. It should summarise the work completed, drawing on the evidence presented in the results obtained, provide a proper discussion, reflection on the research aims stated at the beginning of the thesis, and considerations for future work. Again, this is more of a

restructure than writing new content, but I would encourage the candidate to more completely consider the research questions posed at the beginning of the thesis – that is, the nature of accuracy, efficiency, and accessibility. These are actually very difficult questions to ask, and so to answer, and in fact have not been properly addressed in this work (e.g. accuracy might consider listening tests as well as objective measures; efficiency should include algorithmic performance metrics; accessibility might include user tests) I am not suggesting that these research questions should be changed or more work done towards them at this stage, but the candidate should consider them more carefully, and what – if appropriate – might be more appropriate questions for this work, and again in the context of an MA thesis.

Everything from the "Analysis" section of the Evaluation chapter forwards has been moved to its own "Conclusion" chapter. This chapter begins with a new discussion of the testing procedure, and its shortcomings. It was not feasible to conduct further work towards the research goals, and instead ideas for tests which would have allowed more conclusive results to be drawn have been suggested in section 10.1.1. The "Simulation Method" section has been renamed to "Evaluation of Project Goals" (10.1.2) and updated to individually consider each of the research goals. The section on plausibility (10.1.2.1) has been adjusted to discuss the interplay between different components of the simulation. The paragraphs about efficiency (10.1.2.2) have been rewritten, and no longer discuss use in architectural contexts which are not relevant to the project aims. A completely new section on accessibility (10.1.2.3) has been added, which covers the design goals of the interface, some background on the implementation, and suggests future improvements. Finally, the summary (10.2) has been slightly extended to explain that the project scope was too broad, and that Wayverb's main contribution is a starting point for future projects, rather than a tool which is useful in its own right.

---

p8: KHz is used throughout instead of kHz – please correct.

Fixed throughout

---

p12 - Project Aims: although I am not suggesting that these should be rewritten, their scope might be more focused to enable a more complete analysis and reflection in the conclusions. Also it states that this work should not required any specialist training in 'acoustics or programming' yet the thesis that follows is absolutely founded on both, and this assumption has not actually been properly tested as stated above.

1.5: Slightly reworded these. In particular, "accuracy" was changed to "plausibility" as this seems more useful in creative contexts. Rewrote the discussion of plausibility, to focus it more on the aspects of the program which are actually tested. Clarified the need for a sliding scale between quality and speed. Added sentence to explain why accessibility is less important than the first two goals. The "specialist training" comment was supposed to refer to users of the software, so this has been made clearer.

---

p18: Audibility checking example above paragraph starting 'The validation….' should map to Figure 2.2 for clarity.

3.1.2 last paragraph: added reference to fig. 3.2

---

p19: what does the author mean by 'reasonable time'? Refer back to project aims/research questions. In general, such loosely defined terms are used at quite a few points in the thesis and mean very little.

3.1.3 1st paragraph: Added concrete example of exponential simulation time.

---

p23: How has this replacement of the Dirac function with the truncated IR of a low-pass filter been tested? In general, it would be expected to see some interim results here of the IS algorithm as presented.

Summary also needed as suggested above.

The need for further tests is now recognised in the conclusion, but it was not feasible to produce new results. Summary has been added (section 3.3).

---

p25: '25s' – units incorrect.

4.1.4 3rd paragraph: now 25µs

---

p26: Vorlander = Vorländer (please correct throughout)

Fixed throughout.

---

p27: define the scattering coefficient 's' (some of this detail comes later in the report, hence the need, in part, for a literature/theory chapter earlier).

This is now discussed in section 2.2.2. The Theory chapter is referenced in the paragraph following eq. 4.3.

---

p28: Reflections are considered here – but no consideration is given regarding octave band considerations. Should they be? Please see comments about a 'new' literature review/background chapter.

4.2.2: Clarified that there is a histogram per frequency band.

---

p36: Please define 'accuracy' in the context of your discussion regarding mesh topologies. What constitutes an 'accurate' result?

5.2.1 3rd paragraph: Defined the accuracy criterion for mesh topologies.

---

p40: [jeong_source_2012] please fix reference

5.2.2.2 4th paragraph: fixed

…like and obvious = like an

5.2.2.2 4th paragraph: fixed

Figure 4.3 – y-axis magnitudes of frequency domain plots should have dB scale?

Linear scales are required in order to display the similarity between the shape of each pulse in the time and frequency domains. Also, pulse spectra are given on a linear scale in fig. 1 of the following paper. Therefore, I believe this approach is valid.

]J. Sheaffer, M. van Walstijn, and B. Fazenda, 'Physical and numerical constraints in source modeling for finite difference simulation of room acousticsa)', The Journal of the Acoustical Society of America, vol. 135, no. 1, pp. 251–261, 2014.

---

p43: Regarding stability, this paper should be considered and cited:

J. Botts and L. Savioja, "Spectral and Pseudospectral Properties of Finite Difference Models Used in Audio and Room Acoustics," IEEE/ACM Trans. Audio Speech Lang. Process., vol. 22, no. 9, pp. 1403–1412, Sep. 2014.

5.2.2.2 fourth-from-last paragraph: Added this paragraph which explains that numerical error is a plausible cause of simulation growth (based on the results in that paper), but that increasing numerical precision is not viable.

---

p44: "Now the inner…." Please rephrase the start of this section for readability.

5.3.2: Changed first sentence.

Figure 4.6: This should be discussed or referred to in the main body of the text.

5.3.3 last paragraph: Refers to what is now fig. 5.7

---

p51: Discussion is given as to hybrid crossover frequency but never formally cited or justified. Please do so.

6.2.2 last paragraph: added references supporting the approach in Wayverb.

---

p53: 'and at the to end' - ?

6.3.2 2nd paragraph: fixed

---

p56: T20 and T30 used before definition in later chapter. See comments on a proper literature review.

Discussion of reverb time given in Chapter 2.

---

p61: This achieved by taking = This is achieved

7.5 last paragraph: fixed

---

p62: Hacıhabiboglu = Hacıhabiboğlu

Fixed throughout

---

p63: Does mesh oversampling always need to be used? Cannot interpolation be applied in some cases? Please Comment.

7.6.1.1 4th paragraph: Added reference to paper dealing with interpolation for receivers in 2D meshes, and explanation of why this technique is not viable in 3D.

---

p67: 'At low frequencies' – define what you mean by 'low'. Nothing is presented for below 500Hz, which is where your DWM approach will be applied? Hence these results are invalid in the context of your work? What do the colours in Figs 6.1 – 6.3 refer to? This is not explained. If the black/darker colour is a reference, from where has it been obtained?

7.6.3 4th paragraph: “low frequencies” changed to “the lower frequencies shown”. Results re-computed and graphed starting at 100Hz. Graph colours explained in the 3rd paragraph of this section.

Chapter 7: Some repetition here of Image Source section. See comments made earlier.

All background information moved to Chapter 2.

---

p68: Define ‘$’

This was a typesetting error, now fixed.

---

p74: Define what you mean by ‘realism’ in the context of this section.

8.3.2: reworded to “physically-based behaviour” and added explanation that the LRS technique considers edges as well as flat surfaces.

---

p78: Define what you mean by ‘reasonably accurate’ in the context of this section.

8.4.3 1st paragraph: reworded to remove ambiguous claim.

---

p79: Fig 7.6. I am not convinced that the higher frequency deviations from the norm are down to dispersion error. This needs to be considered more carefully, and I would inspect your raw IRs to see what is contributing to this. Could be edge reflection effects from your surface, hard source reflections, incorrect windowing? Does not look like dispersion error. Where are the predicted results sourced from? The plots need to be improved – the scales hide some of the detail. Would a difference/error plot be a better metric?

Now fig. 8.4: I checked the IRs, and they had very low amplitude. I re-ran the tests using a higher-amplitude initial signal, and the generated graphs became much smoother, leading me to believe that the magnitude of the signal was too small relative to the floating-point epsilon in the original test, leading to a loss of precision which caused noise in the graphed results. Figure 8.4 now shows the results obtained using a higher-amplitude input. The scales have not been adjusted, as they were chosen to show the entire range of outputs across all tests, and having different scales on each graph may be misleading. Error plots have not been supplied, as these were not given in the paper by Kowalczyk upon which this boundary test is based.

I didn’t previously discuss the source injection method for this test. This has been added to the 2nd paragraph of 8.4.1, along with a rationale.

The first paragraph of the evaluation (8.4.3) has been rewritten to take more possible causes of error into account.

---

p80: ‘Little low frequency error’ – again, please define what you mean here. More investigation should be carried out into the source of the error or limitations of this approach clearly discussed and the consequence for subsequent results.

This paragraph was removed, as it did not add to the discussion/was already covered in the re-written paragraph. Test limitations are covered in that paragraph. Implications for subsequent results are covered in the new chapter summary (8.5).

---

p81: This is the first time that the nature of the 3D geometry files used to define your rooms has been discussed at all. There is further discussion later, in relation to the limitations of the software, but stating up front that you will use a file importer from existing 3D CAD would be much better.

Added a sentence to the explanation of Assimp in 1.6.2 explaining that users can load CAD files.

---

p82: Define RT60 in relation to T20 and T30. Your assumption that RT60 = 2xT30 = 2xT20 is incorrect. Please refer to ISO3382 and related versions of this standard for a proper definition of how T20 and T30 are determined from a RIR.

This is done in Chapter 2.

---

p84: What are the limitations of the Sabine equation? You should discuss this and it may help with the interpretation of your results later? Is the cuboid room you test a valid example for a Sabine estimation? In this test you should also consider and average over more source/receiver positions to arrive at a better estimation of RT60 for the room, which is what a Sabine estimation gives (rather that something specific to a given source/receiver pair).

How have you arrived at the results in the table? Please discuss. %-error might be a better metric to include, and also reference to JND for RT60. Octave band results should also be included, rather than overall RT60, as this will hide some possibly important detail.

If you are not able to re-do results, discuss the limitations of the test instead.

This is where your results start to reveal that what has seemed 'easy and obvious' based on the choices made in earlier chapters, do not give the best results, giving rise to more questions – the number of choices and decisions you have made are all highly complex and interact with each other when producing the final result. Why should your room resonate at 400Hz? What physical reason is there for this? It could be down to the crossover of the filters used in the hybrid model and how you have implemented them. Again, some better discussion needed.

Limitations of the Sabine equation are discussed in section 2.3.1, and in the third-from-last paragraph of section 9.2.1. The penultimate paragraph of section 9.2.1 was added, discussing the averaging of several source/receiver positions. The final two paragraphs of the section discuss the limitations of the test in more detail.

9.2.1 4th paragraph: Added that reverb times have been computed in accordance with the ISO 3382 specification. Percentage errors have been added to the table. Reference to JND added in the 5th paragraph. Octave-band results are given in fig. 9.2, and table 9.2 has been added showing the percentage differences in reverb times between the bands with the longest and shortest reverb times for each room size.

---

p85: 'In effect the waveguide exhibits a high noise-floor under some conditions.' Where have you shown this? How or why is this? More explanation needed.

The "noise floor" was supposed to refer to the continued oscillation at 400Hz, 40dB below the level of the initial contribution, which was clearly shown in the spectrograms in figure 9.3. However, "noise floor" was an incorrect description and this sentence did not add to the discussion, so it has been removed.

Fig 8.2 – What are Octave band RT30 measurements? Be clear about what you are plotting. If you are plotting octave band results we need to see the actual data points you have, not just a connected line that joins them up. Please also mark what aspect of you model has contributed to each plot point

(e.g. DWM, ISM, etc) 'were placed' = were replaced? You should also check octave band results vs broadband results if possible.

Fig 9.2: corrected RT30 to T30. It should be clear from the 4th paragraph of 9.2.1 that the T30 measurements are derived in accordance with the ISO specification. The data points have been marked more clearly, and the lines replaced with dotted lines. The model crossover frequency has been marked. The broadband results are given in table 9.1.

9.2.1 10th paragraph: placed -> replaced

---

p87: 'show that wall impedances are accurately modelled' = define accurately, and are they actually 'accurate'?

9.2.1 11th paragraph: replaced "impedances are accurately modelled" with "reflectances generally match predicted values to within 1dB for three material types and angles of incidence, at least in the band below 0.15 of the mesh sampling rate".

General point - A better approach to testing a study of this nature is presented in reference [26].

10.1.1 1st paragraph: the shortcomings of the testing procedure have been discussed, and the method in [26] has been put forward as a better alternative.

Your comment on the 'issue with shorter low-frequency decay times' needs more detail – we have no data to go on in relation to what these decay times are. Give the data (octave bands) and discuss, or more speculative comments should be removed.

Fig 9.4: octave-band T30 measurements have been given. Table 9.4, showing the difference between the minimum and maximum reverb time across all frequency bands, has been added. In section 9.2.2 it has been made clear that low-frequency bands have lower reverb times than high-frequency bands. The final paragraph of 9.2.2 has been added to put forward evidence that the low-frequency resonance at 400Hz seen in some tests is due to a deficiency specifically in the boundary model.

---

p89: 'Therefore it is obvious' – I'm not sure the results presented here mean that the conclusions you draw are in fact obvious. Energy metrics such as C50/C80 or event EDT could have been used instead.

Table 9.5: this table has been added, showing T20, T30, and EDT measurements, along with percentage differences. The 2nd paragraph of 9.2.3 has been added, explaining the purpose of the EDT measurement, and how the results can be interpreted.

---

p92: Please make clear where diffraction is indicated in Fig. 8.7 'The low frequencies have a fast onset… extended decay' I am not sure what you mean by this, or that the conclusions you draw follow on as a consequence. It is more likely that there are limitations with your implementation that have not as yet been explored and tested fully.

9.2.4 last paragraph: Replaced the "fast onset…" sentence with a few sentences taking concrete readings from fig. 9.8.

---

p94, Fig 8.9 – I would suggest adding the plot for the cardioid microphone pointing to the source here in another colour. Would make the case very clearly.

Fig: 9.10 has been adjusted to include responses from two cardioid capsules, one facing towards and one away from the receiver. Section 9.2.6 has been adjusted accordingly.

---

p95: 'A concert hall is simulated…' – where and what are the details of this model? Please give the details of this test. I would improve Fig 8.10, to make the differences clearer, used a scale that enables ITD values to be highlighted.

9.2.7 1st paragraph: More details of the model are given here, and the schematics are shown in fig. 9.11.

Fig 9.12 has been adjusted to cover a shorter time period, which should make the ITD more obvious.

Analysis – move this and what follows to a proper Conclusion Chapter as discussed above.

Done

'precise stereo effects' – strictly, you have not demonstrated this. Listening tests would be needed. Points about no diffraction in the ray tracer are fair, but the mis-match between modelling approaches and reverb times is probably down to errors or assumptions that could be fixed or improved?

10.1.2.1 1st paragraph: the word "precise" has been removed. It has been noted in 10.1.1 that proper listening tests are required, although it's evident that some kind of stereo effect is possible from the results in fig 9.12.

---

p98: The candidate starts off in the early chapters by making a case for an efficient model, and that the time is right for this – but by this point in the report, goes back to established arguments that two more years computing advancements will help to solve the computational limits of the current implementation!

10.1.2.1 last paragraph: This argument has been supplemented by an explanation that Wayverb is well-placed to benefit from current advances in processor technology, which tend to increase the number of processor cores rather than the overall clock speed.

'than may be expected' – but what is expected of the user interface and how has this been tested?

10.1.3.2 1st paragraph: Made it clear that the comparison is to the features of other software, rather than to user expectations.

10.1.2.2 5th paragraph: explained that usability tests have not been conducted and why they are necessary.

---

p101: 'the finished project is sufficiently accessible' – as determined by whom?

This sentence has been removed, and a more complete discussion of accessibility can be found in section 10.1.2.2.

---

p102: Valimaki = Välimäki

Fixed

---

p107: complete reference for [72] needed – where was this work presented?

Now reference [80]: added “in Proc. EuroNoise 2015”

Misc

Added section numbers to make the document easier to navigate.

3.1.2 Added the last three paragraphs, giving a basic overview of some of the maths involved in the image-source model.