UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Non-contact Ultrasonic Guided Wave Inspection of Rails : Next Generation Approach

Permalink

<https://escholarship.org/uc/item/34h106nj>

Author

Mariani, Stefano

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Non-contact Ultrasonic Guided Wave Inspection of Rails:
Next Generation Approach**

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Structural Engineering

by

Stefano Mariani

Committee in Charge:

Professor Francesco Lanza di Scalea, Chair
Professor David J. Benson
Professor Michael J. Buckingham
Professor William S. Hodgkiss
Professor Chia-Ming Uang

2015

Copyright,
Stefano Mariani, 2015
All rights reserved.

The dissertation of Stefano Mariani is approved, and it is acceptable in
quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2015

DEDICATION

To Bob

TABLE OF CONTENTS

Signature Page	iii
Dedication	iv
Table of Contents	v
List of Abbreviations	ix
List of Figures	xi
List of Tables	xxii
Acknowledgments	xxiv
Vita	xvii
Abstract of the Dissertation	xxix
1 Introduction	1
1.1 Non Destructive Evaluation of railroad tracks. Introductory discussions and motivations for the research	1
1.2 Outline of the numerical and experimental approach of the dissertation	3
2 State of the art review	5
2.1 Development of structural defects in rails	5
2.1.1 Transverse defects in the rail head	8
2.1.2 Longitudinal defects in the rail head	16
2.1.3 Web defects	19
2.1.4 Base defects	22
2.1.5 Rolling contact fatigue conditions	23
2.1.6 Defective welds	31
2.2 Current rail inspection methods	32
2.2.1 Historical background	32

2.2.2	Currently used methods	34
2.3	Ultrasonic guided waves for defect detection in rails	41
3	Concepts of the proposed rail defect detection system	47
3.1	Introduction	47
3.2	Former UCSD rail inspection system based on a hybrid laser/air-coupled configuration	48
3.3	Newly developed rail inspection system solely based on air-coupled transducers	50
3.3.1	Current transducers configuration	50
3.3.2	Alternative transducers configurations tested over the course of the research project	59
3.4	Conclusions	61
3.5	Acknowledgments	62
4	Numerical modeling of ultrasonic guided wave propagation: theoretical fundamentals and application to the rail case.....	63
4.1	Introduction and theoretical fundamentals	63
4.1.1	Ultrasonic waves propagation in an unbounded medium	66
4.1.2	Guided waves propagation in a bounded medium	68
4.2	Local Interaction Simulation Approach (LISA).....	76
4.2.1	Introduction	76
4.2.2	Proof of concept of the proposed rail defect detection system.....	78
4.2.3	Distribution of wave energy on the rail top surface	86
4.2.4	Parametric study of the prototype sensitivity to different defect conditions.	99
4.2.4.1	Case I: signals not affected by noise (ideal condition).....	109
4.2.4.2	Case II: noise artificially added to the signals (realistic condition).....	121

4.2.4.3 Case II: probabilistic analysis through computation of Receiver Operating Characteristics curves (for “high” SNR levels).....	139
4.2.4.4 Case II: probabilistic analysis through computation of Receiver Operating Characteristics curves (for “low” SNR levels).....	148
4.2.4.5 Case II: comparison with results obtainable using coherent array techniques.....	154
4.2.5 Future works.....	169
4.3 Finite element analysis of the complete experimental setup using LS-DYNA....	170
4.3.1 Introduction	170
4.3.2 Finite element analysis of the pressure field generated in air by a cylindrically focused air-coupled transmitter	172
4.3.3 Simulation of the complete ultrasonic wave propagation through air and rail	186
4.3.4 LS-DYNA aid in the design of a dedicated holder for a newer transmitter	205
4.4 Conclusions	218
4.5 Acknowledgments	219
5 Development of the rail inspection prototype.....	221
5.1 Introduction	221
5.2 UCSD/FRA Rail Defect Testing Facility	223
5.3 Prototype parts design	224
5.4 Electronic hardware layout.....	226
5.5 Signal processing techniques employed to enhance the signal-to-noise ratio of the acquired waveforms	230
5.5.1 Digital filtering techniques (used in former versions of the prototype)	231
5.5.2 Electrical impedance matching networks.....	233
5.6 Statistical algorithm designed to enhance the rail defect detection performance	237

5.6.1	Defect-sensitive features extraction	242
5.7	On the issue affecting higher speed testing performance	250
5.8	Field test at Transportation Technology Center	256
5.8.1	Introduction	256
5.8.2	Operational parameters used during the execution of the tests and test site lay-out	258
5.8.3	Field test timeline	261
5.8.4	Field test results: Damage Index traces	263
5.8.5	Field test results: Receiver Operating Characteristic curves	266
5.8.6	Conclusions	270
5.9	Further studies conducted on the field test data	271
5.10	Conclusions	274
5.11	Acknowledgments	276
6	Conclusions	277
6.1	Review of the dissertation and summary of the results	277
6.2	Recommendations for future work	280
	References	282

LIST OF ABBREVIATIONS

2D-FFT	Two Dimensional Fast Fourier Transform
AREMA	American Railway Engineering and Maintenance-of-Way Association
AUC	Area Under the Curve
CFR	Code of Federal Regulations
CH	Crushed Head
CUDA	Compute Unified Device Architecture
DF	Detail Fracture
DI	Damage Index
EMAT	Electro-Magnetic Acoustic Transducer
FRA	Federal Railroad Administration
GPU	Graphics Processing Unit
HSD	Head Surface Damage
HSH	Horizontal Split Head
LISA	Local Interaction Simulation Approach
NDE	Non Destructive Evaluation
NDT	Non Destructive Testing
NRBC	Non Reflecting Boundary Conditions
NTC	Number of Threshold Crossings
PD	Probability of Detection
PFA	Probability of False Alarm
PML	Perfectly Matched Layer
PSB	Passenger Services Building

RCF	Rolling Contact Fatigue
RDTF	Rail Defect Testing Facility
RMS	Root Mean Square
ROC	Receiver Operating Characteristics
R&D	Research and Development
SAW	Surface Acoustic Wave
SDH	Side Drilled Hole
SHM	Structural Health Monitoring
SNR	Signal to Noise Ratio
TD	Transverse Defect
TTC	Transportation Technology Center
UCSD	University of California, San Diego
VSH	Vertical Split Head
% H.A.	Percentage of the rail Head Area

LIST OF FIGURES

Figure 2.1: Terminology used to identify defects development planes in relation to the rail section	6
Figure 2.2: Train accidents due to rail, joint bar and rail anchoring defects in the 2005–2014 decade from FRA Safety Statistics Data.....	8
Figure 2.3: Identification of a transverse defect based on its cross-sectional area	9
Figure 2.4: Transverse defects. On the right, case of coupling with surface shelling	9
Figure 2.5: Transverse fissures	10
Figure 2.6: Compound fissure	11
Figure 2.7: Detail fracture originating from a visible shell	12
Figure 2.8: Detail fracture from head check	13
Figure 2.9: Reverse detail fracture.....	14
Figure 2.10: Engine burn fracture showing significant growth	15
Figure 2.11: Rail failure fracture face showing no transverse defects.....	16
Figure 2.12: Horizontal split head originating from an internal seam.....	17
Figure 2.13: Side view of a horizontal split head	17
Figure 2.14: Vertical split head.....	18
Figure 2.15: Head and web separation defect associated with rail joint.....	19
Figure 2.16: Web failure resulting from high residual stresses	20
Figure 2.17: Piped rail showing significant rail collapse.....	21
Figure 2.18: Bolt hole crack.....	21
Figure 2.19: Half moon shaped broken base	22
Figure 2.20: Base fracture showing nick with transverse development	23
Figure 2.21: Wheel contact on a curve with “high” rail and “low” rail [24]	24
Figure 2.22: Contact stresses on tight curved track	24
Figure 2.23: Gage side shell showing severe parent metal decay.....	26
Figure 2.24: Centralized flaking condition showing chipping of parent metal	27
Figure 2.25: Thermal cracks on burned rail stock rail	27

Figure 2.26: Gage side head checking and flaking	28
Figure 2.27: Flattened rail head showing displacement of parent metal (spalling)	29
Figure 2.28: Sketch of a crushed head [4]	30
Figure 2.29: General appearance of rail wear	30
Figure 2.30: Defective weld.....	31
Figure 2.31: Scheme of the contact induction technique	35
Figure 2.32: Eddy current sensor magnetic field pattern	36
Figure 2.33: Example of ultrasonic wheel testing system	37
Figure 2.34: Ultrasonic probe wheel transducers arrangement	37
Figure 2.35: Ultrasonic beam coverage of three 70° probes (left) versus a single 70° probe (right) [27].....	38
Figure 2.36: Ultrasonic method handheld [28] (left) and vehicle-mounted version (right)	39
Figure 2.37: TDs detection rates based on flaw cross-sectional sizes. Industry averages vs AREMA recommendations.....	41
Figure 2.38: Comparison between traditional bulk waves and guided waves inspection (FBS Inc.).....	42
Figure 2.39: Dispersion curves for a traction-free aluminum plate [36]	43
Figure 2.40: Wave structures for various points (f_d = frequency*thickness) on the S0 mode of an aluminum plate, showing the in-plane (u, solid line) and out-of-plane (w, dashed line) displacement profiles across the thickness of the plate [36]	44
Figure 2.41: Comb transducer excitation of guided waves in a typical 115-lb AREMA rail [22].....	45
Figure 3.1: Schematic of the proposed system for rail defect detection (side view). The lift-off distance between rail top surface and transmitter is dictated by its focal distance (for the one currently in use is 76 mm). The lift-off distances of the receivers can be as high as 5 to 10 cm before incurring in significant SNRs deteriorations	55
Figure 3.2: Schematic of the proposed system for rail defect detection (top view). The drawing is not to scale. The two red lines indicate the hypothetical locations of a VSH (left) and a TD (right) internal cracks	55
Figure 3.3: Frequency spectrum of noise detected by an accelerometer as a train approached it traveling at ~40 km/h [52].....	58

Figure 3.4: Alternative sensor configuration studied in the past. Attenuations due to defects was again the main effect targeted by this system (only one receiver per position is sketched to improve the clarity of the drawing; a group of 4 receivers was actually deployed in each receiver position)	60
Figure 3.5: Alternative sensor configuration studied in the past. Reflections caused by defects was the main effect targeted by this system (only one receiver is sketched to improve the clarity of the drawing; a group of 4 receivers was actually deployed)	61
Figure 4.1: Displacement field for Rayleigh wave	64
Figure 4.2: Displacement field for Love wave	65
Figure 4.3: Displacement fields for Lamb waves. Symmetric mode (left) and antisymmetric mode (right).....	65
Figure 4.4: P wave (top) and S wave (bottom) propagation	68
Figure 4.5: Guided waves generation	69
Figure 4.6: Guided wave propagation within a thin isotropic plate.....	70
Figure 4.7: Rayleigh wave elliptical particle motion [36]	75
Figure 4.8: Normalized displacements for a Rayleigh surface wave as a function of depth in an aluminum half-space. Curves normalized with respect to u_3 value at $x_3=0$. λ_R is the Rayleigh surface wavelength [36]	75
Figure 4.9: AREMA 136RE railroad track geometrical details.....	79
Figure 4.10: Modeled portion of rail 136RE	79
Figure 4.11: Transmitted signal	81
Figure 4.12: Amplitude modulation along the excited line	81
Figure 4.13: Section cuts of the LISA models of the 136RE rail with 10% H.A. transverse defect (on the left) and 30% H.A. transverse defect (on the right).....	82
Figure 4.14: Vertical displacement field (in color scale) pertaining to the rail model (a) without defect, (b) with 10% H.A. transverse defect and (c) with 30% H.A. transverse defect after 60 μ sec of simulation. Views from above the rail top surface	84
Figure 4.15: Vertical displacements at the top of the railhead for the rail model (a) without defect, (b) with 10% H.A. TD and (c) with 30% H.A. TD.....	85
Figure 4.16: Top view (left) and 3D view (right) of the contour plot showing the pressure field produced by the modeled air-coupled transmitter at its focal distance	87
Figure 4.17: Contour plot of vertical displacement field after 1.5 μ sec (a) and 31 μ sec (b), respectively	88

Figure 4.18: Contour plot of vertical displacement field after 60 μ sec (c) and 90 μ sec (d), respectively	89
Figure 4.19: 3D views that help to visualize where the energy plots shown in the next figures are located in the rail. Due to the symmetric bi-directionality of the excitation, nodes were selected and monitored on only one side around the transmitter, as correctly shown in the bottom plot.....	92
Figure 4.20: Peak-to-peak amplitude distribution on the rail top surface for pristine rail conditions. The black rectangles approximate the sensors positions during the field tests in Colorado. The aspect ratio is not exact, the plot is stretched along the y-dimension...	94
Figure 4.21: Peak-to-peak amplitude distributions on the rail top surface for model with 25% H.A. TD on the gage side as the transmitter gets further away from the defect (3.5(a), 7(b), 10.5(c) cm). The aspect ratio is distorted: the plot is stretched along the y-dimension.....	96
Figure 4.22: Peak-to-peak amplitude distributions on the rail top surface for model with 25% H.A. TD on the gage side as the transmitter gets further away from the defect (14(d), 17.5(e), 21(f) cm). The aspect ratio is distorted: the plot is stretched along the y-dimension.....	97
Figure 4.23: 2D-FFT from the same cuLISA3D simulation results used to obtain the plot in Figure 4.20. A line being 16 mm apart (toward the gage side) from the rail top surface central axis was the one considered to produce this plot	99
Figure 4.24: Cross-sectional view of VSH (left) and HSH having depths of 7.5 mm (center) and 15 mm (right) defects. Studies of the influence of the defect length.....	105
Figure 4.25: 10% H.A. TDs on the gage side. Study of the influence of the depth. Depths shown (from top to bottom-right): 2 (zoomed-in), 12, 22 mm, as measured at the highest tip of each defect	106
Figure 4.26: TDs on the gage side. Study of the influence of the cross-sectional dimensions (depth held constant). Dimensions shown: 5, 10, 15, 30, 40, 50 % H.A. (from the top-left to the bottom-right)	107
Figure 4.27: 25% H.A. TDs on the gage side. Study of the influence of the depth. Depths shown (from left to right): 1, 6, 11 mm, as measured at the highest tip of each defect..	107
Figure 4.28: Case I (no noise). TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack).....	116
Figure 4.29: Case I (no noise). TD 10% H.A. in the gage side. Study over the influence of the depth.....	117

Figure 4.30: Case I (no noise). TD 25% H.A. in the gage side. Study over the influence of the depth.....	117
Figure 4.31: Case I (no noise). TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack).....	118
Figure 4.32: Case I (no noise). VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm).....	118
Figure 4.33: Case I (no noise). HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm).....	119
Figure 4.34: Case I (no noise). HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm).....	119
Figure 4.35: Typical noise shape	122
Figure 4.36: Super-imposition of 5 signals acquired by 3 air-coupled receivers in a pristine part of the rail. The right plots are zoom-ins from the areas marked as black rectangles in the corresponding left figures. Values of the ‘Normalized Maximum Amplitude RMS’ are given for each sensor in the respective title	125
Figure 4.37: Values of the ‘Normalized Maximum Amplitude RMS’ for each available simulated sensor. Red circled are the ones considered to simulate the actual air-coupled receivers	126
Figure 4.38: Damage Index for rail in pristine conditions. Comparison between numerical simulations results and experimental results.....	127
Figure 4.39: Case II (added noise). TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	129
Figure 4.40: Case II (added noise). TD 10% H.A. in the gage side. Study over the influence of the depth.....	129
Figure 4.41: Case II (added noise). TD 25% H.A. in the gage side. Study over the influence of the depth.....	130
Figure 4.42: Case II (added noise). TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	130
Figure 4.43: Case II (added noise). VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm).....	131

Figure 4.44: Case II (added noise). HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm).....	131
Figure 4.45: Case II (added noise). HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm).....	132
Figure 4.46: Example showing the degree of variability introduced by noise affecting the signals	135
Figure 4.47: Comparison between experimental and numerical results. DIs showing the detection of a TD. The top plot (a) shows a zoomed-out view to appreciate the low noise floor around the TD indications. In (b) a zoomed-in version is furnished to facilitate the comparison with plot (c)	137
Figure 4.48: Typical DI traces shapes for different types of defect, indicating the possibility to estimate the detected type of defect simply measuring the highest peaks distance	138
Figure 4.49: Case II ($f = 1.5$), ROC. TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	143
Figure 4.50: Case II ($f = 1.5$), ROC. TD 10% H.A. in the gage side. Study over the influence of the depth.....	144
Figure 4.51: Case II ($f = 1.5$), ROC. TD 25% H.A. in the gage side. Study over the influence of the depth.....	144
Figure 4.52: Case II ($f = 1.5$), ROC. TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	145
Figure 4.53: Case II ($f = 1.5$), ROC. VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm)	145
Figure 4.54: Case II ($f = 1.5$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm).....	146
Figure 4.55: Case II ($f = 1.5$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm).....	146
Figure 4.56: Case II ($f = 3$), ROC. TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	149
Figure 4.57: Case II ($f = 3$), ROC. TD 10% H.A. in the gage side. Study over the influence of the depth.....	150

Figure 4.58: Case II ($f = 3$), ROC. TD 25% H.A. in the gage side. Study over the influence of the depth.....	150
Figure 4.59: Case II ($f = 3$), ROC. TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)	151
Figure 4.60: Case II ($f = 3$), ROC. VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm).....	151
Figure 4.61: Case II ($f = 3$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm).....	152
Figure 4.62: Case II ($f = 3$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm).....	152
Figure 4.63: Alignment of LISA signals.....	156
Figure 4.64: Alignment of experimental signals, based on the ones at the test starting point	159
Figure 4.65: Examples of signals randomly moving away from the initial aligned positions	160
Figure 4.66: Examples of variations on the initially perfect alignment of LISA signals induced to simulate realistic detrimental effects.....	163
Figure 4.67: Comparisons over the performances offered by the 3 considered types of analysis (10% TDs in the gage side at various depths).....	164
Figure 4.68: Comparisons over the performances offered by the 3 considered types of analysis (TDs in the field side at depth of 2 mm)	165
Figure 4.69: Comparisons over the performances offered by the 3 considered types of analysis (VSH defects).....	166
Figure 4.70: Finite element model to study the pressure field generated by the cylindrically focused transmitter. In light blue: air (the only visible face is the one corresponding to the focused active surface). In green: PML (there is a PML layer also in the hidden face, the one on the opposite side with respect of the active surface)	173
Figure 4.71: Air part salient dimensions.....	176
Figure 4.72: The transmitting transducer, having 25×25 mm active surface and 38 mm focal distance	177
Figure 4.73: Highlighted in white are the faces of the solid elements subjected to the pressure time-history depicted in Figure 4.74 in order to mimic the ultrasonic waves released by the cylindrically focused air-coupled transmitter	177

Figure 4.74: Pressure time-history applied to the elements shown in Figure 4.73	177
Figure 4.75: Sets of elements considered in the analysis.....	178
Figure 4.76: Top view of the contour plots showing pressure fields pertaining to 5 planes lying normally to the wave propagation direction	180
Figure 4.77: 3D view of the contour plots showing pressure fields pertaining to 5 planes lying normally to the wave propagation direction	181
Figure 4.78: Variation of pressure along lines directed as the y-direction at 12.5 mm if considering Figure 4.76	182
Figure 4.79: Variation of pressure along lines directed as the x-direction at 12.5 mm if considering Figure 4.76	182
Figure 4.80: Pressure variation moving away from the transmitter active surface. LS-DYNA simulation results.....	184
Figure 4.81: Pressure variation moving away from the transmitter active surface. Experimental results using a PICO sensor as receiver.....	184
Figure 4.82: Experimental test to assess the pressure field generated by the transmitter	185
Figure 4.83: Experimental test to assess the pressure field generated by the transmitter	185
Figure 4.84: LS-DYNA (left) vs experimental test (right) results. Pressure field in a plane 38 mm way from the transmitter.....	185
Figure 4.85: LS-DYNA model. In red: rail. In light blue: PML material. In green: air. Note: no PML was used at the air boundaries	186
Figure 4.86: Boundary conditions applied at the nodes on the outer surfaces of the PML layers	187
Figure 4.87: Finite element mesh of AREMA 136RE rail cross-section	188
Figure 4.88: Finite element mesh quality. Maximum side length in the rail cross-section	189
Figure 4.89: Finite element mesh quality. Aspect ratio distribution in the rail cross-section	189
Figure 4.90: Finite element mesh quality. Jacobian distribution in the rail cross-section	190
Figure 4.91: 15% H.A. TD - Defect modeling details	191
Figure 4.92: 15% H.A. TD located in the middle of the modeled rail.....	191

Figure 4.93: Air column below the transmitter and in contact with the top surface of the rail	193
Figure 4.94: Highlighted in white are the faces of the solid element subjected to the pressure time-history as in Figure 4.74 in order to mimic the signal released by the cylindrically focused air-coupled transmitter	193
Figure 4.95: Air column below the receivers.....	195
Figure 4.96: Highlighted in white are the elements on the top surface of the air-part that mimic the extension of the active surface of an air-coupled receiver.....	195
Figure 4.97: Relative distances between transmitter, defect and receiver	196
Figure 4.98: Section cut made on the model to highlight the salient features of the results. In greyscale is the removed portion of the model.....	197
Figure 4.99: Contour plot of pressure after (a) 50 μ sec, (b) 125 μ sec and (c) 180 μ sec.	202
Figure 4.100: Contour plot of pressure after (d) 205 μ sec, (e) 215 μ sec and (f) 225 μ sec	203
Figure 4.101: Waveforms received from the simulated receiver in the models with and without defect.....	204
Figure 4.102: Zoomed version of Figure 4.101 to highlight the salient part of the waveforms.....	204
Figure 4.103: (a) Least efficient method to keep a 25mm-long line of excitation and (b) final solution	206
Figure 4.104: (a) AUTOCAD [®] design and (b) 3D-printed transmitter adapter/holder. The three holes shown on the left side allow for a secure connection of the holder to the other parts through the use of screws, while still permitting to rotate the transmitter as wanted.	207
Figure 4.105: (a) Model ‘a’: the green part is air, and so are all the other parts as well. (b) Model ‘b’: the black part is ABS, all the others are air	209
Figure 4.106: Model ‘a’: top view (left) and side view (right).....	209
Figure 4.107: Model ‘b’: top view (left) and front view (right). Some salient dimensions of the models are shown here.....	210
Figure 4.108: Model ‘a’: contour plots of pressure after 60, 140, 200, 260 μ sec. Top view (section cut operated halfway through the height of the model)	212
Figure 4.109: Model ‘b’: contour plots of pressure after 60, 140, 200, 260 μ sec. Top view (section cut operated halfway through the height of the model)	213

Figure 4.110: Model ‘a’: contour plots of pressure after 60, 140, 200, 260 μ sec. Side view (section cut operated in correspondence of the tip of the active surface)	214
Figure 4.111: Model ‘b’: contour plots of pressure after 60, 140, 200, 260 μ sec. Side view (section cut operated in correspondence of the tip of the active surface)	215
Figure 4.112: Comparison of the air-pressure waveforms generated by the modeled 25x25mm transmitter and the modeled 50x50mm one, when used with or without the adapter	217
Figure 4.113: Comparison of the air-pressure waveforms generated by the modeled 25x25mm transmitter and the modeled 50x50mm one, when used with or without the adapter. The three signals have been aligned in time to facilitate the comparison	217
Figure 5.1: UCSD/FRA Rail Defect Testing Facility	224
Figure 5.2: Design of the prototype parts in AutoCAD [®] environment	225
Figure 5.3: Prototype parts attached to the cart rolling on rails	226
Figure 5.4: Schematic of the hardware layout deployed in the current version of the prototype	227
Figure 5.5: The cart hosting the prototype parts and the hardware components at Camp Elliott RDTF	228
Figure 5.6: Close-up look at the encoder wheel	229
Figure 5.7: Photo of the inside of the grey box shown on top of the cart hosting the prototype in the previous two figures. The impedance matching networks realized for the available 8-channels as described in Section 5.5.2 are visible	229
Figure 5.8: Digital filtering techniques used in former versions of the prototype (both bandpass filter and matched filter are applied directly to the raw signal, i.e. the matched filter on the bottom sub-plot was not applied to the bandpassed signal shown in the central sub-plot)	232
Figure 5.9: Electrical circuit “preamplifier-inductor-preamplifier” in PSpice software (Student version)	235
Figure 5.10: Results of “AC sweep” type of analysis done in PSpice software	236
Figure 5.11: Comparison of signals received without (top) and with (bottom) the L-matching network from one of the air-coupled receivers used in the prototype	236
Figure 5.12: Example of ROC curves computed to evaluate the performances of different sensors pairs/features/filters combinations	248

Figure 5.13: Part of a Microsoft Excel® sheet that was made to compare the performances of few examined sensors pairs/features/filters combinations. The numerical values shown in the cells are the AUC of the ROC curve computed for each combination	249
Figure 5.14: Signal recorded by an air-coupled receiver that shows successive wave arrivals after the meaningful signal leaked from the rail top surface. Three different x- or y-scales are used to plot the same waveform in the three sub-plots	254
Figure 5.15: Signal recorded by the same air-coupled transducer when the sponge is not being used (red) vs when it is being used (blue). Two different y-scales are used to plot the same two waveforms in the two sub-plots	255
Figure 5.16: Rail Defect Test Facility (RDTF) at TTC	256
Figure 5.17: The prototype mounted on the cart. The sensors are behind the grey box which is attached to the cart.....	257
Figure 5.18: The prototype mounted on the cart was towed by the FRA’s hy-railer	257
Figure 5.19: TTC’s RDTF map showing details of the testing location.....	259
Figure 5.20: Approximate current coverage of the rail defect detection prototype	261
Figure 5.21: Damage Index traces. Red lines: locations of defects included in both “List 1” and “List 2”. Green lines: locations of defects included only in “List 1”. Cyan lines: locations of welds. Black lines: locations of joints.....	265
Figure 5.22: ROC curves computation procedure	267
Figure 5.23: ROC curves computed for each of the four runs in the Technology Development Section of TTC’s defect farm	269
Figure 5.24: “Cumulative” ROC curves for the four runs (top) and super-imposition of all the results in two plots, one for each Defect List considered (bottom)	270
Figure 5.25: Comparison of the best results achieved with each of the 3 different techniques	273

LIST OF TABLES

Table 2.1: Advantages and disadvantages of different rail inspection methods.....	46
Table 4.1: Map of models including TD defects on the gage side	103
Table 4.2: Map of models including TD defects on the field side.....	104
Table 4.3: Map of models including VSH defects (located on the vertical plane passing through the rail central running axis).....	104
Table 4.4: Map of models including HSH defects.....	105
Table 4.5: Features used to analyze the simulations results	109
Table 4.6: Summary of the salient (non-statistical) Damage Index values for the various models (case I: signals not affected by noise)	120
Table 4.7: Summary of the salient (averaged) Damage Index values for the various models (case II: noise is artificially added to the simulation results; “high” SNR level ($f = 1.5$)).....	133
Table 4.8: Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow	147
Table 4.9: Summary of the salient ROC curves results for the various models. Case of “low” SNR level ($f = 3$). The values lower than 100% are highlighted in yellow.....	153
Table 4.10: Sensors treated as arrays in the “ideal” case: perfectly coherent sums (Figure 4.63). Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow	167
Table 4.11: Sensors treated as arrays in the “realistic” case: induced normally distributed dephasing with standard deviation 15° (Figure 4.66). Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow	168
Table 4.12: Effect of temperature on properties of air (for an atmospheric pressure of 1 atm)	171
Table 4.13: Consistent system of units adopted for the LS-DYNA simulations	172
Table 5.1: List showing the 9 features proposed for the system. In yellow background are the ones currently in use	245
Table 5.2: Map of defects agreed upon with Volpe and TTC personnel from survey of October 29, 2015.....	260

Table 5.3: Summary of the results obtained with the three techniques. Area Under the Curves of ROC plots computed for each run are given. The best results obtained for each run/defect list are highlighted in yellow	274
--	-----

ACKNOWLEDGEMENTS

My deepest thanks go to my advisor, Prof. Francesco Lanza di Scalea, who believed in me and brought me in his research group at the NDE/SHM Laboratory of UCSD. I would also like to thank all the past and present colleagues who I had the opportunity to meet during this journey. Special thanks go to Thompson Nguyen and Xuan “Peter” Zhu. We started this experience together, and now we are all getting close to wrap it up. One of my principal mentors, before becoming a friend, as I was approaching this new chapter of my life has certainly been Dr. Robert Phillips. I would like to dedicate this dissertation to him, and to his memory.

The research presented in this dissertation was supported by the Federal Railroad Administration (FRA) of the U.S. Department of Transportation, under university research grant FR-RRD-0027-11-01. The author is particularly grateful to Mahmood Fateh, who is the Program Manager for this grant from the FRA Office of Research and Development.

Special thanks go to Ph.D. candidate Thompson Nguyen, who has shared with me most, if not all, of the experimental testing experience made on the prototype object of this dissertation. Furthermore, he is the investigator who developed the LabVIEW[®] software components that governs the system in real-time. Thanks are extended to Peter Zhu, who has also helped in many stages and in different aspects of the prototype development.

The dissertation author would also like to thank Piotr Kijanka, Dr. Pawel Packo and Prof. Wieslaw Jerzy Staszewski from the AGH University of Science and

Technology, Krakow, Poland, who provided the software based on Local Interaction Simulation Approach that has been used to simulate wave propagations in rail. In particular, Piotr Kijanka collaborated on the development of Section 4.2.2 of this work.

Special thanks go to my friend Marco Pigazzini who provided fundamental suggestions and tips during the setting up of the LS-DYNA numerical model.

Thanks are extended to Bob Parks, Dr. Christopher Latham and to all the past and present crews working at UCSD Powell Labs and at Camp Elliott. Their help and guidance have been crucial in different parts of prototype development.

Several individuals and organizations were instrumental for the preparation and the conduction of the field test at the Transportation Technology Center, in Colorado, in October 2014. In particular, special thanks go to Eric Sherrock and his crew at ENSCO for field test support and hy-railer operation, and to Dr. Robert Wilson of the Volpe Transportation Center for his technical advices and assistance with the post-test result analysis.

Chapter 4, in part, has been published in the *Structural Health Monitoring Journal*, Mariani, S., Nguyen, T., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., Staszewski, W. J., Fateh, M., and Carr, G. (2013). The title of this paper is “Non-contact air-coupled ultrasonic guided wave inspection of rails”. The dissertation author was the primary investigator and primary author of this paper.

Chapter 3 and 4, in part, will be submitted for publication in *Transportation Research Record: Journal of the Transportation Research Board*, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Air-coupled wave propagation in

steel rail”. The dissertation author will be the primary investigator and author of this paper.

Chapter 4 and 5, in part, will be submitted for publication to the *Structural Health Monitoring* Journal, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Defect detection performance of the non-contact air-coupled ultrasonic guided wave inspection of rails prototype”. The dissertation author will be the primary investigator and author of this paper.

Chapter 4 and 5, in part, will be submitted for publication to the *Experimental Mechanics* Journal, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Signal to Noise Ratio enhancement for air-coupled wave propagation in steel rail”. The dissertation author will be the primary investigator and author of this paper.

VITA

- 2007 Bachelor of Science, University of Perugia, Italy
- 2010 Master of Science, University of Perugia, Italy
- 2011 - 2015 Research Assistant, University of California, San Diego, USA
- 2015 Doctor of Philosophy, University of California, San Diego, USA

JOURNAL ARTICLES

1. Mariani, S., Nguyen, T., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., Staszewski, W. J., Fateh, M., and Carr, G. (2013). Non-contact air-coupled ultrasonic guided wave inspection of rails. Structural Health Monitoring, 1475921713498533.

CONFERENCE PROCEEDINGS

1. Mariani, S., Nguyen, T. V., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., and Staszewski, W. J., "Non-contact ultrasonic guided wave inspection of rails," Proc. SPIE 8692, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2013, 86921L, 2013
2. Mariani, S., Nguyen, T. V., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., and Staszewski, W. J., "Non-contact ultrasonic guided wave inspection of rails," IWSHM, Stanford, September 10-12, 2013.
3. Nguyen, T. V., Mariani, S., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., and Staszewski, W. J., "Non-Contact Ultrasonic Guided Wave Inspection of Rails," In ASME 2013 International Mechanical Engineering Congress and Exposition (pp.

V011T06A012-V011T06A012). American Society of Mechanical Engineers, 2013

4. Nucera, C., Phillips, R., Zhu, P., Mariani, S., Lanza di Scalea, F., Fateh, M., Carr, G., “Nonlinear guided waves for neutral temperature measurement in continuous welded rails: results from laboratory and field tests,” Proc. SPIE 8695, Health Monitoring of Structural and Biological Systems 2013, 869528, 2013
5. Nguyen, T. V., Mariani, S., Lanza di Scalea, F., "Non-contact ultrasonic guided wave inspections of rails," Proc. SPIE 9064, Health Monitoring of Structural and Biological Systems 2014, 90640V, 2014
6. Mariani, S., Nguyen, T. V., Lanza di Scalea, F., and Fateh, M., “High Speed Non-Contact Ultrasonic Guided Wave Inspections of Rails,” In 2014 Joint Rail Conference (pp. V001T01A008-V001T01A008). American Society of Mechanical Engineers, 2014
7. Mariani, S., Nguyen, T. V., Zhu, X., Lanza di Scalea, F., Fateh, M., “Non-contact ultrasonic guided wave inspection of rails: field test results and updates,” Proc. SPIE 9435, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2015, 94351O, 2015

ABSTRACT OF THE DISSERTATION

Non-contact Ultrasonic Guided Wave Inspection of Rails: Next Generation Approach

by

Stefano Mariani

Doctor of Philosophy in Structural Engineering

University of California, San Diego, 2015

Professor Francesco Lanza di Scalea, Chair

In the last ten years, over 1,000 train derailments that occurred in the US railroad system were caused by undetected rail internal defects that suddenly and dramatically emerged as track breakages. The total cost of those accidents is quantifiable in a few hundred millions of dollars, without considering the associated tragic losses of life and bodily injuries. While there already exist a few methods for the detection of rail internal flaws, several well-known limitations prevent each of them from detecting all of the critical flaws.

As a proposal to address this issue, the object of this dissertation has been the development of a novel rail inspection system based on ultrasonic guided waves propagating through rails. Both the generation and the detection of these waves are achieved in a non-contact manner through the use of piezoelectric air-coupled transducers. One of the advantages of employing such non-contact method is the potential to perform tests as the train or inspector car travels at high speed along the railroad.

Nevertheless, the main drawback of using air-coupled transducers on steel rail is represented by the significant energy losses occurring at the interface between air and steel due to the large acoustic impedance mismatch between these media. As a result, the signal to noise ratio available when analyzing the data is severely penalized. In an attempt to overcome this limitation, very effective electrical impedance networks have been designed. In parallel, a statistical analysis method based on multivariate outlier detection has been implemented to enhance the defect-sensitivity of the system.

Numerical analyses of the ultrasonic wave propagation and interaction with different types of rail internal defects have been carried out using both a finite difference method, based on the Local Interaction Simulation Approach (LISA), and a commercial finite element method software. The results of these analyses were instrumental in understanding salient aspects of the guided wave propagation phenomenon in rails and throughout the ever ending decision-making process for the definition of the many system operating parameters involved.

A prototype based on these technologies has been built and tested both in-house at the UCSD Rail Defect Test Facility located at Camp Elliott, and in the field at the

Transportation Technology Center located in Colorado, in October 2014. Receiver Operating Characteristics curves were used to characterize the performance of the defect detection based on the trade-off between defect detection rate and false alarm rate. In particular, the results of the field test were found to be quite satisfactory and perfectly in line with the predictions of the numerical analyses. Future work should be aimed at improving the prototype performance, particularly in terms of test speed, based on the lessons learnt from the October 2014 field tests and in the laboratory.

1 Introduction

1.1 Non Destructive Evaluation of railroad tracks. Introductory discussions and motivations for the research

Safety statistics data from the Federal Railroad Administration (FRA) [1] show that, in the 2005-2014 decade, a total of 2,653 train accidents were caused by track failures, with a total reportable direct cost of \$651M. Among the causes of such accidents, Transverse/Compound Fissures and Detail Fractures were the most relevant, having caused 1,033 derailments accountable for \$276M in costs. Transverse Fissures and Detail Fractures both belong to the Transverse Defect (TD) category, which, as the term implies, include cracks primarily growing in a plane perpendicular to the rail running direction [2, 3, 4, 5]. Since the 1960s, ultrasonic defect detection has been used to detect internal defects in rails [6]. Currently, one of the most widely used rail inspection techniques is based on a fluid-filled wheel probe inside which piezoelectric transducers are placed to ensure coupling to the top of the rail [7, 8, 9]. However, the main drawback of wheel-based systems is that surface shelling cracks, or shelling, may mask internal transverse defects. This limitation was the cause of some major derailments in the recent past, such as in Superior, WI in 1992 and Oneida, NY in 2007, where hazardous material spillage caused the evacuation of 40,000 people. In response to those accidents and others, the National Transportation Safety Board issued Safety Recommendations to the FRA for improving the effectiveness of rail inspection technologies in the detection of rail internal defects, especially under shelling [10]. Other drawbacks of the conventional rail inspection methods include speed limitations

(typically less than 25 km/hour) as well as some difficulties in detecting Vertical Split Head defects.

In the last 20 years, efforts have been focused on non-contact rail inspection. Theoretically, a non-contact method can lead to higher testing speeds and can be less sensitive to surface-conditions or non-effective coupling. One of the investigated non-contact inspection technique is based on Electro-Magnetic Acoustic Transducers (EMATs) [11, 12, 13, 14]. Drawbacks of this technique include a low efficiency and the needing for large magnets to be sufficiently effective. Another non-contact rail inspection technique was developed at the University of California at San Diego (UCSD) [15, 16] and, in parallel, at the Transportation Technology Center (TTC) [17], and it is based on a hybrid laser transmission and air-coupled detection. The UCSD system features a statistical analysis of the data collected by the receivers that increases the defect detectability and minimizes false positives [18]. The drawback is that the presence of the laser makes these systems quite costly and potentially hard to maintain. Wooh [19, 20] proposed a different method based on air-coupled transmission and detection, where the presence of defects was revealed through a Doppler shift due to the relative motion between the moving inspection vehicle and the defect. This technique was only tested in the laboratory and it is only effective for surface-breaking cracks.

In this dissertation a new method is presented, based on guided wave air-coupled testing of rails in pair with a statistical detection. A prototype based on this technology has been developed and thoroughly tested under a Federal Railroad Administration R&D grant. In particular, tests were performed both in-house at the UCSD Rail Defect Testing

Facility located at Camp Elliott, and in the field at the Transportation Technology Center located in Colorado, in October 2014.

1.2 Outline of the numerical and experimental approach of the dissertation

In this dissertation the emphasis is on the use of ultrasonic guided waves for performing non-destructive inspection of railroad tracks. In particular, the current state of knowledge of guided waves rail inspection is extended both from a fundamental wave propagation viewpoint and from an experimental prototype development viewpoint.

Chapter 2 discusses the motivations behind the research presented in this dissertation. In particular, an overview of the various types of rail defects is given, along with a summary of the historical background of rail inspection in the United States. Currently used rail inspection systems are reviewed pointing out both their salient features and their main limitations. The guided wave approach is then presented as a promising technique with potential advantages compared to the traditional methods.

A unique part of the research has been the development of a prototype for rail defect detection based on air-coupled ultrasonic guided waves technique, which is first introduced in chapter 3. In this introductory chapter, a general overview of the main concepts is given, along with the reasons behind the choice of many fundamental operational parameters.

Chapter 4 presents the salient results obtained using two different numerical analysis techniques to study the wave propagation phenomena occurring during the execution of the experimental tests. In particular, a newly designed software based on the

Local Interaction Simulation Approach has been used to get insights on peculiar aspects of the experimentally induced guided waves propagation in rail and to thoroughly characterize the defect sensitivity of the system based on the currently available signal-to-noise ratios. On the other hand, a commercial finite element software has been used to exploit the complete ultrasonic wave propagation occurring during experimental tests in both air and steel media.

Chapter 5 presents a detailed description of the prototype hardware design and the statistical algorithm to enhance the defect-sensitivity. Results of the field test performed in October 2014 at TTC testing facility for the Federal Railroad Administration are also presented.

Chapter 6 summarizes and discusses the key results of this research, and provides recommendations for future work.

2 State of the art review

2.1 Development of structural defects in rails

The rail service life is primarily affected by the chemical composition of the rail, the track maintenance programs and the speed and tonnage of the vehicles running on it. All of these factors have an influence on the development of rail defects [3].

Understandably, rail stresses are the predominant factor in the creation and development of internal rail defects. Gross vehicle load is the main source of bending and shear stresses affecting the rail. Furthermore, the rail is also subjected to thermal and residual stresses.

One way of classifying rail defects is based on their origin [2]. Following this method, there are three families of defects: defects caused by improper use or handling, rail manufacturing defects and rail wear and fatigue. Defects deriving from improper use or handling of rails are mostly due to sudden train brakes or spinning of train wheels. Rail manufacturing defects are usually due to inclusions of non-metallic origin or wrong local mixings of the rail steel components that, under operative loads, generate local concentration of stresses which in turns can originate the creation of cracks [21]. The last class of defects includes the ones caused by wearing mechanisms of the rolling surface and/or by fatigue. Not all rail defects are considered critical. A critical one is a rail defect that may affect the safety of train operations. Non-critical defects are defects that occur in the rail but do not affect its structural integrity or the safety of the trains operating over it [22].

A different method of rail defects classification is based on the direction of propagation of the flaws under operative loads [3]. Two types of damages are found in

this classification method: transverse (extending primarily in the rail transverse plane) and longitudinal defects (extending either in the rail vertical or horizontal plane) (Figure 2.1).

Figure 2.1: Terminology used to identify defects development planes in relation to the rail section

The Title 49 of the Code of Federal Regulations (CFR) Part 213.339 states that [23]:

“(a) A continuous search for internal defects shall be made of all rail in track at least twice annually with not less than 120 days between inspections.

...

(c) Each defective rail shall be marked with a highly visible marking on both sides of the web and base.

...”

Upon detection of an internal defect, the Title 49 of the CFR Part 213.337 prescribes the remedial actions required based on the defect type:

“(a) When an owner of track to which this part applies learns, through inspection or otherwise, that a rail in that track contains any of the defects listed in the following table, a person designated under § 213.305 shall determine whether or not the track may continue in use. If the person determines that the track may continue in use, operation over the defective rail is not permitted until—

- (1) The rail is replaced; or
- (2) The remedial action prescribed in the table is initiated
- ...

A list of remedial actions is provided in case the owner of the track decides not to remove the defective part from the track.

Statistics related to accident reports in U.S. are available through the Federal Railroad Administration Office of Safety Analysis website [1]. Reports relative to any desired period of time can be readily generated. Figure 2.2 shows a list of train accidents caused by rail defects in the 2005-2014 decade.

The first cause of accidents were “Transverse/compound fissures”, that were responsible for the 21.7% of rail failures in the considered timeframe (575 accidents), resulting in \$113M of direct costs. Those costs are calculated considering train delays, costs of inspection, costs of remedial and preemptive treatments, loss of business confidence and customer support.

“Detail fractures” were the second prevalent cause, with 458 accidents (17.3% of the total) resulting in even higher direct costs, totaling \$163M.

“Vertical split heads” come third, with 319 accidents and \$46M costs.

TRAIN ACCIDENTS BY CAUSE FROM FORM FRA F 6180.54

MAJOR CAUSE= Track

Selections: Railroad - ALL
Region - All Regions
State - ALL, County - ALL

ALL ACCIDENT TYPES / All TRACK TYPES / T - Rail, Joint Bar and Rail Anchoring
Time Frame: Jan 2005 To Dec 2014

Specific causes:	Total		Type of Accident			Reportable Damage		Casualty	
	Cnt	%	Coll	Der	Othr	Amount	%	Kld	Nonf
T201- Bolt hole crack or break	104	3.9	1	102	1	30,535,238	4.7	0	2
T202- Broken base of rail	215	8.1	1	214	-	35,483,799	5.4	0	38
T203- Broken weld (plant)	12	0.5	-	12	-	3,764,535	0.6	0	0
T204- Broken weld (field)	58	2.2	-	58	-	33,596,419	5.2	0	4
T205- Defective or missing crossties	126	4.7	-	124	2	7,466,623	1.1	0	1
T206- Defect/missing spike-oth rail fastener	56	2.1	-	56	-	11,779,484	1.8	0	1
T207- Detail fracture - shelling/head check	458	17.3	2	455	1	163,340,077	25.1	2	14
T208- Engine burn fracture	8	0.3	-	8	-	2,271,098	0.3	0	0
T210- Head and web sep(outside jt bar limit)	257	9.7	2	254	1	24,713,793	3.8	1	6
T211- Head & web separation-in jt bar limit	49	1.8	-	49	-	8,374,971	1.3	0	0
T212- Horizontal split head	70	2.6	-	70	-	5,765,287	0.9	0	1
T213- Joint bar broken (compromise)	24	0.9	-	24	-	32,483,223	5.0	0	132
T214- Joint bar broken (insulated)	18	0.7	-	18	-	19,016,470	2.9	0	0
T215- Joint bar broken (noninsulated)	41	1.5	-	41	-	27,024,787	4.1	0	2
T216- Joint bolts, broken, or missing	21	0.8	-	21	-	5,781,350	0.9	0	0
T217- Mismatched rail-head contour	56	2.1	-	56	-	2,597,471	0.4	0	0
T218- Piped rail	5	0.2	-	5	-	331,562	0.1	0	0
T219- Rail defect with joint bar repair	4	0.2	-	4	-	844,992	0.1	0	0
T220- Transverse/compound fissure	575	21.7	-	572	3	113,160,573	17.4	0	5
T221- Vertical split head	319	12.0	1	317	1	46,240,083	7.1	0	0
T222- Worn rail	61	2.3	-	60	1	7,303,209	1.1	0	0
T223- Rail Condition - Dry rail, freshly ground	4	0.2	-	4	-	121,108	0.0	0	0
T299- Other rail and joint bar defects	112	4.2	1	110	1	69,274,627	10.6	0	3
-- Total	2,653	100	8	2,634	11	651,270,779	100.0	3	209

Figure 2.2: Train accidents due to rail, joint bar and rail anchoring defects in the 2005–2014 decade from FRA Safety Statistics Data

Sections 2.1.1 through 2.1.6 present insights on the different types of rail defects, with a brief description of their mechanical aspects and actual pictures taken from rail tracks. The information provided are mostly based on a booklet recently released by the FRA Office of Railroad Safety that presents a comprehensive description of the different types of rail defects encountered in rail tracks [3].

2.1.1 Transverse defects in the rail head

A transverse defect (TD) is a crack that has developed in the cross-sectional plane of the rail head. TDs include *transverse/compound fissures*, *detail fractures*, *engine burns fractures*, *welded burn fractures* and *ordinary breaks*. Usually these defects develop over

a period of time in gradual stages. The development is typically progressive and can be uninterrupted.

The exact type of transverse defect cannot be determined until after the rail is broken for examination. It might remain hidden until it reaches an outer surface. After the rail is broken, the transverse defect is measured against the cross-sectional area of the rail head. Sizing of all types of transverse-oriented defects is reported by an approximated percentage of cross-sectional area of the rail head (Figure 2.3). All other defect types are normally reported in inches.

Figure 2.3: Identification of a transverse defect based on its cross-sectional area

Figure 2.4: Transverse defects. On the right, case of coupling with surface shelling

Pictures of two TDs are furnished in Figure 2.4, where the coupling of a TD with a surface *shelling* is also shown.

In the best case scenario (left picture in Figure 2.4), the TD is not hidden by *shelling* and has pronounced growth rings, while in the worst case scenario (right picture in Figure 2.4) the smooth TD hides underneath a shell. *Shelling* is usually caused by wearing mechanisms of the rolling surface, and it is described in Section 2.1.5.

Transverse fissures:

Figure 2.5: Transverse fissures

Transverse fissure (Figure 2.5) indicates a transversal fracture starting from a crystalline center or nucleus in the rail head from which it spreads outward as a smooth circular or oval surface. The main difference with the other types of flaws is the crystalline center or nucleus and the nearly smooth surface of the development that surrounds it.

Typically a transverse fissure will remain hidden in the rail head for some time without further development. Development is highly influenced by wheel impact and rail

bending stresses, and growth is normally slow until a size encompassing 20–25% of the cross-sectional area of the rail head. Once the defect reaches that size, growth is normally more accelerated. Multiple fissures are often present in one rail length.

Transverse fissure defects are inherent from the manufacturing process and are found predominantly in non-control-cooled rail made prior to the mid-1930s. However, they can develop in more modern high-chrome rails from hydrogen imperfections. Identification of this type of defect is not accurately made until the rail section is broken and the size determined by the cross-sectional area of rail head affected.

Compound fissures:

Figure 2.6: Compound fissure

Compound fissure (Figure 2.6) indicates a progressive flaw branching out from a *horizontal split head*, which turns up or down in the head of the rail as a smooth surface progressing until it is substantially transverse to the length of the rail.

Typically this defect originates as a horizontal separation from an internal longitudinal seam, segregation, or inclusion inherent from the manufacturing process. It then develops longitudinally prior to transverse progression upward, downward, or in both directions in relation to the transverse plane of the rail section.

Until reaching a size of about 30–35% the development remains slow. The compound fissure is considered a hazardous rail defect.

Detail fractures:

Figure 2.7: Detail fracture originating from a visible shell

Detail fracture (Figure 2.7) indicates a progressive flaw originating at or near the surface of the rail head and developing transversely. It often originates from shelly spots, head checks, or flaking. The latter is the main difference with transverse and compound fissures, or other defects, which have internal origins.

The detail fracture is usually associated with the presence of a longitudinal seam or streak near the running surface on the gage side. Unlike the transverse fissure, no nucleus is present.

The development is usually slow until a size of about 10-15% of the cross-section of the rail head. After that, growth can become rapid and/or sudden, prior to complete failure. More than one detail fracture may originate in a close by area where the conditions that initiate their development, such as shelling or head checking, are present.

Figure 2.8: Detail fracture from head check

Figure 2.8 shows a detail fracture originated by a *head check condition* located at the upper gage corner of the rail, which is normally associated with concentrated loading that cold works the steel. This can also be referred to as a thermal crack.

Similarly, the reverse detail fracture (Figure 2.9) is a progressive transverse fracture that originates at the bottom corner of the gage side of the rail head. The origin is a notching condition on the cold rolled lip located on the bottom corner of the rail head

which act as a source of load concentrations. The cold rolled lip condition is typically associated with severely worn rail and high axle loadings. In a similar fashion to the detail fracture, the development is normal to a size of about 10% and is often rapid or sudden prior to complete failure of the rail section. However, complete rail failure can arise at a size much less than that of a typical detail fracture type defect.

Figure 2.9: Reverse detail fracture

Engine burn fractures:

The engine burn fractures (Figure 2.10) are progressive fractures that originate in spots where driving wheels have slipped on top of the rail head. After the interested portion of rail surface is heated, rapid cooling forms thermal cracks. Impact from wheels over the affected burned area initiates a slight horizontal separation of the burned metal from the parent rail metal and develops a flat spot. Transverse separation may start from a thermal crack in the region of the burn at any time.

Figure 2.10: Engine burn fracture showing significant growth

Typically, more than one engine burn fracture is located within a short distance. The development is slow until a size of about 10–15% of the rail head area; after that, the growth can quickly accelerate.

Welded burn fractures:

The welded burn fracture is a transverse flaw which develops from an engine burn that has been resurfaced by welding. An insufficient cleaning of the old engine burn prior to resurfacing is usually the cause. This essentially fails to eliminate thermal cracks created by the original driver burn. Improper cooling of resurfaced burn can also create new thermal cracks.

Ordinary breaks:

When a partial or complete rail break is found, and none of the other defects described in this section are identified, the term ordinary break (Figure 2.11) is used.

Figure 2.11: Rail failure fracture face showing no transverse defects

This type of rail fracture can occur as a result of a significant wheel impact from a flat or broken wheel, particularly in very cold weather or when an unevenly supported base is present.

2.1.2 Longitudinal defects in the rail head

Progressive fractures having a longitudinal separation are referred to as longitudinal defects. This category includes *horizontal split heads* and *vertical split heads*.

Horizontal split heads:

Horizontal split head (Figure 2.12) indicates a horizontal progressive fracture in the rail head, typically one-quarter of an inch or more below the rail top surface. When the fracture reaches the side of the rail head, it appears as a crack lengthwise of the rail.

Figure 2.12: Horizontal split head originating from an internal seam

Figure 2.13: Side view of a horizontal split head

The origin of horizontal split head is an inclusion, an internal longitudinal seam, or segregation inherent from the manufacturing process. The horizontal fracture will develop longitudinally and horizontally and is normally rapid in growth. If a wheel impact initiates a transverse separation the defect is classified as a compound fissure. The horizontal separation may be present in several locations within the same rail section.

Vertical split heads:

Figure 2.14: Vertical split head

Vertical split head (Figure 2.14) indicates a vertical fracture through or close to the middle of the head, propagating along the rail direction. A crack or rust streak may show under the head, close to the web, in what is called the fillet area. The vertical split head defect can also produce a widening of the rail head for the length of the fracture or a dark streak on the running surface.

As for the horizontal split head case, the cause is an inclusion, an internal longitudinal seam, or segregation inherent from the manufacturing process. The vertical fracture will develop longitudinally and vertically, and may gradually turn toward the gage or field side of the rail head. Typically, a portion of a vertical split head will develop toward the gage side of the rail head while the other end develops toward the field side.

The development is normally very rapid once the seam or separation has opened up anywhere along its length.

2.1.3 Web defects

A web defect is any progressive fracture occurring in the web of the rail having, primarily, a longitudinal separation. This general classification includes four types of defects: *head and web separation*, *split web*, *pipel rail* and *bolt hole crack*.

Head and web separation:

Figure 2.15: Head and web separation defect associated with rail joint

Head and web separation (Figure 2.15) indicates a progressive fracture longitudinally separating the head and web of the rail at the fillet under the head.

A corrosion fatigue at the head/web fillet area may initiate due to an acidic action from some asphalt-based fill material used in road crossing. Excessive speed on curves, gravel in crossings or incorrect canting of the rail can cause eccentric loading of the rail head and initiate the development. The fracture can appear externally as a slight horizontal cracking under the head or as a rust-colored “rail strain” in the fillet area. These defects are also found at the jointed rail end as a result of extreme stress conditions often created by pumping or swinging joints.

Split web:

Figure 2.16: Web failure resulting from high residual stresses

A split web (Figure 2.16) is a crack along the side of the web in the rail direction, extending into or through it. The cause can be a mechanical damage or a seam. Also, it is not uncommon for the split web to develop at locations where heat numbers are stamped into the web. In other cases split webs are due to high residual stresses from the rail welding, roller straightening process, or joint application.

The development can be very rapid once the crack extends through the web, and it can be accelerated by heavy axle loading. The defect appears externally as a rust-colored bleeding along the crack development.

Piped rail:

A piped rail (Figure 2.17) is a vertical split in a rail, usually in the web, due to failure of the shrinkage cavity in the ingot to unite in rolling.

A longitudinal seam or a cavity inside the web inherent from the manufacturing process are the two possible causes of a piped rail. Once development initiates, the seam

will develop vertically toward the head and base of the rail. Heavy axle loading can result in the seam spreading or opening up in a crosswise direction, resulting in a bulge in the web. This type of defect is relatively uncommon in modern rail manufacturing technology.

Figure 2.17: Piped rail showing significant rail collapse

Bolt hole cracks:

Figure 2.18: Bolt hole crack

A bolt hole crack (Figure 2.18) is a crack originating from a bolt hole and progressing in a direction pointing either upward toward the rail head or downward

toward the base. Fully developed bolt hole cracks may continue in a longitudinal and horizontal direction along the head/web or base/web fillet, or they may progress into and through the head or base to separate a piece of the rail end from the rail.

The possible origins of a bolt hole cracks are: excess stresses due to pumping or swinging joints, improper drilling, excessively worn joint bars, abnormal rail end impacts from rolling stock, or unchamfered holes. Often the rail will break even from a very small defect when the rail end is subjected to unusual stresses.

2.1.4 Base defects

A base defect is any progressive fracture originating in the base of the rail. This general classification covers two types of defects: *broken base* and *base fracture*.

Broken base:

Figure 2.19: Half moon shaped broken base

A broken base is usually restricted to the flange area of the rail base and is usually an oval-shaped break referred to as a “half moon” break (Figure 2.19). Typical causes of such break are segregations, seams or improper bearings on the tie plate.

Base fracture:

Figure 2.20: Base fracture showing nick with transverse development

Base fracture (Figure 2.20) indicates a progressive fracture originating on the outer edge of the base of the rail, which usually develops in a transverse plane. It can result in complete transverse failure of the rail section. Typical causes of base fractures are nicks on or blows to the edge of the base, which results in an indentation or similar damage. Damage of this nature is sometimes caused by improper rail handling.

The growth is normally slow until the defect has progressed some distance into the rail section. However, a complete and sudden transverse rupture of the rail can occur with minimal transverse progression.

2.1.5 Rolling contact fatigue conditions

Rolling Contact Fatigue (RCF) conditions may develop in rails at the wheel/rail interface. One of the danger associated with such condition is that it can act as an obstacle in the detection of an underlying rail defect. For this reason, when a detector car

operator is in doubt about the integrity of the test due to surface conditions, he has the option to record the rail section as an invalid test and report the location to the railroad owner.

Even if cracking can be found in the head of all types of track, it is mostly found on highly canted curves (Figure 2.21) where high stresses arise due to extra pressure and wear of the wheel on the “high” rail (Figure 2.22).

Figure 2.21: Wheel contact on a curve with “high” rail and “low” rail [24]

Figure 2.22: Contact stresses on tight curved track

Hidden transverse defects may originate from RCF. For example, a disastrous derailment that occurred at Hatfield in the UK in October 2000 causing four deaths and

more than thirty injuries was originated from RCF damage. In the United States, a derailment most likely caused by RCF damage occurred near Superior, in Wisconsin, and caused hazardous material spillage, resulting in the evacuation of 40,000 people. A National Transportation Safety Board report states that the rail was inspected six weeks before the disaster and the ultrasonic inspection car operator reported that the rail contained shelling but did not consider the conditions severe enough to warrant an exception report. Additional ultrasonic inspections were performed with handheld equipment, but the operator considered the rail to be free from internal defects based on his evaluation of the test and his experience.

RCF damage initiates on or very close to the rail rolling surface. The crack is initially microscopic in size and then propagates through the heavily deformed orthotropic layers of steel at an angle of about 10° respect to the rail running surface, until it reaches a depth where the steel retains its original isotropic properties. At this stage the crack is a few millimeters deep into the rail head and it might simply lead to *spalling* of material from the rail surface. For reasons that have not been clearly understood, isolated cracks can turn down into the rail and cause the rail to break, if not timely detected [2, 24].

RCF occurrence is increasing especially on heavy haul railways or on high speed passenger lines. Normally, expensive rail grinding is performed in the attempt to remove an RCF condition. Its incidence can be reduced by lubricating rolling surfaces. However, fluid entrapment in the metal is one of the most common causes for speeding up the surface initiated crack growth [2].

Some of the most critical types of rail surface conditions are described below.

Shelling:

Figure 2.23: Gage side shell showing severe parent metal decay

A shell (Figure 2.23) is as progressive horizontal separations, generally on the gage side of the rail head, which may crack out at any level, most of the times at the upper gage corner. Shelling may also turn down to form a detail fracture. Uncapped or gutted shells will result in the dislodgement of parent metal from the rail section.

Flaking:

Flaking (Figure 2.24) originates at the rail head rolling surface and is usually found near the stock rail area of a switch where concentrated loading cold works the steel. Flaking progresses as horizontal separation with chipping of small segments of parent metal.

Figure 2.24: Centralized flaking condition showing chipping of parent metal

Burned rail:

Figure 2.25: Thermal cracks on burned rail stock rail

Burned rail (Figure 2.25) indicates a rail head condition resulting from friction due to slipping locomotive drivers. The damaged area can gradually chip out and roughen

under repeated traffic. Transverse defects can develop from thermal cracks associated with the burned area. Once the surface condition produces a large enough displacement of the rail head surface material, the detection of a potential underlying rail flaw usually results obstructed.

Head checking:

Figure 2.26: Gage side head checking and flaking

With head checking (Figure 2.26) is indicated a slight separation of metal on the gage side of the rail head, usually found in the high side of curves. Other times is found in switch areas, due to the lateral force induced on the rail head from wheel displacement through turnouts. Head checking can turn down and develop into a transverse separation.

Spalling:

With spalling (Figure 2.27) is indicated the displacement of parent metal from the rail head due to high contact stresses associated with cyclical loading.

Figure 2.27: Flattened rail head showing displacement of parent metal (spalling)

Crushed head:

Crushed head (Figure 2.28) refers to the flattening of several inches of rail head, usually accompanied by a crushing down of the metal. Small cracks in a depression on the running surface are also produced. In advanced stages, a bleeding crack at the fillet under the head may appear.

The cause of crushed head is usually a softer spot in the steel of the head area, which gives way under heavy wheels loads, or an advanced rail surface collapse condition [5].

The crushed head is not considered a hazardous defect. However, it is normally removed from high speed track because it causes rough riding and because it acts as a point of concentrated loads from which internal defects may develop [4].

Figure 2.28: Sketch of a crushed head [4]

Rail wear:

Rail wear indicates the loss of material from the running rail surface and from the side of the rail head due to the passage of wheels over the rail. It appears as a rounding of the running surface of the rail head, particularly on the gage side (Figure 2.29).

Figure 2.29: General appearance of rail wear

2.1.6 Defective welds

Figure 2.30: Defective weld

A defective weld (Figure 2.30) is a field or plant weld which contains any pockets or discontinuities exceeding either 5% of the rail head area individually or 10% in the aggregate. The origin can be in the rail head, web, or base, and in some cases cracks may progress from the defect into either or both adjoining rail ends.

Possible causes of defective welds are: lack of fusion between weld and rail end metal, incomplete penetration of the weld metal between the rail ends, underbead or other shrinkage cracking, entrapment of slag or sand, or fatigue cracking. Plant welds are caused by shearing processes used to remove excessive weld material. Both types of welds can also fail at an angular or oblique direction from an anomaly associated with the web area such as shear gouges, trapped oxides, or improper heat.

2.2 Current rail inspection methods

2.2.1 Historical background

The first rail inspection method ever used was the visual inspection. A sand mirror inspector could inspect one mile of rail per day. However, many external defects as well as visible indications of internal defects were easily missed with that approach.

The first patent showing an effort to detect rail defects through an automated system is dated 1877 [25]. The idea was to energize the rail with flux and then to pick up variations in residual magnetism.

Awareness of the danger caused by transverse fissures was raised as a result of a derailment at Manchester, NY in 1911, in which 29 persons were killed and over 60 seriously injured [4]. Private investigations were performed by most of the railroad owners to determine the prevalence of transverse fissures in their rails. The results of those investigations showed that transverse fissures were wide-spread.

In response to that derailment, the railroads of the country requested an aid from the U. S. Bureau of Standards in the development of an inspection method able to locate and measure the size of such defects in rails.

Dr. Elmer A. Sperry was the first to develop a dedicated rail inspection car to detect transverse fissures in railroad tracks. The method was based on electromagnetism. The car employed an induction method: a heavy current was induced through the rail to be tested via brushes in contact with the rail. Search coils were moved through the resultant magnetic field to find perturbations in this field due to the presence of defects [21].

One of Sperry's inspection cars detected a large transverse fissure and recommended the railroad owner to remove and replace the section of rail. Although annoyed by the recommendation, the section of rail was removed and taken to Dr. Sperry for further inspection. Once the rail was broken, the defect was shown and the ability of the rail inspection cars became evident. The railroads were convinced that the testing was worthwhile. By 1930 Sperry's service fleet numbered 10 cars.

Sperry's first ultrasonic testing car was deployed by 1959 for the New York City subway system. In the 1950s continuously welded track started to replace joint bars. This resulted in a shift of defect types found in rails, away from joint defects. In the following twenty years the average age of the rail naturally increased as well as the average load of the rail cars. Then, in the 1980s, with increased use of inspection cars through North America, the general trend was a lower defect detection rate. About twenty years ago typical rail life was 800 million gross tons, while today 1.5 billion gross tons is common.

In these days there are great pressures to increase traffic in order to gain efficiencies and therefore greater financial returns. This results in higher train speeds and heavier axle loads. Due to higher train speeds, there are shorter work windows for conducting inspections. At the same time, federal regulations dictate immediate remediation of detected defects, regardless of the type, size, or quantity of traffic that the rail carries.

Today ultrasonic inspection technology is the most widely used rail inspection system in North America. Most of the major contractors run rail testing cars solely based on ultrasonic inspection methods. Some other companies run units employing both ultrasonic and magnetic induction technologies. Currently, the ultrasonic rail testing cars

can operate at an average speed of 30 mph in a nonstop testing mode. However, in practice, the inspection vehicles average only 6 to 8 mph due to frequent stops to hand verify indications from the rail testing system.

2.2.2 Currently used methods

Currently, the Non Destructive Testing methods employed on railways around the world are based on:

- Visual inspection
- Contact induction technique inspection
- Non-contact induction technique inspection, based on eddy current
- Traditional ultrasonic inspection

Insights on these methods are given below.

The visual inspection method is still widely used. However, it is now becoming widely accepted that even surface cracking often cannot be seen by naked eye. Recently, automated visual inspection/detection systems have been developed by different research groups in the world. These automated visual inspection systems use state of the art computer vision and pattern recognition techniques to assess the condition of the rail surface. They can be used in parallel with other internal rail defect detection systems for a better assessment of the rail conditions.

The induction methods are based on the introduction of high-level direct currents in the rail head through brushes in contact with the rail surface in order to establish a magnetic field around the rail head (Figure 2.31). Once the magnetic field is established, it remains constant in strength and shape as long as rail head contour, rail weight and

current flow remain also constant. A specially designed group of sensors placed on the inspection car is then passed through the magnetic field [21]. When the vehicle starts running along the rail, the sensor unit starts moving through the magnetic field. If a distortion is detected in the concentric lines of force in the established magnetic field an electromotive force is induced, which is measurable as a voltage. In modern rail weights, only the head and the top part of the web are “saturated” with current. Unfortunately, the induction methods are quite sensitive to the presence of rail head surface anomalies. For these reasons, test results require significant evaluation by the operator to determine the presence of a defective condition.

Figure 2.31: Scheme of the contact induction technique

Recently, a non-contact induction technique based on eddy currents was developed [26]. With this method no brushes are needed to close the circuit with the rail. A magnetic core is used to induce eddy currents in the rail, which losses are correlated to the damage presence. In Figure 2.32 the orientation of the magnetic field from the eddy current sensor is shown. Once the magnetic field is induced transversely to the running direction of the rail, variations in the active and reactive components of the field are tracked. One of the largest limitations of the eddy current method is the limited lift off distance, usually set to 2-4 mm in order to achieve consistent measurements. If this

distance is employed, the depth of the field penetration is limited to 4-6 mm. This implies that only damages on the surface or close to it can be detected.

Figure 2.32: Eddy current sensor magnetic field pattern

A common drawback of all induction methods is the pronounced sensitivity to electromagnetic noise. Another disadvantage is the sensitivity to the presence of joints, bars and holes in the rail, resulting in false detections that need to be purged from the results.

Ultrasonic testing is the most common method used for rail inspections. It consists of scanning the rail head through ultrasonic beams and tracking the detection of the return of reflected or scattered energy using ultrasonic transducers [6]. A traditional pulse-echo method is employed, where standard ultrasonic piezo-electric transducers are mounted into fluid filled special wheels (Figure 2.33) or sleds. A coupling fluid, usually oil- or water- based, is needed to ensure proper transmission of the ultrasonic wave from the wheel to the rail.

Figure 2.33: Example of ultrasonic wheel testing system

Figure 2.34: Ultrasonic probe wheel transducers arrangement

Inside each wheel three to six transducers are usually mounted, with fixed orientations. In some cases, as many as 12 transducers are hosted per wheel or sled. Normally two wheels per rail are employed. A typical configuration (Figure 2.34) uses transducer orientations able to generate ultrasonic beams propagating at normal incidence (0°), 45° and 70° from the normal to the rail surface. The 45° and 70° probes are mounted

both in forward both in reverse directions for a complete scan of the rail. Snell's law dictates the transducers orientation angles for achieving the desired propagation directions. The 0° probe targets mostly horizontal cracks. The 45° probe is used to target discontinuities such as bolt hole cracks and defective welds. The 70° probe is particularly sensitive to transverse defects. Usually more than one 70° probes are employed together for a better coverage of the rail head area (Figure 2.35) [27]. The wheel also accommodates a side looking transducer (Figure 2.34) that is able to generate beams in the transverse plane of the track aimed to detection of discontinuities such as vertical split heads. The side looking transducer is typically oriented at 45° with respect to the vertical plane. The amplitude of the reflections and their arrival times indicate the presence, the location and the severity of the damage [28].

Figure 2.35: Ultrasonic beam coverage of three 70° probes (left) versus a single 70° probe (right) [27]

The ultrasonic detection systems may be either operated by pedestrians either mounted on a vehicle (Figure 2.36).

Figure 2.36: Ultrasonic method handheld [28] (left) and vehicle-mounted version (right)

Although ultrasonic testing is capable of inspecting the whole railhead, it has several drawbacks such as:

- RCF conditions, such as head checks, flaking and shelling, can mask underlying internal defects by reflecting the ultrasonic beams. It is partly for this reason that rail grinding techniques have been developed to remove certain types of RCF cracks [2].
- Typically the current systems are able to detect 65% to 80% of the actual defects. In fact, in addition to the masking of internal defects due to rail surface ones, ultrasonic wheels are also susceptible to surface conditions. Materials and debris on the rail surface have adverse effects to ultrasonic signals. Grease, icing and even leaves can completely mask the signals.
- Ultrasonic inspection car have limited speeds [2]. One of the limiting factors is represented by the needing to always ensure good coupling conditions between

the special-wheel and the rail top surface, which have to be realized through a continuous deployment of a “good” layer of coupling fluid; the problem is that as the speed increases the coupling conditions tend to get worse. A second limiting factor has to do with a tendency of ovalization of the fluid filled wheel due to increasing fluid internal pressure as the speed increases. The resulting test speed is usually slower than the line speed, compelling the inspection operation to be carried out outside the commercial track periods in order to avoid disruption of the normal train timetables. Theoretical operational speeds are between 25 and 45 mph. However, in practice, test speeds reduce to an average of less than 10 mph, because if damage is detected (correctly or wrongly) the inspection crew has to check immediately the nature and the severity of the damage.

- The last drawback is worsened by the high false defect detection rates. Current data reports that from 70% to 80% of the detected defects reveal to be false after hand test verification. This contributes to slowing down the inspection operations.

Figure 2.37 shows transverse defects detection rates based on flaw cross-sectional sizes as reported in literature [15]. Industry averages are compared to guidelines recommended by AREMA. Unfortunately, those recommendations do not specify the allowed false positive rates.

Other methods for rail inspection are currently under investigations, such as the *magnetic particle* and the *acoustic emission techniques* [29]. However, they both are in the early stages of research.

In the last decade, ultrasonic guided wave methods for rail defect detection have gained the attention of different NDT research groups around the world [30, 31, 32]. A detailed description on the use of ultrasonic guided waves for structural health monitoring of rails is presented in the next section.

Figure 2.37: TDs detection rates based on flaw cross-sectional sizes. Industry averages vs AREMA recommendations

2.3 Ultrasonic guided waves for defect detection in rails

Stress waves propagating through an elongated structure (such as pipe, rod, railroad track, beam, plate, wing skin, etc.) are constrained between its geometric boundaries and experience multiple reflections. The interaction between waves and boundaries of the waveguide produces continuous reflections, refractions and mode conversions, which in turns results in the generation of a complex mixture of constructive

and destructive interferences. As a result, so-called “guided waves” are generated. Compared to bulk waves, these waves can travel very long distances (even up to hundreds of meters) with little loss in energy, ensuring a complete coverage of the cross-section of the structure acting as waveguide (Figure 2.38). In practical terms this enables a rapid screening from a single transducer position and remote inspection of physically inaccessible areas of the structure, such as hidden structures under water, coatings, insulations, etc.

Figure 2.38: Comparison between traditional bulk waves and guided waves inspection (FBS Inc.)

From an economical point of view, the high cost effectiveness of guided wave-based techniques is related to large coverage areas and potential high testing speeds that are achievable in case of non-contact probing. However, the use of guided waves for SHM is usually quite challenging, due to their dispersive nature and to the existence of multiple modes propagating simultaneously. Because of dispersion phenomena, the energy in a wave-packet propagates at different speeds depending on the frequency

content. This results in a spreading of the wave-packet in space and time as it propagates through the structure, which in turns produces an increase in the temporal duration of the wave-packet and a decrease in amplitude. These effects are undesirable in long range guided wave testing because they reduce the resolution and the sensitivity of the testing system.

Sensitive mode-frequency combinations can be selected according to the specific application. For example, when testing a large structure, it is usually beneficial to excite modes in a non-dispersive region to avoid changes in shape of the wave packet as it propagates along the structure. Low dispersion can be achieved by testing at frequency-thickness products close to the maxima in the group velocity dispersion curves (Figure 2.39) [33, 34, 35]. In addition, it is possible to select modes more sensitive to the defect types of interest. The wave structures (Figure 2.40) give a good indication of the likely sensitivity to defects at different positions through the waveguide's cross-section.

Figure 2.39: Dispersion curves for a traction-free aluminum plate [36]

Figure 2.40: Wave structures for various points (fd = frequency*thickness) on the S_0 mode of an aluminum plate, showing the in-plane (u , solid line) and out-of-plane (w , dashed line) displacement profiles across the thickness of the plate [36]

In terms of excitation and/or sensing means, contact or non-contact guided wave methods can be employed. In the formers, the probe is in contact with the surface of the tested structure. Some coupling media is needed to ensure good contact conditions. Relatively inexpensive, unobtrusive, piezoelectric ceramics such as PZT [35] or piezoelectric composites like the MFC (Macro Fiber Composite) [37] or AFC (Active Fiber Composite) [38] transducers can be used. The variable-angle beam transducer and phased arrays are also common. A more sophisticated transducer system for generating guided waves is a “comb” transducer [39]. Comb transducers are made of bulk waves transducers (usually longitudinal waves) that are pasted on comb structures, and allow selection of individual modes in excitation. To this scope, transducer element size and spacing dimensions are chosen to set a desired wavelength, therefore exciting points along a straight line on the phase velocity dispersion curves diagram (Figure 2.41).

Figure 2.41: Comb transducer excitation of guided waves in a typical 115-lb AREMA rail [22]

While contact methods have the advantage of simplicity of operation, their application ranges are limited due to the needing of strict surfaces and coupling fluids. For example they are unsuitable for the inspection of curved or rough surfaces, complex structures or structures that prohibit the use of couplants.

Non-contact guided wave excitation is performed mainly using air-coupled transducers, laser or electromagnetic acoustic transducers (EMATs). The sensing can be achieved using a variety of different sensors, such as magnetostrictive type sensors [40], EMATs [41], capacitive [42] or piezoelectric air-coupled transducers [43].

Typically, the ultrasonic guided wave features employed for damage detection are related to wave energy or shape. Quantification of these features can be done within the time [44], frequency [45] or joint time-frequency domain [46, 47] of the reflected [48], diffracted [49] or transmitted waves [50].

Table 2.1 summarizes the features of the different techniques described in this chapter, pointing out their advantages and disadvantages.

Table 2.1: Advantages and disadvantages of different rail inspection methods

Inspection method	Advantages	Disadvantages
Visual inspection method	<ul style="list-style-type: none"> • Easy to perform • High speed if automated visual inspection is employed 	<ul style="list-style-type: none"> • Not reliable, large number of defects are missed
Contact induction method	<ul style="list-style-type: none"> • Reliable for surface defects detection 	<ul style="list-style-type: none"> • Limited testing speed • Unreliable for internal deep defect detection • Contact is required • High sensitivity to electromagnetic noise • Sensitivity to joints, bars and bolt holes
Non-contact induction method based on eddy current	<ul style="list-style-type: none"> • Reliable for surface defect detection • Non-contact deployment • High testing speed 	<ul style="list-style-type: none"> • Unreliable for internal deep defect detection • Small sensor lift-off • High sensitivity to electromagnetic noise • Sensitivity to joints, bars and bolt holes
Traditional ultrasonic testing	<ul style="list-style-type: none"> • Well established • Reliable for detection of many internal and surface defects 	<ul style="list-style-type: none"> • Difficulty of detecting transverse defects under shelling • Limited testing speed • Contact required
Ultrasonic guided waves based methods	<ul style="list-style-type: none"> • Potentially reliable for surface and internal defect detection • Potentially reliable for the detection of transverse defects covered by shelling • Non-contact deployment • High testing speed 	<ul style="list-style-type: none"> • Signal processing techniques required to increase SNR • Lower frequencies may reduce sensitivity to smaller defects • Reliable deployment needs to be demonstrated in the field

3 Concepts of the proposed rail defect detection system

3.1 Introduction

The initial objectives of this research were to develop a rail defect detection system that could allow higher speed of testing execution and with an increased reliability in terms of defect detection rate when compared to current rail inspection methods (which are described in Section 2.2.2).

A general description of the system that is proposed for the accomplishment of these tasks is presented in this chapter. The reasons behind the choice of some fundamental testing parameters, such as the types of means used for guided wave generation and detection, excitation frequency and sensors arrangement, will be discussed. Details about the current hardware layout of the prototype that has been developed, the filtering techniques that were analyzed and implemented, and the data processing/analysis algorithm that has been designed to enhance the defect detection are given in Chapter 5.

Numerical analyses and experimental testing were performed hand in hand over the different stages of prototype development. The formers have been instrumental for the understanding of peculiar aspects of the system and to pinpoint ways to increase its efficacy. Nevertheless, in this dissertation those two fields are separated and presented in two different chapters in the attempt to maintain a higher clarity in presenting the salient research findings. As such, this chapter represents a necessary introduction to either of those two chapters.

The salient aspects regarding the development of the dedicated software that handles in real-time the signal generation, data acquisition and data processing of the

prototype, which was deployed in National Instruments® LabVIEW® environment, will not be described/anticipated in this dissertation. Sincere thanks go to Thompson Nguyen who is the investigator who developed such software and who will have the honor to present the results of his work.

3.2 Former UCSD rail inspection system based on a hybrid laser/air-coupled configuration

A previous non-contact rail inspection system was developed by former students in the NDE&SHM research group of UCSD [15, 16, 22]. The method was based on hybrid laser transmission and air-coupled detection. The system featured a specialized statistical analysis of the data collected by the receivers in order to increase the defect detectability and to minimize false positives [18]. A prototype was built and field tested in various locations in the US achieving highly satisfactory results. A comprehensive report which discusses the design and the development of the system as well as the results of the seven field tests that were conducted can be found in [15]. The latest field test of June 2010, in particular, demonstrated an excellent reliability of internal rail flaws detection, which compared favorably to both American Railway Engineering and Maintenance-of-Way Association (AREMA) Standards and Industry Average performance (Figure 2.37).

The main drawback of that system was the presence of the laser, which is quite costly and potentially hard to maintain. Furthermore, the limited repetition rate of the laser excitation (maximum of 120 shots per second, due to duty cycle limitations for the

laser employed in that research) represented a probably insurmountable obstacle toward the increase of the testing speed to values greater than ~10 mph.

A natural solution to these issues was represented by the substitution of the laser with some other excitation mean. For this task, a piezoelectric air-coupled focused transmitter was employed, which is able to produce a somehow similar ultrasound excitation on the rail top surface. The obvious advantages of such choice were represented by:

- a significant reduction in overall cost of the system;
- the disappearance of any danger connected with the use of a highly powerful tool as a laser generator is;
- the ability to indefinitely increase the repetition rate of signals transmission;
- the easiness of operation and maintenance (large power supply and water-cooling system were needed to support the laser).

On the other hand, two significant drawbacks were introduced by the use of an air-coupled ultrasonic generation:

- the loss of a wideband excitation, with all the benefits that that involves (above all, the possibility of interrogating the specimen at different frequencies, leading to crack depth measurement). In fact, while a laser pulse carries substantial energy spanning over hundreds of kHz, the tone-burst produced by a piezoelectric air-coupled transducer is strictly narrowband;

- a considerable reduction of energy transmitted into the rail head, which makes the wave sensing quite challenging, if not impossible, without employing sophisticated filtering methods to raise up the SNR to acceptable levels.

Given the differences, and in particular the new challenges, brought by the replacement of the excitation mean, many aspects of the system had to be completely rethought, both in terms of hardware layout, software configuration and data analysis method to ensure robust defect detection.

3.3 Newly developed rail inspection system solely based on air-coupled transducers

3.3.1 Current transducers configuration

Figure 3.1 (side view) along with Figure 3.2 (top view) show a schematic of the proposed non-contact rail defect detection system solely based on air-coupled transducers. A cylindrically focused air-coupled piezoelectric transmitter sends a narrowband ultrasonic tone-burst which propagates through air and focuses its energy on a ~25mm-long line on the rail top surface. Once the air pressure wave impinges on the steel surface, part of its energy is effectively transmitted inside the rail head. Unfortunately, due to the large acoustic impedance mismatch between air and steel, most of the energy gets reflected back in air (and this is the cause of an important issue described in Section 5.7). The transmitted energy generates a complex wave propagation traveling symmetrically through the rail toward both rail running directions. Such propagation mostly remains confined in the rail head. Because the wavelength of such

ultrasonic waves (a discussion about the chosen excitation frequency is held before the end of this section) is of the same order of magnitude as the typical dimensions of some radii of curvatures of the rail head, they exhibit characteristics of “guided” waves, and in particular of Surface Acoustic Waves, due to constructive/destructive interferences of multiple surface modes (Section 4.2.3). The transmitter is held off-centered above the rail top surface (Figure 4.103) in order to produce an excitation line located toward the gage side of the rail, as shown in Figure 3.2. The reason behind this choice is explained later in this section. As those SAWs travel through the rail head, a small part of their energy leaks back into air at an angle, as measured from the vertical, of approximately $\sim 6.5^\circ$, which is dictated by Snell’s law (details are given in Section 4.3.3). To be precise, leakages will occur only from those modes whose structure is characterized by at least some out-of-plane particles motion, the only one capable of producing pressure waves in air; in fact, modes whose particles motion takes place entirely in-plane on the rail top surface will not produce any effect in air, which as any other fluid cannot carry shear type waves. Few air-coupled receivers (up to eight at the current stage of development) are placed symmetrically around the transmitter and are inclined in a way to best capture those leakages (Figure 3.1), thus effectively sensing the waves propagating in the rail top surface.

This whole process is repeated with a spatial frequency that can be set by the user as the prototype moves along the rail track. As a matter of example, 48 tone-bursts per foot, which equate in one test every 6.35 mm, was the resolution used during the field tests at TTC (the results of which are detailed in Section 5.8). For each testing point, if the conditions of the portion of rail comprised between the furthest receivers are pristine,

a symmetric propagation is expected (toward left and right directions around the transmitter as seen in Figure 3.1). Therefore, the signals captured by receivers symmetrically located on either side of the transmitter (in Figure 3.2 R_1 - R_8 , R_2 - R_7 , and so on) are (ideally) identical. When instead the test takes place in a defective portion of the rail head, complex wave/defect interactions arise, which modify the wave propagation on both sides. As a result, the signals acquired by each pair of symmetrically located receivers will present differences, both in terms of shapes and amplitudes. Features are extracted from the waveforms that quantify those differences, permitting to infer the presence of a defect in real-time. A good degree of redundancy is available considering that every time the test is repeated the portion of rail under test is roughly the one comprised between the furthest receivers deployed in the system. Therefore, if a defect is present in a specific point of the rail, all the measurements performed with the crack being located in between the furthest receivers could potentially reveal it. In reality, Section 4.2.4 shows how for each type of defect there exist only few specific relative positions of transmitter/defect/receivers for which the defect sensitivity is greatly enhanced in a way to be detectable on top of the measurements typical noise floor.

Let us consider, for example, how the waves would interact with the internal TD located at the gage side on the right of the transmitter in Figure 3.2. The two main effects of such interaction are:

- reflections traveling backward (toward the left in the figure) are produced, which will interfere with the waves propagating straight from the excitation, also traveling in the same direction;
- an attenuation of the waves past the defect location.

At the current stage of development, the prototype mainly relies on the attenuation effect to infer on the possible presence of defects. A sensor configuration specifically designed to target reflections coming back from defects was tested (Figure 3.5), but the results were not as satisfactory.

Theoretically the effects of wave/defect interactions are much more pronounced when the wave propagation direction is normal to the plane of development of the crack, and gradually diminish as this orthogonality condition worsens. This is intuitive: when the crack is orthogonal, it opposes a larger area of discontinuity to the propagation of the incoming waves. For this reason ultrasonic waves traveling only along the rail running direction would be insensitive to VSH defects, which usually originates toward the center of the rail and develops along the rail running direction as well. This explains the choice of using an off-centered excitation line as the one in Figure 3.2. As shown in figure, the wave components traveling diagonally from the excitation should provide good sensitivity to these types of defect.

By extension, this type of waves propagation should ensure sensitivity to cracks inclined at all the possible angles respect to the rail running direction, ranging from the orthogonal ones (as is the case of most TDs) to the parallel ones (which is the already mentioned case of VSHs). Furthermore, the detection should be ensured even when the crack does not lie on a vertical plane (i.e. is rotated around a horizontal axis), where the case limits is represented by HSHs, which are defects lying on horizontal planes. The last conclusion may not appear as intuitive, but arises from the fact that the fundamental modes generated by such excitation exhibit characteristics of Surface Acoustic Waves (as will be discussed further in Section 4.2.3), and these types of waves “feel” an acoustic

impedance mismatch as they propagate through a layer of reduced thickness (which is represented by the portion of steel comprised between the rail top surface and the discontinuity offered by the HSH). Thus, once again, complex interactions with the defect occur (details for a similar case in a composite plate can be found for example in [51]), which break the symmetry of wave propagations originated from the transmitter location.

The available air-coupled receivers are placed in a way to maximize their sensitivity to the various defects of interest. As Figure 3.2 suggests, the two pairs of receivers located toward the field side are expected to be the ones more sensitive to VSHs as well as to defects located toward the field side. The other two pairs, which are located between rail central axis and gage side, are expected to be the ones more sensitive to defects developed toward the gage side (which represent the majority of the cracks encountered in actual railroads).

As a result, the proposed system is capable of targeting at least all the three types of defects that were ranked as the first three causes of train accidents in the last decade according to FRA statistics (Figure 2.2), namely “Transverse fissures”, “Detail fractures” and “Vertical Split Heads”, with the addition of HSHs and possibly others. An extensive work based on numerical simulations was carried out in order to define the limits in terms of both size and depth of the defects that can be effectively detected by the current system when the available SNR level is taken into account. The salient results are presented in Section 4.2.4.

Figure 3.1: Schematic of the proposed system for rail defect detection (side view). The lift-off distance between rail top surface and transmitter is dictated by its focal distance (for the one currently in use is 76 mm). The lift-off distances of the receivers can be as high as 5 to 10 cm before incurring in significant SNRs deteriorations

Figure 3.2: Schematic of the proposed system for rail defect detection (top view). The drawing is not to scale. The two red lines indicate the hypothetical locations of a VSH (left) and a TD (right) internal cracks

It is important to stress that, based on the description above, the system is fully non-contact with respect to the rail, with all the benefits that follow. In particular, the potential to execute the tests at much higher speeds than the ones achievable with the currently widely used system based on the fluid filled wheel (details were given in Section 2.2.2), having removed all the speed limiting factors of that system. However, an

issue arose during recent experimental testing on the proposed air-coupled system that could potentially limit the achievable speeds. Details on the issue are given in Section 5.7. Ongoing efforts are devoted to overcome such limitation.

One question that might arise at this point regards the choice of using two groups of sensors on the two sides around the transmitter with the consequent extraction of features quantifying the relative differences between waveforms collected by the two groups instead of absolute variations in waveforms amplitudes, shapes, or others, as it is normally done in NDE methods based on ultrasonic transducers employed in pitch-catch configurations. The reasons behind this choice are based on the fact that the pitch-catch tests employed in this proposed rail defect detection system will have to be performed while the sensors travel at (potentially) high speed in harsh conditions as the ones encountered in railroad tracks. Therefore, there exist numerous environmental causes that can play a role in diminishing (or sometime increasing) the energy of the waves effectively transmitted from the air-coupled transmitter into the rail. Among those we can list:

- RCF conditions;
- debris or other external material on top of the rail partially blocking the ultrasonic beam;
- rough rail top surface;
- environmental conditions such as rain or snow;
- vibrations induced on the cart holding the transmitter that hence will experience movements from the set position at the focal distance from the rail top surface;

- different levels of rail wear, which changes the rail top surface geometry at the gage side where the ultrasonic beam impinges;
- changes of rail type under test (RE136, RE141, etc.) that modifies the guided waves propagation pattern;
- shot to shot variations of the energy released by the transmitter (although this seems to be of minor relevance, because after extensive testing on the air-coupled transmitter in use a great consistency of the released signals energy has been noticed).

Hence, in order to avoid numerous potential false alarms that would be caused by changes in the acquired waveforms solely due to the action of one of those causes, and by no means due to actual discontinuities in the rail, the idea of this relative comparison arose. Certainly this does not eliminate every possible source of false alarms, but plays a fundamental role in narrowing them down dramatically.

One last point to discuss regards the steps that brought to the choice of the specific narrowband frequency used to perform the tests. From one side, the chosen frequency could not be too low, so to stand well above the maximum frequencies of the typical noise produced by passing trains, which otherwise would interact with the generated signal in unpredictable ways. Works by Rose et al. [52] showed that the ultrasonic energy of passing trains is confined to frequencies up to about 100 KHz (Figure 3.3). Above those values the final choice on the used frequency resulted from the following compromise arising from different requirements based on defect detection standpoints. Considering the types of guided wave modes being generated in the rail head by the air-coupled transmission, which exhibit SAW modes characteristics, as a first

approximation the wave penetration is expected to be confined in an upper layer of the rail head until depths on the order of one wavelength, just as in the case of a pure Rayleigh mode. Therefore, a very high frequency wave would not be sensitive to internal defects whose upper tip is located well below the rail running surface. On the contrary, as the frequency decreases and the wavelength increases, so does the sensitivity to deeper defects. However, at the same time the sensitivity to smaller defects (the ones having dimensions of fractions of the wavelength) decreases as well. To some extents, this is a positive outcome, because an insensitivity to RCF conditions such as shelling, or simply to rough top surfaces, is certainly welcomed. However, when this starts to signify that defects of significant extensions are missed, it becomes a drawback.

The frequency that was chosen in the attempt to satisfy at once all of the above requirements is low enough to guarantee a sufficient insonification depth into the rail head. The approximate resulting enlightened area inside the rail head is shown in Figure 5.20. Most of the defects encountered in railroads lie, at least partially, in that area, so this should guarantee a good defect detection rate.

Figure 3.3: Frequency spectrum of noise detected by an accelerometer as a train approached it traveling at ~40 km/h [52]

3.3.2 Alternative transducers configurations tested over the course of the research project

During the different stages of prototype development, alternative sensors configurations were proposed and their possible benefits were evaluated. Similarly to the one currently in use, the configuration sketched in Figure 3.4 was aimed at targeting defects mostly exploiting their attenuation effects on the wave propagation. Here the excitation is located on the same side respect to the two groups of sensors, which allows for an inclination of the transmitter at approximately the same angle ($\sim 6.5^\circ$) as the one at which receivers are also inclined, with the advantage of increasing the energy effectively transmitted from air into the rail head (again, based on the physics described by Snell's law) which in turns increases the SNR of the acquired waveforms. This system worked in a similar way as the one described in the previous section: a baseline of waveforms acquired from pristine rail locations was stored and differences arising from this baseline were monitored to infer about the presence of defects. The portion of rail under test would be mainly the one comprised between the two groups of receivers, even if also when the defect is located between transmitter and the closest group of receivers the wave propagation field could be affected in a way to produce detectable relative modifications in the acquired waveforms. However, the gain in terms of SNR was not as high as it was initially hoped (probably in the order of 10%) while the resulting defect sensitivity worsened, as emerged from extensive experimental tests carried out in Camp Elliott (which is a research test facility described in Section 5.2), suggesting that this configuration is not as effective/reliable as the one shown in Figure 3.1 in terms of defect detection.

Figure 3.4: Alternative sensor configuration studied in the past. Attenuations due to defects was again the main effect targeted by this system (only one receiver per position is sketched to improve the clarity of the drawing; a group of 4 receivers was actually deployed in each receiver position)

The second alternative sensors configuration that was tested is sketched in Figure 3.5. The fundamental idea was to exploit the reflection effect resulting from waves/defects interaction. In this case the presence of a defect would be readily inferred every time a reflection is sensed by either of the deployed receivers. However, the results of the experimental tests performed using this configuration were not satisfactory. In fact, unfortunately, these reflections are usually not as powerful as needed for the receivers to sense them above the noise floor.

All the discussions and the results that are shared in this dissertation are solely based on the sensors configuration described in Section 3.3.1, which is the one more deeply investigated and the one currently in use, as it was found to be the most effective and reliable for rail defect detection purposes.

Figure 3.5: Alternative sensor configuration studied in the past. Reflections caused by defects was the main effect targeted by this system (only one receiver is sketched to improve the clarity of the drawing; a group of 4 receivers was actually deployed)

3.4 Conclusions

The general concepts behind the system which is proposed to accomplish high-speed rail defect detection with high reliability were detailed in this chapter. As it has been pointed out, the system can be seen as an upgrade over a previous prototype that was developed by former investigators in the NDE&SHM research group of UCSD. Specifically, the new system replaced the ultrasound generation laser with an air-coupled ultrasound transmitter, while retaining the air-coupled ultrasound receivers. Therefore the system is completely based on (non-contact) air-coupled transduction of the ultrasonic waves into and from the rail. The key advantages of the new system over the previous one are the potential for higher testing speeds, a significant cost reduction and an increased easiness of operation and maintenance. The main drawback is a significant

decrease in transmitted energy, for which a specially-designed statistical algorithm and very effective filtering techniques have been implemented, as will be described in Chapter 5.

The key aspects of the complex guided waves propagation phenomena occurring in rails during experimental testing have been discussed, and at the same time the system sensitivity to different defect types has been predicted.

The next chapter corroborates these predictions and discussions by showing results of numerical analyses that were performed using two different software/techniques, and that were confirmed by comparisons with actual experimental testing results.

3.5 Acknowledgments

Part of this chapter will be submitted for publication in *Transportation Research Record: Journal of the Transportation Research Board*, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Air-coupled wave propagation in steel rail”. The dissertation author will be the primary investigator and author of this paper.

4 Numerical modeling of ultrasonic guided wave propagation: theoretical fundamentals and application to the rail case

4.1 Introduction and theoretical fundamentals

Although ultrasonic testing experienced its greatest development in the late 19th and 20th centuries, our understanding of sound waves is much older. In 240 B.C., the Greek philosopher Chrysippus speculated that sound took the form of waves after observing waves in water. However, we need to wait until the late 16th and early 17th centuries for the development of the first laws governing sound by Galileo Galilei (the ‘‘Father of Acoustics’’) and Marin Mersenne. At the end of the 17th century, Sir Isaac Newton developed the first mathematical theory of sound, considering it as a series of pressure pulses transmitted between particles. Later expansion of Newton’s theory by Euler, Lagrange, and d’Alembert led to the development of the wave equation, which allowed sound waves to be represented mathematically [53].

A first theoretical description of guided wave propagation in linear elastic regime is due to Pochhammer and his work on the vibration phenomena occurring in semi-infinite elastic cylindrical waveguides, at the end of the 19th century [54]. A few years later, Chree treated the same problem independently [55]. However, detailed calculation of the roots did not appear until the middle of the 20th century, due to the complexity of the governing problem. Bancroft was the first to study the lowest branch of the roots of the Pochhammer-Chree frequency equation and to evaluate the relation between phase velocity and wavenumbers [56]. In the second half of the century, Gazis developed his theoretical analysis on the propagation phenomena for hollow, single layer, elastic

circular cylinders in vacuum [57], while Zemanek was one of the first authors to present a complete analytical and experimental study of the same phenomena [58].

Due to a lack of understanding of elastic wave phenomena during the first recorded earthquake seismograms in the early 1880s, Lord Rayleigh studied flat layered waveguides and derived the equation for waves traveling along the free surface of a semi-infinite elastic half space (Figure 4.1). He also showed how the energy of this particular type of waves decreases rapidly with depth and how the velocity of propagation is smaller than that of bulk waves [59]. At the beginning of the 20th century Love demonstrated that transverse modes were also possible in a half-space covered by a layer of finite thickness and different elastic properties, and that they propagate with a shear type motion in the plane of the layer (Figure 4.2). A generalization by Stoneley over the single interface problem studied by Rayleigh involved a study of interface waves propagating without leakage at the boundary between two solid half spaces [60].

Figure 4.1: Displacement field for Rayleigh wave

Figure 4.2: Displacement field for Love wave

At the beginning of the 20th century, a fundamental development of the wave propagation theory in flat waveguides was due to the works of Horace Lamb, who was the first to study in detail [61] the propagation of waves in infinite domains bounded by two surfaces (free plates). He derived two distinct expressions (Rayleigh-Lamb equations) whose roots represent symmetric and antisymmetric plate modes (Figure 4.3). Plotting these roots in the frequency domain, Lamb wave dispersion curves are obtained. The possibility of using Lamb waves for non-destructive testing was demonstrated by the experimental work of Worlton about fifty years later [62].

Figure 4.3: Displacement fields for Lamb waves. Symmetric mode (left) and antisymmetric mode (right)

In the next sections, the basic equations governing bulk and guided waves propagation are summarized for the purpose of reference.

4.1.1 Ultrasonic waves propagation in an unbounded medium

The development of the equation of motion for an elastic isotropic solid is a classical topic covered in many elasticity textbooks [63, 64].

A three-dimensional body having density ρ is considered. If body forces are neglected, according to Newton's second law the equations of motion are:

$$\begin{aligned}\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} &= \rho \frac{\partial^2 u_x}{\partial t^2} \\ \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yz}}{\partial z} &= \rho \frac{\partial^2 u_y}{\partial t^2} \\ \frac{\partial \sigma_z}{\partial z} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{xz}}{\partial x} &= \rho \frac{\partial^2 u_z}{\partial t^2}\end{aligned}\tag{4.1}$$

where u_x , u_y and u_z correspond to the displacement along the x , y and z Cartesian coordinates. These equations hold regardless of the stress-strain behavior of the medium.

The stress and strain in the elastic material are related by the generalized form of Hooke's law expressed as:

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{16} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{Bmatrix}\tag{4.2}$$

where C_{ij} represent the elastic constants of the material. For a material which is anisotropic, 21 independent elastic constants exist (C_{ij} matrix is symmetric). For an

isotropic solid, the number of independent elastic constants reduces to two, denoted as Lamé's constants λ and μ . The Lamé constants are directly related to the Young's modulus E and Poisson's ratio ν of the material.

The stress-strain relations for an isotropic medium can thereby be represented in tensor form as:

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij} \quad (4.3)$$

where δ_{ij} is the Kronecker delta and $\varepsilon_{kk} = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$ is the dilatation. Substitution of the stress-strain relation into the equations of motion described by eq. (4.1) leads to the Navier's equation of motion for an isotropic elastic medium:

$$(\lambda + \mu) \vec{\nabla}(\vec{\nabla} \cdot \vec{u}) + \mu \nabla^2 \vec{u} = \rho \frac{\partial^2 \vec{u}}{\partial t^2} \quad (4.4)$$

where $\vec{u}$ is the displacement vector, $\vec{\nabla}$ represents the divergence vector, while $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace differential operator. Based on Clebsch theorem, the displacement $\vec{u}$ can be decomposed into dilatation and rotation using the scalar and vector potentials ϕ and $\vec{\psi}$ such that:

$$\vec{u} = \vec{\nabla}\phi + \vec{\nabla} \times \vec{\psi} \quad \text{with} \quad \vec{\nabla} \cdot \vec{\psi} = 0 \quad (4.5)$$

If this expression for $\vec{u}$ is substituted into Navier's equation, the following expression is obtained:

$$\vec{\nabla} \left[(\lambda + 2\mu) \nabla^2 \phi - \rho \frac{\partial^2 \phi}{\partial t^2} \right] + \vec{\nabla} \times \left[\mu \nabla^2 \vec{\psi} - \rho \frac{\partial^2 \vec{\psi}}{\partial t^2} \right] = 0 \quad (4.6)$$

Equation (4.6) is satisfied only when both the scalar and vector potential vanish, which results in the following two decoupled wave equation:

$$c_L^2 \nabla^2 \phi = \frac{\partial^2 \phi}{\partial t^2}, \quad c_T^2 \nabla^2 \vec{\psi} = \frac{\partial^2 \vec{\psi}}{\partial t^2} \quad (4.7)$$

where:

$$c_L = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \quad c_T = \sqrt{\frac{\mu}{\rho}} \quad (4.8)$$

Therefore, two types of bulk waves exist. One is called longitudinal (or P) wave and travels with a constant speed c_L , while the other is called shear (or S) wave and travels with a constant speed c_T (Figure 4.4).

Figure 4.4: P wave (top) and S wave (bottom) propagation

4.1.2 Guided waves propagation in a bounded medium

A body elongated in one direction and having a cross-section of finite dimension represents a waveguide. This is also valid for a structure having two dimensions much larger than the third one. Typical examples of waveguides in structural mechanics include plates, rods, pipes, shells, strands, rails, etc. When stress waves interact with the

boundaries of a waveguide, multiple reflections and mode conversions take place giving rise to a complex mixture of constructive and destructive interference. After a certain distance from the source, the superposition of reflected, refracted and converted longitudinal and shear waves will form coherent wave packets i.e. ultrasonic guided waves (Figure 4.5). These waves are characterized by multimodal dispersive behavior, meaning that infinite modes coexist at any given frequency, which are characterized by frequency-dependent velocity and attenuation. The description of the three dimensional motion of an elastic waveguide is quite challenging even in the linear elastic regime, especially when dealing with complex geometries and/or material properties. Classical treatises on guided wave propagation in isotropic media can be found in literature [36, 65, 66, 67, 68].

Figure 4.5: Guided waves generation

Surface Acoustic Waves (SAWs) are guided waves that propagate along the surface of a medium. Examples of such waves are Rayleigh, Scholte, Stoneley, Love, and

creeping waves [68]. Lord Rayleigh [59] was the first to investigate upon such waves. He demonstrated that their amplitude decays rapidly with depth and that their velocity of propagation is smaller than that of bulk waves. For the purpose of reference, a general derivation of the equation governing surface wave propagation is furnished in the following (based on [36]).

A homogenous, isotropic, linear elastic half-space having a free surface is considered. Such medium follows Navier's equation of motion defined in eq. (4.4). As already seen, after a Helmholtz decomposition of the displacement vector into ϕ and $\vec{\psi}$, the equation of motion can be separated into dilatational and rotational waves, furnished in eq. (4.7).

Assuming a plane strain formulation, the displacement vector takes this form:

$$\vec{u} = (u_1, 0, u_3) \quad (4.9)$$

Figure 4.6: Guided wave propagation within a thin isotropic plate

Therefore, the function ϕ results to be:

$$\phi = \phi(x_1, x_3, t) \quad (4.10)$$

while ψ_1 and ψ_3 are either zero or constant due to $u_2 = 0$, given eq. (4.5). Thus:

$$\psi = \psi_2 = \psi(x_1, x_3, t) \quad (4.11)$$

where ψ_2 motion is only in the (x_1, x_3) -plane. Rewriting eq. (4.5) in Lamé form we obtain:

$$\begin{aligned} u_1 &= \phi_{,1} - \psi_{,3} \\ u_3 &= \phi_{,3} + \psi_{,1} \end{aligned} \quad (4.12)$$

Considering that we are seeking harmonic solutions for eq. (4.7), representing harmonic waves traveling in the x_1 direction, we can write:

$$\begin{aligned} \phi &= D_1(x_3)e^{i(kx_1 - \omega t)} \\ \psi &= D_2(x_3)e^{i(kx_1 - \omega t)} \end{aligned} \quad (4.13)$$

Upon substitution of eq. (4.13) into the dilatational wave equation of eq. (4.7) we obtain:

$$\nabla^2 \phi - \frac{1}{c_L^2} \frac{\partial^2 \phi}{\partial t^2} = \left[-k^2 D_1(x_3) + D_1(x_3)_{,33} + \frac{\omega^2}{c_L^2} D_1(x_3) \right] e^{i(kx_1 - \omega t)} = 0 \quad (4.14)$$

which can be simplified to:

$$D_1(x_3)\xi^2 = D_1(x_3)_{,33} \quad (4.15)$$

where:

$$\xi^2 = k^2 \left(1 - \frac{c^2}{c_L^2} \right) \quad (4.16)$$

A solution of the differential equation can be:

$$D_1(x_3) = A_1 e^{-\xi x_3} + A_2 e^{\xi x_3} \quad (4.17)$$

The second term of the solution is discarded, because it does not attenuate for increasing x_3 . In a similar fashion we can obtain the expression for ψ , yielding:

$$\begin{aligned} \phi &= A_1 e^{-kq x_3} e^{ik(x_1 - ct)} \\ \psi &= B_1 e^{-ks x_3} e^{ik(x_1 - ct)} \end{aligned} \quad (4.18)$$

where A_1 and B_1 are arbitrary constants and:

$$q = \sqrt{1 - \left(\frac{c}{c_L}\right)^2}, \quad s = \sqrt{1 - \left(\frac{c}{c_T}\right)^2}, \quad c = \frac{\omega}{k} \quad (4.19)$$

Substituting eq. (4.18) into eq. (4.12) we obtain the displacement expressions:

$$\begin{aligned} u_1 = \phi_{,1} - \psi_{,3} &= k(iA_1 e^{-kqx_3} + sB_1 e^{-ksx_3})e^{ik(x_1 - ct)} \\ u_3 = \phi_{,3} + \psi_{,1} &= k(-qA_1 e^{-kqx_3} + iB_1 e^{-ksx_3})e^{ik(x_1 - ct)} \end{aligned} \quad (4.20)$$

From here, in order to achieve the stresses in the material due to Rayleigh surface waves, one should use the constitutive eq. (4.3) valid for homogeneous isotropic material, after having computed the strain components as:

$$\begin{aligned} \varepsilon_{11} &= u_{1,1} \\ \varepsilon_{22} &= 0 \\ \varepsilon_{33} &= u_{3,3} \\ \varepsilon_{12} &= 0 \\ \varepsilon_{23} &= 0 \\ \varepsilon_{13} &= \frac{1}{2}(u_{1,3} + u_{3,1}) \end{aligned} \quad (4.21)$$

Once the stresses are obtained in terms of the non-zero strains in eq. (4.21), the following boundary conditions can be imposed:

$$\sigma_{33} = \sigma_{13} = 0 \quad \text{for } x_3 = 0 \quad (4.22)$$

These operations yield:

$$\begin{aligned} \sigma_{33}|_{x_3=0} &= rA_1 - 2isB_1 = 0 \\ \sigma_{13}|_{x_3=0} &= 2iqA_1 - rB_1 = 0 \end{aligned} \quad (4.23)$$

where $r = 2 - c^2/c_T^2$.

From eq. (4.23) the following relation between A_1 and B_1 is obtained:

$$A_1 = \frac{2is}{r} B_1 = \frac{ir}{2q} B_1 \quad (4.24)$$

Upon substitution of eq. (4.24) into eq. (4.20) we obtain the forms of u_1 and u_3 that satisfy the boundary conditions:

$$\begin{aligned} u_1 &= A[re^{-kqx_3} - 2sqe^{-ksx_3}]e^{ik(x_1-ct)} \\ u_3 &= iAq[re^{-kqx_3} - 2e^{-ksx_3}]e^{ik(x_1-ct)} \end{aligned} \quad (4.25)$$

where $A = -k B_1/2q$. Considering that u_1 and u_3 have to be real, being components of the displacement vector, we can rewrite:

$$\begin{aligned} \tilde{u}_1 &= \frac{u_1}{A} = (re^{-kqx_3} - 2sqe^{-ksx_3}) \cos k(x - ct) \\ \tilde{u}_3 &= \frac{u_3}{A} = q(-re^{-kqx_3} + 2e^{-ksx_3}) \sin k(x - ct) \end{aligned} \quad (4.26)$$

where $\tilde{u}_1$ and $\tilde{u}_3$ indicate normalized displacements with respect to the unknown constant A .

Furthermore, eq. (4.23) yields the following equation for the unknown phase velocity c :

$$r^2 - 4sq = 0 \quad (4.27)$$

The latter is the characteristic equation for this surface wave problem. A solution can be obtained introducing the new variables:

$$\eta = \frac{k_T}{k} = \frac{c}{c_T}, \quad \zeta = \frac{k_L}{k_T} = \frac{c_T}{c_L} \quad (4.28)$$

After few algebraic manipulations we obtain:

$$\eta^6 - 8\eta^4 + 8\eta^2(3 - 2\zeta^2) + 16(\zeta^2 - 1) = 0 \quad (4.29)$$

which is the Rayleigh wave velocity equation.

Considering that:

$$\zeta = \frac{c_T}{c_L} = \sqrt{\frac{1-2\nu}{2(1-\nu)}} \quad (4.30)$$

then we have three roots for η which are only function of Poisson's ratio ν . From each root $\eta = c/c_T$, a Rayleigh surface wave velocity $c = c_R$ can be obtained. However, only one of the three roots is real and realistic, therefore, only one Rayleigh surface velocity exists for each material, which only depends on ν . Note that c_R does not depend on frequency, thus Rayleigh surface waves are non-dispersive.

Viktorov [68] in 1967 furnished an approximate solution for η which is frequently used:

$$\eta \cong \frac{(0.87 + 1.12\nu)}{1 + \nu} \quad (4.31)$$

At this point let's go back to analyze the expressions for the normalized displacement components $\tilde{u}_1$ and $\tilde{u}_3$ (eq. (4.26)). After a simple algebraic manipulation follows that for any value of x_3 the vector of coordinates $(\tilde{u}_1, \tilde{u}_3)$ moves along an ellipse (Figure 4.7). Further analysis reveals that this elliptical particle motion is a function of depth, showing that the motion direction reverses below a particular depth. Figure 4.8 shows how the displacement components vary with depth.

Figure 4.7: Rayleigh wave elliptical particle motion [36]

Figure 4.8: Normalized displacements for a Rayleigh surface wave as a function of depth in an aluminum half-space. Curves normalized with respect to u_3 value at $x_3=0$. λ_R is the Rayleigh surface wavelength [36]

4.2 Local Interaction Simulation Approach (LISA)

4.2.1 Introduction

Wave propagation analyses can be performed using a number of different numerical computational techniques. Among these we recall: boundary elements methods [69], pseudospectral element methods [70] such as “k-Wave” [71], global matrix approaches [72], finite strip element methods [73], finite difference methods, Semi-Analytical Finite Elements methods (SAFE) [22, 74], mass-spring lattice models [75], Finite Elements (FE) analysis [76] and the Local Interaction Simulation Approach (LISA), which is a variant of the finite difference method.

A large part of the numerical analyses performed over the course of this research project were carried out using *cuLISA3D*[®], a software that implements the LISA method and that was recently developed at the Department of Robotics and Mechatronics at AGH University of Science and Technology, Poland. The software is now commercialized from MONIT SHM[®], a company based in Krakow, Poland. The system is based on the MATLAB[®] environment and the CUDA technology offered by modern graphic processors. Thanks to the latter, complex models can be computed in a very fast manner.

In parallel with the former, a commercial finite element method software (*LS-DYNA*[®]) was used to simulate the propagation of ultrasonic waves also in the air domain surrounding the rail structure.

The Local Interaction Simulation Approach was first presented by Delsanto in the early 1990s [77, 78, 79]. His goal was to develop a computer simulation technique based on parallel processing able to describe ultrasonic wave propagation phenomena in media of arbitrary complexity and with sharp impedance changes.

LISA algorithm is a powerful tool when modeling wave interaction with structural damage is required. In fact, it can reduce if not avoid numerical errors associated with other techniques when the model contains boundaries or sharp discontinuities [80].

The algorithm is derived from the elasto-dynamic wave equations for homogeneous media given in eq. (4.4). The examined structure is discretized into a grid of cubic cells having edge length c , while the time is subdivided into discrete time steps of size τ . The material properties are assumed to be constant within each cell, but may vary between cells. After the elasto-dynamic equations are discretized (details in [81]) a series of iterative equations are obtained having the form:

$$D^{t+1} = f(S^D \subseteq \{D^{t'} : t' \leq t\}) \quad (4.32)$$

where D denotes a vector of field variables and t is time. The algorithm can be parallelized very efficiently considering that the future time step is calculated as a combination of the same quantities already calculated in the previous time steps. Therefore, parallel computation technology offered by modern Graphics Processing Units (GPUs) and Compute Unified Device Architecture (CUDA) used in low-cost graphical cards was used for computation of the LISA equations. More details about the implementation of such equations can be found for example in [80, 82, 83].

Since LISA belongs to the finite difference group of methods, it inherits numerical limitations of this group. The most important of these is the Courant–Friedrichs–Lewy stability condition, defined as:

$$\frac{c}{\tau} \geq \sqrt{v_L^2 + v_T^2} \quad (4.33)$$

where v_L and v_T are longitudinal and shear bulk wave velocity, respectively. Eq. (4.33) must be fulfilled for stable simulation.

Another limitation, which is typical of most numerical methods, is the requirement of about 10 grid points per wavelength in order to achieve an acceptable accuracy [84]. Some studies recommend even a more stringent condition of 20 nodes per wavelength [85].

4.2.2 Proof of concept of the proposed rail defect detection system

A first set of simulations was performed using an older version of cuLISA3D software, namely version 0.8.5.0, while a newer version, the 0.8.5.1, was used to perform a second, more extensive, set of simulations (described in Section 4.2.3 and Section 4.2.4). The goal of this first set of simulations was mainly to produce a proof of concept of the proposed rail defect detection system to demonstrate its feasibility.

An AREMA 136RE rail (Figure 4.9), which is a type of rail commonly used in the US railway system, was modeled and discretized into cubic cells. The length of the modeled portion of rail was set to 40 cm, a value chosen in order to avoid spurious ends of model reflections during the desired simulation timeframe. Furthermore, only the rail head and part of the web were modeled, in order to reduce the computational effort (Figure 4.10). Some unwanted small reflections (of bulk waves) from the bottom end of the modeled rail portion appear during the simulation timeframes. However, those unwanted small contributions arrive only toward the end of the wave-packet of interest (constituted by the first wave-packet arrival of surface waves), hence they do not significantly influence the results that are presented in this section.

Figure 4.9: AREMA 136RE railroad track geometrical details

Figure 4.10: Modeled portion of rail 136RE

The rail steel material was modeled as an isotropic medium having density ρ of 7,850 kg/m³, elastic modulus E of 200 GPa and Poisson's ratio ν of 0.29. Based on such material properties, the longitudinal and shear bulk wave velocities as well as the Rayleigh velocity can be calculated as:

$$\begin{aligned}
 c_L &= \sqrt{\frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)}} = 5,778 \text{ m/s} \\
 c_T &= \sqrt{\frac{E}{2\rho(1+\nu)}} = 3,142 \text{ m/s} \\
 c_R &\cong \frac{0.87 + 1.12\nu}{1 + \nu} c_T = 2,911 \text{ m/s}
 \end{aligned} \tag{4.34}$$

Temporal discretization and cubic cell size for the LISA algorithm were set in order to satisfy the stability condition expressed in eq. (4.33) as well as the accuracy condition. The cell size was set to 0.5 mm, resulting in ~13 million cells in the model, while the time step was set to 0.05 μ sec. The accuracy condition is largely fulfilled, offering more than 30 elements for the smallest excited wavelength. The choice of such small cells size was moved by the needing to model the curved part of the model with a good approximation, especially in the definition of the defects shapes. On this topic, the results of a study quantifying the error introduced by the so-called “staircase” problem in LISA models can be find for example in [86].

Because the propagation of waves in air was not included in the model, a simplified modeling of wave excitation was used. In the attempt to simulate the excitation exercised by the focused air-coupled transmitter used in the experimental prototype, a set of nodes along a 25mm-long line on the rail top surface was selected and excited by a

time-history displacement as the one shown in Figure 4.11 normally to the rail top surface. Based upon the results of an experimental test carried out to characterize the pressure field generated by the available air-coupled transmitter at its focal distance, which are shown toward the end of Section 4.3.2 (right plot in Figure 4.84), the amplitudes of the signals assigned to each point along the 25mm-long excitation line were scaled according to the values shown in Figure 4.12.

Figure 4.11: Transmitted signal

Figure 4.12: Amplitude modulation along the excited line

The tone-burst signal used in the experimental prototype is a 10-cycles sinusoid centered at the resonant frequency of the air-coupled transmitter. Although the electrical

signal delivered to the transducer is not windowed (and this is in order to maximize the energy effectively released in air by the transmitter), the response of the transmitter itself produces a pressure wave with a window-like shape. For this reason, the time-histories assigned to the nodes of the model acting as actuators were modulated using a Hanning window (Figure 4.11).

Two defects were simulated in order to study the sensitivity of the excited guided waves to internal structural damage. Therefore, the results from three models are presented. In the first model the rail is in pristine conditions. In each of the other two models, a transverse defect was simulated having area equal to $\sim 10\%$ and $\sim 30\%$ of the rail Head Area (% H.A.), respectively. Both defects are located in the rail gage side (the one where rail wheels run by). The crack opening was set as the minimum possible when using this simulation method, i.e. one cube edge length (0.5 mm). Section cuts of the two defective models in correspondence of the TDs are depicted in Figure 4.13. The defects are approximately 1 cm away from the excitation signal.

Figure 4.13: Section cuts of the LISA models of the 136RE rail with 10% H.A. transverse defect (on the left) and 30% H.A. transverse defect (on the right)

The calculations were performed with a PC featuring an Intel® Core™ i7-3930K @ 3.20 GHz processor, 32 GB of RAM and NVIDIA GeForce GTX 670 GPU. Double precision 64-bit floating point numbers were used. Each simulation was iterated until 120 μ sec, which resulted in only 6 minutes of analysis time.

Figure 4.14 shows how the wave propagation is strongly influenced by the presence of a discontinuity in the rail head. In particular, the reduced transmission coefficient due to presence of defects is the main effect that is measured by the experimental prototype. The screenshots show a top view of the vertical displacement field pertaining to each of the three models after 60 μ sec of simulation. The excitation line is placed in the middle of each rail model, all the way toward the gage side, to mimic actual experimental testing conditions. The reason for such asymmetric placement is discussed in Section 3.3.1. A picture of the colorbar used to color the figures is not available in cuLISA3D software, however, the highest values of vertical displacements are characterized by hot colors (red being the maximum), while the lowest values by cold colors (blue being the minimum). In the screenshots, the defects are placed on the left of the excitation line, and their presence strongly influenced the wave propagation, due to scattering of the incident waves. In both defective models, the amplitude of the wave travelling toward the left side significantly drops with respect to the non-defective model case. The waves travelling toward the right results modified as well, due to interferences with reflections from the crack which also travel toward the right side. Plots of the vertical displacement histories experienced by the same node on each of the three models were generated in order to have a better quantitative understanding of the defect influence on the signals that can be received by the air-coupled receivers. For example, Figure 4.15

shows the waveforms recorded by the point lying at the center of the rail top rail surface and located 10 cm away to the left of the excitation line. The signal energy drops significantly when the ~10% H.A. defect is present, and it gets even weaker as the defect size grows to ~30% H.A.. Root Mean Square values computed from each waveform are shown in the legend of each subplot. As already discussed, it is this reduced transmission coefficient due to the defect scattering that is used in the experimental prototype to assess the rail condition.

Figure 4.14: Vertical displacement field (in color scale) pertaining to the rail model (a) without defect, (b) with 10% H.A. transverse defect and (c) with 30% H.A. transverse defect after 60 μ sec of simulation. Views from above the rail top surface

Figure 4.15: Vertical displacements at the top of the railhead for the rail model (a) without defect, (b) with 10% H.A. TD and (c) with 30% H.A. TD

4.2.3 Distribution of wave energy on the rail top surface

A second, more extensive, set of simulations was performed using an updated version of cuLISA3D software, namely version 0.8.5.1, and the salient results of those analyses are presented in this section and in the next one. As this upgraded version of cuLISA3D was capable of handling bigger models than the previous one, for these simulations the whole cross-section of a 136RE rail was modeled and the length of the modeled rail was extended to 81 cm, which allowed the execution of all the simulations presented in this dissertation without interferences of unwanted waves reflected by the ends of the model within the simulations timeframes.

The same steel material properties used in the models described in the previous section were utilized: $\rho = 7,850 \text{ kg/m}^3$, $E = 200 \text{ GPa}$ and $\nu = 0.29$. Cells size and analysis time step are also the same: 0.5 mm (resulting in ~61 million cubes) and 0.05 μsec , respectively.

The excitation exercised by the air-coupled transmitter was simulated exciting a set of nodes lying on an $25 \times 25 \text{ mm}$ area of the rail top surface by a time-history displacement as the one shown in Figure 4.11 (Hanning-windowed 10-cycles tone-burst), thus in a similar way as for the simulations described in the previous section. However, in this case the amplitudes of the signals assigned to each node in the area were modulated (scaled) according to the results of the LS-DYNA simulations described in Section 4.3.2. In particular, this modulation is based on the simulated pressure field generated by the air-coupled transmitter at its focal distance, namely 38 mm, which is showed in Figure 4.16 (these two sub-plots are extracted from Figure 4.76 and Figure 4.77). This should

guarantee a very accurate simulation of the real excitation produced in rails by the actual prototype.

Figure 4.16: Top view (left) and 3D view (right) of the contour plot showing the pressure field produced by the modeled air-coupled transmitter at its focal distance

Figure 4.17 and Figure 4.18 show the vertical displacement fields as contour plots for different instants in time. Plot (a) shows the field at 1.5 μsec , corresponding to the time at which there is a maximum of the first cycle of the excitation (according to Figure 4.11). A zoomed-in view toward the excitation area (not shown here for the sake of brevity) showed the same contour plot as the one in Figure 4.16, confirming the successful implementation of the LS-DYNA results into the LISA model. The other plots reveal that surface acoustic waves generated by the air-coupled transmitter will propagate mainly along the rail top surface, but in every other direction as well. Nevertheless, the energy of interest for the purpose of the proposed prototype is only the one that produces out-of-plane (vertical) displacement on the top surface, being the only one that, after leaking in air, can be sensed by the air-coupled receivers. While those plots are already useful to understand where the energy tends to focus the most, the next paragraph takes one more step further in the process of determining the best positions to place the air-coupled receivers used in the prototype.

Figure 4.17: Contour plot of vertical displacement field after 1.5 μsec (a) and 31 μsec (b), respectively

Figure 4.18: Contour plot of vertical displacement field after 60 μsec (c) and 90 μsec (d), respectively

Contour plots of dedicated features to find “hot spots” in the rail top surface

As was briefly discussed in Section 3.3.1, when the energy released by the air-coupled transmitter is introduced in the rail through the rail top surface, a complex wave propagation phenomenon showing characteristics of guided SAWs takes place (mostly) in the rail head. A simplified interpretation of the signals sensed by the receivers as single Rayleigh wave arrivals instead of as a superposition of multiple surface wave modes would lead to wrong analyses. The same conclusions were drawn by the authors of [87], where the existence of multiple surface wave modes in rails at similar frequencies as the ones excited by the proposed system were demonstrated by plotting the maximum of the signal envelopes measured using a laser vibrometer at different distances from their narrowband excitation source. In fact, while a gradually decaying amplitude is expected for a single Rayleigh wave mode, their plot showed an heavy modulation that could be only caused by constructive and destructive interferences of multiple modes. A similar plot was obtained by Coccia [22] using a more broadband excitation source, a laser, in pair with an air-coupled receiver.

Based on a similar idea, simulations were performed using the cuLISA3D model described above to understand where, in the rail top surface, the energy of the waves excited by the prototype would focus the most. This provides useful information in terms of understanding where would be more indicated to place the receivers deployed in the prototype. In fact, due to the low available SNRs typical of any air-coupled ultrasound transduction method in steel, it would be ideal to place the receivers where they can sense the highest possible energy leaking from the rail. However, this does not strictly mean that those positions are also the most effective ones in terms of defects-sensitivity. This

topic will be briefly addressed before the end of this section, and more extensively in the next one.

To obtain the plots that will be shown below, a grid of nodes were selected on the rail top surface of the model, and the vertical displacement experienced by those nodes were recorded over the duration of the simulation, which was fixed to 200 μsec . The area covered by those nodes is a rectangle of 40×5 cm. Such rectangle is centered on top of the rail (when considering the transversal direction) and starts from the excitation (to be precise, from the center of the excitation area, corresponding to a vertical line at 12.5 mm in the left sub-plot of Figure 4.16). To save computational effort, only one every four nodes was selected (along both planar directions), thus there is one measurement point every 2 mm in both planar directions. Figure 4.19 helps to visualize the extents of the area being considered on the rail top surface. The rail shown in the figure is a superimposed CAD model, i.e. is not the actual cuLISA3D model used to perform these simulations. In the bottom plot are also shown (in black rectangles) the positions where the four pairs of receivers were deployed during the field tests at TTC facility in Colorado. The energy plot shown in figure is the one corresponding to the pristine model, and is the same as the one shown in Figure 4.20, which will be commented below. Due to the bi-directionality and symmetry of the excited wave propagation phenomena, only nodes on one side of the transmitter were monitored, as correctly shown in the bottom plot of Figure 4.19. The top plot was simply made by mirroring the available side around the center of the excitation with the only intent to provide a complete view of the generated wave propagation phenomenon in rail.

Figure 4.19: 3D views that help to visualize where the energy plots shown in the next figures are located in the rail. Due to the symmetric bi-directionality of the excitation, nodes were selected and monitored on only one side around the transmitter, as correctly shown in the bottom plot. The top plot was made by mirroring the available side with respect to the center of the excitation

Once the waveforms acquired by each node are available, different features can be extracted and plotted as contour plots. The figures shared in this dissertation show

contour plots of the peak-to-peak amplitude of the first wave-packet arriving at each sensor location, which is the same wave-packet monitored and considered by the rail defect detection prototype to perform its analyses. Similar plots were obtained after computing the RMS of the same wave-packet arriving at each location; however, the results are very similar, hence those plots are not shown here, for the sake of brevity.

The peak-to-peak amplitude distribution is plotted in Figure 4.20, where dimensions and colorbar are also given. This is the same plot as the ones super-imposed with the rail CAD model in previous Figure 4.19, where the correct/realistic aspect ratio was shown. Once again, the black rectangles try to approximate the actual dimensions of the active surface of the receivers, which are actually 19×19 mm squares, looking like elongated rectangles because of the distorted aspect ratio used in figure. In the post-processing of the signals collected during the Colorado field test it turned out that one of the four pairs of receivers was affected by a very low SNR, thus that pair was not used, i.e. was not fed into the algorithm that performs the rail conditions assessment. This pair is the one constituted by the receivers placed as the top right one in Figure 4.20, which in fact is located in a very poorly “energized” area of the rail top surface. The other three pairs were instead characterized by good SNRs, as correctly predicted by these simulation results. In fact, very good agreements were found after comparing these results with experimental testing results, as will be reported in the next Section 4.2.4.2.

Figure 4.20: Peak-to-peak amplitude distribution on the rail top surface for pristine rail conditions. The black rectangles approximate the sensors positions during the field tests in Colorado. The aspect ratio is not exact, the plot is stretched along the y-dimension

Similar plots were obtained using models where defects were simulated. For example, Figure 4.21 and Figure 4.22 show the contour plots of the same feature considered in Figure 4.20 (the colorbar, not shown here, is also the same) for a model including a TD on the gage side. The area of the simulated crack (shaped as a circle) in the rail cross-section is 25% of H.A.. Its depth as measured considering its higher tip is 2 mm. More details about the modeling of this crack and of many others are given in the next section (this defect is the one denoted as “model number 5” in Table 4.1). The two figures show the results obtained from six different simulations, one where the excitation was 3.5 cm away from the crack, the others where this distance was increased by 3.5 cm each time. The goal of this set of simulations was to understand if there exist specific relative locations for the ensemble transmitter/defect/receiver for which the defect sensitivity is enhanced. From these sub-plots, it can be seen that when the transmitter is

very close to the crack (sub-plot (a), 3.5 cm away), the latter will block and reflect back a big portion of the incoming wave energy; thus, it seems that for this defect type the detection would be greatly enhanced when the transmission occurs at close distance to the crack (this behavior will find confirmations in the analyses results shown in next section, in particular in Figure 4.39). Then, as the transmitter gets further, the waves will have more possibilities to pass over the defect location (and thus to energize the area behind it) because other paths are less affected by the discontinuity. For example, for this defect case, the wave energy traveling up toward the field side will find soon an almost undisturbed path. One more interesting detail to notice, it appears that after the first wave arrivals impinge on the defect, their reflections will interact with further incoming waves, giving rise to the formations of those scattered highly energetic thin areas seen as dark red colors for example at about 12 cm in sub-plot (d).

Figure 4.21: Peak-to-peak amplitude distributions on the rail top surface for model with 25% H.A. TD on the gage side as the transmitter gets further away from the defect (3.5(a), 7(b), 10.5(c) cm). The aspect ratio is distorted: the plot is stretched along the y-dimension (even if less stretched than Figure 4.20, which can generate some confusion in the interpretation). Colorbar is as in Figure 4.20

Figure 4.22: Peak-to-peak amplitude distributions on the rail top surface for model with 25% H.A. TD on the gage side as the transmitter gets further away from the defect (14(d), 17.5(e), 21(f) cm). The aspect ratio is distorted: the plot is stretched along the y-dimension (even if less stretched than Figure 4.20, which can generate some confusion in the interpretation). Colorbar is as in Figure 4.20

2D-FFT plot confirming the excitation of multiple guided wave modes

Using the same data utilized to obtain the plot in Figure 4.20, i.e. the waveforms recorded in correspondence of a 2 by 2 mm grid of nodes on the rail top surface, it is possible to produce two-dimensional fast Fourier transform (2D-FFT) plots, following a procedure originally proposed by Alleyne and Cawley in [84], to confirm the presence of multiple surface wave modes that are excited by the air-coupled transmitter and that propagate along the rail running direction. In this case, due to the asymmetric excitation which is produced in the rail top surface, there is some freedom in the choice of which line (along the rail running direction) to consider as the main propagation direction.

For example, after picking a line that passes approximately through the middle of the excitation segment (16 mm apart from the rail central axis, going toward the gage side) the resulting 2D-FFT plot is shown in Figure 4.23. In figure, the normalized modulus of the result is plotted as grayscale against wavenumber and frequency. The presence of different modes (revealed by the gray patches) exhibiting different dispersion relations in terms of frequency vs wavenumber is confirmed. Two highly pronounced regions are found (the two more toward the right side of the plot), which correspond to signals carrying the main part of the energy; other excited areas at lower wavenumbers carrying lower energies are also found. The energy is obviously very focused in a narrow frequency region, thus only a small area of the dispersion curves is visible. At least 6 or 7 modes are discernible; however, there might be more, hidden by possible overlaps due to resolution limits.

Figure 4.23: 2D-FFT from the same cuLISA3D simulation results used to obtain the plot in Figure 4.20. A line being 16 mm apart (toward the gage side) from the rail top surface central axis was the one considered to produce this plot

4.2.4 Parametric study of the prototype sensitivity to different defect conditions

Based on the same model and parameters used for the simulations described in the last section, an extensive series of studies were performed to characterize the sensitivity of the proposed system to different types/sizes/depths of defects. Aside from the 81cm-long pristine rail model, which in the following will be denoted as “model number 0”, Table 4.1 through Table 4.4 give a list of all the models created by simulating a defect inside “model number 0”. As it was done for the TDs simulated in Section 4.2.2, each crack has been created by removing one layer of cubic cells in correspondence of the

desired defect shape (therefore the crack opening is 0.5 mm, which is the minimum allowed by a LISA algorithm). The TDs were shaped as circles, while VSHs and HSHs as rectangles. The characteristic dimensions of each simulated defect are given in the tables.

The types of defects considered in this study covers basically all the possible rail head internal cracks encountered in actual railroads. However, each simulated defect lies purely in either a vertical (TD, VSH) or horizontal (HSH) plane. Other intermediate inclinations were not considered in this set of models.

Based on the relative positions of transmitter and receivers as used in the current version of the prototype, the idea was to simulate the outcomes (i.e. the Damage Index trace) of an actual test where the prototype is moved along the rail track and repeatedly performs a test every ' x ' mm, where ' x ' can be set by the operator (for example it was set to 6.35 mm in the Colorado field tests). In order to obtain such results, for each model/defect a series of different simulations have to be performed where the ensemble receivers/transmitter is sequentially moved with respect to the crack which instead is held in the same position. Each simulation has to be carried out for a duration long enough to allow the generated waves to travel toward the defect, interact with it, and then travel toward the receivers. Considering that the relative distances between transmitter and closer/further receivers are 14 and 18 cm respectively (Figure 3.2-Figure 4.20), the end of each simulation timeframe was set conservatively to 150 μ sec (the first wave-packet, which is the one currently monitored by the prototype's algorithm, will transit by the furthest receivers roughly between 60 and 120 μ sec of simulation).

In theory, a defect can be detected by the prototype only when it lies in a position in between the furthest deployed symmetric pairs of receiver (or, at most, when it is right past the outer receivers, due to reflections that can possibly modify enough the wave-packet monitored by the system). However, in order to save computational effort, the distance considered for the setup of these simulations is the one in between the inner pairs of receivers, which is $14 \times 2 = 28$ cm. Furthermore, due to symmetry around the excitation location, this distance can be cut in half, resulting in a 14cm-long portion of the modeled rail along which the ensemble transmitter/receivers has to be moved while maintaining each defect in a fixed position to study its detectability (i.e. to plot the resulting Damage Index trace). In particular, starting with the transmitter being right on top of the considered defect (let us call this “position 0 cm”), simulations were performed as it was sequentially moved by consecutive steps of 0.5 cm (10 cubes of the LISA model) from “position 0 cm” to “position 14 cm” (where the inner receivers are on top of the defect). This results in 29 different simulations to perform in order to characterize each defective-model. Thanks to the already mentioned symmetry around the excitation, the 28 simulations corresponding to the transmitter going in the opposite direction from “position -0.5 cm” to “position -14 cm” are not necessary. To clarify, in order to maintain conditions of symmetry for the VSH and HSH defects, which have a longitudinal extension along the rail running direction, “position 0 cm” is the one where the transmitter is located right at the middle of such extension.

In order to ensure that spurious end of model reflections did not interfere with the wave arrivals at the receivers locations during the simulations timeframe, “position 0 cm” (which in other words is the position at which each defect was simulated in its respective

model) was set 33.5 cm away from one end of the model. After a little math, considering all the different excitation points, this equates to “worst case scenarios” minimum distances of 15.5 cm between each end of the model and the receivers furthest away from the transmitter (understandably, the two worst case scenarios are “position 0 cm” and “position 14 cm”). Because surface acoustic waves traveling at about 3,000 m/s would need slightly more than 150 μ sec to cover the path “*transmitter – end of the model – receivers*” in the worst case scenario, this ensures that unwanted reflections do not arrive at the simulated receivers locations during the chosen simulation timeframe (150 μ sec).

The simulations were performed with the same PC described in Section 4.2.2 (Intel® Core™ i7-3930K @ 3.20 GHz processor, 32 GB of RAM and NVIDIA GeForce GTX 670 GPU). Single precision 32-bit floating point numbers were used (before proceeding with 32-bit numbers, validation tests were performed to verify that results obtained using single precision did not yield appreciable differences when compared to the ones obtained using double precision). Each analysis required ~15 minutes to iterate the wave propagation until 150 μ sec, which translates in ~7 hours to perform the 29 simulations required to fully characterize each defect/model.

Table 4.1: Map of models including TD defects on the gage side

Model Nr.	Type	Cross-sectional dimension (% of H.A.)	Depth (measured at the highest tip, in mm)	Figure
Study of the influence of the cross-sectional dimensions (depth held constant)				
1	TD (gage)	5	2	Figure 4.26
2	TD (gage)	10	2	Figure 4.26
3	TD (gage)	15	2	Figure 4.26
4	TD (gage)	20	2	-
5	TD (gage)	25	2	-
6	TD (gage)	30	2	Figure 4.26
7	TD (gage)	35	2	-
8	TD (gage)	40	2	Figure 4.26
9	TD (gage)	45	2	-
10	TD (gage)	50	2	Figure 4.26
Study of the influence of the depth (on a TD 10% H.A.)				
2	TD (gage)	10	2	Figure 4.25
11	TD (gage)	10	7	-
12	TD (gage)	10	12	Figure 4.25
13	TD (gage)	10	17	-
14	TD (gage)	10	22	Figure 4.25
Study of the influence of the depth (on a TD 25% H.A.)				
15	TD (gage)	25	1	Figure 4.27
5	TD (gage)	25	2	-
16	TD (gage)	25	3	-
17	TD (gage)	25	4	-
18	TD (gage)	25	5	-
19	TD (gage)	25	6	Figure 4.27
20	TD (gage)	25	7	-
21	TD (gage)	25	8	-
22	TD (gage)	25	9	-
23	TD (gage)	25	10	-
24	TD (gage)	25	11	Figure 4.27

Table 4.2: Map of models including TD defects on the field side

Model Nr.	Type	Cross-sectional dimension (% of H.A.)	Depth (measured at the highest tip, in mm)	Figure
Study of the influence of the cross-sectional dimensions (depth held constant)				
25	TD (field)	5	2	Figure 4.26 (mirror image of #1)
26	TD (field)	10	2	Figure 4.26 (mirror image of #2)
27	TD (field)	15	2	Figure 4.26 (mirror image of #3)
28	TD (field)	20	2	-
29	TD (field)	25	2	-
30	TD (field)	30	2	Figure 4.26 (mirror image of #6)
31	TD (field)	35	2	-
32	TD (field)	40	2	Figure 4.26 (mirror image of #8)
33	TD (field)	45	2	-
34	TD (field)	50	2	Figure 4.26 (mirror image of #10)

Table 4.3: Map of models including VSH defects (located on the vertical plane passing through the rail central running axis)

Model Nr.	Type	Length (in mm)	Height (in mm)	Depth (measured at the highest tip, in mm)	Figure
Study of the influence of the length (height and depth held constants)					
35	VSH	10	30	5	Figure 4.24
36	VSH	20	30	5	Figure 4.24
37	VSH	40	30	5	Figure 4.24
38	VSH	80	30	5	Figure 4.24
39	VSH	160	30	5	Figure 4.24

Table 4.4: Map of models including HSH defects

Model Nr.	Type	Length (in mm)	Width (in mm)	Depth (measured at the highest tip, in mm)	Figure
Study of the influence of the length (for depths of 7.5 mm)					
40	HSH	10	40	7.5	Figure 4.24
41	HSH	20	40	7.5	Figure 4.24
42	HSH	40	40	7.5	Figure 4.24
43	HSH	80	40	7.5	Figure 4.24
44	HSH	160	40	7.5	Figure 4.24
Study of the influence of the length (for depths of 15 mm)					
45	HSH	10	40	15	Figure 4.24
46	HSH	20	40	15	Figure 4.24
47	HSH	40	40	15	Figure 4.24
48	HSH	80	40	15	Figure 4.24
49	HSH	160	40	15	Figure 4.24

Figure 4.24: Cross-sectional view of VSH (left) and HSH having depths of 7.5 mm (center) and 15 mm (right) defects. Studies of the influence of the defect length

Figure 4.25: 10% H.A. TDs on the gage side. Study of the influence of the depth. Depths shown (from top to bottom-right): 2 (zoomed-in), 12, 22 mm, as measured at the highest tip of each defect

Figure 4.26: TDs on the gage side. Study of the influence of the cross-sectional dimensions (depth held constant). Dimensions shown: 5, 10, 15, 30, 40, 50 % H.A. (from the top-left to the bottom-right)

Figure 4.27: 25% H.A. TDs on the gage side. Study of the influence of the depth. Depths shown (from left to right): 1, 6, 11 mm, as measured at the highest tip of each defect

Defect-sensitivity based on current receivers locations

The results described in the following were obtained considering simulated sensors deployed approximately in the same positions where the actual receivers are placed in the current version of the defect detection prototype, which is the one that was tested in the field at TTC in Colorado (Section 5.8). Both Figure 3.2 and Figure 4.20 show the approximate rail surface area scanned by each air-coupled receiver. The term ‘approximate’ is used, given the uncertainties in producing very accurate measurements due to the lift-off distance of ~6 cm between rail top surface and active surface of the receivers, which are held at an angle, making exact measurements difficult to accomplish. Out of the four available pairs of receivers, only three were actually used (i.e. their waveforms were fed to the prototype algorithm) during the processing of the field test data. In fact, the one shown in top-right position in Figure 4.20 was discarded because affected by very low levels of SNRs (as it is also confirmed by the results of the numerical simulation shown in the same figure). Therefore, only the other three pairs of receivers were used to perform the numerical study here presented. The sensors were simulated by using the vertical displacement histories experienced by the nodes of the model located exactly at the center of each active surface. Those nodes have coordinates (14,-1.125), (18,-1.125) and (14,2.125), measured in cm and using the reference system shown in Figure 4.20.

As previously stated, 29 simulations/measurement points were available for each defect, corresponding to the transmitter ranging from being right on top of the defect (in the middle for VSH and HSH cases) to being 14 cm away on one side of it, with steps of

0.5 cm. Symmetry allows to mirror the results to obtain the case of the transmitter going in the opposite direction respect to the defect.

The same algorithm as the one implemented in the defect detection prototype was used to process the simulation results (except in Case I in Section 4.2.4.1, where modifications had to be made). Details on such algorithm specifically designed to enhance the defect-sensitivity of the prototype are given in Section 5.6. The features extracted from each symmetric pair of simulated receivers are the ones currently used by the prototype, given in yellow background in Table 5.1, and also listed in Table 4.5 to facilitate the reading.

Table 4.5: Features used to analyze the simulations results

DI Max	$Max\left(\frac{Max x_1 }{Max x_2 }, \frac{Max x_2 }{Max x_1 }\right)$
Peak to Peak New	$Max\left(\frac{Ppk x_1 }{Ppk x_2 }, \frac{Ppk x_2 }{Ppk x_1 }\right)$
Ratio RMS	$\frac{RMS(x_2)}{RMS(x_1)}$

4.2.4.1 Case I: signals not affected by noise (ideal condition)

The numerical simulations offer an “ideal” case where $SNR = \infty$, due to absence of noise sources that in actual experiments are represented by: noise introduced by sensors, by other electrical components, environmental changes such as temperature or humidity, changes of rail geometry and/or top surface due to different wearing, etc.

Therefore, analyses of the results obtained in non-noisy conditions cannot yield definitive conclusions about whether or not the actual prototype would be able to detect each considered defect in real experiments, where noise can mask the signals variations caused by cracks. This task is addressed in the next paragraphs (Case II), where noise is artificially added to the numerical results in order to yield SNR values comparable to the ones available in the actual prototype.

Nevertheless, the advantage of having “pure” signals, not corrupted by noise, is that deterministic conclusions can be drawn regarding the intrinsic sensitivity of the proposed system to the different crack types/sizes/depths. Also, it allows to understand at which relative distances between crack and transmitter the defect detection is enhanced, distances that, how will be shortly discussed, are type-of-defect dependent. Lastly, it represents the limit case: the more the SNR is managed to be increased in the actual prototype (for example, by increasing the voltage delivered to the transmitter, by replacing the transmitter/receivers with more efficient/sensitive ones, by replacing the electrical components such as preamplifiers/filters with more efficient ones, etc.), the more the results will tend to these ones.

In case of absence of noise, the Damage Index cannot be based on a “Mahalanobis squared distance” (eq. (5.1)) as the one used in the prototype algorithm (the covariance matrix would be singular, being composed of a sequence of zeroes). Thus, a different metric is used in this case. First, it can be recognized that for pristine rails each feature in Table 4.5 would yield values of 1, for each pair of sensors (this follows straight from the fact that a symmetric wave propagation will take place toward both side around the excitation, which in turns means that each pair of symmetric

simulated sensors will sense identical waveforms). Therefore, a meaningful DI can be computed simply by measuring and summing the deviations of each feature from the unity, as follows:

$$DI = \sqrt{\sum_{ij} (f_{ij} - 1)^2} \quad (4.39)$$

where f_{ij} is the j -th feature (out of three) relative to the i -th pair of sensors (out of three).

Figure 4.28 through Figure 4.34 show the DI traces obtained using eq. (4.39) to the simulations results of each considered model. Every figure shows a different case study of the ones listed in Table 4.1 to Table 4.4. Table 4.6 tries to summarize the salient results, to facilitate comparisons between all the considered defects. In the table, for each model are listed (1) the summation of the respective DI values (in the 29 considered transmitter-defect relative positions), (2) the maximum DI value and (3) the position at which this occurred.

Figure 4.28 considers the case of TDs in the gage side with increasing cross-dimensional areas, ranging from 5% H.A. to 50% H.A. (Figure 4.26). The DI trace obtained from the pristine case (which is a sequence of zeroes) is also plotted in the figure for reference purposes. Several things are noteworthy in this figure. First, for every cross-sectional dimension the defect detection is enhanced when the transmitter is located close by the defect location. In particular, when transmitter and defect are spaced about 1.5 cm apart the highest DI values are obtained. Second, as the cross-sectional area increases from 5% up to 35% H.A. the maximum DI value of each defect monotonically increases, as expected, but starting at 40% H.A. the trend is inverted (this can be

appreciated more easily in the lower zoomed-in plot). However, if one considers the whole traces, approximately every DI values for distances greater than 2 cm are monotonically increasing as one moves from 5% to 50%. Nevertheless, while as the cross-sectional dimension increases from 5% up to 20% the sensitivity gets significantly higher step after step, for TDs greater or equal to 25% H.A. the variations are much smaller. This is intuitive, in fact the relative increase in defect area gets monotonically smaller at every increment (from 5% to 10% the area is doubled, from 10% to 15% is 1.5 times bigger, and so on). In conclusion, while every defect, including the smallest here considered (5% H.A.), appears to be detectable from this plot, final conclusions on the detectability cannot be drawn until noise is added to the waveforms (as will be done in next sections).

Figure 4.29 and Figure 4.30 show the case studies of TDs with 10% and 25% H.A., respectively, whose depth is gradually increased. In Figure 4.25 and Figure 4.27 are shown the extreme cases, as well as the middle one, for both case studies. As expected, in both cases as the depth increases the detectability decreases. Model 14 (shown in Figure 4.25) includes a 10% TD whose depth (22 mm) is below the theoretical detectability limit that would be valid for pure Rayleigh waves (Section 3.3.1). The Rayleigh wavelength in steel is in fact about 20 mm long. However, as discussed in Section 4.2.3, multiple modes are excited in this wave propagation phenomenon (neither of which is purely a Rayleigh wave) and this probably explain why the detectability is still possible at these depths. Nevertheless, once again, after the introduction of noise, detecting these very deep defects will be much more complicated.

One important aspect to notice when comparing the first three figures is that the simultaneous influence of both defect size and defect depth on the DI traces makes almost impossible to guess either of those by the only knowledge of the DI trace computed by the prototype in experimental tests. Also, those two influences have comparable magnitudes, considering that -as it can be seen in Figure 4.28 and Figure 4.30- defects whose size is comprised between 5% and 50% H.A. and whose depth is 2 mm have DIs ranging from about 0.5 to 6, while the same 25% TD generates DI values in the same range as the depth is increased from 1 mm to 11 mm. In other words, while the proposed prototype can effectively detect the presence of a crack, once this happens, then a different method has to be used to establish its severity. Nevertheless, there might be a possibility for the prototype to estimate at least the type of crack being detected, based on typical DI traces shapes, as will be discussed at the end of Section 4.2.4.2.

Figure 4.31 shows the case of TDs on the field side having increasing sizes. The defects geometry/locations are as in Figure 4.26, after mirroring each crack around the rail central running axis. For this type of defects, the detectability is enhanced when the transmitter is far from the flaw. This is intuitive: considering that the transmitter sends its signal toward the gage side, when the defect is in the same location but in the opposite side, the wave propagation toward both directions is not affected because the path does not cross the crack. The arrows in Figure 3.2 clarify this concept. Until the TD is located (approximately) in the triangle formed by the top horizontal line (indicating the rail head side) and the two closer arrows (indicating the wave propagation), its presence is not enlightened by the ultrasound. In Figure 4.31 it can be seen that the moment in which those TDs start to be detected is when the transmitter is about 4 cm away from them, for

the bigger ones, and much later (at about 11/12 cm) for the smaller ones. The highest DI values are all found at 11 cm for the defects greater or equal to 35%, and at 13-13.5 cm for all the smaller ones. The DI values of the smallest ones (up until about 15%) are quite low, suggesting that when noise will be introduced their detectability will be at risk.

Figure 4.32 shows the case of Vertical Split Head defects having height of 30 mm, depth (measured at the highest tip) of 5 mm and longitudinal extensions increasing from a minimum of 10 to a maximum of 160 mm. Figure 4.24, left sub-plot, shows the cracks geometry in the rail cross-sectional plane. As briefly discussed before, at distance 0 the transmitter is right at the center of the VSH longitudinal extension, so symmetry is maintained, and the results can be exactly mirrored for a transmitter running toward the opposite direction. The DI traces suggest that until the VSH does not reach lengths of at least 40 to 80 mm its detectability would probably be problematic. Then, once a length of 80 mm is reached, high DI values are generated. The values are probably monotonically increasing up to the next considered case, which is 160 mm. It appears that the higher DI values are found when the transmitter is right at the tip of the crack (distances of ~4 and ~8 cm for the 8 and 16 cm cases respectively). This can be explained by the fact that when the transmitter is at the tip, approximately the entire energy can travel undisturbed to the side of receivers in the non-defective direction, while almost all the waves propagation toward the other direction are “trapped” between the longitudinal discontinuity and the side of the rail head.

Lastly, Figure 4.33 and Figure 4.34 show the cases of Horizontal Split Heads having width of 40 mm, depths of 7.5 and 15 mm (respectively) and increasing longitudinal dimensions ranging from 10 to 160 mm. Figure 4.24 (central and right sub-

plots) shows the two cracks geometries in the rail cross-sectional plane. For both cases, the DI values are much higher than the ones relative to the VSH. In other words, it appears that detecting HSH would be much easier than detecting VSH. Even the shorter extensions here considered, 10 mm, offer DIs of significant magnitudes. Generally, as expected, the values for the 7.5 mm depth case are higher than the ones for the 15 mm case. Surprisingly, the highest DI values among all the considered HSH are obtained with the 80mm-long one at the shallowest depth, instead of with the 160 mm one. However, except for this particular case, in both plots as the defect length is increased, increasing DIs are obtained. Furthermore, just as noticed for the VSH case, in almost every model the highest DI is obtained when the transmitter is at the tip of the respective crack. The same explanation as the one given for the VSH case can be provided.

a) TD gage side - Study over the influence of the cross-sectional dimensions (depth = 2 mm)

b) ZOOM-IN

Figure 4.28: Case I (no noise). TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.29: Case I (no noise). TD 10% H.A. in the gage side. Study over the influence of the depth

Figure 4.30: Case I (no noise). TD 25% H.A. in the gage side. Study over the influence of the depth

TD field side - Study over the influence of the cross-sectional dimensions (depth = 2 mm)

Figure 4.31: Case I (no noise). TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.32: Case I (no noise). VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm)

Figure 4.33: Case I (no noise). HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm)

Figure 4.34: Case I (no noise). HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm)

Table 4.6: Summary of the salient (non-statistical) Damage Index values for the various models (case I: signals not affected by noise)

Model	Sum(DI)	Max(DI)	Distance
0	0.00	0.00	-
1	18.57	1.21	1
2	37.86	2.58	1
3	48.71	3.73	1
4	53.48	4.54	1.5
5	54.61	5.23	1.5
6	58.50	5.93	1.5
7	63.73	6.10	1.5
8	66.76	6.05	1.5
9	69.21	5.68	1.5
10	72.74	5.39	1.5
11	21.04	1.20	1
12	11.40	0.97	0.5
13	6.92	0.68	0.5
14	4.45	0.46	4.5
15	66.97	6.50	1.5
16	45.91	4.46	1.5
17	39.44	3.88	1.5
18	34.37	3.35	1.5
19	30.27	2.87	1.5
20	26.85	2.44	1.5
21	23.89	2.03	1.5
22	21.31	1.67	1.5
23	19.04	1.62	3.5
24	17.08	1.62	3.5

Model	Sum(DI)	Max(DI)	Distance
25	5.32	0.86	13
26	3.93	0.52	13
27	5.73	0.84	13.5
28	7.53	1.52	13.5
29	9.36	1.87	13.5
30	12.36	1.55	13.5
31	16.89	1.54	11.5
32	19.32	1.86	11
33	21.66	2.14	11
34	23.62	2.39	11
35	1.16	0.13	0.5
36	2.58	0.19	11.5
37	5.87	0.40	11.5
38	13.49	0.78	9.5
39	28.74	2.11	7.5
40	7.86	0.52	14
41	17.81	1.44	2
42	45.92	3.23	1.5
43	125.62	7.64	3.5
44	86.43	5.50	13
45	6.26	0.80	2.5
46	11.00	1.26	13.5
47	20.09	2.03	2
48	25.45	2.38	4
49	42.18	3.52	8

4.2.4.2 Case II: noise artificially added to the signals (realistic condition)

In order to accurately reproduce the actual conditions in which the prototype operates while scanning the rail-track, noise was artificially added to the signals obtained from the numerical models. Both the (I) shape/frequency content and the (II) RMS of the added noise were chosen in such a way to replicate the noise affecting the signals in the current version of the prototype, which features highly efficient impedance matching networks as described in Section 5.5.2.

(I) To preserve actual shapes and frequency contents, 6,000 waveforms were recorded from an air-coupled receiver used in the prototype when the amplifier feeding the transmitter was turned off. The same instrumentations and environmental conditions (acquisitions performed at the Camp Elliott RDTF) as used in actual tests were retained. Figure 4.35 shows one of those waveforms.

Every time an analysis of the LISA simulated signals was performed, each signal was corrupted by randomly selected noisy waveforms as the one in Figure 4.35. In particular, the MATLAB[®] function *randi* was used to randomly pick one of the 6,000 recorded noisy waveforms.

(II) A factor $f = 1.5$ was applied to those noise-waveforms in a way to achieve comparable SNRs as the ones available in actual experimental conditions. The value assigned to f was determined by computing:

- (1) the $\text{RMS}_{\text{NOISE}}$ from all the noise-waveforms (similar value of the one shown in the legend in Figure 4.35);
- (2) the mean of the peak-to-peak amplitudes $\text{AMP}_{\text{EXP-SIGNAL}(1)}$ among a set of 100 signals sensed by an air-coupled receivers (the one positioned as the

bottom-left rectangle in Figure 4.20, here denoted as number ‘1’) during experimental testing over a pristine piece of rail;

- (3) the peak-to-peak amplitude $AMP_{LISA-SIGNAL(1)}$ of the waveform acquired by the corresponding simulated sensor in the LISA pristine rail model.

Figure 4.35: Typical noise shape

Then, computing the available experimental SNR at receiver # 1 as

$$\frac{AMP_{EXP.-SIGNAL(1)}}{RMS_{NOISE}} = SNR_{EXP(1)}, \text{ the factor } f \text{ was determined such that } \frac{AMP_{LISA-SIGNAL(1)}}{f \times RMS_{NOISE}} =$$

$SNR_{LISA(1)} = SNR_{EXP(1)}$. Looking at Figure 4.20, it appears that for the other two considered sensors (the one at the bottom-right rectangle is here denoted # 2, the one at the top-left # 3, while the one at the top-right as discussed above is not used), which are aimed to less-energized portions of the rail top surface, the respective $SNR_{(2)}$ and $SNR_{(3)}$ will be both smaller than $SNR_{(1)}$. In fact, f and RMS_{NOISE} does not change from sensor to sensor, while the signal amplitudes sensed by those two sensors are smaller.

A confirmation that using the factor $f = 1.5$ yields accurate simulated results for each of the three considered sensors can be found in Figure 4.36 and Figure 4.37. In the sub-plots of Figure 4.36 are super-imposed five signals acquired by each of the 3 air-coupled receivers (the ones symmetrically located on the other side of the transmitter are not shown, however, they provide similar results) while the prototype was scanning a pristine part of the rail. The zoomed-in plots on the right show the extent of the variations produced to the signals amplitude by noise. While only five signals are shown in figure in order to facilitate the visualization, 100 signals per sensor were actually considered. First, from each signal the maximum amplitude value $MAX_{(i)}$ was extracted, then the mean of the maxima $MEAN_{max}$ was calculated, and finally a ‘Normalized Maximum Amplitude RMS’ was calculated per each sensor by computing the RMS of the deviations ($MAX_{(i)} - MEAN_{max}$) and dividing it by $MEAN_{max}$. This feature provides an indication of how much the noise makes the otherwise constant signals taken in pristine rail conditions vary in relation to the signal strength that can be sensed by each sensor. For signals much stronger than the noise (or, inversely, for very low noise respect to signal strength) this feature tends to 0, which is therefore the desired value. The same procedure was used to calculate the same feature for the LISA simulated sensors, after corrupting 100 pristine-rail constant waveforms with randomly picked noise-waveforms and using the factor f . Figure 4.37 plots the obtained feature values for all the available simulated sensors at transversal (to the rail) lines 14 and 18 cm away from the transmitter. The three sensors used in the simulations described in this section are circled in red and their values compare really well with the ones given in Figure 4.36 for the actual sensors. As expected after analyzing Figure 4.20, the two sensors toward the rail gage side can count

on higher signal amplitudes than the one at the field side and hence suffer less variability due to noise. Among the two, sensor #1 performs slightly better than sensor #2. The fact that those values compare so well with the experimental ones is also a confirmation that the LISA simulations as the one that was used to plot Figure 4.20 are faithful representation of the actual wave propagation taking place in the rail head after the excitation released by the air-coupled transmitter.

Figure 4.36: Super-imposition of 5 signals acquired by 3 air-coupled receivers in a pristine part of the rail. The right plots are zoom-ins from the areas marked as black rectangles in the corresponding left figures. Values of the 'Normalized Maximum Amplitude RMS' are given for each sensor in the respective title. Those values were calculated using 100 signals, instead of only the 5 shown here

Figure 4.37: Values of the ‘Normalized Maximum Amplitude RMS’ for each available simulated sensor. Red circled are the ones considered to simulate the actual air-coupled receivers

Implementation of the algorithm for defect-detection

Once the noisy conditions were simulated, the same algorithm for defect detection as the one used in the actual prototype was utilized to analyze the results of the numerical simulations. The details on the algorithm are given in Section 5.6. To summarize the procedure, a baseline formed by 170 noise-corrupted acquisitions made on the pristine rail model was formed (N.B.: in the pristine-rail model, which was defined as model ‘0’, each of the 29 simulations made moving the transmitter in different rail locations are equivalent, due to absence of any defect in the model). Then, the features listed in Table

4.5 were extracted from the signals associated with each of the three pairs of symmetric sensors in those 170 corrupted acquisitions. Mean vector and covariance matrix were computed. At that point, Damage Index values using “Mahalanobis squared distances” (expressed as in eq. (5.1)) could be computed for each desired model/defect, including model ‘0’, whose DIs represent the noise floor to which DIs obtained from every other defective-model have to be related according to the statistical framework.

Figure 4.38: Damage Index for rail in pristine conditions. Comparison between numerical simulations results and experimental results

Figure 4.38 shows a comparison of Damage Index traces furnished (a) by the numerical pristine-rail model, and (b) by the actual experimental testing in pristine rails.

The LISA data shown in the figure were obtained lining up 1,000 different observations obtained by adding the randomly picked noise. The experimental data were instead acquired at the TTC facility in Colorado. Mean and RMS of both traces are provided in each respective legend. The agreement is found to be very satisfactory.

Due to the randomness introduced by the presence of noise altering the signals, for each defect/model a DI trace was computed 1,000 times through successive iterations of the same procedure. After that, averaged DI traces were obtained simply taking the mean of the available 1,000 outcomes for each defect. Those averaged DIs are shown in Figure 4.39 through Figure 4.45. While the shapes of the traces relative to the different defects are similar the ones obtained in the ideal case of absence of noise (Case I, previous paragraph), here one more important piece of information is conveyed. In fact, the y-axis values, i.e. the statistical Damage Indexes, can be directly compared to the values of the pristine-rail case shown in Figure 4.38 to get a first idea on how likely is that each specific defect can be detected by the actual prototype, given this available SNR. Similarly to previous Table 4.6, in Table 4.7 the salient results from those averaged DIs are summarized.

Figure 4.39: Case II (added noise). TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.40: Case II (added noise). TD 10% H.A. in the gage side. Study over the influence of the depth

Figure 4.41: Case II (added noise). TD 25% H.A. in the gage side. Study over the influence of the depth

Figure 4.42: Case II (added noise). TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.43: Case II (added noise). VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm)

Figure 4.44: Case II (added noise). HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm)

Figure 4.45: Case II (added noise). HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm)

Table 4.7: Summary of the salient (averaged) Damage Index values for the various models (case II: noise is artificially added to the simulation results; “high” SNR level ($f = 1.5$))

Model	Sum(DI)	Max(DI)	Distance
0	145	5	-
1	964	104	1
2	5874	776	1
3	9609	1671	1
4	11590	1756	1
5	11912	1869	1
6	13393	2307	1.5
7	15209	2426	1.5
8	16144	2360	1.5
9	16567	2062	1.5
10	17755	1882	1.5
11	1502	156	1
12	379	37	0.5
13	141	18	0.5
14	107	6	4.5
15	19339	3217	1
16	7932	1233	1.5
17	5577	905	1.5
18	4047	662	1.5
19	2992	479	1.5
20	2245	346	1.5
21	1692	238	1.5
22	1293	156	1.5
23	1006	137	3.5
24	799	132	3.5

Model	Sum(DI)	Max(DI)	Distance
25	181	29	13
26	119	9	13
27	149	28	13.5
28	270	105	13.5
29	404	166	13.5
30	447	110	13.5
31	685	108	11.5
32	870	164	11
33	1070	221	11
34	1268	278	11
35	116	5	0
36	96	5	0
37	107	6	4
38	350	31	9.5
39	1942	228	7.5
40	180	26	14
41	1057	202	2
42	7282	917	1.5
43	68305	8432	3.5
44	32004	4402	11.5
45	154	32	14
46	621	185	13.5
47	2190	364	13
48	2367	488	4
49	7909	1206	8

Since the previous figures only show averaged DIs, definitive conclusions on the defects detectability cannot be drawn without further information. Figure 4.46 gives an idea about the degree of variability introduced by noise affecting the signals. Results from 5 different iterations on model #5 as well as the mean of the 1,000 available iterations are super-imposed. The maximum DI value of the averaged plot is just below 2,000, achieved at a distance of 1 cm, while the variation of the DI at the same point corresponding to the 5 shown traces is in the range [1350:3000]. Minimum and maximum DI values obtained at the same distance in the 1,000 available iterations were also computed, yielding a new enlarged range of [560:5940]. Also, the overall maximum value in this set of iterations when considering the whole 14 cm range is actually 7,250, thus is achieved in a different position. As it appears from these numbers, the variability produced by noise is significant; nevertheless, in this case, even the minimum value of the range, 560, is quite above the typical pristine-case DI values that, as shown in Figure 4.38, seem to be bounded below ~ 100 . In conclusion, the result is that this defect would always be detected. Nevertheless, this is not always the case when analyzing the other defects, suggesting that a probabilistic approach is necessary. Such probabilistic analysis was tackled through computation of Receiver Operating Characteristics curves, presented in next paragraph.

Figure 4.46: Example showing the degree of variability introduced by noise affecting the signals

Before moving to the analysis of the ROC curves, a few more considerations can be made based on the shapes of the presented averaged DIs. First, Figure 4.47 compares the typical DI shape indicating the presence of a TD in the gage side as resulting from LISA simulations (c) with one experimental trace obtained at TTC facility in Colorado, where a TD listed as 5% H.A. (Table 5.2) was well detected. The top plot (a) is provided only to show the low noise floor around the indications of the defect. The comparison between plots (b) and (c) confirms the fact that the TD (gage side) detection is enhanced when the transmitter is very close to the crack. The different heights of the two peaks in the experimental plots are due to the noise variations, as can be seen in the numerical results plotted in Figure 4.46. Agreements between experimental and numerical results

are also noticeable focusing the attention on the 2 smaller peaks in each side toward the outer extremes.

A last observation is made with the aid of Figure 4.48, about the typical shapes obtained when considering the different types of defect. The suggestion is that there might be a possibility to estimate at least the type of defect being detected by simply measuring the distance between the detected symmetric higher peaks. Closer peaks (in the order of 5 cm) would indicate a TD in the gage side, wider peak, as wide as the distance between the symmetric pair of receivers, would indicate a TD in the field side, while intermediate distances could indicate a VSH or a HSH. However, a systematic study to prove such observations still needs to be carried out. Estimates based on the amplitudes of the highest peaks can also be attempted, mostly considering that generally a TD in the gage side generates much higher DIs than a correspondent one in the field side. However, caution has to be taken if using amplitudes, considering that several factors intervene contemporarily that can dramatically modify the peaks heights, such as: defect depth below the rail top surface, defect area, variations simply due to noise randomly affecting the signals, effects of different defects inclinations from the orthogonal to the rail running direction considered in this study.

Figure 4.47: Comparison between experimental and numerical results. DIs showing the detection of a TD. The top plot (a) shows a zoomed-out view to appreciate the low noise floor around the TD indications. In (b) a zoomed-in version is furnished to facilitate the comparison with plot (c), where the (averaged) DI obtained from LISA model #4 is shown, mirrored respect to the defect location

Figure 4.48: Typical DI traces shapes for different types of defect, indicating the possibility to estimate the detected type of defect simply measuring the highest peaks distance

4.2.4.3 Case II: probabilistic analysis through computation of Receiver Operating Characteristics curves (for “high” SNR levels)

Receiver Operating Characteristic (ROC) curves are useful tools that allows to readily characterize the performances offered by a detector. These curves graph the Probability of Detection (PD) rate over the Probability of False Alarm (PFA) rate offered by the detector when the value of the threshold in the Damage Index trace is gradually increased [88]. As the term implies, when applied to the results of these simulations the PD measures the number of time a defect is detected over the total number of observations available for the specific defect. Similarly, the PFA measures the number of time a pristine-case is wrongly detected as being defective over the total number of available observations for the pristine-rail case.

For each defect/model, including the pristine case, 1,000 observations were produced by iterating the procedure of randomly adding noise to the numerical signals and then computing the corresponding DI traces, as described in last paragraph. However, as already pointed out, each of those observations contains a DI trace corresponding to the transmitter being moved 14 cm away in only one direction respect to the defect location, thanks to the symmetry of the problem. Nevertheless, in real applications the prototype has the opportunity to scan the defected rail portion twice that length, with the transmitter being at first on one side of the defect, and then on the other side. As a result, this increases its probability of detecting the defect. Thus, in order to adhere to real conditions, from the initially available 1,000 analysis outcomes a total of 500 “complete” observations were obtained by lining up each time the DI traces corresponding to pairs of “half-long” observation (and, to be precise, for each pair only

one of the two outcomes available for the 0 cm transmitter position was retained, randomly picked among the two). In summary, the resulting DI trace forming one observation to be tested by the algorithm performing ROC curves computation is formed by 57 DI values, corresponding to the transmitter being moved from -14 cm to +14 cm around each defect, with steps of 0.5 cm. The same applies for the pristine-case, for which each observation is then formed by 57 DI values resulting from the transmitter being moved for 28 cm in non-defective rail.

Considerations have to be made on the conditions that have to be met for the actual prototype to call a defect. Looking at Figure 4.38, which shows that for pristine rails some higher peak can randomly be activated by the action of noise, it appears that might be advantageous in terms of keeping low PFA values to avoid calling a defect if only one (or maybe two, three, etc..) Damage Index exceeds the threshold within a short piece of rail having length on the order of the distance between the furthest deployed receivers. In fact, when a defect is really present, multiple high DI peaks are expected to be seen within that rail length. Nevertheless, Figure 4.39 through Figure 4.45 showed that there exist only few relative positions of the ensemble transmitter-defect-receivers (positions that are type-of-defect-dependent) for which the highest Damage Indexes are activated. Therefore, an analysis was carried out as the parameter which can be denoted as Number of Threshold Crossing (NTC) was increased from 1 to 10, looking for the one capable of yielding the best detection performances according to the resulting ROC curves. As a result of such analysis, the overall best performances were achievable when the condition to call a defect is that at least 3 Damage Indexes (out of the 57 forming each observation) exceed the threshold. Obviously, this value is strictly dependent on the

spatial resolution at which the tests are performed (5 mm in these simulations), which determines how many DI points are available as the prototype is moved over a defective position. For example, if finer resolutions are possible, probably better performances can be achieved by setting NTC higher than this value, while for poorer resolutions is probably necessary to reduce it. Interestingly enough, as the same type of analysis was (independently) performed over the experimental data collected at the TTC field tests in Colorado (details are in Section 5.8.5), where the spatial resolution at which the prototype was operating was very similar (~6.35 mm), it was found that the best performances were yield by setting the same NTC of 3, which provide further confirmation of good agreements between experimental and simulated results.

The ROC curves resulting from such analysis are shown in Figure 4.49 through Figure 4.55. In order to guarantee a sufficient resolution in plotting those curves, the threshold was linearly increased through 500 steps starting from a value of 0 (where both PD and PFA are obviously 1, or 100%) up to the maximum DI value among the 500 pristine-case DI traces (once that value is reached, the PFA drops to 0, and further threshold increases would only produce the PD dropping along the vertical axis at 0-PFA all the way till 0-PD). Once those curves are available, one way to assess the quality of the performances offered by the detector is to evaluate their area under the curves (AUC), which are shown in the plots legends. An AUC of 50% is the one that would be achieved by random guessing, while an AUC of 100% means that there exist thresholds for which all the defects are detected while not including any false alarm. Furthermore, Table 4.8 shows other interesting parameters that were extracted from the curves, namely the PD

values in correspondence to some “significant” PFAs. In particular, PFAs of 0%, 1%, 5% and 10% are considered.

Figure 4.49 through Figure 4.51 show a very good efficacy in the detection of Transverse Defects in the gage side. Among the ones considered in these simulations the only two that did not yield a perfect 100% AUC are the deepest TD 10% H.A., at a depth of 17 and 22 mm respectively. Among those two, while in the 17 mm case good PD values are still possible for low PFAs, the performance at low PFAs of the 22 mm case is quite poor. These results suggest that would be interesting to setup another set of simulations on a different set of gage sided TD defects being smaller than 25% H.A. and located at different depths and/or transversal positions in the rail cross-sectional plane, to define more precisely the limits on the detectability.

Figure 4.52 shows that, quite surprisingly, the detection of the field sided TD 5% H.A. outperforms the detection of the larger 10% and 15% defects. In fact, looking at both Figure 4.31 and Figure 4.42 the highest DI peak among these three defects are found for the 5% defect. Probably this has something to do with the specific lateral position in the rail cross-sectional plane of the 5% defect respect to the other two. Further analyses on the same shaped defect as it is slightly moved in the transversal direction will be performed in the future. Other than that, a perfect AUC of 100% is achieved in the field side for defects bigger than 20% H.A.

Figure 4.53 shows that the detection performance of VSH shorter than 40 mm is quite poor (at least, when the depth and height are the ones considered here). However, when those defects reach lengths around 40 mm the prototype would start to detect them, and at longer lengths it will be 100% successful.

Finally, as it already indicated by the highest DI values shown in the figures of the last two paragraphs, Figure 4.54 and Figure 4.55 confirms that the detection of HSHs is easier than that of VSHs. Except for the smallest length (10 mm) among the ones considered, all the longer ones show 100% AUCs.

Figure 4.49: Case II ($f = 1.5$), ROC. TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.50: Case II ($f = 1.5$), ROC. TD 10% H.A. in the gage side. Study over the influence of the depth

Figure 4.51: Case II ($f = 1.5$), ROC. TD 25% H.A. in the gage side. Study over the influence of the depth

Figure 4.52: Case II ($f = 1.5$), ROC. TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.53: Case II ($f = 1.5$), ROC. VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm)

Figure 4.54: Case II ($f = 1.5$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm)

Figure 4.55: Case II ($f = 1.5$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm)

Table 4.8: Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow

Model #	PD achieved in correspondence of “significant” PFA values				AUC
	0% PFA	1% PFA	5% PFA	10% PFA	
1	100	100	100	100	100
2	100	100	100	100	100
3	100	100	100	100	100
4	100	100	100	100	100
5	100	100	100	100	100
6	100	100	100	100	100
7	100	100	100	100	100
8	100	100	100	100	100
9	100	100	100	100	100
10	100	100	100	100	100
11	100	100	100	100	100
12	100	100	100	100	100
13	42.2	79.0	89.2	94.8	98.3
14	14.4	36.2	50.4	65.0	88.4
15	100	100	100	100	100
16	100	100	100	100	100
17	100	100	100	100	100
18	100	100	100	100	100
19	100	100	100	100	100
20	100	100	100	100	100
21	100	100	100	100	100
22	100	100	100	100	100
23	100	100	100	100	100
24	100	100	100	100	100
25	56.2	80.0	88.8	93.2	97.3
26	9.8	30.0	46.0	58.2	84.7
27	35.0	70.4	83.0	89.2	96.6
28	85.6	96.0	98.2	99.8	99.8
29	99.2	100	100	100	100
30	100	100	100	100	100
31	100	100	100	100	100
32	100	100	100	100	100
33	100	100	100	100	100
34	100	100	100	100	100
35	0.2	3.4	6.4	12.0	55.8
36	0.4	6.0	12.4	22.4	64.9
37	15.0	45.4	64.4	76.2	92.7
38	99.4	100	100	100	100
39	100	100	100	100	100
40	62.2	91.4	97.4	99.0	99.5
41	100	100	100	100	100
42	100	100	100	100	100
43	100	100	100	100	100
44	100	100	100	100	100
45	63.2	87.6	95.0	98.4	99.2
46	100	100	100	100	100
47	100	100	100	100	100
48	100	100	100	100	100
49	100	100	100	100	100

4.2.4.4 Case II: probabilistic analysis through computation of Receiver Operating Characteristics curves (for “low” SNR levels)

The results described in the previous section shows a quite satisfactory performance achievable by the prototype. However, those results describe the performance of which the proposed prototype would be capable only if “normal” noisy conditions are maintained throughout the entire execution of experimental testing. Unfortunately, in actual testing conditions, there are a number of factors, of environmental, electrical, or physical origin, that can yield an increase of the background noise sensed by the air-coupled receivers, or, conversely, a decrease of the ultrasonic signals strengths. As the SNR drops, the performance is expected to deteriorate as well, because the alterations solely produced by noise on the signals will gain increasing importance when compared to the alterations produced by the ultrasonic waves interacting with defects.

To quantify the negative effects produced on the prototype performances by a possible drop of SNR, the same analyses described in the last two sections were iterated, after setting a new factor $f' = 2 \times f = 3$ used to double the noise incidence (i.e. the SNR was halved). Figure 4.56 through Figure 4.62 and Table 4.9 show the corresponding results.

The detection performance in finding TDs on the gage side does not deteriorate dramatically, at least considering the shallowest ones. However, the ones at the highest depths (TD 25% H.A. at 10 mm or below and TD 10% H.A. at 12 mm or below) experienced noticeable drops. In particular, the detection of the 17 and 22 mm deep TDs is almost compromised.

Nevertheless, the detection performance when considering the shallow field sided TDs drops dramatically. Only the bigger ones can still be detected without having to include a big number of false alarms.

The VSH detection remains possible at lengths of 8 cm or more, however is compromised the possibility to detect the ones shorter than 8 cm.

Finally, for the HSHs the situation does not worsen significantly, considering that only the 10 mm long case and partially the 20 mm one experience significant performance drops.

These results indicate that maintaining at least the current “high” SNR level (or even better, increasing it, if possible) throughout the whole development of experimental testing is crucial for the detection of certain defect types.

Figure 4.56: Case II ($f = 3$), ROC. TD in the gage side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.57: Case II ($f = 3$), ROC. TD 10% H.A. in the gage side. Study over the influence of the depth

Figure 4.58: Case II ($f = 3$), ROC. TD 25% H.A. in the gage side. Study over the influence of the depth

Figure 4.59: Case II ($f = 3$), ROC. TD in the field side. Study over the influence of the cross-sectional dimensions (depth held constant at 2 mm, as measured at the highest tip of the crack)

Figure 4.60: Case II ($f = 3$), ROC. VSH. Study over the influence of the length (height = 30 mm, depth = 5 mm)

Figure 4.61: Case II ($f = 3$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 7.5 mm)

Figure 4.62: Case II ($f = 3$), ROC. HSH. Study over the influence of the length (width = 40 mm, depth = 15 mm)

Table 4.9: Summary of the salient ROC curves results for the various models. Case of “low” SNR level ($f = 3$). The values lower than 100% are highlighted in yellow

Model #	PD achieved in correspondence of “significant” PFA values				AUC
	0% PFA	1% PFA	5% PFA	10% PFA	
1	84.8	96.4	99.4	100	99.9
2	100	100	100	100	100
3	100	100	100	100	100
4	100	100	100	100	100
5	100	100	100	100	100
6	100	100	100	100	100
7	100	100	100	100	100
8	100	100	100	100	100
9	100	100	100	100	100
10	100	100	100	100	100
11	98.6	100	100	100	100
12	16.8	40.6	65.6	83.6	94.2
13	2.8	7.0	15.8	27.8	69.9
14	1.0	5.4	11.4	17.4	61.8
15	100	100	100	100	100
16	100	100	100	100	100
17	100	100	100	100	100
18	100	100	100	100	100
19	100	100	100	100	100
20	99.8	100	100	100	100
21	98.8	100	100	100	100
22	89.8	98.4	99.6	99.8	99.9
23	69.4	88.4	95.6	98.8	99.2
24	52.0	75.4	92.2	96.4	98.4
25	1.6	4.4	10.8	20.2	63.8
26	0.6	2.2	5.6	11.6	55.7
27	0.4	4.4	9.6	16.0	59.1
28	1.8	8.0	19.2	29.0	70.0
29	4.8	13.0	26.2	41.6	78.0
30	12.4	25.2	43.0	57.2	84.8
31	25.0	46.0	65.2	78.6	93.5
32	37.6	61.6	78.4	87.4	96.3
33	55.4	74.4	91.4	94.6	98.1
34	63.2	83.6	94.6	97.4	98.9
35	0.2	1.2	4.4	8.6	49.5
36	0.2	0.8	3.2	8.2	51.1
37	0.8	3.6	11.2	17.6	62.0
38	6.4	19.0	37.4	51.0	84.1
39	56.0	81.0	92.4	97.6	98.9
40	5.6	16.0	32.0	49.4	83.3
41	85.0	96.2	98.8	99.8	99.8
42	100	100	100	100	100
43	100	100	100	100	100
44	100	100	100	100	100
45	4.0	10.2	23.6	35.4	74.8
46	52.6	72.2	87.4	95.4	98.1
47	99.8	100	100	100	100
48	100	100	100	100	100
49	100	100	100	100	100

4.2.4.5 Case II: comparison with results obtainable using coherent array techniques

Considering that more than one receiver is available in each side of the transmitter, and within close distances between each other, the natural question that arose is whether or not treating the information sensed by each sensor as an ensemble instead of separately (as it is done in the current version of the prototype, and as was described in the last sections) could yield benefits toward defect detection performances. Several times during the prototype's development stages, attempts were made to use different coherent array processing techniques while analyzing experimental results collected from both the UCSD RDTF facility at Camp Elliott, both the TTC facility in Colorado (some results obtained from the TTC data are given in Section 5.9). However, each time, when comparing the best results achieved using coherent array techniques against the ones obtained by extracting features from each available symmetric pair of sensors treated as separate entities, the results furnished by the latter were superior. In theory, if coherent sums are perfectly made among the signals acquired by each receiver located on the same side of the transmitter, the SNR would increase, and that is a clear fact. However, even temporarily overlooking the fact that "perfectly coherent" sums are difficult to be performed in a system operating at high speeds and in continuously changing testing conditions (fact that will be treated soon), the trade-off is that localized observations taken from distinct parts of the rail top surface are lost, or better, are averaged in a cumulative waveform representing a much larger observed area. As will be shown later in this section, this averaging can yield slightly better results for cases such as the one depicted in Figure 4.21(a), i.e. of a Transverse Defect in the gage side when it is right in

front of the excitation, which would block almost entirely the wave propagation toward one side. However, for almost all the other cases, where the wave propagation toward one side is not totally prevented, but only partially disturbed (*), it is actually more effective to capture the arisen variations in the distribution of the waves energy, which when compared to the distribution on the other side of the transmitter in a point-by-point manner greatly enhance the defect-induced asymmetries. While those variations on the energy distribution are certainly caused by waves-defect interactions, they are then greatly enhanced by the fact that constructive and destructive interferences of the multiple propagating modes in the bounded medium are modified, in unpredictable ways. In any case, if those punctual observations are averaged before comparing to the other side, the induced asymmetries are reduced.

(*) Basically those cases include all the types of defects other than gage sided TDs, which are never capable (for any relative position transmitter-crack) of entirely blocking the wave transmission toward one rail direction, but also include many situations of the same TDs on the gage side, as it can be seen for example in the instants caught in Figure 4.21(b) through Figure 4.22(f).

These considerations found confirmations on the results that are presented in this section (and, independently, on the study described in Section 5.9). To produce such results, the delay-and-sum technique was implemented on the numerical signals. The delays needed to correctly align the signals received by the three considered simulated receivers were set using the simulation results of the pristine-rail model (not corrupted by the artificially added noise). Figure 4.63 shows the result of the alignment. After noise is introduced in the models, summing up together the three signals in such coherent fashion

would certainly decrease the SNR, as the incoherent noise contributions are averaged toward 0.

Figure 4.63: Alignment of LISA signals

A similar procedure can be performed in setting up actual experimental testing. Basically, once the sensors are deployed over the rail top surface, the exact delays to align the signals can be simply calculated and set in the algorithm. However, in the real case, these initially calculated delays will not remain accurate throughout the tests. In fact, as it was already briefly introduced, ever-changing conditions are encountered as the actual prototype travels along the rail-track, whose effects deteriorate the ideal results achievable if “perfect” coherent sums were always performed. A certainly not comprehensive list of the different causes responsible of these detrimental effects is provided below.

- (1) The “specimen” being tested is continuously changing, sometime gradually, some other abruptly. Those changes are both in terms of the geometry, both of the mechanical properties. In both cases, the waves propagation is varied, therefore de-phasing of signals sensed by different receivers arises. For example: different rail sections are encountered (RE136, RE141, etc.); different levels of wears of top surface toward the gage side; different mechanical properties due to steel plastic hardening on the upper rail head portion, etc.
- (2) The “transfer function” constituted by the ultrasonic wave propagation in air between the leakages from rail top surface and the sensing by the receivers active surface can change dramatically. In fact, considering that the wavelength in air is ~ 2 mm, even changes on the order of 1 mm in the distance between rail top surface and each receiver active surface will shift the received signal completely in anti-phase. Vibrations affecting the cart holding the prototype sensors as traveling at high speeds, especially while accelerating or decelerating, or when wheels hit asperities such as rail joints or similar, will make each receiver oscillate in ranges well higher than 1 mm, possibly causing sudden de-phasing of the signals acquired by the various receivers. Furthermore, different rail top surface wear levels encountered as traveling on the track are also within ranges well wider than that value, therefore causing changes in the distance traveled by sound in air before reaching the receivers.
- (3) Changing of various environmental conditions, which can bring a number of detrimental effects.

- (4) As it appears from signals acquired by the different air-coupled sensors used in the prototype, the exact frequency at which every sensor is put in motion when hit by the waves leaking from the rail might slightly differ from sensor to sensor. For example, Figure 4.64(b) shows the red signal (receiver # 2) centered at slightly higher frequency than the blue one (receiver # 1). This might be due to the exact internal resonant frequency value of each specific transducer, or might be caused by the impedance matching networks to which each sensor is connected, which obviously cannot be all easily tuned at the same exact frequency. Thus, in order to maintain coherency among the various waveforms, signals would have to be “stretched” or “contracted” accordingly.

In theory, procedures can be implemented to recalculate at each new acquisition the updated delays, in order to mitigate the continuous and randomly occurring de-phasing of the signals, shown for example in Figure 4.64 and Figure 4.65 (only 2 receivers are compared in figures, for clarity; in the second figure only zoomed-in portion of the signals are shown, to highlight the dis-alignments; more than 10 sinusoidal-cycles are seen in the received experimental waveforms, even if 10-cycles are fed to the transmitter, explanations on this phenomenon are given in 4.3.2). For example, cross-correlations between the different signals can be computed, and the measured lags at which peaks are obtained can be used. However, those procedure would have to be performed in real time as the system is operating at high speeds, thus at very high testing repetition rates, with the risk of slowing down the real-time processing of the data.

Furthermore, other detrimental effects, such as the one listed as # (4), and clearly seen in Figure 4.64(b), are much more complicated to mitigate.

a) Position 1 (starting point). Alignment of experimental signals sensed by rec. #1 and rec. #2

b) ZOOM-IN

Figure 4.64: Alignment of experimental signals, based on the ones at the test starting point

a) Pos. 2 (~10 cm ahead of starting point). Random dis-alignment of the exp. signals (ZOOM-IN)

b) Pos. 3 (~1.5 m ahead of starting point). Random dis-alignment of the exp. signals (ZOOM-IN)

c) Pos. 4 (~2.5 cm ahead of starting point). Random dis-alignment of the exp. signals (ZOOM-IN)

Figure 4.65: Examples of signals randomly moving away from the initial aligned positions

Based on these considerations, in an attempt to mimic the role of all the detrimental effects occurring in actual testing conditions, a second set of LISA delay-and-sum-type analysis was performed after introducing small random variations of the exact delays calculated as in in Figure 4.63. Those variations were randomly set following a normal distribution with mean 0 and standard deviation 15° -phase-change. For example, Figure 4.66 show the first three outcomes of the modifications induced on the signal alignments by this process.

In both cases, once the waveforms sensed by the three receivers on each side were summed up together, from the newly formed pair of signals the features listed in Table 4.5 were extracted and fed into the algorithm for defect detection, in a similar fashion as for the case where the different pairs of sensors were treated separately. The analyses were performed using the “high” SNR factor $f = 1.5$, thus the results have to be compared to the ones discussed in Section 4.2.4.3.

For the sake of brevity, only the interesting results are plotted in comparison to the results described in Section 4.2.4.3. In particular, the ROC curves of the case studies of the TDs in the gage side at depths of 2 mm as well as the 25% H.A. TDs at increasing depths are not plotted, because using each of the three types of analysis 100% AUC have been achieved for those defects. Similarly, the HSH case is not shown because except for the 10 mm long defect, the others all had 100% AUC. However, in Table 4.10 and Table 4.11 all the details are given for the two delay-and-sum analyses, in a similar fashion as Table 4.8 that summarized the case described in Section 4.2.4.3.

As seen in the plots in Figure 4.67, a slight benefit is obtained using the “ideal” array type of analysis in the detection of TDs in the gage side. However, if the ideal case

is not respected (i.e. if slight misalignments of the signals occur) the performances drop below the level reached treating each pairs of sensors in an independent manner.

On the other side, for both TDs in the field side and VSH types of defects, averaging the information of the various sensors into one cumulative signal has the detrimental effect of decreasing the system detection performances, even when considering the “ideal” case.

Both the intrinsic decrease in performance regarding defects other than gage sided TDs, both the further decrease due to non-perfect alignments of the signals, for which also the case of TDs in the gage side deteriorate, suggests that using the available sensors as different entities is probably the best option. The results of a similar analysis carried out on the experimental data collected during the field tests in Colorado, whose results are presented in Section 5.9, confirms these conclusions.

Figure 4.66: Examples of variations on the initially perfect alignment of LISA signals induced to simulate realistic detrimental effects

Figure 4.67: Comparisons over the performances offered by the 3 considered types of analysis (10% TDs in the gage side at various depths)

Figure 4.68: Comparisons over the performances offered by the 3 considered types of analysis (TDs in the field side at depth of 2 mm)

Figure 4.69: Comparisons over the performances offered by the 3 considered types of analysis (VSH defects)

Table 4.10: Sensors treated as arrays in the “ideal” case: perfectly coherent sums (Figure 4.63). Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow

Model #	PD achieved in correspondence of “significant” PFA values				AUC
	0% PFA	1% PFA	5% PFA	10% PFA	
1	100	100	100	100	100
2	100	100	100	100	100
3	100	100	100	100	100
4	100	100	100	100	100
5	100	100	100	100	100
6	100	100	100	100	100
7	100	100	100	100	100
8	100	100	100	100	100
9	100	100	100	100	100
10	100	100	100	100	100
11	100	100	100	100	100
12	100	100	100	100	100
13	58.6	91	96.6	98.4	99.3
14	13.6	42.8	60.4	76.2	91.5
15	100	100	100	100	100
16	100	100	100	100	100
17	100	100	100	100	100
18	100	100	100	100	100
19	100	100	100	100	100
20	100	100	100	100	100
21	100	100	100	100	100
22	100	100	100	100	100
23	100	100	100	100	100
24	100	100	100	100	100
25	9.4	39.2	57	68.8	88.5
26	2.6	13.6	23.8	37	72.9
27	9.6	40.8	56	71.2	89.7
28	50.4	82	90.6	95.6	98.2
29	75.8	93.6	96.6	98.6	99.4
30	89	98.8	99.8	100	99.9
31	99.8	100	100	100	100.0
32	99.6	100	100	100	100.0
33	100	100	100	100	100
34	100	100	100	100	100
35	0.8	2.8	7.4	14	53.5
36	0.4	5.4	8.8	20.4	59.5
37	2.2	16.2	28.4	42.2	75.9
38	8.6	37	53.6	69.8	88.7
39	22.8	59.6	77.2	88.6	96.3
40	33.6	70.6	84.2	91.8	96.7
41	100	100	100	100	100
42	100	100	100	100	100
43	100	100	100	100	100
44	100	100	100	100	100
45	81.8	97.8	98.8	99.4	99.8
46	100	100	100	100	100
47	100	100	100	100	100
48	100	100	100	100	100
49	100	100	100	100	100

Table 4.11: Sensors treated as arrays in the “realistic” case: induced normally distributed dephasing with standard deviation 15° (Figure 4.66). Summary of the salient ROC curves results for the various models. Case of “high” SNR level ($f = 1.5$). The values lower than 100% are highlighted in yellow

Model #	PD achieved in correspondence of “significant” PFA values				AUC
	0% PFA	1% PFA	5% PFA	10% PFA	
1	100	100	100	100	100
2	100	100	100	100	100
3	100	100	100	100	100
4	100	100	100	100	100
5	100	100	100	100	100
6	100	100	100	100	100
7	100	100	100	100	100
8	100	100	100	100	100
9	100	100	100	100	100
10	100	100	100	100	100
11	100	100	100	100	100
12	100	100	100	100	100
13	50.2	75.4	90.6	95.6	98.0
14	14.4	29.8	56.8	69.8	88.0
15	100	100	100	100	100
16	100	100	100	100	100
17	100	100	100	100	100
18	100	100	100	100	100
19	100	100	100	100	100
20	100	100	100	100	100
21	100	100	100	100	100
22	100	100	100	100	100
23	100	100	100	100	100
24	100	100	100	100	100
25	8.2	27.0	49.0	61.8	85.4
26	2.0	8.8	26.4	38.6	72.7
27	13.0	27.6	54.4	69.4	89.0
28	49.4	70.4	88.2	92.8	96.7
29	65.4	85.0	93.8	97.0	98.9
30	78.0	90.6	98.2	99.0	99.5
31	94.2	97.6	99.6	99.8	99.9
32	98.2	99.4	100	100	100
33	99.6	100	100	100	100
34	100	100	100	100	100
35	0.6	2.2	9.2	14.4	53.0
36	1.0	3.4	11.4	18.8	58.0
37	2.2	9.2	26.6	42.0	75.0
38	13.4	30.6	57.8	70.6	90.5
39	30.2	51.2	72.6	83.4	93.4
40	28.8	57.4	84.4	90.0	96.5
41	99.2	100	100	100	100
42	100	100	100	100	100
43	100	100	100	100	100
44	100	100	100	100	100
45	68.4	89.4	98.4	99.6	99.6
46	100	100	100	100	100
47	100	100	100	100	100
48	100	100	100	100	100
49	100	100	100	100	100

4.2.5 Future works

The analyses regarding Transverse Defects at the gage side showed that the prototype's performance when considering this type of defects located at shallow depths is very satisfactory. As expected, when analyzing TDs deeper than ~10 mm the results start to worsen. It would be interesting to generate a new set of TDs located at high depths, both in the gage and in the field side, varying both their dimensions and their transversal positions on the rail cross-sectional plane, to characterize in a more detailed fashion their detectability. It would also be interesting to incline those defects, and see how this affects the results. However, the resolution at which those inclined defects can be created cannot be very accurate, because it is dependent on the grid size used to discretize the rail domain. At the moment, cubes of 0.5 mm edges are used to discretize the rail model, and the resulting model dimensions are already quite big (more than 60 million of cubes).

Considering that, from the simulations already performed, data has been saved for the waveforms collected by simulated sensors placed in many other positions (a 2.5×2.5 mm grid on top of the rail surface), a research can be conducted to understand if deploying the available receivers in positions other than the ones currently used (and which resulted from months of analyses over experimentally collected data) could increase the overall defect detection performance. Similarly, also the position of the excitation can be varied. Furthermore, other features than the ones considered up to this point can be evaluated.

For example, a genetic algorithm can be designed for features selection based on the already available data, in conjunction with the research of better sensors positions, to

form an overall optimization process, where the constant check on the goodness of the many parameters involved would be based on ROC curves as the ones presented in this chapter.

Furthermore, based both on LISA-generated numerical results, both on experimental results, pattern recognition algorithms can be also implemented to study the feasibility of not only detecting defects, but also quantifying them, in terms of predicting the flaw type, position, size and depth.

4.3 Finite element analysis of the complete experimental setup using LS-DYNA

4.3.1 Introduction

In order to simulate the complete ultrasonic wave propagation that takes place every time the prototype sends a tone-burst signal through the transmitter and receive it through one of its receivers, the air surrounding the examined rail portion has to be simulated as well. However, because the speed of sound in air is equal to 343 m/s (at a temperature of 20 °C and a pressure of 1 atm, according to Table 4.12) (that is longitudinal speed, a fluid does not support shear wave propagation), in order to fulfill stability and accuracy requirements a much finer cubic grid would have to be used in LISA, given the much reduced wavelength respect to the ones in the rail domain. cuLISA3D does not allow to model one material (rail steel) with a different cubic grid dimension than the other (air). This means that, in order to model even only a limited portion of air, the whole rail/air model would have to be discretized using the smallest cube dimension resulting from stability and accuracy conditions on the two materials.

Therefore, it is impractical to use cuLISA3D for such modeling, considering that the model would soon become too big to be handled.

At the same time, a much finer cubic grid would also be needed to model the curvature of the focused active surface of the transmitter with a good enough approximation.

To overcome these limitations, finite element analyses using the commercially available software LS-DYNA[®] were performed.

Table 4.12: Effect of temperature on properties of air (for an atmospheric pressure of 1 atm)

Temperature T (°C)	Speed of sound c (m/s)	Density of air ρ (kg/m³)	Characteristic specific acoustic impedance z_0 (Pa·s/m)
35	351.88	1.145,5	403.2
30	349.02	1.164,4	406.5
25	346.13	1.183,9	409.4
20	343.21	1.204,1	413.3
15	340.27	1.225,0	416.9
10	337.31	1.246,6	420.5
5	334.32	1.269,0	424.3
0	331.3	1.292,2	428
−5	328.25	1.316,3	432.1
−10	325.18	1.341,3	436.1
−15	322.07	1.367,3	440.3
−20	318.94	1.394,3	444.6

An explicit dynamic analysis was preferred over an implicit one. In fact, explicit analysis is more computationally efficient in simulating ultrasound wave propagations

problems. A good description of the factors in favor of the explicit integration can be found for example in [89].

While finite element implicit analyses are unconditionally stable, in explicit analyses numerical instabilities can arise if the integration time step is not small enough to correctly track the dynamic event evolution. If this requirement is not fulfilled, unstable solutions characterized by unrealistic displacement fields that usually oscillate with increasing amplitude are achieved. The time step size roughly corresponds to the transient time of an acoustic wave through an element using the shortest characteristic distance [90]. The LS-DYNA solver automatically determines the maximum time step through which it should iterate the solution, so this requirement did not present any issue.

The consistent system of units adopted through the analyses is shown in Table 4.13.

Table 4.13: Consistent system of units adopted for the LS-DYNA simulations

MASS	LENGTH	TIME	FORCE	STRESS
ton	mm	sec	N	MPa

The degree of freedom of the finite elements used to model air is the acoustic pressure, which is the local pressure deviation from the ambient pressure caused by an ultrasonic wave [91].

4.3.2 Finite element analysis of the pressure field generated in air by a cylindrically focused air-coupled transmitter

The pressure field as generated in air by an available cylindrically focused air-coupled transmitter was simulated and analyzed, and the salient results are presented in

this section. The final goal of such analysis was to obtain a highly accurate representation of the excitation exercised by the air-coupled transmitter to the rail top surface, to be used in analyses done at a later stage using cuLISA3D, which are described in Section 4.2.4.

Due to the short wavelength in air, it is important to limit the amount of air-domain modeled in the analysis, otherwise the computational cost would easily exceed the resources available. For this reason, only a column of air in front of the transmitter active surface was modeled. Perfectly Matched Layer (PML) material was put all around the column of air (thus at every face except the one corresponding to the focused active surface) in order to avoid unwanted wave reflections, thanks to a complete absorption of the ultrasonic waves traveling outside of the modeled air region. Figure 4.70 shows the complete model.

Figure 4.70: Finite element model to study the pressure field generated by the cylindrically focused transmitter. In light blue: air (the only visible face is the one corresponding to the focused active surface). In green: PML (there is a PML layer also in the hidden face, the one on the opposite side with respect of the active surface)

The PML technique was introduced in 1994 by Berenger for electromagnetism [92], and later has been extended to other fields such as acoustics [93] and other elastic waves [94, 95, 96]. PML elements perfectly match the impedance of the material of study, so that waves enter in PML without reflections (at least in theory; in practice, studies have shown that some numerical reflections occur, mostly due to insufficient numbers of element per wavelength employed in the model [97]). Once the waves enter the PML material, they exponentially decay. As those decaying waves reach the outer boundary of the PML, they are reflected back toward the area of study, while continuing to decay. Finally they will re-enter the area under test. Thus, the PML material has to be defined so that those reflections are negligible. For this reason, following the recommendation expressed in the LS-DYNA manual [98], 10 elements were modeled through the PML thickness. Furthermore, as expressly required by the same manual, all the nodes on the outer surfaces of the PML layer were fully constrained.

A typical element dimension of 0.2 mm was employed. The material type assigned to the air part is the *MAT_90*, which is a linear Eulerian element that was specifically formulated for tracking low pressure stress wave in an acoustic medium such as air or water. It can be used only with an acoustic pressure element formulation, either *ELFORM 8*, which is a 1-point integrated volume element, or *ELFORM 14*, which is the fully integrated version of the previous one. Only the pressure information is available to the user with respect to such elements [99]. 1-point integrated *ELFORM 8* hexaedral elements were employed throughout the air-domain.

MAT_231 was assigned to the PML part, which is a material that requires the use of the *ELFORM 14* element formulation. The material properties of the PML were set identical as the ones assigned to air, which are described below.

Among the parameters that the user needs to set when using the *MAT_90*, there are mass density and sound speed, which were set equal to the values at 20 °C in Table 4.12 (1.204 kg/m³ and 343 m/s, respectively), the atmospheric pressure that was set to 0 (thus the obtained pressure values can be read as relative to atmospheric pressure), and a damping factor which can range from 0 to 1. The LS-DYNA documentation does not specify exactly how this factor is measured. After a few iterations of simulations the damping factor value was set to 0.01, based upon comparison against some experimental tests that are reported at the end of this section.

Figure 4.71 shows the modeled column of air in front of the cylindrically focused transmitter, from a side view that highlights the curvature of the active surface. The curved surface on the left is shaped exactly as the active surface of the cylindrically focused transmitter (Figure 4.72). The outer face of the elements on the curved surface were selected (Figure 4.73) and subjected to a pressure time-history as the one depicted in Figure 4.74. This provides a good simulation of the actual excitation provided by the transmitter in the experimental tests. However, the actual piezoelectric air-coupled transducer is characterized by an internal “ringing” effect due to resonance at the center frequency of the transducer, so that when a voltage signal is fed to the transmitter, the actual active surface oscillations will present a short tail after the transmitted voltage signal has ceased. Although this effect cannot be captured exactly by the proposed modeling, the differences caused by this behavior are not substantial for the purpose of

the simulations. This behavior cannot be exactly captured if not employing an advanced modeling of the actual transducer, which would have to take into account: the piezoelectric material properties; all the impedance matching layers placed in front of the piezoelectric element; the outer environmental protection layer (more details on air-coupled transducers fabrication can be found in [100]).

The efficiency of an air-coupled transmitter is defined as the pressure produced in air, measured in Pascals per Volt applied to the transducer. The transmitter efficiency was not available from the manufacturer. As a first approximation, without loss of generality, and based on efficiency values found in literature for similar air-coupled transducers, a peak value of 2,000 Pa was assigned at the pressure time-history applied to the air-part top surface (Figure 4.74). A procedure that allows a measurement of the pressure field produced by an air-coupled transducer is reported in [101], but was not employed in this study.

Figure 4.71: Air part salient dimensions

Figure 4.72: The transmitting transducer, having 25×25 mm active surface and 38 mm focal distance

Figure 4.73: Highlighted in white are the faces of the solid elements subjected to the pressure time-history depicted in Figure 4.74 in order to mimic the ultrasonic waves released by the cylindrically focused air-coupled transmitter

Figure 4.74: Pressure time-history applied to the elements shown in Figure 4.73

Six sets of elements were selected, as shown in Figure 4.75, and the pressure time-histories experienced by the correspondent elements were recorded.

Figure 4.75: Sets of elements considered in the analysis

Five planes normal to the wave propagation direction (at distances from the tip of the active surface of 14, 23, 35, 38 and 41 mm) were considered to visualize how the wave energy gets focused toward a line about 25 mm long (which is the width of the active surface) and about 38 mm distant from the active surface (considering that 38 mm is the focusing distance). The contour plots in Figure 4.76 and Figure 4.77 show top views and 3D views of the peak-to-peak pressure as recorded by the elements in those planes. The shape at 38 mm is the most relevant for this study, as that it is the distance at which the transmitter is held above the rail top surface in the experimental tests. The contour plot shows the presence of two peaks symmetrically located at a distance of about 5.5 mm from the central lower peak. This can be better appreciated in Figure 4.78. Figure 4.79 clearly shows how at the focus distance (and also at its proximity) the energy

is almost all contained in a beam having width of about 3 to 4 mm. Very low ripples are present outside the main beam. No significant pressure variation is seen as long as one remains in the proximity of the focus distance (from 35 to 41 mm, i.e. ± 3 mm from the focus), which is a good outcome considering that when the transmitter is put in motion while it is held above the rail, dynamic vibrations of the cart holding the sensor, as well as possible changes in shape of the rail top surface (mostly due to wear induced by passing wheels), will continuously slightly change the distance between active surface and rail top surface, deviating it from the exact focal distance. To investigate further about this aspect, the elements lying on the plane containing the tip of the active surface and oriented parallel to the wave propagation were selected, as shown in Figure 4.75, and their pressure time-histories were stored during the analysis. Then, the average pressure acting on the elements lying on each line parallel to the tip of the active surface was computed, and the resulting peak-to-peak values are plotted in Figure 4.80. As long as one remains between ~ 32 mm to ~ 41 mm away from the transmitter, the maximum pressure variation is only on the order of 5%. The figure reveals that the “near-field” region of the transducer ends at a distance of about 30 mm. This is the region where the ultrasound beam is greatly affected by constructive and destructive wave interferences that lead to significant fluctuations in the pressure. Once the near-field ends, the so-called “far-field” begins, which is the region where the sound pressure, after reaching its maximum (which seems to be at approximately 36 mm in Figure 4.80), gradually drops to zero.

Figure 4.76: Top view of the contour plots showing pressure fields pertaining to 5 planes lying normally to the wave propagation direction

Figure 4.77: 3D view of the contour plots showing pressure fields pertaining to 5 planes lying normally to the wave propagation direction

Figure 4.78: Variation of pressure along lines directed as the y-direction at 12.5 mm if considering Figure 4.76

Figure 4.79: Variation of pressure along lines directed as the x-direction at 12.5 mm if considering Figure 4.76

An experimental test was performed to validate the simulation. A PICO sensor from Physical Acoustics was used to receive the signal generated from the air-coupled

transmitter. Due to the finite physical dimensions of the PICO receiver (the diameter of the active surface is equal to ~ 4.8 mm) and the cumbersomeness associated with acquiring many acquisitions while maintaining the required accuracy on positioning the PICO sensor, the spatial resolution of the experimental test is sensibly smaller than the one available from the simulation. Figure 4.81 shows the peak-to-peak of the signals acquired as the receiver was moved away from the transmitter, with steps of 5 mm. A similar behavior as the one depicted in Figure 4.80 can be noticed. However, a finer spatial resolution would be needed to correctly capture the bumps in the ‘near-field’ region.

The damping factor assigned to the air material *MAT_230* was determined comparing the simulation results (as the one plotted in Figure 4.80, which is valid for a damping factor of 0.01, the value that was finally chosen) to the experimental plot reported in Figure 4.81. Lower damping factors would make the pressure values in the plateau in proximity of the focus distance to increase more with respect to the initial value (approx. 4,000 Pa peak-to-peak, which is the pressure applied to each element forming the transmitter active surface), while higher damping factors will results in the plateau getting lower. The attenuation in air depends by the frequency, the temperature and the atmospheric pressure.

Figure 4.84 shows a comparison between the pressure fields on the plane 38 mm away from the transmitter as obtained through the LS-DYNA simulation and an experimental test (Figure 4.82 and Figure 4.83). Peak-to-peak voltage values were recorded as the PICO sensor was moved in a 2.5×2.5 mm grid, for a total of 100 spatial points. Excellent agreement is found. The two higher peaks symmetrically placed around

the lower central peak match considerably well. The beam width is also very similar, even if the 2.5 mm spatial resolution used in the experimental test does not allow an exact comparison. The same consideration applies for the small ripples outside the main beam.

Figure 4.80: Pressure variation moving away from the transmitter active surface. LS-DYNA simulation results

Figure 4.81: Pressure variation moving away from the transmitter active surface. Experimental results using a PICO sensor as receiver

Figure 4.82: Experimental test to assess the pressure field generated by the transmitter

Figure 4.83: Experimental test to assess the pressure field generated by the transmitter

Figure 4.84: LS-DYNA (left) vs experimental test (right) results. Pressure field in a plane 38 mm away from the transmitter

4.3.3 Simulation of the complete ultrasonic wave propagation through air and rail

Just as for the LISA model in Section 4.2.2, to reduce the computational effort only the rail head and part of the web were modelled (Figure 4.10). However, here the length of the model was set to 30 cm. PML material, with 10 elements through the layer thickness, was placed at the boundaries of the rail model (Figure 4.85) to absorb the ultrasonic waves traveling outside of the modeled rail region. The nodes on the outer surfaces of the PML layers were fully constrained (Figure 4.86).

Rail material properties

The isotropic hypoelastic material *MAT_001* was assigned to the rail, specifying the same values of density, elastic modulus and Poisson's ratio as the ones used in Section 4.2.2 for LISA analyses (7,850 kg/m³, 200 GPa and 0.29, respectively). Identical values were assigned to the PML material (*MAT_230*).

Figure 4.85: LS-DYNA model. In red: rail. In light blue: PML material. In green: air. Note: no PML was used at the air boundaries

Figure 4.86: Boundary conditions applied at the nodes on the outer surfaces of the PML layers

Rail mesh geometry and elements formulation

As already stated in previous sections, in order to achieve accurate results the mesh dimension should not exceed $\lambda_{min}/10$. Therefore, a typical element dimension of 1.5 mm was employed for the rail material. Previous studies have shown how the use of highly irregular elements characterized by different shapes or sizes leads to a change of acoustic impedance that generates unwanted numerical wave reflections through the model [97]. This observation led to the choice of using linear hexaedral elements trying to keep the highest possible regularity while still allowing the modeling of the desired defect shapes. Namely, the rail mesh was developed using 8-node linear hexaedral elements with constant stress *ELFORM 1*, which is an efficient and accurate element formulation. For PML, the mesh geometry was kept identical to the rail one, but a fully integrated *ELFORM 2* element formulation was used, as required by the LS-DYNA

documentation [98] for PML *MAT_230*. The mesh was first deployed in the cross-sectional plane and then it was extruded along the rail direction, using 1.5 mm intervals all the way up to the total length of 30 cm. To allow for a precise modeling of a round 15% H.A. Transverse Defect, the mesh at both gage and field side presents circular patterns as depicted in Figure 4.87. Figure 4.88 shows the maximum side length distribution in the rail cross-section. The figure shows that only few elements slightly exceed the requirement of 10 elements per smallest wavelength. Figure 4.89 and Figure 4.90 illustrate the mesh quality by plotting the Aspect Ratio and the Jacobian distribution in the rail cross-section respectively. The Jacobian is always above a value of 0.6, which indicates a good quality of the proposed mesh.

Figure 4.87: Finite element mesh of AREMA 136RE rail cross-section

Figure 4.88: Finite element mesh quality. Maximum side length in the rail cross-section

Figure 4.89: Finite element mesh quality. Aspect ratio distribution in the rail cross-section

Figure 4.90: Finite element mesh quality. Jacobian distribution in the rail cross-section

Defect modelling

A 15% H.A. TD located in the gage side was created by disconnecting the adjacent elements at the nodes along the desired crack shape. Those nodes are indicated in Figure 4.91. This approach for modeling cracks has been reported to be reliable in previous studies of guided wave interactions with defects [102]. This defect was positioned in the cross-section located in the middle of the modeled rail (which is 15 cm away from both ends) (Figure 4.92). In a similar fashion, the same type/size of defect but located in the field side was created. For the sake of brevity, only the results from the model with the defect in the gage side are presented. A more detailed study that focuses on the sensitivity of the system to different types of cracks is presented in Section 4.2.4, where LISA simulation results from models containing a variety of different types/sizes/location of defects are shown.

Figure 4.91: 15% H.A. TD - Defect modeling details

Figure 4.92: 15% H.A. TD located in the middle of the modeled rail

Air modeling

As already discussed in the previous section, it is important to limit the amount of air modeled in the analysis, to keep a low computational cost. For this reason, only two small portions of air were modeled, a column of air between transmitter and rail top surface, and a column of air between few possible receiver locations and the rail top surface. To save computational cost, the receiver on only one side around the transmitter was modeled (Figure 4.85).

Instead of using PML as for the air model in Section 4.3.2, Non-reflecting boundary conditions (NRBC) [103, 104] were used to avoid reflections of waves traveling outside the air domain, and this allowed for a reduction of the air domain to the minimum possible extent. NRBCs are special boundary conditions used in finite element methods to model wave propagation in unbounded media. The advantage of using them over PML is that the dimensions of the model remain unchanged, only the boundary conditions are changed, with the introduction of extra variables to approximate the infinite expanse of the medium.

The same material (MAT_90) and element formulation (ELFORM_8) of the air model in Section 4.3.2 were used.

Transmitter side

Figure 4.93 shows the modeled column of air below the cylindrically focused transmitter, from a side view that highlights the curvature of the active surface. The only geometric difference of this air part against the one described in Section 4.3.2 is that the length in the direction of the wave propagation has been reduced to 38 mm, counting

from the tip of the active surface to the rail top surface. The top faces of the upper elements of this air part (Figure 4.94) were selected and subjected to a pressure time-history as the one depicted in Figure 4.74.

Figure 4.93: Air column below the transmitter and in contact with the top surface of the rail

Figure 4.94: Highlighted in white are the faces of the solid element subjected to the pressure time-history as in Figure 4.74 in order to mimic the signal released by the cylindrically focused air-coupled transmitter

Receiver side

Figure 4.95 and Figure 4.96 show the modeled air domain below the receivers. The height of such part was limited to 10 mm (this is the averaged value, as shown in figure) to save computational effort; the actual receivers lift-off distances used in the prototype can be much higher (in the current prototype version is ~6 cm). In Figure 4.96 in white are highlighted the actual dimensions of the active surface of one of the air-coupled receiver. This air portion was made significantly longer than what strictly needed in the transversal direction with respect to the rail. The idea behind such choice was that by a rapid change of selected elements it was possible to evaluate the benefit of placing the receiver completely toward the gage side as well as completely toward the field side as well as in any position in between. The top surface of this air domain is inclined by 6.4° with respect to the horizontal in order to mimic the actual position of each receiver above the rail top surface in the experimental tests. The receivers are in fact inclined at a similar angle to best catch the leaky surface waves propagating in the rail head. Explanations on why this is needed are provided later in this section.

In a similar fashion as for the transmitter case, once the active surface of the transducer is put in motion by a pressure wave at the receiver center frequency, internal “ringing” occurs, and the converted electric signal transmitted to the scopes will present a short tail after the ceasing of the actual pressure wave propagating through air, which cannot be exactly captured by this model.

Figure 4.95: Air column below the receivers

Figure 4.96: Highlighted in white are the elements on the top surface of the air-part that mimic the extension of the active surface of an air-coupled receiver

The relative distances between transmitter, defect and receiver used in the model are shown in Figure 4.97.

Figure 4.97: Relative distances between transmitter, defect and receiver

Because the air mesh is much finer than the rail one, the element nodes of the two meshes do not coincide. Therefore, a contact option had to be introduced to link the two meshes. The LS-DYNA *BOUNDARY_ACOUSTIC_COUPLING_MISMATCH* card is specifically made to facilitate contacts between acoustic elements and structural ones. Therefore such contact was employed to ensure contact between both the transmitter-air-domain with the rail-domain, and the receiver-air-domain again with the rail-domain.

Figure 4.85 shows the full domain of the 3D finite element simulation. The model contains 4,588,971 nodes and 4,466,725 elements. The same PC as the one described in Section 4.2.2 was used to run the simulations. Each analysis was iterated until 300 μsec , for a running time of 49 minutes (double precision 64-bit floating point numbers were used).

Figure 4.98: Section cut made on the model to highlight the salient features of the results. In greyscale is the removed portion of the model

Simulations results

Figure 4.99 and Figure 4.100 show some results obtained using the rail model with the 15% H.A. transverse defect on the gage side (as in Figure 4.91 and in Figure 4.92). A section cut was operated as shown in Figure 4.98 to highlight the salient features of the results. The contour plots of the pressure field pertaining to six different instants in time of the simulation are shown. Only positive pressures are plotted to facilitate the readiness of the figures. The scale factor relative to the air-part placed below the receiver had to be reduced by a factor of 10^5 to allow the wave propagation visualization. The values range from 0 to 10^{-3} MPa for both the rail-part and the air-part below the transmitter, while they range from 0 to 10^{-8} MPa for the air-part below the receiver.

The factor 10^5 that is needed to correctly visualize the leaky “Rayleigh” wave emanating from the rail top surface into the surrounding air is in agreement with what expected based on acoustic impedances of both air and rail. The specific acoustic impedances of air and steel are:

$$\begin{aligned}
 Z_{air} &\cong 420 \text{ Rayls} \\
 Z_{steel} &\cong 45 \text{ MRayls}
 \end{aligned}
 \tag{4.35}$$

This large impedance mismatch between those two materials is responsible for the big loss of energy at the interfaces between the two media. The expressions (in terms of particles velocity) for the reflection and transmission coefficients valid for a plane acoustic wave impinging normally on a straight boundary between two media are:

$$\begin{aligned}
 R &= \frac{Z_2 - Z_1}{Z_2 + Z_1} \\
 T &= \frac{2 \cdot Z_2}{Z_2 + Z_1}
 \end{aligned}
 \tag{4.36}$$

where Z_1 is and Z_2 are the specific acoustic impedances of the first and second medium respectively. Thus, the reflection and transmission coefficients when the ultrasonic wave impinges on the rail top surface coming from the air and the ones pertaining to the continuous leakage from the rail top surface back into air are (the values for the steel→air path are only approximate because the given formula is derived for cases of wave normal incidence to the boundary between the two materials):

$$\begin{aligned}
 R_{air \rightarrow steel} &= \frac{Z_{steel} - Z_{air}}{Z_{steel} + Z_{air}} \cong 1 \\
 R_{steel \rightarrow air} &\cong \frac{Z_{air} - Z_{steel}}{Z_{air} + Z_{steel}} \cong -1 \\
 T_{air \rightarrow steel} &= \frac{2 \cdot Z_{steel}}{Z_{steel} + Z_{air}} \cong 2 \\
 T_{steel \rightarrow air} &\cong \frac{2 \cdot Z_{air}}{Z_{air} + Z_{steel}} \cong 1.86 * 10^{-5}
 \end{aligned}
 \tag{4.37}$$

Focusing the attention on the first transmission coefficient ($T_{air \rightarrow steel}$), this means that once the waves coming through air from the transmitter impinge on the rail top surface, the pressures of the waves generated in steel rail are of the same order of magnitude as the air pressures. Recalling that the peak pressure applied to the air elements mimicking the transmitter was equal to 0.002 MPa, this explains why the range $0-10^{-3}$ MPa is a good fit for both rail-part and air-part below the transmitter. After that, due to the second transmission coefficient ($T_{steel \rightarrow air}$), the waves leaking back into air from the rail top surface have a magnitude 10^5 times smaller than the former ones. Thus this explains why the range $0-10^{-8}$ MPa is a good fit for the air-part below the receiver.

After 50 μ sec of simulation (Figure 4.99a) almost all of the applied tone-burst has been released by the transmitter (N.B. the signal in Figure 4.74 ends at ~ 59 μ sec). It is shown how the curved wavefront is focusing the energy toward a line at the center of the air-part and normal to the plane of the figure. The red arrows in this plot and in the following ones are added to avoid confusions regarding the direction of propagation of the different trains of waves.

The wave traveling in air at a speed of 343 m/s takes about 110 μ sec to cover the 38 mm between top and bottom of the air-part. Thus after 125 μ sec (Figure 4.99b) part of the wave has already entered the rail top surface. From those points of contacts both bulk waves (longitudinal and shear) and guided waves are generated, mostly Surface Acoustic Waves.

The surface wave traveling in steel at a speed of roughly 2,900-3,000 m/s takes about 40 μ sec to cover the 12 cm between the point where the sound waves entered in rail and the defect location. Once the wave encounters the crack, interaction with the latter

occurs. There exists a vast literature (for example [105, 106, 107]) about the interaction phenomena between Rayleigh waves and cracks in solids. Part of the wave is transmitted through the crack location; part is reflected back in opposite direction (clearly visible in Figure 4.100d-f); part propagates along the edges of the crack (visible in Figure 4.99c); finally, mode conversions will occur, and bulk waves are generated from the tips of the crack which act as secondary sources (visible in Figure 4.99c and Figure 4.100d).

As the surface waves travel along the boundaries of the rail, small leakages into surrounding air occur. This is well captured by the simulation, as shown in Figure 4.99c to Figure 4.100f, in correspondence to the modeled portion of air. Snell's law dictates the angle at which such leakage occurs. The well-known formula states that the ratio of the sines of the angles of incidence (θ_1) and refraction (θ_2) for a sound wave passing through a boundary between two different media is equivalent to the ratio of phase velocities in the two media, i.e.:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2} \quad (4.38)$$

where θ_1 and θ_2 are measured from the normal at the boundary between the two media. Substituting the actual values, which are 90° for θ_1 , 2,911 m/s for v_1 (this is the velocity of the Rayleigh wave in steel as computed in eq. (4.34), which can be used here as a good approximation) and 343 m/s for v_2 yields a value of θ_2 of about 6.7° . The exact value will slightly differ due to the fact that the speed of the surface waves traveling at or below the rail top surface won't be exactly the Rayleigh velocity. After a few iterations in order to find the best surface inclination being as parallel as possible to the wavefront, a value of 6.4° for the inclination of the air-part top surface was determined and used.

Other ~ 30 μsec are needed by the ultrasonic waves to cover the ~ 10 mm between the rail top surface and the inclined top surface of the air column below the receiver.

Summing up the calculated times needed by the wave to propagate along the three paths air-rail-air, the first arrival as seen by the receiver is expected after about $110+40+30=180$ μsec , as correctly shown in Figure 4.101.

One last thing to be noted in Figure 4.99c and Figure 4.100d-f is that after the train of waves launched by the transmitter impinges on the rail top surface it gets reflected almost in its entirety, due to a reflection coefficient in terms of pressure approximately equal to 1 (from eq (4.37)). This creates an issue that is discussed in Section 5.7.

As already mentioned above, the active surface of an air-coupled receiver was simulated through selection of a specific set of elements in the top surface of the air-part. As a first approximation, the mean pressure on such set of elements simulates what would be recorded in actual experimental tests. In fact, the output of an actual receiver is a voltage signal proportional to the acoustic pressure sensed by its active surface.

For the sake of brevity, only the results of the analyses performed when the receiver was placed in the exact position shown in Figure 4.96 (all the way toward the gage side) are presented. Figure 4.101 (zoomed version in Figure 4.102) shows a superimposition of the waveforms received when the pristine rail model (in blue) and the model with the 15% H.A. defect (in red) are used. As a direct result of an attenuation of the transmitted wave traveling through the defect, the calculated Root Mean Square (RMS) value of the ‘defective’ waveform is about 70% of the RMS value of the ‘pristine’ waveform ($\sim 1.22 \times 10^{-9}$ MPa vs $\sim 1.76 \times 10^{-9}$ MPa).

Figure 4.99: Contour plot of pressure after (a) 50 μsec , (b) 125 μsec and (c) 180 μsec

Figure 4.100: Contour plot of pressure after (d) 205 μsec , (e) 215 μsec and (f) 225 μsec

Figure 4.101: Waveforms received from the simulated receiver in the models with and without defect

Figure 4.102: Zoomed version of Figure 4.101 to highlight the salient part of the waveforms

4.3.4 LS-DYNA aid in the design of a dedicated holder for a newer transmitter

As the TTC field test was approaching, one of the upgrades made to the prototype in order to increase the available SNR was the purchase of a bigger cylindrically focused transmitter. The focal distance was doubled at 76 mm vs 38 mm and, more importantly, its active surface was 4 times bigger than the previously used transmitter. The latter is a key upgrade in terms of excitation energy released by the transmitter and delivered to the rail top surface. The reason is self-evident: as a first approximation, supposing, for simplicity, that through the entire active surface of each transmitter a uniform ultrasonic pressure is generated on the adjacent particles of air, and admitting an identical efficiency of the two transmitters (in terms of pressure generated in air - per unit area of active surface - per voltage), then the integration through the areas of each of the two active surfaces yields a total pressure in air (per voltage) four times bigger when using the larger transmitter in place of the smaller one (actually it will be slightly less than four, due to the longer distance traveled toward the focal distance, 76 mm vs 38 mm, which exposes the ultrasounds to higher absorptions). A drawback is that now the line over which the signal is focused is roughly doubled in length, from ~25 mm (at a distance of ~38 mm) to ~50 mm (at a distance of ~76 mm). Efforts were aimed at reducing this line back to a length of about 25 mm, in order to keep a similar non-centered excitation on the rail top surface (for the reasons given in Section 3.3.1), and at the same time trying to avoid any loss of precious available sound pressure.

As a matter of example, the easiest (and least efficient) method for keeping the excitation on a ~25mm-long line would be to simply place the transmitter approximately

halfway out of sight of the rail top surface, as depicted in Figure 4.103a. Obviously, a big portion of the available energy (in terms of air ultrasonic pressure) would be lost in this way. For this reason, LS-DYNA finite element analyses were performed ultimately leading to the design of a special adapter for the transmitter, which was then finalized in AUTOCAD[®] environment (Figure 4.104a). From the 3D-geometry, the adapter was 3D-printed in ABS material (Figure 4.104b) using an AW3D HDX printer from Airwolf 3D[®], a company located in Costa Mesa, CA. Two AUTOCAD views of the transmitter adapter/holder are shown in Figure 4.103b.

Figure 4.103: (a) Least efficient method to keep a 25mm-long line of excitation and (b) final solution

Figure 4.104: (a) AUTOCAD® design and (b) 3D-printed transmitter adapter/holder. The three holes shown on the left side allow for a secure connection of the holder to the other parts through the use of screws, while still permitting to rotate the transmitter as wanted. The single hole in the other face is used to firmly secure the transmitter once it is put inside the holder

LS-DYNA models description and simulation results

In this section the simulation results from two models are presented. The first model (in the following denoted as model ‘a’) simulates the ultrasonic wave propagation field as generated by the air-coupled transmitter as-is, without the special adapter (Figure 4.105a). The second model (denoted as ‘b’) shows the beneficial effects of the adapter when it is applied to the transmitter (Figure 4.105b).

Acoustic hexaedral *ELFORM* 8 elements were employed to model the air parts in the two models. *MAT_230* was assigned to those elements, with the same material properties as the ones reported in the previous sections (density of 1.204 kg/m^3 , speed of sound equal to 343 m/s , damping factor of 0.01). NRBCs were employed at the outer boundaries of the air parts to allow for the ultrasonic wave to exit out of the model

without significant reflections coming back inside. In model 'b', the adapter was modeled using hexaedral elements *ELFORM 1* and elastic material *MAT_001* assigning material property values typical of ABS plastic (a more advanced elasto-plastic modeling for ABS was not needed in these simulations where the adapter was not subjected to any significant stress). To exclude unwanted possible numerical contributions to the wave propagation due to the employed meshes, which could affect the two models in an unpredictable and dissimilar way, the mesh geometry was maintained identical in the two models, to the highest possible extent. For example, the green and blue parts shown in Figure 4.105a have the same mesh as the black and blue parts shown in Figure 4.105b. However, air material is assigned to the green part of model 'a', while ABS material is assigned to the black part of model 'b'. The parts seen in transparent style in the two figures are air parts, and are plotted in this style only to facilitate the readiness of the models. The selected elements (in white) in the two models are the same ones. Those are the elements lying on a 25×10 mm rectangle located 76 mm away from the tip of the active surface (this and other salient dimensions are shown in Figure 4.107). The active surface, which is subjected to the pressure time-history shown in Figure 4.74, is the blue surface visible in both plots of Figure 4.105, and have lateral dimensions of 50×50 mm. The air part visible right in front of the ABS adapter in Figure 4.105b is also modeled in model 'a' with identical mesh. However, in model 'b' its extension was prolonged in the direction of the wave propagation. That is because the pressure field at further distances from the active surface was studied in that model. In model 'a', one additional outer layer of air was modeled all around the other air parts in order to make sure that a correct propagation of all the wave contributions traveling from the active surface toward every

element in the white selection was allowed. In particular, the top view in Figure 4.106 shows that, without the extra layer, the direct path between some elements on the edges of the active surface and some on the edges of the white selection would be missing. It is important to stress that, although the different air parts are represented with different colors, and lines of separation are shown in the various plots, really there are no boundaries in between those parts. As explained before, the only reason why these many parts were created was to maintain an identical mesh between the two models, to the highest possible extent.

Figure 4.105: (a) Model 'a': the green part is air, and so are all the other parts as well. (b) Model 'b': the black part is ABS, all the others are air

Figure 4.106: Model 'a': top view (left) and side view (right)

Figure 4.107: Model ‘b’: top view (left) and front view (right). Some salient dimensions of the models are shown here

Simulation results are shown in Figure 4.108 through Figure 4.111. In particular, the pressure fields obtained in the two models at four instants in time (namely, after 60, 140, 200, 260 μsec since the start of simulation) are shown. Two different points of view are used, a top view and a side view, which are defined for example in Figure 4.106. However, here section cuts were operated in order to highlight salient aspects of the results. Basically, for each view and for each model, the half of the model pointing in the direction of the camera has been removed. Therefore, the pressure fields seen in the figures are the ones of the planes passing by the center of the active surface. The contour plots share the same colorbar, ranging from cold colors for the lower values to hot colors for the higher values, where the chosen extreme values are 0 and 2,000 Pa respectively (thus only positive pressures are plotted to facilitate the readiness).

A comparison of the top views in Figure 4.108 and Figure 4.109 reveals the beneficial effects of the two walls of the adapter visible in these plots (the lateral walls). Each wall runs from the corresponding active surface edge, at an orthogonal direction respect to the active surface itself. While in the absence of such walls (model ‘a’, Figure

4.108) part of the generated air pressure will inevitably and constantly leak away from the focusing beam as the ultrasonic wave propagate toward the focal distance, when the adapter is in place (model 'b', Figure 4.109) those contributions are not lost, and will reinforce the beam. The walls are interrupted 14 mm before reaching the focal distance of 76 mm (Figure 4.107), to avoid any contact between adapter and rail top surface in the experimental tests. The colors in the plots readily suggest a significant gain of amplitude of the ultrasound reaching the focal distance in the two cases.

In a similar manner, Figure 4.110 and Figure 4.111 show the beneficial effects of the other two walls of the adapter (top and bottom wall). Without the adapter, part of the available energy would go lost above and below the beam as the wave propagates toward the focal distance. Here the inclination of the walls is such that the waves are 'guided' to fit in a line 25mm-long at the focal distance, as it was originally desired.

Figure 4.108: Model 'a': contour plots of pressure after 60, 140, 200, 260 μs . Top view (section cut operated halfway through the height of the model)

Figure 4.109: Model 'b': contour plots of pressure after 60, 140, 200, 260 μ sec. Top view (section cut operated halfway through the height of the model)

Figure 4.110: Model 'a': contour plots of pressure after 60, 140, 200, 260 μsec . Side view (section cut operated in correspondence of the tip of the active surface)

Figure 4.111: Model 'b': contour plots of pressure after 60, 140, 200, 260 μsec . Side view (section cut operated in correspondence of the tip of the active surface)

To precisely quantify the gain obtainable using the 50×50 mm transmitter over the 25×25 mm one, and in turns, the additional gain when using the proposed adapter, the time varying mean pressure in the elements forming the white selection shown in Figure 4.105 through Figure 4.107 (rectangle of dimensions 25×10 mm, as already expressed above) was computed and plotted in Figure 4.112. To obtain the black signal shown in the graph, in the model of the 25×25 mm transmitter, described in Section 4.3.2, the elements lying in an identical 25×10 mm rectangle located at the focal distance (38 mm in that case) were selected and in the same manner the mean pressure was computed. This signal arrives earlier than the other two (~110 vs ~220 μ sec), given the path between active surface and focal distance being exactly half. To facilitate the comparisons, the three signals are aligned and superimposed in Figure 4.113. The peak-to-peak was computed for each plot and then normalized over the value relative to the black waveform. It appears that a substantial gain of ~220% is achieved when using the 50×50 mm transmitter with adapter attached, in place of the 25×25 mm transducer, which makes it a key upgrade toward a consistent defect detection.

Once the larger transmitter was purchased, and once the adapter was 3D-printed in ABS plastic, this predicted SNR gain was confirmed, with the waveforms sensed by the air-coupled receivers that essentially doubled their peak-to-peak values as the 50×50 mm transmitter with its adapter was used in place of the 25×25 one. Unfortunately, it is not trivial to make a rigorous comparison through experimental testing as the one numerically done and plotted in Figure 4.113. It is in fact very difficult to ensure that as one switches from a transmitter to the other while performing the experiments, the exactly identical conditions are kept. In particular, when trying to perform the

comparison pointing each transmitter to the same specimen (as could be the rail top surface), or alternatively toward the active surface of a transducer receiving the signal directly through air, it is quite difficult to make sure that the ultrasonic beam is hitting exactly the same region, exactly at the focal distance, and with the specimen surface being exactly orthogonal to the oncoming beam.

Figure 4.112: Comparison of the air-pressure waveforms generated by the modeled 25x25mm transmitter and the modeled 50x50mm one, when used with or without the adapter

Figure 4.113: Comparison of the air-pressure waveforms generated by the modeled 25x25mm transmitter and the modeled 50x50mm one, when used with or without the adapter. The three signals have been aligned in time to facilitate the comparison

4.4 Conclusions

Two different numerical computational techniques were used to analyze the ultrasonic wave propagation occurring each time the experimental prototype executes a test. In particular, a software based on Local Interaction Simulation Approach (cuLISA3D) was utilized to get insights about the multi-modal wave propagation in the rail head, and the results showed that the air-coupled transmitter used in the prototype excites several Surface Acoustic Wave modes interacting with each other. The same software was also used to investigate the limits of the detectability of different types of defects, according to their sizes and depths, and when considering the currently available signal-to-noise ratios. The results of such analyses are already satisfactory when considering the sensors positions and the statistical analysis details used in the current version of the prototype. Nevertheless, further studies can be performed on similar datasets to pinpoint ways to maximize the system's defects-sensitivity.

The second approach that has been used to simulate air-coupled ultrasonic wave propagation phenomena is based on Finite Element Analysis, performed with a commercially available software (LS-DYNA). This software was chosen because of its well-known capabilities in performing explicit analysis, which are more convenient than implicit ones when dealing with ultrasonic simulations involving high frequencies. Taking advantage of this software's capabilities, the complete wave propagation occurring not only in the rail under test, but also in the surrounding air domains was simulated. Further studies were carried out both to fully characterize the type of excitation provided by the air-coupled transmitter in use, and to maximize the energy transferred from its active surface to the rail top surface.

Comparisons between numerical results and experimental results from laboratory and field tests have been systematically carried out and consistently confirmed the correctness of the models predictions. In fact, during the course of the research project, many of the prototype's operational parameters were chosen not only from analyzing experimental test results, but also from utilizing predictions and suggestions provided by these numerical models.

The next chapter will describe in detail the current stage of prototype's development, and will discuss the results achieved in the first field test done out of UCSD in October 2014.

4.5 Acknowledgments

This chapter, in part, has been published in the *Structural Health Monitoring Journal*, Mariani, S., Nguyen, T., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., Staszewski, W. J., Fateh, M., and Carr, G. (2013). The title of this paper is "Non-contact air-coupled ultrasonic guided wave inspection of rails". The dissertation author was the primary investigator and primary author of this paper.

Part of this chapter will be submitted for publication in *Transportation Research Record: Journal of the Transportation Research Board*, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is "Air-coupled wave propagation in steel rail". The dissertation author will be the primary investigator and author of this paper.

Part of this chapter will be submitted for publication to the *Structural Health Monitoring Journal*, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is "Defect detection performance of the non-contact air-coupled ultrasonic guided

wave inspection of rails prototype”. The dissertation author will be the primary investigator and author of this paper.

Part of this chapter will be submitted for publication to the *Experimental Mechanics* Journal, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Signal to Noise Ratio enhancement for air-coupled wave propagation in steel rail”. The dissertation author will be the primary investigator and author of this paper.

5 Development of the rail inspection prototype

5.1 Introduction

This chapter provides a description of the rail inspection prototype that has been developed over the course of about three years of work at the NDE&SHM laboratory of UCSD. The prototype technology embeds all the theoretical results achieved in the field of rail inspection not only by the current UCSD SHM&NDE group of investigators, but also by former investigators who worked at the same lab in the last decade or so. Ongoing efforts are currently devoted to solve some further remaining issues, the most important of which currently represents a limiting factor toward the maximum speeds at which the tests can be performed, at least before starting to sensibly decrease the defect detection performance. Insights about such issue as well as on one proposed solution that will be tested in the near future will be given in the remainder of the chapter.

Inevitably, as it is always the case in the design and development of a complex system of this type, countless modifications/refinements were be made/iterated to basically every aspect of the system, which can be tentatively summarized as:

- the specifics of each hardware electronic components and their global layout;
- the mechanical parts (mainly made in plastic materials) whose scope was both to hold the sensors firmly in place as the system was moved along the rail track once the desired sensors positions was set, both to allow flexibility in terms of moving or rotating each sensor around independently from the others (both transmitter and receivers), in the almost endless process of finding the most effective sensors positions and overall sensors configuration;
- the specifics and the types of excitation;

- the de-noising techniques applied to the acquired waveforms in the attempts to enhance the SNR to the highest possible extent;
- the most defect-sensitive features to extract from the acquired waveforms;
- the type and the specifics of the data processing technique employed to analyze the signal features and in turns to infer about the possible presence of defects;
- the ways in which the desired signal excitation, waveforms acquisition, features extraction and data processing could be most effectively handled and implemented in the LabVIEW[®]-based software governing the system in real-time (where speed of execution can be a real issue).

For the sake of brevity and clarity of exposition, only the details pertaining to the final (i.e. the current) version of the prototype will be described through the chapter, with only a few exceptions. Furthermore, for the reasons already expressed in Section 3.1, details of the LabVIEW implementation will not be given in this dissertation.

Starting from the very earliest stages of prototype development, countless sessions of tests were performed at the UCSD/FRA Rail Defect Testing Facility in order to carefully evaluate every proposed upgrade to the system. Step by step, the results of those tests, in pair with the relevant numerical analyses results, were instrumental in order to make decisions on whether or not integrating the modification being tested to the system. Once again, for the sake of brevity, only a minimal part of these test results will be presented in this dissertation. The focus will be instead on the results achieved in the first (and last, at the moment) field test.

In fact, in October 2014 the prototype was packaged and shipped to be tested at the Transportation Technology Center (TTC) in Pueblo, Colorado, USA. The salient results of the field test will be shown and discussed. Ongoing efforts are devoted toward the solution of the issue that limited the testing speeds at which the field tests were performed, and which will be described in the chapter.

5.2 UCSD/FRA Rail Defect Testing Facility

Through all the different stages of prototype development, with the exception of few issues that could be tackled at the NDE&SHM lab at the UCSD main campus, the vast majority of the experimental tests were performed at the UCSD Rail Defect Testing Facility (Figure 5.1). The possibility of testing the prototype at an UCSD in-house facility has been an invaluable tool toward the development of the current technology employed in the prototype.

The facility is located at UCSD's Camp Elliott Field Station Laboratory, approximately 15 kilometers from the UCSD main campus. It hosts a class I track that was constructed in 2009 by Sopac Rail Inc. under FRA sponsorship with in-kind donation of materials (in particular, rails, most of the ties and ballast) by the Burlington Northern Santa Fe (BNSF) railway. The 75-meters-long track is composed of a mix of different rail types (RE136, RE141, ...) featuring a number of natural and artificial defects, including, among the others: Detail Fractures, Transverse Fissures, Rolling Contact Fatigue defects, Vertical and Horizontal Split Heads. The track is formed by a tangent portion (approximately 37.5m-long) and an 8° curved portion (approximately 37.5m-long). The

facility is also available for testing to other developers of rail inspection technology on a recharge basis.

Figure 5.1: UCSD/FRA Rail Defect Testing Facility

5.3 Prototype parts design

During the different stages of prototype development different parts were designed and built to sustain the sensors while allowing the tuning of their positions and inclination angles above the rail top surface. In particular, different sensors configurations were tested, as sketched in Figure 3.1, Figure 3.4 and Figure 3.5, therefore numerous changes had to be made to allow testing in the different conditions.

Figure 5.2 shows a 3D-view of the prototype parts designed in AutoCAD® environment for the current version of the prototype. The material in which those parts were created was Delrin®, which is a highly-crystalline polymer offering high stiffness and strength as well as excellent weather-resistance. 3D and 2D design sheets were submitted to the UCSD Campus Research Machine Shop personnel, who realized the actual parts, as shown in Figure 5.3.

Differently than the other parts, the transmitter holder/adapter was first designed with the aid of numerical analyses carried out in LS-DYNA, as described in Section 4.3.4, and was ultimately 3D-printed in ABS material (Figure 4.104).

The assembled prototype parts were mounted on a cart previously built by ENSCO Inc., a company located in Virginia, to perform testing in Camp Elliott RDTF. The cart is shown in grey in Figure 5.2.

Figure 5.2: Design of the prototype parts in AutoCAD® environment

Figure 5.3: Prototype parts attached to the cart rolling on rails

5.4 Electronic hardware layout

Figure 5.4 shows a layout of the hardware components deployed in the current version of the prototype. At the heart of the system is a National Instruments[®] PXI chassis and CPU running a National Instruments[®] Real Time LabView[®] operating system. The operator can interface with the Real Time chassis through a laptop running a Windows[®] operating system in a wireless manner.

The Real Time PXI executes all the computation, data acquisition, timing and hardware control tasks. An encoder connected to a wheel running on the rail top surface sends triggers at a fixed spatial frequency, which can be set by the operator. When the PXI receives a trigger, it delivers a voltage signal through its function generator to a linear amplifier. After the voltage signal is amplified it is ready to be delivered to the cylindrically focused air-coupled transmitter which converts it in an ultrasonic pressure

wave. At that point, wave propagation phenomena occur as described in Section 3.3.1 and as simulated in Section 4.3.3. Each of the 8 air-coupled receivers deployed in the system sense pressure waves leaking from the rail top surface and converts them to voltage signals which have to be greatly amplified before reaching the digitizer. Both amplification and filtering occur through the hardware components shown on the right in Figure 5.4. In particular, an electrical impedance matching network is employed to enhance the interesting frequencies while filtering out the unwanted ones. Details are given in Section 5.5.2. The amplified and filtered signal is then fed to the A/D digitizer of the PXI chassis, which converts it to a discrete waveform, ready to be analyzed through the statistical algorithm described in Section 5.6.

Figure 5.4: Schematic of the hardware layout deployed in the current version of the prototype

Figure 5.5 shows the cart hosting the different prototype parts at Camp Elliott RDTF. Many of the hardware components discussed above are visible in the figure. Figure 5.6 is a close-up look at the encoder wheel in contact with the rail top surface, while Figure 5.7 shows the inside of the grey box on top of the cart which hosts all the impedance matching networks (described in Section 5.5.2) for the 8-channels available in the current version of the prototype.

Figure 5.5: The cart hosting the prototype parts and the hardware components at Camp Elliott RDTF

Figure 5.6: Close-up look at the encoder wheel

Figure 5.7: Photo of the inside of the grey box shown on top of the cart hosting the prototype in the previous two figures. The impedance matching networks realized for the available 8-channels as described in Section 5.5.2 are visible

5.5 Signal processing techniques employed to enhance the signal-to-noise ratio of the acquired waveforms

As it has already been stressed in Section 3.3 the main drawback of any non-contact ultrasonic testing system is a greatly reduced signal-to-noise ratio of the waveforms acquired by the receivers, when compared to conventional contact testing. The use of signal processing becomes necessary, yet sometime not sufficient to overcome this problem. Several advanced de-noising techniques are available and represent good candidates to be employed for the signal conditioning of ultrasonic waves.

The choice of a particular signal processing technique is also strictly connected with the implementation of an efficient feature extraction procedure, considering that the output of the former is fed as the input to the latter. Among the significant properties that characterize the performance of a signal processing/feature extraction procedure we can include robustness, low computational speed (particularly relevant to systems working in real-time), low false positive and false negative rates.

Different signal processing approaches were originally proposed, implemented, and tested in the process of finding the one able to guarantee the best performance for the defect detection prototype. Those are: discrete wavelet transforms (DWT) [108]; least mean squares (LMS) filters [109]; Wiener filters [110]; fast Fourier transforms (FFT) [111]; lock-in amplifications; digital band-pass filters [112]; matched filters in time and frequency domain [113]; pulse-coding modulation (PCM) techniques [114]; electronic impedance matching networks.

Until the implementation of very efficient electrical impedance matching networks at the preamplifiers of the receivers (introduced to the system around June

2014), a digital bandpass filter and a matched filter were the techniques assuring the best performances overall. Both of them have been extensively tested and for long time used in former versions of the prototype. Nevertheless, at the moment, the prototype is solely based on impedance matching networks, which have been proved to guarantee the best performances in terms of both SNRs and defect detection rates.

For the reasons above, only a few representative experimental results obtained using band-pass and matched filters are briefly presented in Section 5.5.1, while a more detailed description of the implemented impedance matching networks is given in Section 5.5.2.

5.5.1 Digital filtering techniques (used in former versions of the prototype)

Figure 5.8 shows an example of the SNR gain achieved through digital filtering techniques applied to the signals delivered to the A/D digitizer when impedance matching networks were not employed in the system.

The bandpass was employed using a Butterworth filter of order 16 in a very tight frequency region around the frequency of interest.

In regards to the matched filter, the “reference” signal to be used in the cross-correlation matching procedure was obtained by averaging a high number of signals (100) acquired by each deployed receiver when the prototype was held in the same position, in a pristine part of the rail, before the beginning of the actual testing. The number of averages was big enough to filter out the incoherent noise components and hence to reveal the actual signal that each receiver was expected to sense, at least in

pristine parts of rail. Alternatively, attempts were made to update the “reference” signal for each receiver in real-time as the prototype was moving on the rail-track by averaging in a similar manner consecutive 100 signals sensed by each receiver. However, no particular benefits were noticed in terms of defect detection rates as provided by this more computational-demanding procedure.

Figure 5.8: Digital filtering techniques used in former versions of the prototype (both bandpass filter and matched filter are applied directly to the raw signal, i.e. the matched filter on the bottom sub-plot was not applied to the bandpassed signal shown in the central sub-plot)

As already stated in the previous section, bandpass filter and matched filter were the digital filtering techniques used for long time in previous versions of the prototype

being the ones providing the most reliable results. Those techniques were not employed at the same time, i.e., in Figure 5.8, the matched filter is applied directly on the raw signal on top, not on the bandpassed signal in the central sub-plot. In other words, these two techniques were used alternatively and their respective performances were monitored and compared while working on choosing the definitive one. However, when impedance matching networks were employed, the performances of both of them in terms of SNR gain were beaten, thus studies on both of them were soon interrupted. Attempts were made of applying matched filtering on the signals resulting from the impedance matching networks; however, the system performance when both of these two techniques were applied at the same time was proved to actually deteriorate.

5.5.2 Electrical impedance matching networks

Considering the fact that the tone-burst signal used to excite the rail is narrowband, electrical impedance “L”-matching networks could be employed to enhance the SNR levels. Typically, impedance matching networks are realized at the transmitter side to maximize the voltage delivered to the sensor to be converted in ultrasonic motions (for instance [115]). The design of an “L”-matching network for the transmitter side has been carried out in PSpice[®] Student version 9.1 software after measuring the air-coupled transmitter impedance for a range of frequencies using an HP 4192A impedance analyzer. However, such network has never been realized in practice, due to the fact that according to manufacturer specifications the voltage level currently feed to the air-coupled transmitter through the linear amplifier in use is already close to the piezoelectric breakdown value for that specific sensor.

However, similar concepts of impedance matching were applied at the receivers side to increase the available SNR. In particular, using the same software and the same impedance analyzer to measure impedances of each air-coupled receiver, “L”-networks were designed to create resonant conditions at the frequency of interest in the circuits formed by each receiver and its respective preamplifier. In order to carry on these designs, the input resistance of the ultra-low-noise preamplifier in use (model 5660C from Olympus[®]) was measured experimentally and found to be quite lower than the “nominal” one furnished by the manufacturer (30 k Ω vs 1M Ω). On the contrary, all the other values furnished by the manufacturer (both for the input and for the output) were found to be quite accurate.

A second type of “L”-network was also designed and tested, and found to be capable of providing higher SNR gains respect to the previously considered case, after experimental comparisons. This network is schematized in Figure 5.9. As seen in the figure, in this case RLC resonant conditions are created between two preamplifiers put in series one after the other (the air-coupled receiver is connected with the first one, while the output of the second one is delivered to the A/D digitizer of the PXI as schematized in Figure 5.4). Figure 5.10 shows the results of an “AC sweep” analysis done in PSpice that predict increases of frequency components in a narrow region around the frequency of interest in the order of 8 to 9 times. After experimental testing it was found that those predictions were very accurate. As a matter of example, Figure 5.11 shows a comparison of the signals received with and without the impedance matching network from one of the air-coupled receiver used in the prototype. The two signals were acquired first disconnecting and then reconnecting the network. Understandably, the unwanted noisy

components are also affected by the generated resonant conditions; however, after measuring both the increase of signal and the increase of noise, the SNR gain has been found to be in the order of 4-5 times higher than the initial one. This gain was compared to the one provided by the “L”-network realized between receiver and preamplifier, as well as to the ones achievable using the digital filtering techniques described in the previous sections, and was found to be the highest one among all of those. One more advantage of using the network in this part of the circuit is that the input/output electrical impedances of the preamplifier’s model in use are consistently very similar among the different units, as experimentally confirmed after testing all the available units, while the impedance values of each air-coupled receiver, as measured with the impedance analyzer, present significant variations from unit to unit. Therefore, in terms of repeatability and easiness of unit substitution, using the electrical impedance network in this part of the circuit results particularly convenient. Figure 5.7 shows the realization of 8 of these networks for the 8 channels used in the current version of the prototype.

Figure 5.9: Electrical circuit “preamplifier-inductor-preamplifier” in PSpice software (Student version)

Figure 5.10: Results of “AC sweep” type of analysis done in PSpice software

Figure 5.11: Comparison of signals received without (top) and with (bottom) the L-matching network from one of the air-coupled receivers used in the prototype

5.6 Statistical algorithm designed to enhance the rail defect detection performance

As already discussed in Section 3.3.1, each time the prototype executes a test, the actual inspected portion of rail is the one comprised between the furthest receivers deployed in the system. As a general recap, after the signal is generated through the transmitter, and after the wave propagation occurs in that portion of the rail, up to eight receivers deployed symmetrically around the transmitter (as in Figure 3.2) sense the ultrasounds leaking into air from the rail top surface. Those pressure waves are converted by the piezoelectric membranes into voltage signals, which after the signal conditioning described in Section 5.5.2 are available to be analyzed by some method able to infer about the possible presence of a defect in the rail portion under test. The performance of such method can be assessed based on its ability to increase the defect detectability while minimizing false positives.

The method employed in the current version of the prototype is a novelty detection method of the type described in [18]. As such, it is an unsupervised statistical method based on a multivariate outlier analysis that establishes whether or not the last available testing point is discordant with a “baseline” consisting of n previous testing points that (ideally) describe a pristine rail condition.

The term “unsupervised” indicates that such method does not require any learning cycle on known defects. In other words, there is no needing for any “a priori” knowledge of the characteristics of specific defective conditions. This is considered advantageous, due to the countless parameters that would otherwise have to be taken into account in order to setup enough supervised learning cycles, such as: type of defect; its size and

inclination in all the possible directions; its depth; the type of rail being tested; different levels of wears of the rail under test. However, it is not excluded that, in the future, as the population of test samples will keep growing until forming a large enough database comprising many of the different situation that can be encountered in real railroad tracks, some information about defects characteristics could be added to the algorithm in order to increase its performance as defect detector.

The framework of outlier analysis can be found in classical statistical textbooks [116], while its application to diagnose structural damage was first presented by Worden [117].

A detailed description of how the method was implemented follows. First of all, the dimensions of the available data are reduced through extraction of few meaningful features from the acquired waveform. In the current version of the prototype, three features are extracted from each symmetric pair of receivers (those features are expressed in the next section). Because the eight receivers form four symmetric pairs, up to 12 features are available in total (or less than that, if one wants to exclude one or more pairs of receivers from the analysis). To maintain a certain degree of generality, let's indicate with m the number of features extracted from every tested point.

As the prototype starts moving along the rail track, tests are repeated with a fixed spatial frequency (every 6.35 mm in the Colorado field tests, discussed in Section 5.8). The first n tests outcomes are stored and are used to form a “baseline” which characterizes the type of rail under test.

At this point, a brief discussion over the value of n is due. From a statistical point of view, the sample size n would have to be at least equal or greater than the “minimum

statistically significant size”, which, for this specific case of rail defect detection, is difficult to establish rigorously. In practice, analyses have been carried out with the aid of Receiver Operating Characteristics (ROC) curves [88] (as the ones that are described in Section 5.8.5) as the value of n was systematically varied, performing a so-called “optimization problem” in order to find the value for which the best defect detection performance could be achieved. As a result, in the current version of the prototype a n value of 170 is used. In any case, once the desired value of n is established, the following applies.

It would be beneficial, especially in regards to the inspection of the first few feet or so of rail, to collect the first baseline points in a part of the rail known to be in pristine conditions. However, when this is not possible, considering that usually the conditions of the first section of rail are not known already, the effects of such “bad start” will vanish relatively soon as the prototype moves away, and this will be made clear after reading the whole procedure that is explained below.

The sampling distribution of the m features along the n acquisitions forming the baseline is analyzed through computation of their mean (vector of length m denoted as $\{\bar{x}\}$) and covariance (matrix $m \times m$ denoted as $[K]$), which are stored in the system.

Once the $n + 1$ acquisition is available, the deviation of its corresponding features (vector $\{x_\zeta\}$) from the baseline is quantified through computation of the so-called “Mahalanobis squared distance”, D_ζ :

$$D_\zeta = (\{x_\zeta\} - \{\bar{x}\})^T [K]^{-1} (\{x_\zeta\} - \{\bar{x}\}) \quad (5.1)$$

which in the following will be denoted as Damage Index (DI).

The higher the value of the computed DI, the higher the deviation of the features extracted from the new testing point when compared to the features used to form the baseline, thus the higher the probability of this discordancy being generated by changed condition in the portion of rail under test. In particular, the computed DI is compared to a threshold value: if the former is lower than the latter (“case I”), the rail under test is considered pristine, if it is higher (“case II”), a defect is called in that position.

Clearly, the threshold value is also dependent on both the number of observations forming the baseline (n), and the number of features in each observation (m). A clever method to determine the most advantageous threshold value for the specific $m \times n$ dimensions case is to compute ROC curves, which allow for a straightforward visualization of the variations of Probability of False Alarm (PFA) and Probability of Detection (POD) as the threshold value is gradually increased [88]. Clearly, for the computation of such curves, the prototype has to be tested in a rail track where the exact positions of defects are known. For example, this procedure was followed based on the findings of the Colorado field test, and it is described in Section 5.8.5.

Let’s get back to the discussion of the employed algorithm. If the $(n + 1)$ testing point falls under “case I”, the baseline is updated with the new observation: the first (furthest) observation forming the previous baseline (i.e. the baseline used to calculate DI_{n+1}) is discarded, while the $(n + 1)$ observation is added to form an updated baseline. The following may help to better visualize this updating process:

$$\begin{aligned} &\text{Observations considered to form the baseline to be used for the} \\ &\text{computation of } DI_{n+2} \rightarrow [1, 2, \dots, n, n + 1] \end{aligned} \tag{5.2}$$

Once again, mean vector and covariance matrix are formed, and these new values are used to compute the DI of the $(n + 2)$ testing point.

If, instead, the $(n + 1)$ testing point fell under “case II”, then it would be considered an outlier, so its features would be discarded and the baseline would not be updated prior of the computation of DI_{n+2} .

As such, this employed novelty detection method is defined “exclusive”, because it excludes the outliers from the computation of the baselines, as opposed to “inclusive” methods. Nevertheless, let’s elaborate a little further on the thresholds employed in the algorithm. To facilitate the understanding of the whole process, it was implied that only one threshold is used to both (1) call a defect and (2) discard the new observation from the updating of the baseline. In reality, the algorithm is flexible enough to allow for the introduction of two different threshold values, one to be used for task (1), the other for task (2). The reason is straightforward. Let’s suppose that DI_{n+1} is a rather high value, however, it is not high enough for the prototype to call a defect in that point (as we said above, from the findings of the ROC curves, in the interests of maintaining a low enough PFA rate). Even so, there is no significant harm in discarding the new observation from the formation of an updated baseline, when its origin is in doubt. In other words, it is like having the flexibility to allow for a lower (or, in general, different) admissible PFA toward the decision of whether or not updating the baseline. Nevertheless, a unique threshold value can be still used for the two tasks, if one prefers.

This process is repeated *ad nauseam* as the prototype keeps moving along the rail track, and as new observations $(n + 2)$, $(n + 3)$, etc., are presented as inputs to the algorithm.

It is worth to stress that, as described, the dimension of the observations population forming the baseline at each point during the tests will always be equal to n .

At this point, it should be understandable why a “good start” in a pristine part of the rail during the formation of the first n -long baseline should be preferred, whenever possible; nevertheless, the negative effects of a possible “bad start” would disappear soon as the prototype moves away only few feet or so, as the baseline gets first partially, then completely, rebuilt.

Examples of DI traces that are obtained in real-time as output of this process are shown in Section 5.8.4, where the salient results obtained in the Colorado field test are reported.

5.6.1 Defect-sensitive features extraction

The last section described the implementation of a procedure that, starting from some m features (input) whose main role is to compress the information contained with the different waveforms collected by the receivers deployed in the prototype (usually 8), is able to produce a value which was called DI (output) from which one can directly infer about the possible presence of a defect in the portion of rail under test. However, no information was given on the nature of those features, and also, as it appears from the description, there are no restrictions on the value that could be given to the number m . In practice, the upper limit of m would be represented by the smaller number for which the increased computational effort would prevent a real-time use of the prototype at normal operating speeds. It is also understandable that a connection should exist for which a particular feature might be more indicated/advantageous when it is used in pair with

some filtering methods applied to the raw signals (Section 5.5) and less when it is used with some other of those methods. This leaves the investigator with a practically unlimited number of possibilities to choose from in its effort to determine the best overall system configuration, possibilities that are not solely related to the choice of specific feature formulas to use.

Even restricting the field after fixing at least some parameters, such as the type/location of the excitation, the type of sensor configuration (Figure 3.1, or Figure 3.4, or Figure 3.5) and the maximum number of receivers that could be used at once (8, limit imposed by the available electronic hardware), there are still many other factors that will take a role toward the optimal defect detection, the most important of which are:

- positions to give to the various sensors (both considering their distances from the transmitter, both the transversal location on the rail top surface);
- filtering technique;
- number of features to be extracted from each sensor or pair of sensors (and this number can actually vary from pair to pair, if needed);
- nature of those features.

The endless process of understanding which exact configuration would provide the best and most reliable performance under all the possible circumstances, defects and situations that can be encountered in real railroads by the prototype was carried out all over the duration of this study, and cannot be considered concluded. An invaluable aid in this research is represented by numerical modeling, which allows for quick simulations of large numbers of different defects, giving full control to vary the flaw size, depth, location in the cross-section of the rail and orientations (respect to every possible angle).

Nevertheless, experimental testing is still strictly needed, because of the unexpected problems that always arise when transferring from an “ideal” and noise-less environment as is the numerical one, to the much noisier and unpredictable real-life conditions. The main tool that was repeatedly used over time as a guidance to understand the possible benefits produced by this or that configuration, transmitter and receivers locations, number of cycles in the excitation signal, filtering technique, as well as number and types of features to use, was represented by the computation of ROC curves over the DI traces obtained through extensive testing at the RDTF of UCSD (Section 5.2) and later at the TTC facility in Colorado.

In regards to those factors, the conclusions at which the research is arrived (up to the present), and that were incorporated to the version of the prototype field tested in Colorado in October 2014, are in large part already given in previous sections of this dissertation (type/location of excitation is discussed in Section 3.3.1 and 4.2.2, the sensor configuration is the one showed in Figure 3.1-Figure 3.2, the implemented filtering technique is discussed in Section 5.5.2, while the number of features extracted from each pair of sensors is 3 as briefly expressed during the previous section). It only remains to specify the features used, which is done in this section.

Initially, 8 features were proposed, all of which needs to be extracted from a pair of sensors. Later, a 9th one was added to the list. Below is a table showing formulas and names given to such features.

Table 5.1: List showing the 9 features proposed for the system. In yellow background are the ones currently in use

1	Normalized Variance	$\frac{ Var(x_1) - Var(x_2) }{\sqrt{Var(x_1) * Var(x_2)}}$
2	Normalized RMS	$\frac{ RMS(x_1) - RMS(x_2) }{\sqrt{RMS(x_1) * RMS(x_2)}}$
3	New RMS	$\frac{(RMS(x_1) - RMS(x_2))^2}{RMS(x_1) * RMS(x_2)}$
4	Normalized XCORR	$\frac{max(xcorr(x_1, x_2))}{\sqrt{max(xcorr(x_1)) * max(xcorr(x_2))}}$
5	Max Ratio	$\frac{ Max x_1 - Max x_2 }{\sqrt{Max x_1 * Max x_2 }}$
6	DI Max	$Max\left(\frac{Max x_1 }{Max x_2 }, \frac{Max x_2 }{Max x_1 }\right)$
7	Peak to Peak Ratio	$\frac{ Ppk(x_1) - Ppk(x_2) }{\sqrt{Ppk(x_1) * Ppk(x_2)}}$
8	Peak to Peak New	$Max\left(\frac{Ppk x_1 }{Ppk x_2 }, \frac{Ppk x_2 }{Ppk x_1 }\right)$
9	Ratio RMS	$\frac{RMS(x_2)}{RMS(x_1)}$

In the table above, x_1 and x_2 indicate the values of the two waveforms fed into each formula. Except the number 4, which is based on cross-correlations between the two input waveforms, all the others are purely energy-based, making use of parameters as RMS, variance, “maximum amplitude” or “peak to peak of the amplitudes”. In yellow

background are indicated the 3 features in use in the current version of the prototype, where they are extracted from each symmetric pair of receivers.

In theory, the wave/defect interaction would produce changes in the wave propagation field such that also the relative strength of the signals sensed by two non-symmetrically placed transducers could be modified. Therefore, nothing restricts us from computing these features from the signals acquired by non-symmetric pairs of transducers. This was tested repeatedly in the past. Nevertheless, it never showed to be advantageous, so at the moment the four pairs of sensors considered by the current version of the prototype are the four symmetric ones.

As an example, Figure 5.12 shows some of the ROC curves computed in the past over some of the runs performed at the RTDF in Camp Elliott (tests dated early October 2012), with the intents to compare the performances achievable using by each of the proposed features (the first 8 of Table 5.1, the 9th feature was proposed later), on each of the waveforms obtained using few different filters (some of the ones described in Section 5.5.1) on the raw signals, and which were valid specifically for the sensor configuration (the one as in Figure 3.1) and sensors locations used in those tests. In the plots legend, AUC stands for Area Under the Curve, i.e. the area comprised between the ROC curve, the lower axis (at PD=0) and the right axis (at PFA=1). The AUC is a commonly used parameter that allows condensing the performance described by a ROC curve into a single scalar value, thus facilitating the comparisons between many different curves. The higher the AUC, the better the performance. For example, Figure 5.13 shows part of a Microsoft Excel[®] sheet that was made to readily compare the performances of the various examined sensors pairs/features/filters based on AUC values.

The procedure used to construct those ROC curves (as well as all the others that have been computed over time from tests done in Camp Elliott) is similar to the one that was used to evaluate the prototype performance at the TTC field test, which is detailed in Section 5.8.5. However, there is a fundamental difference in the way the PFA rate was computed, because of which a direct comparison among the ROC curves plotted in Figure 5.12 and the ones in Figure 5.23 would not make sense. Namely, while the PFA rate in the ROC curves computed over the Camp Elliott tests was simply obtained as *“number of DI values exceeding the considered threshold and pertaining to non-defective rail portions”* divided by *“total number of testing points pertaining to non-defective rail portions”*, the procedure that is explained in Section 5.8.5 and that was used to calculate the PFA of the TTC tests is different, and certainly more representative of the performance that would be offered by the prototype in real applications. This should not surprise, in fact, while the ROC curves as the ones in Figure 5.12 were computed with the specific goal of guiding the author of this dissertation toward the determination of the optimal operational parameters to use in the prototype (therefore, “realistic” values for PFA were not needed, as long as a meaningful direct comparison among performances offered by different features/configurations was provided), the ones of Figure 5.23 had the primary goal of accurately characterizing the performance achieved by the current version of the prototype in the TTC tests, and to provide this information to the FRA (among the other parties) so that they could evaluate it. In synthesis, if the same procedure were used to re-calculate the PFAs and thus the ROCs of Figure 5.12, the AUC values shown in the plot legends would sensibly drop, but the same relative order among the quality of the performances would be retained.

Figure 5.12: Example of ROC curves computed to evaluate the performances of different sensors pairs/features/filters combinations

SENSOR PAIR (1-8)					SENSOR PAIR (2-7)				
NORMVAR					NORMVAR				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9703177	0.9595379	0.9731251	0.9676603	PASSBAND	0.964266	0.9551623	0.963332	0.9609201
MATCHED	0.9736756	0.9778741	0.9799921	0.9771806	MATCHED	0.9638275	0.963575	0.9625675	0.9633234
FFT	0.960407	0.9369302	0.9434032	0.9469135	FFT	0.967436	0.9560088	0.9578358	0.9604268
FFT GATED	0.9539042	0.9371126	0.9420983	0.9443717	FFT GATED	0.9568737	0.9501615	0.9545802	0.9538718
NORMRMS					NORMRMS				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9706936	0.9592775	0.9740609	0.9680107	PASSBAND	0.9659324	0.9552795	0.9631079	0.96144
MATCHED	0.9736756	0.9785644	0.9798866	0.9773755	MATCHED	0.9688769	0.9685368	0.9686833	0.968699
FFT	0.9644164	0.9403162	0.9526954	0.952476	FFT	0.9780484	0.9782127	0.9728219	0.976361
FFT GATED	0.9652809	0.9411757	0.9544352	0.9536306	FFT GATED	0.9657194	0.9657889	0.9680111	0.9665065
NEW RMS					NEW RMS				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9702802	0.9588217	0.9742322	0.967778	PASSBAND	0.9502456	0.9482211	0.962963	0.9538099
MATCHED	0.9732998	0.9784472	0.9797417	0.9771629	MATCHED	0.956811	0.9589519	0.9587452	0.9581694
FFT	0.9535158	0.8836407	0.9322657	0.9231407	FFT	0.9498446	0.9345601	0.9444972	0.9429673
FFT GATED	0.9434671	0.9166016	0.9274417	0.9291701	FFT GATED	0.9339448	0.9155858	0.9340187	0.9278498
NORMXCORR					NORMXCORR				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9618228	0.9600589	0.9597997	0.9605604	PASSBAND	0.9833985	0.9772881	0.9625016	0.9743961
MATCHED	0.9419636	0.9274887	0.9277975	0.9324166	MATCHED	0.9696537	0.9659582	0.9531699	0.9629273
FFT	0.9558087	0.9554878	0.9515355	0.9542773	FFT	0.9631384	0.9679377	0.966693	0.965923
FFT GATED	0.9485666	0.9552144	0.9597469	0.9545093	FFT GATED	0.9707187	0.9615956	0.945802	0.9593721
MAXRATIO					MAXRATIO				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9618103	0.960684	0.9719388	0.964811	PASSBAND	0.9643537	0.972092	0.9632529	0.9665662
MATCHED	0.9721846	0.9821456	0.9796758	0.978002	MATCHED	0.9686388	0.9773011	0.9728219	0.9723206
FFT	0.9351601	0.9539121	0.9490576	0.9460433	FFT	0.9771588	0.9733422	0.9682483	0.9729164
FFT GATED	0.9300105	0.9551362	0.9626994	0.949282	FFT GATED	0.9638275	0.9637964	0.9675234	0.9650491
DIMAX					DIMAX				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9604571	0.960723	0.9721893	0.9644565	PASSBAND	0.9629504	0.9651769	0.9592856	0.962471
MATCHED	0.9724478	0.9826145	0.9797417	0.978268	MATCHED	0.9690147	0.9666484	0.9633979	0.9663537
FFT	0.9351977	0.9527921	0.9493607	0.9457835	FFT	0.9697414	0.9456295	0.9477791	0.9543833
FFT GATED	0.9245727	0.955006	0.9631607	0.9475798	FFT GATED	0.9494186	0.9403813	0.9470146	0.9456049
PPRATIO					PPRATIO				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9627124	0.9599547	0.9713589	0.9646753	PASSBAND	0.9643788	0.970933	0.9631079	0.9661399
MATCHED	0.9712199	0.9816768	0.9802689	0.9777219	MATCHED	0.9700672	0.9770407	0.9735469	0.9735516
FFT	0.9351601	0.953899	0.9490576	0.9460389	FFT	0.9770335	0.9733422	0.9682483	0.9728747
FFT GATED	0.9300356	0.9551492	0.9626994	0.9492947	FFT GATED	0.9638275	0.9637964	0.9675761	0.9650667
PPNEW					PPNEW				
	RUN 1	RUN 2	RUN 3	AVERAGE		RUN 1	RUN 2	RUN 3	AVERAGE
PASSBAND	0.9605949	0.9598896	0.9717939	0.9640928	PASSBAND	0.9622989	0.9644215	0.9590484	0.9619229
MATCHED	0.9719466	0.9820545	0.9804402	0.9781471	MATCHED	0.9695409	0.9678335	0.9654672	0.9676139
FFT	0.9351977	0.9527921	0.9493607	0.9457835	FFT	0.9697414	0.9456556	0.9477791	0.954392
FFT GATED	0.9245727	0.955006	0.9631607	0.9475798	FFT GATED	0.9493936	0.9403813	0.9470146	0.9455965

Figure 5.13: Part of a Microsoft Excel® sheet that was made to compare the performances of few examined sensors pairs/features/filters combinations. The numerical values shown in the cells are the AUC of the ROC curve computed for each combination

5.7 On the issue affecting higher speed testing performance

As seen in the simulation results shown in Figure 4.99 and Figure 4.100, once the ultrasonic waves released by the transmitter impinge on the rail top surface, a large fraction of their energy gets reflected in air. From the points of contact with the rail, these reflections are sent back toward the transmitter as well as everywhere else in the surroundings, until they are gradually absorbed through air. Their energy is significantly higher than that of the waves leaking from the rail top surface and sensed by the receivers. Namely, as it can be drawn from the discussion in Section 4.3.3, the amplitudes of these reflections from the steel upper face are roughly 10^5 times higher than the amplitudes of the leaky waves. Luckily, the speed of sound in air is much slower than the one in steel, so by the time those reflections arrive at the receivers the meaningful waves leaking by the rail top surface have already been sensed (this is valid only if a certain distance is allowed between transmitter and each receiver; obviously, if they are put too close, there is not enough time for the separation of these two wave trains).

Figure 5.14 should clarify this argument. Note that, in order to highlight the salient features, the same waveform is plotted in the three sub-plots, where the time scales were varied. Namely, going from the top toward the bottom three increasing timeframe durations are showed, while the voltage range in the top plot is 10 times smaller than the one used in the other two plots. The waveform in Figure 5.14 shows a typical signal as sensed by one of the air-coupled receivers used in the prototype. Let us focus in the zoomed-in sub-plot (a). The first packet of waves (until $\sim 100 \mu\text{sec}$) is simply due to EMI interferences caused by the cable carrying the signal delivered to the transmitter. The meaningful part of the signal, which are waves leaked from the rail top

surface, is the one approximately located between 500 and 700 μsec . However, if one lets the system acquire a longer signal, the reflections discussed above can be seen. In the (b) sub-plot it appears that the first reflections arrive at about 900 μsec , thus not much later after the end of the “meaningful” waves, but luckily, late enough. The reflected signals are “clipped” at ± 15 Volt due to the limited voltage range of the acquisition system, but they are actually much stronger than the leaked waves than what appears because of the clipping. In sub-plot (c) other arrivals can be seen, which are due to other reflections from all the surfaces encountered by the air-waves in the surrounding area. Shortly after 0.01 sec, basically all the reflections have passed by or have been dissipated through air absorption, so that only noise is sensed by the receiver.

Everything has been discussed above does not pose any problem for the execution of the tests, at least until the frequency at which consecutive tests are done is increased above certain values. Conservatively, if the excitation repetition rate is kept smaller than ~ 70 Hz (corresponding to ~ 0.015 sec between each test), those unwanted reflections would not affect the signal leaking from the rail. However, if a repetition rate of, say, 100 Hz is needed (corresponding to 0.01 sec between each test), noisy conditions still left by the previous test will start to affect the new acquisition. At 100 Hz, the test would still be doable, even if a drop of SNR would occur, worsening the quality of the defect detection performance. In fact, even if difficult to see due to the high voltage range plotted in the (c) subplot, the RMS of the noise left by the excitation at about 0.01 sec is still a good fraction of the typical max amplitudes of the meaningful signal. Nevertheless, while a repetition rate of 100 Hz could still be roughly considered the limit at which the tests can be performed while only slightly deteriorating the performance, higher rates are

practically not allowed. For example, if one wants to double the test repetition rate (200 Hz), each new acquisition will occur in a very noisy environment as the one showed at about 0.005 sec in the (c) subplot. As shown in the figure, at that time the RMS of the noise left by the previous shot is much higher than the signal amplitude. Further increases of the repetition rate would just make the conditions worsen.

Understandably, a limitation on the temporal repetition rate at which successive tests can be done translates on a limitation either on the testing spatial repetition rate or on the testing speed. For example, if one sets the spatial repetition rate to 48 tests per foot (which is the value used for the Colorado field tests), the highest speed at which this additional noise will not affect the SNR of the signals (thus the defect detection performance) is ~1 mph (corresponding to 70 Hz at 48 points per foot). This limit can be raised up to about 1.5 mph if one is willing to slightly deteriorate the performance (105 Hz is the resulting frequency considering 48 points per foot and 1.5 mph).

Awareness about the importance of this issue was not raised until only a few weeks before the field test in Colorado. In fact, until then, all the tests were performed at the RDTF in Camp Elliott, where the cart hosting the prototype was always manually pushed. In those tests, considering the cart weight and the significant wheels/rail frictions, the testing speed was always in the order of 0.5-1 mph. Therefore, the detrimental effects of this phenomenon were not noticeable.

Since the time when this issue was discovered, efforts have been devoted to mitigate it. At the moment, the only solution that was found to be completely effective consists in the use of a sponge placed in tight contact with the transmitter (or, as well, with the adapter around the transmitter) and with the rail top surface, whose main effect

is to trap the waves reflected by the rail top surface before they can spread in the surroundings. In theory, a different material can be utilized to such extents in place of the sponge currently in use, as long as its properties allow for: (1) an easy compression in order to effectively close any air gap that would allow the ultrasonic waves to escape from the chamber created between transmitter and rail top surface, (2) a sliding friction on top of the rail that do not introduce noises at the frequency used in the tests, (3) ability to retain those properties when the rail is wet. However, in addition to all the above properties, the selected sponge offers one more advantage. In fact, it is made of an open-cell foam with cells dimensions ranging between ~ 0.5 mm to ~ 2 mm. Given the value of the wavelength of the ultrasound used in the tests, those cells provide excellent absorption of the wave energy as the waves try to come out of the chamber. The effects produced by the use of such sponge are demonstrated in Figure 5.15. The two sub-plots show the same two waveforms, only the y-scale is different. The signals received by the same air-coupled receiver when the sponge was not being used (red) and when it was used (blue) show that a proper use of this solution would completely solve the issue. However, the implementation of the sponge in the prototype has not been yet tested intensively, so other unexpected problems might arise in the future.

Figure 5.14: Signal recorded by an air-coupled receiver that shows successive wave arrivals after the meaningful signal leaked from the rail top surface. Three different x- or y-scales are used to plot the same waveform in the three sub-plots

Figure 5.15: Signal recorded by the same air-coupled transducer when the sponge is not being used (red) vs when it is being used (blue). Two different y-scales are used to plot the same two waveforms in the two sub-plots

5.8 Field test at Transportation Technology Center

5.8.1 Introduction

In October 2014, the rail inspection prototype was packaged and shipped to be tested in the field at the Rail Defect Test Facility (Figure 5.16) located at the Transportation Technology Center (TTC) in Pueblo, Colorado, USA. This facility was specifically developed for research and test purposes to enhance technology and verify capabilities of rail defect detection systems.

Figure 5.16: Rail Defect Test Facility (RDTF) at TTC

The primary objective of the tests was to further develop the proposed rail defect detection system used for the first time outside of the UCSD environment (Figure 5.17). The secondary objective was to evaluate the efficacy of the system.

Instrumental for the realization of the test was the technical support of ENSCO Inc. (Eric Sherrock and others), the Volpe National Transportation Center (in the person of Dr. Robert Wilson), and TTC (Lucas Welander and others). Ensco's support consisted on the design and construction of a cart hosting the prototype's elements and on the operation of the FRA's hy-railer equipped with positioning systems (Figure 5.18). Robert Wilson of Volpe Center provided technical guidance throughout the tests and participated to the evaluation of the test results.

Figure 5.17: The prototype mounted on the cart. The sensors are behind the grey box which is attached to the cart

Figure 5.18: The prototype mounted on the cart was towed by the FRA's hy-railer

5.8.2 Operational parameters used during the execution of the tests and test site lay-out

An encoder provided by ENSCO was used to trigger the wave excitations with a fixed spatial resolution as the prototype was towed along the test track. The encoder resolution was set to approx. 0.25 in (6.35 mm), which yields about 48 testing points per foot. This allowed a good redundancy of measurements over the same portion of the rail.

The testing speed was limited to about 1 mph because of the issue described in Section 5.7.

Based on the discussion in the last paragraph of Section 3.3.1, it was declared to the parties involved that the prototype would have been able to detect only the defects located (entirely or partially) in a certain portion of the rail-head, the one depicted in Figure 5.20.

Tests were performed both on the “Technology Development Section” of TTC’s rail defect farm, where the locations/types of defects were shared with the UCSD personnel, and on the “System Evaluation Section”, where “blind” tests were performed (number, types and locations of defects were, and still are, unknown). Only the results from the Technology Development Section are shown and discussed in this dissertation. In fact, the prototype’s performance on the blind tests on the “System Evaluation Section” is still unknown, pending site recertification by FRA and TTC.

The Technology Development Section consists of a rail track, about 1400 feet long, with several natural and artificially made defects of different types and sizes. Tests were performed on only one of the rails (Figure 5.19). Table 5.2 shows the set of defects in the rail under test that lie (partially or entirely) inside or that are approximately at the

borders of the approximate prototype coverage area shown in Figure 5.20. This set is designated as “List 1”. A subset of this list, which is designated as “List 2”, is formed by the defects cross-marked in the last column of Table 5.2. The main reason behind this differentiation was to study the efficacy of the system in the detection of defects having a relevant size. For this reason, defects smaller than 5% of the Head Area were excluded, as well as a couple of defects most likely located outside of the current prototype coverage area. Explanations about the acronyms used in Table 5.2 can be found in the “List of Abbreviations” provided at page ix in the preliminary pages of this dissertation.

Figure 5.19: TTC's RDTF map showing details of the testing location

Table 5.2: Map of defects agreed upon with Volpe and TTC personnel from survey of October 29, 2015

Flaw Type and Sectional Location	Flaw Size	Position (ft)	List 2?
SDH (drilled through from gage to field side)	0.25" diameter	79.3	X
SDH (drilled through from gage to field side)	0.13" diameter	81.3	X
SDH (drilled from gage side, half head deep)	0.25" diameter	83.4	X
SDH (drilled from gage side, quarter head deep)	0.13" diameter (< 5%)	85.4	
SDH (drilled from gage side, quarter head deep)	0.25" diameter	87.4	X
SDH (drilled from gage side, half head deep)	0.13" diameter (< 5%)	89.4	
TD (lower head gage side)	20%	287.7	
TD (gage side)	< 5%	299.6	
CH with TD	10%	356.7-357.4	X
TD with shelling	5%	579.0	X
CH with TD (6" long)	5%	591.3	X
TD notch (lower head)	0.75"	708.5	
HSD with TD notch (lower head)	10" / 0.75"	774.0	X
DF (TD, under shelling, gage side)	10%	781.1	X
HSD with TD (lower head cut)	10" / 20%	800.4	X
Head Chip (2" long)		857.3	
TD notch (gage side upper head, under shelling)	0.75"	870.5	X
TD notch (gage side upper head, under large chip)	0.75"	876.9	X
TD (under shelling)	75%	883.9	X
TD	8%	895.6	X
HSD with TD notch (lower head)	10" / 0.75"	937.9	X
Shell/Eng Burn (field side/center)		952.5	X
HSD with TD notch (lower head)	10" / 0.75"	959.6	X
TD notch (lower head)	0.75"	972.1	
HSD	10"	979.0	X
TD (small)		991.9	X
HSD with TD notch (lower head)		999.4	X
CH (1' long)		1139.0	X
VSH (artificial cut, 0.75" deep from bottom head, 14" long, gage side)		1182.6-1183.8	
SDH (drilled from field side, 1.5" deep)	0.5" diameter	1201.2	X
SDH (drilled from field side, 0.75" deep)	0.5" diameter	1202.2	
SDH (drilled through from gage to field side)	10 mm diameter	1217.3	X
SDH (drilled through from gage to field side)	10 mm diameter	1218.3	X
SDH (drilled through from gage to field side)	10 mm diameter	1219.3	X

Figure 5.20: Approximate current coverage of the rail defect detection prototype

5.8.3 Field test timeline

The field tests were performed over the course of six days, from October 26th-31th, 2014. The tests timeline is presented below.

Day #1. System setup

Day #1 was spent by UCSD and ENSCO personnel to set up the prototype on the hy-railer in the Passenger Services Building (PSB) of TTC.

The UCSD group did some initial mapping of defects, joints and welds in the Technology Development Section and started the creation of a new features list for the

zone. In the following days, the complete list was compiled and agreed upon with Volpe and TTC personnel (the defects are reported in Table 5.2).

Day #2. System setup and Technology Development Section

Five preliminary test runs were completed before the end of Day #2. As a result of these tests, it was determined that spraying water on both cart and truck wheels during testing resulted in potentially less noise affecting the signals than when the wheel/rail interface was dry. However, further tests will have to be done in the future to verify and quantify the introduction of noise due to wheel/rail interactions in dry conditions. ENSCO encoder appeared to be accurate, even if a small adjustment to the scale factor of the encoder had to be applied to refine distance measurements.

Days #3 and #4. Technology Development Section

Several tests were performed in the Technology Development Section during Days #3 and #4. Some issues were encountered, particularly on Day #4, due to misalignments of the encoder. Some runs were started and soon aborted as the scaling factor was varied in the attempt to find the most accurate value for it.

A few different sensors locations were tested through the different runs. Before the end of Day #4 it was clear to UCSD personnel which sensors locations were the most effective (Figure 3.2 - Figure 4.20). In total, in these two days of testing four complete runs were performed with this “optimal” sensors configuration.

Days #5 and #6. System Evaluation Section

After two more runs performed at the Technology Development Section in order to double check and confirm the accuracy of the encoder with the scaling factor set at the end of the previous day, the hy-railer was moved to the beginning of the System Evaluation Section. Three runs were performed on the “outside” rail during Day #5. Following that, three more runs were carried out on the “inside” rail, two during Day #5, and one at the early morning of Day #6. The optimal sensors configuration, as determined from the previous days of tests, was used for all those six runs, and water was always sprayed on the wheels of both truck and cart.

After the final run performed on Day #6, the caravan truck/cart was moved back inside the PSB building, where UCSD personnel dismounted the system while ENSCO personnel dismounted the cart. At about 12.00 pm the equipment was packed and the tests were concluded.

5.8.4 Field test results: Damage Index traces

A total of four runs were performed in the Technology Development Section using the “optimal” configuration in terms of sensors locations. Only three pairs of receivers (out of the four available) showed good/excellent SNR levels, therefore the fourth pair was not “activated” (its features were not fed into the algorithm for rail defect detection). In fact, the LISA simulations described in Section 4.2.3 (Figure 4.20), revealed that the position of the sensors forming the fourth pair (top right black rectangle) was not “energized” enough by the transmitter excitation, while the other three positions were definitely much more energized. Figure 5.21 shows the Damage Index traces

obtained from each of the four runs for a rail portion located between 880 feet and 1005 feet from the test starting point. This interval is chosen for the sake of brevity and because it is considered representative to show the performance of the prototype. The locations of defects are indicated using either a red or a green line. A red line indicates that the defect is included in both “List 1” and “List 2”. A green line (only one in this portion of the track) indicates that the defect is included only in “List 1”. Cyan lines and black lines indicate the locations of welds and joints, respectively. As it can be seen in Figure 5.21, the joints appear as relatively long sequences of particularly high Damage Index values. Thanks to this specific signature, it is rather easy to distinguish joints from actual defects. However, it is more complicated to distinguish between a weld and a defect. In fact, while sometimes a weld appears as a relatively long sequence of high DI values in a similar fashion as a joint (for example the weld at 969 feet in Figure 5.21), in other cases its signature is not very different from that of a defect; furthermore, in some other cases, a weld will not be detected whatsoever (absence of relevant peaks in the DI trace). Thus it is important to either have a list of welds in the track under inspection, or to capture the positions of welds through, for example, a parallel vision-based system.

From the traces shown in the figure, one can see how the prototype performed in each of the four runs. The plots relative to runs #3 and #4 are particularly noteworthy: all the defects (at least the ones located in this rail portion) stand up well above the noise floor. On the contrary, run #2 shows the worst performance.

In order to get a better quantitative understanding of the performance offered by the prototype, ROC curves were computed and the results are discussed in the next section.

Figure 5.21: Damage Index traces. Red lines: locations of defects included in both “List 1” and “List 2”. Green lines: locations of defects included only in “List 1”. Cyan lines: locations of welds. Black lines: locations of joints

5.8.5 Field test results: Receiver Operating Characteristic curves

As already discussed in previous sections, a common tool used to assess the performance of a detector is represented by the computation of Receiver Operating Characteristic curves. These curves graph the “true detection” rate over the “false alarms” rate obtained by the detector when the value of the threshold is gradually increased [88]. Once those curves are obtained, a common way to assess the quality of the performance of the detector is through the computation of the area under the curve (AUC).

A description of the procedure implemented for the computation of the proposed ROC curves follows. Considering the gage length of the prototype (i.e. the distance between the furthest deployed sensors), the spatial resolution of the system can be considered to be about 2 feet long (conservatively). Based on this observation, the system will call a defect when the threshold of the Damage Index trace is exceeded at least 3 times in a 2-feet-long rail segment. This condition is meant to eliminate “isolated spikes” that can result from electrical noise or other noise sources. In fact, if a defect is present, the system should detect it multiple times within the gage length (especially at the 0.25” encoder resolution used in the tests).

For the computation of the True Positive rates (i.e. the defect detection rates), the “defective” population could be built by knowing the footage of every defect included in the lists. The Damage Index values included in +/- 1 foot segment around each defect footage indication was stored. This “defective” population consisted of 34 defects per run considering the Defect List #1, and 25 defects per run considering the Defect List #2.

In regards to the False Positive rates, the “clean” (non-defective) population of 2-foot-long rail segments was built by removing from the Damage Index traces +/-3 feet of

data around any feature listed in the defects list as well as in the welds and joints lists. This ± 3 feet is a precautionary value to protect against (1) encoder inaccuracies, (2) possible footage inaccuracies during the survey of the track, and (3) length of weld, joint or defect along the rail running direction. The goal was solely to avoid contamination of the “clean” population with measurements actually coming from a true positive (a defect) or from joints/welds. After this purging procedure, the “clean” population was still very large, namely composed of ~ 350 2-foot-long rail segments per run.

Once the “defective” and the “clean” populations were formed and stored as described above, ROC curves were computed point-by-point by sequentially increasing the value of the threshold (Figure 5.22 gives a qualitative explanation of this procedure). The curves shown in Figure 5.23 and Figure 5.24 were obtained by considering 500 increasing threshold values ranging from 0 up to 100% of the maximum Damage Index value in the specific run under evaluation.

Figure 5.22: ROC curves computation procedure

Figure 5.23 shows the resulting ROC curves relative to each of the four runs in the Technology Development Section. Similarly, Figure 5.24 plots a “cumulative” ROC curve computed using the Damage Index trace obtained by lining up the four traces, one after the other. By doing so, the goal was to obtain a single ROC curve capable of capturing the overall prototype performance at the field test. The two lower sub-plots in the same figure show a super-imposition of the results from each of the four runs together with the “cumulative” one. The Areas Under the Curve are given in the legend of each plot.

As expected, for each run the performance offered by the system with regards to Defect List #2 was significantly better than the one obtained considering the full Defect List #1. Also, runs #3 and #4 achieved the best performances while run #2 the worst one, as it was already qualitatively noticeable from the portions of the Damage Index traces shown in Figure 5.21. However, the performances of the four runs are relatively comparable, which suggests that the system performs with reasonably good repeatability.

ROC curves as the ones shown here can be used to find a threshold that produces a desired Probability of Detection while having a still acceptable Probability of False Positives (i.e. False Alarms) on non-defective rails. Then the prototype can be “set” to operate at the selected threshold value to produce desirable results. Basically, the prototype would be operating at one point on the ROC curve.

For example, looking at the “cumulative” ROC curve in Figure 5.24, and considering the data pertaining to Defect List #2, one could choose the threshold value corresponding to the point where the curve has a first sharp bend toward the right, which would yields PD of ~71% and PFA of ~3%. The corresponding threshold is 80 (not

shown in the plots). Otherwise, if one wants to increase the PD, as willing to accept a higher PFA affecting the results, a wise choice could be to select the threshold corresponding to the second sharp bend (this threshold value is 52), which yields PD of ~87% and PFA of ~13%.

Figure 5.23: ROC curves computed for each of the four runs in the Technology Development Section of TTC's defect farm

Figure 5.24: “Cumulative” ROC curves for the four runs (top) and super-imposition of all the results in two plots, one for each Defect List considered (bottom)

5.8.6 Conclusions

The prototype’s performance in the Technology Development Section of TTC’s rail defect farm was found to be quite encouraging. Based on ROC curves, which quantify the prototype’s detection performance, a detection threshold value can be chosen to maximize the Probability of Detection given an acceptably low Probability of False Positives. As seen in the detection results shown for Defect List #2 (containing defects

larger than 5% of the rail head area and within the stated coverage area of the system), a high Probability of Detection is indeed achievable with an acceptably low Probability of False Positives. For example, from one of the thresholds on the ROC curve of run #3, Defect List #2, the performance offered by the prototype was 80% PD with 5% PFA on pristine portions of rail. Otherwise, from the same ROC curve, a slightly higher threshold yields 72% PD with only 1% PFA. The detection performance is still encouraging when considering the full Defect List #1. It should be noted that the tests were conducted at very low speeds (~1 mph).

5.9 Further studies conducted on the field test data

The experimental data collected during field testing at TTC facility in Colorado constitute an invaluable resource that can be used to conduct further studies as different operational parameters are changed and as the performance offered by those changes can be readily evaluated by computation of ROC curves.

Among the tests performed on these data, one has been the attempt to treat the signals acquired by the deployed receivers using coherent array techniques. The salient results of this analysis are worth to be presented here. The outcomes of the analysis are in agreement with the conclusions drawn in Section 4.2.4.5, where a similar study was independently performed on the results of numerical simulations done in LISA.

Focusing the attention on the receivers on only one side of the transmitter, the closer receiver on the gage side is denoted as #1, the further as #2, and the one toward field side as #3 (same notation used in Section 4.2.4.5). The same applies for the receivers on the other side, thus by extension the correspondent pairs of sensors are

denoted pair #1, #2 and #3. The results here presented were achieved using the delay-and-sum technique. In Section 4.2.4.5 are discussed the difficulties associated with performing “perfectly” coherent sums of the signals in real-time. On this matter, Figure 4.64 and Figure 4.65 which are shown in that section were actually originated from the study reported in this section.

Two possibilities were taken into account. The first, which in the plots legends is denoted as “1+2, 3”, is to sum together only the signals sensed by the two receivers toward the gage side, namely receiver #1 and #2. Then, the usual 3 features would be extracted both from the newly formed signal both from the signal sensed by receiver #3, thus obtaining a total of 6 features per testing point.

The second possibility, which is denoted as “1+2+3”, is to sum up all the three available waveforms (per side) and to extract the usual 3 features on the newly formed signal (this is the same procedure used to obtain the results in Section 4.2.4.5).

In Figure 5.25 are plotted the best results (based on ROC curves computed per each run) that were obtained using each of the two techniques, considering both Defect List #1 both List #2. In the same figure are also plotted the results achieved using the “normal” technique, which in the legends is denoted as “1, 2, 3”, and which are the same as the ones shown in Figure 5.23. The AUC values relative to each curve are both shown in the plots legends, both summarized in Table 5.3, for a quicker comparison.

As seen either in the table or in the figure, the “1, 2, 3” technique guaranteed the best performances over each run and each defect list. Explanations on the reasons behind the fact that treating the sensors in a cumulative/averaged fashion results to be less effective in terms of defect detection are given in Section 4.2.4.5.

Figure 5.25: Comparison of the best results achieved with each of the 3 different techniques

Table 5.3: Summary of the results obtained with the three techniques. Area Under the Curves of ROC plots computed for each run are given. The best results obtained for each run/defect list are highlighted in yellow

DEFECT LIST #1				
	RUN #1	RUN #2	RUN #3	RUN #4
1, 2, 3	0.76692	0.74084	0.79652	0.829
1+2, 3	0.71163	0.72371	0.76077	0.74492
1+2+3	0.66516	0.64979	0.78042	0.66702
DEFECT LIST #2				
	RUN #1	RUN #2	RUN #3	RUN #4
1, 2, 3	0.93053	0.90976	0.95893	0.9425
1+2, 3	0.90118	0.8718	0.90095	0.91101
1+2+3	0.83118	0.8088	0.90887	0.84143

5.10 Conclusions

This chapter presented the development and testing of a rail defect detection prototype that is based on an air-coupled ultrasonic guided waves approach. Several experimental tests have been carried out over time in-house at the UCSD Rail Defect Testing Facility located at Camp Elliott. Many of the system's operational parameters were selected according to the results obtained from those tests. The main issue affecting the system has been represented by low levels of signal-to-noise ratios in the waveforms acquired by the receivers, which is the inherent challenge of air-coupled ultrasound transduction in steel because of the severe acoustic impedance mismatch at the air-steel and steel-air interfaces. Several filtering techniques have been examined during the

different stages of prototype's development and have been tested in terms of attainable SNR gain. In the current version of the system, an acceptable SNR level could be obtained after implementing efficient electrical impedance matching networks that greatly amplify the frequency components excited by the transmitter. An unsupervised statistical algorithm based on Multivariate Outlier Analysis was also designed to take into account the effects of random noise in the acquired signals and to enhance the system's defects-sensitivity.

The system was tested for the first time out of UCSD in October 2014 at the Transportation Technology Center (TTC) in Colorado. The field tests were conducted both at TTC's Technology Development Section and Technology Evaluation Section ("Blind" Section). While test results were obtained in both test sections, only those from the Technology Development Section, where the defect map was known and agreed upon by all participating parties, were presented in this dissertation. The results from the blind tests at the Technology Evaluation Section will be shared with the FRA following the recertification of the site that is planned for the summer of 2015. The prototype's performance in the field tests was found to be quite encouraging. Based on ROC curves, which quantify the prototype's detection performance, a detection threshold value can be chosen to maximize the Probability of Detection given an acceptable Probability of False Positives. Nevertheless, the field tests were conducted at very low speeds (~1 mph). Future efforts should be devoted to improving the system's hardware and software in order to allow for testing at much higher speeds, at values at least comparable to current ultrasonic rail inspections (~10-15 mph). The main issue that prevented the execution of the first field test at higher speeds has been discussed. A temporary solution involving a

sponge-like material that greatly accelerates the damping-out of residual ultrasound energy from reverberations within air has been proposed and seems to be very effective. Clearly, other possible solutions to the problem can be proposed and their efficacy should be investigated. A second field test is currently scheduled for next fall 2015 where testing is planned to be performed at higher speeds.

5.11 Acknowledgments

Part of this chapter will be submitted for publication to the *Structural Health Monitoring* Journal, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Defect detection performance of the non-contact air-coupled ultrasonic guided wave inspection of rails prototype”. The dissertation author will be the primary investigator and author of this paper.

Part of this chapter will be submitted for publication to the *Experimental Mechanics* Journal, Mariani, S., Lanza di Scalea, F.. The current running title of this paper is “Signal to Noise Ratio enhancement for air-coupled wave propagation in steel rail”. The dissertation author will be the primary investigator and author of this paper.

6 Conclusions

6.1 Review of the dissertation and summary of the results

The research discussed in this dissertation investigated the use of air-coupled ultrasonic guided waves for performing non-destructive inspection of railroad tracks. The current state of knowledge of guided waves rail inspection was extended both from a fundamental wave propagation viewpoint and from an experimental prototype development viewpoint.

A unique part of the research has been represented by the development of an experimental prototype for performing rail defect detection. The proposed system can be seen as an upgrade over a previous prototype that was developed by former investigators in the NDE&SHM research group of UCSD. Specifically, the new system replaced the laser-based ultrasound generation with an air-coupled ultrasound transmitter, while retaining the air-coupled ultrasound receivers. Therefore the system is completely based on (non-contact) air-coupled transduction of the ultrasonic waves into and from the rail. The key advantages of the new system over the previous one are the potential for higher testing speeds, a significant cost reduction and an increased easiness of operation and maintenance. The main drawback is a significant decrease in transmitted energy, for which a specially-designed statistical algorithm and very effective filtering techniques have been implemented.

Two different numerical computational techniques were used to analyze the ultrasonic wave propagation occurring each time the experimental prototype executes a test. In particular, a software based on the Local Interaction Simulation Approach (cuLISA3D) was utilized to get insights about the multi-modal wave propagation in the

rail head, and the results showed that the air-coupled transmitter used in the prototype excites several Surface Acoustic Wave modes interacting with each other. The same software was also used to investigate the limits of the detectability of different types of defects, according to their sizes and depths, and when considering the currently available signal-to-noise ratios.

The second approach that has been used to simulate air-coupled ultrasonic wave propagation in the rail is based on Finite Element Analysis, performed with a commercially available software (LS-DYNA). This software was chosen because of its well-known capabilities in performing explicit analysis, which are more convenient than implicit ones when dealing with ultrasonic simulations involving high frequencies. Taking advantage of this software's capabilities, the complete wave propagation occurring not only in the rail under test, but also in the surrounding air domains was simulated. Further studies were carried out both to fully characterize the type of excitation provided by the air-coupled transmitter in use, and to maximize the energy transferred from its active surface to the rail top surface.

Comparisons between numerical results and experimental results from laboratory and field tests have been systematically carried out and consistently confirmed the correctness of the models predictions. In fact, during the course of the research project, many of the prototype's operational parameters were chosen not only from analyzing experimental test results, but also from utilizing predictions and suggestions provided by these numerical models.

The development and testing of the rail defect detection prototype based on the physics thoroughly investigated with those numerical models have been presented in

Chapter 5. In particular, several experimental tests have been carried out over time in-house at the UCSD Rail Defect Testing Facility located at Camp Elliott. Many of the system's operational parameters were selected according to the results obtained from those tests. The main issue affecting the system has been represented by low levels of signal-to-noise ratios (SNR) in the waveforms acquired by the receivers, which is the inherent challenge of air-coupled ultrasound transduction in steel because of the severe acoustic impedance mismatch at the air-steel and steel-air interfaces. Several filtering techniques have been examined during the different stages of prototype's development and have been tested in terms of attainable SNR gain. In the current version of the system, an acceptable SNR level could be obtained after implementing efficient electrical impedance matching networks that greatly amplify the frequency components excited by the transmitter. An unsupervised statistical algorithm based on Multivariate Outlier Analysis was also designed to take into account the effects of random noise in the acquired signals and to enhance the system's defects-sensitivity.

The system was tested for the first time in the field in October 2014 at the Transportation Technology Center (TTC) in Colorado. The field tests were conducted both at TTC's Technology Development Section and Technology Evaluation Section ("Blind" Section). While test results were obtained in both test sections, only those from the Technology Development Section, where the defect map was known and agreed upon by all participating parties, were presented in this dissertation. The results from the blind tests at the Technology Evaluation Section will be shared with the FRA following the recertification of the site that is planned for the summer of 2015. The prototype's performance in the field tests was found to be quite encouraging. Based on ROC curves,

which quantify the prototype's detection performance, a detection threshold value can be chosen to maximize the Probability of Detection given an acceptable Probability of False Positives. Nevertheless, the field tests were conducted at very low speeds (~1 mph).

6.2 Recommendations for future work

The TTC October 2014 field tests were conducted at very low speed. Future efforts should be devoted to improving the system's hardware and software in order to allow for testing at much higher speeds, in particular, at values at least comparable to current ultrasonic rail inspections (~10-15 mph). The main issue that prevented the execution of the first field test at higher speeds has been discussed. A temporary solution involving a sponge-like material that greatly accelerates the damping-out of residual ultrasound energy from reverberations within air has been proposed and seems to be very effective. Clearly, other possible solutions to the problem can be proposed and their efficacy should be investigated. A second field test is currently scheduled for fall 2015 where testing is planned to be performed at higher speeds.

Further analyses can be performed on the datasets produced from the LISA simulations on the models containing different rail internal defects in the effort to identify possible ways to maximize the prototype's sensitivity to defects. In particular, additional analyses should focus on varying the relative positions transmitter - receivers, as well as on finding other statistical features to extract from the acquired waveforms capable of providing increased Probability of Detections while decreasing the Probability of False Positives. One possible way of carrying out such analyses would be to implement genetic algorithms capable of constantly varying the different parameters involved with the

computation of Damage Indexes and to use ROC curves to assess the goodness of each parameter's selection.

Based on the same LISA-generated numerical results as well as on experimental results, pattern recognition algorithms can be also implemented to study the feasibility of not only detecting defects, but also quantifying them, in terms of predicting the flaw type, position, size and depth.

Increasing the SNR level to the possible maximum extents has been proved to be crucial to ensure a consistent defect detection of the air-coupled system. Therefore, other studies can focus on understanding possible other ways of increasing either the energy delivered by the transmitter, or the quality of filtering at the receivers.

References

- [1] Office of Safety Analysis, Federal Railroad Administration, [Online]. Available: <http://safetydata.fra.dot.gov/officeofsafety/default.aspx>.
- [2] D. F. Cannon, K.-O. Edel, S. L. Grassie and K. Sawley, "Rail Defects: an overview," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 26, pp. 865-886, 2003.
- [3] Federal Railroad Administration, Office of Railroad Safety, "Track Inspector Rail Defect Reference Manual," 2011, August.
- [4] Sperry Rail Service, Rail Defect Manual, Danbury, CT, 1968.
- [5] Nordco, Nordco Rail Flaw Identification Handbook.
- [6] F. Lanza di Scalea, "Chapter 15: Special Applications of Ultrasonic Testing, in Non-destructive Testing Handbook, 3rd edition," in *Ultrasonic Testing Applications in the Railroad Industry*, P.O. Moore, ed., American Society for Nondestructive Testing, 2007, pp. 535-552.
- [7] F. Anon, "Rail-flaw Detection. A Science that Works," *Railway Track and Structures*, vol. 86(5), pp. 30-32, 1990.
- [8] D. S. Grewal, "Improved Ultrasonic Testing of Railroad Rail for Transverse Discontinuities in the Rail Head Using Higher Order Rayleigh (M21) Waves," *Materials Evaluation*, vol. 54(9), pp. 983-986, 1996.
- [9] D. E. Bray and M. Najm, "The Effect of Material Deformation on the Velocity of Critically Refracted Shear Waves in Railroad Rail," in *Proc. Ultrasonics International Conference*, Halifax, Nova Scotia, 1983.
- [10] National Transportation Safety Board (NTSB), Reports HZM-94/01 and RAB-08/05.
- [11] S. B. Palmer, S. Dixon, R. S. Edwards and X. Jian, "Transverse and longitudinal crack detection in the head of rail tracks using Rayleigh wave-liked wideband guided ultrasonic waves," in *Nondestructive Evaluation and Health Monitoring of Aerospace Materials, Composites and Civil Infrastructure IV*, SPIE 5767, 2005.
- [12] G. A. Alers and A. Manzanares, "Use of surface skimming SH wave to measure thermal and residual stress in installed railroad tracks," *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 9, p. 1757, 1990.
- [13] G. A. Alers, "Railroad Rail Flaw Detection System Based on Electromagnetic Acoustic Transducers," U.S. Department of Transportation Report DOTIFRA/ORD-88/09, 1988.
- [14] J. L. Rose, C. M. Lee, T. R. Hay, Y. Cho and I. K. Park, "Rail inspection with guided waves," in *Proceedings of the 12th Asia-Pacific Conference on NDT*, Auckland, New Zealand, 2006, November.
- [15] S. Coccia, R. Phillips, I. Bartoli, S. Salamone, P. Rizzo and F. Lanza di Scalea, "On-line High-speed Rail Defect Detection - part II," Report No. DOT/FRA/ORD-

12/02, 2012, March.

- [16] P. Rizzo, S. Coccia, I. Bartoli and F. Lanza di Scalea, "Non-Contact Rail Monitoring by Ultrasonic Guided Waves," in *Chapter 145 of Encyclopedia of Structural Health Monitoring*, C. Boller, F. Chang and Y. Fujino, Eds., Chichester, Johns Wiley & Sons, 2009, pp. 2397-2410.
- [17] S. Kenderian, B. B. Djordjevic, D. Cerniglia and G. Garcia, "Dynamic railroad inspection using the laser-air hybrid ultrasonic technique," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 48(6), pp. 336-341, 2006.
- [18] F. Lanza di Scalea, S. Coccia, I. Bartoli, S. Salamone and P. Rizzo, "Defect Detection In Objects Using Statistical Approaches". Patent Application Publication US-2011-0238336-A1.
- [19] S.-C. Wooh. Washington, DC: U.S. Patent and Trademark Office Patent No. 6,324,912, 2001.
- [20] S.-C. Wooh, C. C. Kim and C. Wei, Real-time Processing of Continuous Doppler Signals for High-Speed Monitoring of Rail Tracks, D. O. Thompson and D. E. Chimenti, Eds., New York: Plenum Press, 1998, pp. 1681-1688.
- [21] V. V. Murav'ev and E. V. Boyarkin, "Nondestructive testing of the structural-mechanical state of currently produced rails on the basis of the ultrasonic wave," *Russian Journal of Nondestructive Testing*, vol. 39, pp. 24-33, 2003.
- [22] S. Coccia, Ultrasonic guided waves for structural health monitoring and application to rail inspection prototype for the Federal Railroad Administration, University of California, San Diego: ProQuest Dissertations and Theses, 2007.
- [23] Title 49 Code of Federal Regulations Part 213, Track Safety Standards.
- [24] R. Care, S. Clark, M. Dembosky and A. Doherty, "Why rails crack, gauge corner cracking on the British network: analysis," *The Arup Journal*, vol. 1, pp. 16-19, 2006.
- [25] A. Herring, "Improvement in modes of detecting defects in railroad-rails". Patent 189,858, 1887.
- [26] L. Oukhellou, P. Aknin and J. P. Perrin, "Dedicated sensor and classifier of rail head defects," *Control Engineering Practice*, vol. 7(1), pp. 57-61, 1999.
- [27] J. J. Marais and K. C. Mistry, "Rail integrity management by means of ultrasonic testing," *Fatigue and Fracture of Engineering Materials and Structures*, vol. 26, pp. 931-938, 2003.
- [28] C. Robin, "Rail flaw detection: overview and needs for future developments," *NDT & E International*, vol. 37(2), pp. 111-118, 2004.
- [29] K. Bruzelius and D. Mba, "An initial investigation on the potential applicability of Acoustic Emission to rail track fault detection," *NDT&E International*, vol. 37, pp. 507-516, 2004.
- [30] P. Wilcox, M. Evans, B. Pavlakovic, D. Alleyne, K. Vine, P. Cawley and M. Lowe, "Guided wave testing of rail," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 45(6), pp. 413-420, 2003.

- [31] F. Lanza di Scalea and J. McNamara, "Ultrasonic NDE of railroad tracks: air-coupled cross-sectional inspection and long-range inspection," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 45(6), pp. 394-401, 2003.
- [32] P. Cawley, M. J. S. Lowe, D. N. Alleyne, B. Pavlakovic and P. Wilcox, "Practical long range guided wave inspection - applications to pipes and rail," *Materials evaluation*, vol. 61(1), pp. 66-74, 2003.
- [33] Z. Zhou, Z. Fengl, Y. Gao and J. Bai, "An outline of applications of ultrasonic guided waves in non-destructive testing of large structures," in *Asia-Pacific Conference on NDT, 5th – 10th*, Auckland, New Zealand, 2006.
- [34] P. Cawley and D. Alleyne, "The use of Lamb waves for the long range inspection of large structures," *Ultrasonics*, vol. 34, pp. 287-290, 1996.
- [35] J. L. Rose, "A baseline and vision of ultrasonic guided wave inspection potential," *Journal of Pressure Vessel Technology*, vol. 124, pp. 273-282, 2002.
- [36] J. L. Rose, *Ultrasonic Waves in Solid Media*, New York: Cambridge University Press, 2014.
- [37] W. Wilkie, R. Bryant, J. High, R. Fox, R. Hellbaum, A. Jalink, B. Little and P. Mirick, "Low-cost piezocomposite actuator for structural control applications," in *SPIE 7th Annual Int. Symp. on Smart Structures and Materials*, Newport Beach, CA, 2000.
- [38] A. A. Bent, *Active fiber composites for structural actuation*, PhD Dissertation Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 1997.
- [39] J. L. Rose, S. P. Pelts and M. J. Quarry, "A comb transducer for mode control in guided wave NDE," in *IEEE Ultrasonics Symposium*, 1997.
- [40] P. Rizzo and F. Lanza di Scalea, "Monitoring in Cable Stays via Guided Wave Magnetostrictive Ultrasonics," *Materials Evaluation*, vol. 62, no. 10, pp. 1057-1065, 2004.
- [41] P. D. Wilcox, "Omni-directional guided wave transducer arrays for the rapid inspection of large areas of plate structures," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, pp. 699-709, 2003.
- [42] M. Luukkala and P. Meriläinen, "Metal plate testing using airborne ultrasound," *Ultrasonics*, vol. 11, no. 5, pp. 218-221, 1973.
- [43] H. Loertscher, B. Grandia, J. Strycek and W. A. Grandia, "Airscan transducers, technique and applications," *NDT.net*, vol. 1, no. 9, pp. 1-4, 1996.
- [44] H. M. Matt, I. Bartoli and F. Lanza di Scalea, "Ultrasonic guided wave monitoring of composite wing skin-to-spar bonded joints in aerospace structures," *Journal of the Acoustical Society of America*, vol. 118, pp. 2240-2252, 2005.
- [45] C. A. Paget, S. Grondel, K. Levin and C. Delebarre, "Damage assessment in composites by Lamb waves and wavelet coefficients," *Smart Materials and Structures*, vol. 12, pp. 393-402, 2003.
- [46] H. Sohn, G. Park, J. Wait, N. Limback and C. R. Farrar, "Wavelet-based active sensing for delamination detection in composite structures," *Smart Materials and*

- Structures*, vol. 13, pp. 153-160, 2003.
- [47] F. Lanza di Scalea, H. M. Matt, I. Bartoli, S. Coccia, G. Park and C. R. Farrar, "Health monitoring of UAV wing skin-to-spar joints using guided waves and macro fiber composite transducers," *Journal of Intelligent Material Systems and Structures*, vol. 18, pp. 373-388, 2007.
 - [48] S. S. Kessler, S. M. Spearing and C. Soutis, "Damage detection in composite materials using Lamb wave methods," *Smart Materials and Structures*, vol. 11, pp. 269-278, 2002.
 - [49] C. S. Wang and F.-K. Chang, "Diagnosis of impact damage in composite structures with built-in piezoelectrics network," in *Proceedings of SPIE*, 3990, pp. 13-19, 2000.
 - [50] A. K. Mal, S. Banerjee and F. Ricci, "Automated structural health monitoring system using acoustic emission and modal data," in *Proceedings of SPIE*, 5394, pp. 1-10, 2004.
 - [51] F. Ricci, E. Monaco, H. Baid and A. Mal, "Guided waves in a disbanded honeycomb composite structure," in *IWSHM*, Stanford, 2013.
 - [52] J. L. Rose, M. J. Avioli, P. Mudge and R. Sanderson, "Guided wave inspection potential of defects in rail," *NDT & E International*, vol. 37(2), pp. 153-161, 2004.
 - [53] P. J. Shull, *Nondestructive Evaluation: Theory, Techniques, and Applications*, CRC Press, 2002.
 - [54] L. Pochhammer, "Ueber die Fortpflanzungsgeschwindigkeiten kleiner Schwingungen in einem unbegrenzten isotropen Kreiscylinder," *Journal für die reine und angewandte Mathematik*, vol. 81, pp. 324-336, 1876.
 - [55] C. Chree, "The Equations of an Isotropic Elastic Solid in Polar and Cylindrical Coordinates their Solution and Application," *Transactions of the Cambridge Philosophical Society*, vol. 14, p. 250, 1889.
 - [56] D. Bancroft, "The velocity of longitudinal waves in cylindrical bars," *Physical Review*, vol. 59(7), p. 588, 1941.
 - [57] D. C. Gazis, "Three-Dimensional Investigation of the Propagation of Waves in Hollow Circular Cylinders. I. Analytical Foundation," *The journal of the Acoustical Society of America*, vol. 31(5), pp. 568-573, 1959.
 - [58] J. Zemanek Jr, "An experimental and theoretical investigation of elastic wave propagation in a cylinder," *The Journal of the Acoustical Society of America*, vol. 51(1B), pp. 265-283, 1972.
 - [59] J. W. Strutt (Lord Rayleigh), "On waves propagated along the plane surface of an elastic solid," *Proceedings of the London Mathematical Society*, vol. 17, pp. 4-1, 1885.
 - [60] R. Stoneley, "Elastic waves at the surface of separation of two solids," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, pp. 416-428, 1924.
 - [61] H. Lamb, "On Waves in an Elastic Plate," *Proc. of the Royal Soc. (London)*, vol. Ser. A 93, p. 114, 1917.

- [62] D. C. Worlton, "Experimental confirmation of Lamb waves at megacycle frequencies," *Journal of Applied Physics*, vol. 32(6), pp. 967-971, 1961.
- [63] H. Kolsky, *Stress Waves in Solids*, New York: Dover, 1963.
- [64] H. F. Pollard, *Sound Waves in Solids*, London: Pion Ltd, 1977.
- [65] J. D. Achenbach, *Wave propagation in elastic solids*, Amsterdam: North-Holland Pub. Co., 1973.
- [66] B. A. Auld, *Acoustic fields and waves in solids*, 2nd ed. R.E. Krieger, Malabar, Fla, 1990.
- [67] K. F. Graff, *Wave motion in elastic solids*, New York: Dover, 1991.
- [68] I. A. Viktorov, *Rayleigh and Lamb waves: physical theory and applications*, New York: Plenum Press, 1967.
- [69] Y. Cho and J. L. Rose, "A boundary element solution for a mode conversion study on the edge reflection of Lamb waves," *The Journal of the Acoustical Society of America*, vol. 99, pp. 2097-2109, 1996.
- [70] B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge: Cambridge University Press, 1998.
- [71] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave-fields," *J. Biomed. Opt.*, vol. 15, no. 2, p. 021314, 2010.
- [72] M. J. S. Lowe, *Plate waves for the NDT of diffusion bonded titanium*, PhD Thesis, University of London, 1993.
- [73] G. Liu and J. D. Achenbach, "Strip element method to analyze wave scattering by cracks in anisotropic laminated plates," *J. Appl. Mech.*, vol. 62, pp. 607-613, 1995.
- [74] I. Bartoli, *Structural health monitoring by ultrasonic guided waves*, University of California, San Diego: ProQuest Dissertations and Theses, 2007.
- [75] P. P. Delsanto and M. Scalerandi, "A spring model for the simulation of the propagation of ultrasonic pulses through imperfect contact interfaces," *The Journal of the Acoustical Society of America*, vol. 104 (5), pp. 2584-2591, 1998.
- [76] D. N. Alleyne and P. Cawley, "Optimization of Lamb wave inspection techniques," *NDT&E Int.*, vol. 25, pp. 11-22, 1992.
- [77] P. P. Delsanto, T. Whitcombe, H. H. Chaskelis and R. B. Mignogna, "Connection machine simulation of ultrasonic wave propagation in materials. I: the one-dimensional case," *Wave Motion*, vol. 16, pp. 65-80, 1992.
- [78] P. P. Delsanto, R. S. Schechter, H. H. Chaskelis, R. B. Mignogna and R. B. Kline, "Connection machine simulation of ultrasonic wave propagation in materials. II: the two-dimensional case," *Wave Motion*, vol. 20, pp. 295-314, 1994.
- [79] P. P. Delsanto, R. S. Schechter and R. B. Mignogna, "Connection machine simulation of ultrasonic wave propagation in materials. III: the three-dimensional case," *Wave Motion*, vol. 26, pp. 329-339, 1997.

- [80] B. C. Lee and W. J. Staszewski, "Lamb wave propagation modelling for damage detection: I. Two-dimensional analysis," *Smart Materials and Structures*, vol. 16(2), pp. 249-259, 2007.
- [81] MONIT SHM, cuLISA3D Theory and User Information.
- [82] B. C. Lee and W. J. Staszewski, "Lamb wave propagation modelling for damage detection: II: Damage monitoring strategy," *Smart Materials and Structures*, vol. 16, pp. 260-274, 2007.
- [83] P. Paćko, T. Bielak, A. B. Spencer, W. J. Staszewski, T. Uhl and K. Worden, "Lamb wave propagation modelling and simulation using parallel processing architecture and graphical cards," *Smart Materials and Structures*, vol. 21(7), p. 075001, 2012.
- [84] D. Alleyne and P. Cawley, "A two-dimensional Fourier transform method for the measurement of propagating multimode signals," *The Journal of the Acoustical Society of America*, vol. 89(3), pp. 1159-1168, 1991.
- [85] F. Moser, L. J. Jacobs and J. Qu, "Modeling elastic wave propagation in waveguides with the finite element method," *Ndt & E International*, vol. 32(4), pp. 225-234, 1999.
- [86] P. Packo, T. Bielak, A. B. Spencer, T. Uhl, W. J. Staszewski, K. Worden, T. Barszcz, P. Russek and K. Wiatr, "Numerical simulations of elastic wave propagation using graphical processing units—Comparative study of high-performance computing capabilities," *Computer Methods in Applied Mechanics and Engineering*, vol. 290, pp. 98-126, 2015.
- [87] D. Hesse and P. Cawley, "Surface wave modes in rails," *The Journal of the Acoustical Society of America*, vol. 120(2), pp. 733-740, 2006.
- [88] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, 1st edition, Englewood Cliffs, NJ: Prentice Hall, 1998.
- [89] A. Pantano and D. Cerniglia, "Simulation of laser generated ultrasound with application to defect detection," *Applied Physics A*, vol. 91(3), pp. 521-528, 2008.
- [90] Livermore Software Technology Corporation (LSTC), *LS-DYNA Keyword User's Manual Volume I*, 2014.
- [91] A. Pantano and D. Cerniglia, "Simulation of laser-generated ultrasonic wave propagation in solid media and air with application to NDE," *Applied Physics A*, vol. 98(2), pp. 327-336, 2010.
- [92] J. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114(2), pp. 185-200, 1994.
- [93] Q. Qi and T. L. Geers, "Evaluation of the perfectly matched layer for computational acoustics," *Journal of Computational Physics*, vol. 139(1), pp. 166-183, 1998.
- [94] U. Basu and A. K. Chopra, "Perfectly matched layers for transient elastodynamics of unbounded domains," *International Journal for Numerical Methods in Engineering*, vol. 59(8), pp. 1039-1074, 2004.

- [95] U. Basu, "Explicit finite element perfectly matched layer for transient three-dimensional elastic waves," *International Journal for Numerical Methods in Engineering*, vol. 77(2), pp. 151-176, 2009.
- [96] A. Fathi, B. Poursartip and L. Kallivokas, "Time-domain hybrid formulations for wave simulations in three-dimensional PML-truncated heterogeneous media," *International Journal for Numerical Methods in Engineering*, vol. 101(3), pp. 165-198, 2015.
- [97] M. B. Drozd, Efficient finite element modelling of ultrasound waves in elastic media, Doctoral dissertation, Imperial College London, 2008.
- [98] Livermore Software Technology Corporation (LSTC), LS-DYNA Keyword User's Manual Volume II Materials Models, 2014.
- [99] L. Hammond and R. Grzebieta, "The requirement for hydrostatic initialisation in LS-DYNA/USA finite element models," *Shock and Vibration*, vol. 7(2), pp. 57-65, 2000.
- [100] M. C. Bhardwaj, "Ultrasonic transducer for high transduction in gases and method for non-contact ultrasound transmission into solid materials". Washington, DC Patent U.S. Patent No. 6,311,573, 6 November 2001.
- [101] C. Biateau, B. Hosten and D. Roziere, "Measurement of air-coupled transducer characteristics for ultrasonic non-destructive evaluation," *Quantitative Nondestructive Evaluation*, vol. 615(1), pp. 921-928, 2002.
- [102] M. J. Lowe, P. Cawley, J. Y. Kao and O. Diligent, "The low frequency reflection characteristics of the fundamental antisymmetric Lamb wave a_0 from a rectangular notch in a plate," *The Journal of the Acoustical Society of America*, vol. 112(6), pp. 2612-2622, 2002.
- [103] L. L. Thompson and R. Huan, "Implementation of exact non-reflecting boundary conditions in the finite element method for the time-dependent wave equation," *Computer Methods in Applied Mechanics and Engineering*, vol. 187(1), pp. 137-159, 2000.
- [104] D. Givoli and J. B. Keller, "Non-reflecting boundary conditions for elastic waves," *Wave Motion*, vol. 12, pp. 261-279, 1990.
- [105] Z. You and W. Lord, "Finite element study of elastic wave interaction with cracks," *Review of Progress in Quantitative Nondestructive Evaluation*, vol. Springer US, pp. 109-116, 1989.
- [106] X. Jian, S. Dixon, N. Guo and R. Edwards, "Rayleigh wave interaction with surface-breaking cracks," *Journal of Applied physics*, vol. 101(6), p. 064906, 2007.
- [107] K. G. Hall, "Crack depth measurement in rail steel by Rayleigh waves aided by photoelastic visualization," *Non-destructive Testing*, vol. 9(3), pp. 121-126, 1976.
- [108] A. N. Akansu and R. A. Haddad, Multiresolution signal decomposition: transforms, subbands, and wavelets, Boston, MA: Academic Press, 1992.
- [109] B. Widrow, J. M. McCool, M. Larimore and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proceedings of the IEEE*, vol. 64(8), pp. 1151-1162, 1976.

- [110] N. Wiener, Extrapolation, interpolation, and smoothing of stationary time series (Vol. 2), Cambridge, MA: MIT press, 1949.
- [111] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Mathematics of computation*, vol. 19(90), pp. 297-301, 1965.
- [112] A. V. Oppenheim and R. W. Schaffer, Discrete-Time Signal Processing, Upper Saddle River, NJ: Prentice-Hall, 1999.
- [113] G. Turin, "An introduction to matched filters," *Information Theory, IRE Transactions on*, vol. 6(3), pp. 311-329, 1960.
- [114] B. M. Oliver, J. R. Pierce and C. E. Shannon, "The philosophy of PCM," *Proceedings of the IRE*, vol. 36(11), pp. 1324-1331, 1948.
- [115] G. Petersen, "L-matching the output of a RITEC gated amplifier to an arbitrary load," RITEC Inc., USA, 2006.
- [116] V. Barnett and T. Lewis, Outliers in Statistical Data, New York: Wiley, 1994.
- [117] K. Worden, "Structural fault detection using a novelty measure," *Journal of Sound and Vibration*, vol. 201, pp. 85-101, 1997.
- [118] S. Coccia, R. Phillips, I. Bartoli, S. Salamone, F. Lanza di Scalea, M. Fateh and G. Carr, "Ultrasonic Guided Wave Detection of Rail defects: Prototype Development and Field Testing for the Federal Railroad Administration," in *Proceedings of 7th International Workshop on Structural Health Monitoring*, Stanford University, September 2009.
- [119] S. Coccia, I. Bartoli, R. Phillips, S. Salamone and F. Lanza di Scalea, "Noncontact Rail Defect Detection: Fourth Field Test," Technical Report No. SSRP-09/04 to the Federal Railroad Administration, University of California, San Diego, 2009.
- [120] S. Coccia, I. Bartoli, F. Lanza di Scalea and P. Rizzo, "Non-contact rail defect detection: first and second field tests," UCSD Report SSRP-07/15, 2007.
- [121] S. Coccia, R. Phillips, I. Bartoli, S. Salamone and F. Lanza di Scalea, "Non-contact rail defect detection: third field test," UCSD Report SSRP-08/02, 2008.