

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Ultrasonic guided waves for structural health monitoring and application to rail inspection prototype for the Federal Railroad Administration

Permalink

<https://escholarship.org/uc/item/8cp2j7mn>

Author

Coccia, Stefano

Publication Date

2007

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Ultrasonic Guided Waves for Structural Health Monitoring and Application to
Rail Inspection Prototype for the Federal Railroad Administration**

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

Doctor of Philosophy

in

Structural Engineering

by

Stefano Coccia

Committee in Charge:

Professor Francesco Lanza di Scalea, Chair
Professor William S. Hodgkiss
Professor Stefan Savage
Professor Michael D. Todd
Professor Qiang Zhu

2007

Copyright,
Stefano Coccia, 2007
All rights reserved.

This dissertation of Stefano Coccia is approved, and it is
acceptable in quality and form for publication on microfilm

Chair

University of California, San Diego

2007

DEDICATION

To all the people who have been supportive over the years of research work and especially my family. Without them I would not have been able to accomplish this important task.

EPIGRAPH

*“The truth will set you free,
but first it will make you miserable.”*

J. A. Garfield

TABLE OF CONTENTS

Signature Pageiii
Dedication.....	iv
Epigraph....	v
Table of Contents	vi
List of Symbols.....	xi
List of Figures.....	xiv
List of Tables	xxv
Acknowledgments	xxvii
Vita	xxx
Abstract	xxxiv
1 Introduction	1
1.1 Health monitoring of railroad tracks. Introductory discussion and motivation for research.....	1
1.2 Outline of the numerical and experimental approach of the dissertation.....	2
2 State of the art review.....	5
2.1 Description of structural defect in rails	5
2.1.1 Rail manufacturing defects.....	8
2.1.2 Rail defects caused by improper use or bad maintenance.....	14
2.1.3 Rail defects due to wearing mechanisms of the rolling surfaces and to fatigue (Rolling Contact Fatigue – RCF defects)	16

2.2	Current rail inspection methods.....	26
2.2.1	Historical background	26
2.2.2	Current inspection systems.....	27
2.3	Ultrasonic guided waves for defect detection in rails	33
2.3.1	Advantages of guided waves for rail inspection.....	37
2.3.2	Laser generation of ultrasonic guided waves	38
2.4	Modeling guided wave propagation in rails	40
3	Modeling ultrasonic wave propagation in waveguides of arbitrary cross-section.....	43
3.1	Introduction	43
3.2	Guided Waves in an unbounded medium.....	43
3.3	Guided waves in a bounded medium. Case of a single isotropic plate	46
3.4	Semi Analytical Finite Element Method	50
3.4.1	SAFE Mathematical Framework.....	52
3.4.2	Modeling the attenuation in the rail.....	59
3.4.3	Rail as undamped waveguide	61
3.4.4	Group and energy velocity	63
3.4.5	Methods to accelerate the SAFE solution process	64
3.4.6	Mode tracking algorithm	69
3.5	Unforced solution of guided waves in rails.....	71
3.6	Forced solution for guided waves in rails.....	85
3.7	Response of the rail to a laser excitation	91
3.7.1	Distribution of in-plane strain energy in the rail cross-section	102

3.7.2	Influence of the excitation pattern on the propagating mode	109
3.7.3	Effect of the frequency on the penetration depth of the guided wave.....	122
3.8	Conclusions	125
4	Defect-sensitive feature extraction from ultrasonic waves	127
4.1	Introduction	127
4.2	Continuous and discrete wavelet analysis	128
4.2.1	Theory behind wavelet analysis	129
4.2.2	Discrete Wavelet Transform (DWT) analysis for signal decomposition.....	133
4.2.3	DWT for Data De-noising and High Speed Data Manipulation.....	141
4.3	Digital processing of ultrasonic signals.....	145
4.3.1	Digital filtering	145
4.3.2	Spatial averaging algorithm for enhancement of defect detection in “reflection mode”.....	146
4.4	Feature extraction in ultrasonic rail inspection	158
4.4.1	Joint time-frequency domain feature extraction.....	158
4.4.2	DWT-based feature extraction. A synopsis of univariate and multivariate approaches.....	168
4.4.3	Feature extraction for D.I. computation	171
4.5	Validation of signal processing for defect detection in rails	175
4.5.1	Introduction	175
4.5.2	Parametric study of Damage Index computation using the DWT	178

4.5.3	Parametric study of Damage Index computation using Digital Filters	186
4.6	Conclusions	191
5	Development of Rail Inspection Prototype for the Federal Railroad Administration	193
5.1	Introduction	193
5.2	Stage I: Hardware and Software	194
5.2.1	Q-Switched laser generator and optical elements	194
5.2.2	Air-coupled Transducers and Preamplifiers	197
5.2.3	D.I. Computation	199
5.2.4	Hardware Controller and DAQ Boards	201
5.2.5	Software implementation	205
5.3	Stage II. Prototype deployment	222
5.3.1	Hardware layout	222
5.3.2	Software implementation	225
5.4	Stage III. Prototype upgrade	235
5.4.1	New cart	235
5.4.2	Improved sensor arrangement	236
5.4.3	Software upgrade	239
5.5	Field testing	241
5.5.1	First field test	241
5.5.2	Second field test	258
5.6	Conclusions	265

6	Conclusions	268
6.1	Review of the dissertation and summary of the results	268
6.2	Recommendations for future work	270
7	References	271

LIST OF SYMBOLS

A	Eigensystem matrix
$A^{(e)}$	Area of triangular element stiffness matrix
$A_{1,2}$	Eigencoefficients of Isotropic Plate Solution
A_n	Level n approximations
α_m	Modal participation factor
B	Eigensystem matrix
$B_{1,2}$	Eigencoefficients of Isotropic Plate Solution
C	Elastic stiffness matrices
$\tilde{\mathbf{C}}$	Complex stiffness matrix
$\hat{\mathbf{C}}_e$	Reduced element stiffness matrix
c_1	Dilatation Wave Speed in Isotropic Plate
c_2	Rotational Wave Speed in Isotropic Plate
c_{gr}	Group velocity
C_{ij}	Elastic constant
$Xcorr$	Cross correlation
c_p	Phase velocity
$D.I.$	Damage Index
D_n	Level n details
δ_{ij}	Kronecker delta
∇	Divergence vector
Δf	Bandwidth
E	Young's Modulus
e_k	Particle kinetic energy
e_p	Particle potential energy
ε_{ij}	Normal strain
$\hat{\varepsilon}^{(e)}$	Reduced element strain
f	Frequency
F	Fourier transform
$f(t)$	Time dependant function
\bar{f}	Fourier transform of f
f_s	Signal sampling rate
$\varphi(t)$	Mother wavelet scaling function
$\Phi_m^{\mathbf{R}}$	Right eigenvector
$\Phi_m^{\mathbf{L}}$	Left eigenvector
$\Phi_m^{\mathbf{Rup}}$	Upper Right eigenvector

Φ_m	m -th eigenvector
G	Shear Modulus
G_s	Gabor shaping factor
γ_{ij}	Shear strain
h	Plate half thickness
HT	Hilbert Transform
i	Imaginary unit
ID	Internal Defect
Im	Imaginary component
J	Joint
K	Kinetic Energy
K_L	Longitudinal Waves Attenuation
K_T	Transverse Waves Attenuation
$\mathbf{k}_1^{(e)}, \mathbf{k}_2^{(e)}, \mathbf{k}_3^{(e)}$	Element stiffness matrices
$\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3$	Assembled global stiffness matrices
η	Mother wavelet central frequency
LU P	Matrix operators for LUP decomposition
L_1	Distance Laser Source Air-coupled Sensor
L_d	Distance Laser Source Crack
$\mathbf{L}_{X,Y,Z}$	Compatibility Equation Matrices
$\mathbf{m}^{(e)}$	Element mass matrix
M	Number of total degrees of freedom of the system
\mathbf{M}	Assembled global mass matrix
n	Number of largest-in-magnitude eigenvalues
N	Sensing positions in the spatial averaging process
$\mathbf{N}(y, z)$	Shape functions matrix
\mathbf{P}	Poynting vector
\mathbf{p}	Point load
$\hat{\mathbf{p}}$	Spectral amplitudes of nodal load
POD	Probability of Detection
$Pwf(u, s)$	Energy density spectrum of CWT
\mathbf{Q}	Eigensystem solution vector
$\mathbf{q}^{(e)}$	Element nodal displacement vector
Re	Real component
RMS	Root mean square
σ	Stress vector
σ_{ij}	Stress
σ	Eigenvalue shift
σ_ω	Spread of motherwavelet frequency window
σ_t	Spread of motherwavelet time window
s	Scaling parameter

s_{MAX}	Max wavelet scale
S_s	Sensor spacing
t	Time
\mathbf{t}	Traction vector
\mathbf{T}	Matricial operator
\mathbf{u}	Displacement vector
u	Translation parameter
\mathbf{U}	Global vector of displacements
$\dot{\mathbf{U}}$	Global vector of velocities
$\ddot{\mathbf{U}}$	Global vector of accelerations
ξ	Wavenumber
$\bar{\psi}$	Vector potential of displacement in isotropic plate
$\psi_{u,s}$	Scaled and shifted mother wavelet function
\mathbf{VI}	Virtual Instrument
wc	Wavelet coefficients
$Wx(u,s)$	CWT coefficient
x,y,z	Cartesian coordinates
x_s	Distance of excitation source
$\Delta\omega$	Frequency range
η	Viscosity tensor
γ	Mother wavelet central frequency
θ	Angle of wave propagation direction with respect to y axis
λ	Wavelength
λ, μ	Lamé constants
ν	Poisson's ratio
\mathbf{v}	Shift-inverted system eigenvalue
ρ	Density of material
Φ	Scalar potential of displacement vector in isotropic plate
Ψ	Vector potential of displacement vector in isotropic plate
ω	Angular frequency
Ω	Waveguide domain
∞	Infinity

LIST OF FIGURES

Figure 2.1: Rail defects, operational restrictions and remedial actions for the U.S. Department of Army	6
Figure 2.2: Relative position of planes through a rail	8
Figure 2.3: Appearance of different cases of transverse defects	9
Figure 2.4: Transverse defects. Case of coupling with shelling	10
Figure 2.5: General appearance of an horizontal split head	11
Figure 2.6: General appearance of a vertical split head	12
Figure 2.7: Case of a head/web separation	12
Figure 2.8: General appearance of a piped rail.....	13
Figure 2.9: Sketch of cases of mill defects.....	13
Figure 2.10: General appearance of a split web	14
Figure 2.11: Sketch of an engine burn.....	15
Figure 2.12: General appearance of engine burn fractures.....	15
Figure 2.13: Sketch of a crushed head.....	16
Figure 2.14: Sketch of a case of torched cut rail	16
Figure 2.15: Wheel contact on a curve with the “high” rail and the “low” rail	17
Figure 2.16: Contact stresses on tight curved track.....	17
Figure 2.17: Contact patch pressure and forces causing surface elastic deformation of the rail	18
Figure 2.18: Example of the broken base defect	19
Figure 2.19: Sketches of bolt hole cracks.....	20
Figure 2.20: Complete rail break (broken rail).....	21

Figure 2.21: Sketch of the corrugation phenomenon	22
Figure 2.22: General appearance of the end batter defect	22
Figure 2.23: Sketch of the flaking phenomenon	23
Figure 2.24: General appearance of rail wear	24
Figure 2.25: Example of shelling	25
Figure 2.26: Trapped fluids causing cracking to worsen	25
Figure 2.27: Ultrasonic probe wheel transducer arrangements	29
Figure 2.28: Scheme of the contact induction technique	30
Figure 2.29: Guided waves propagating within a confined geometry.....	33
Figure 2.30: Comb transducer excitation of guided waves in a typical 115-lb A.R.E.M.A. rail	35
Figure 2.31: Angle beam ultrasonics for guided waves excitation in a typical 115-lb A.R.E.M.A. rail	36
Figure 2.32: Rail defect detection: “long-range method”	37
Figure 2.33: Laser generation of ultrasound in (a) thermoelastic regime and (b) ablative regime.....	39
Figure 3.1: Guided wave propagation within a thin isotropic plate	46
Figure 3.2: Waveguide in vacuum: infinitely wide plate.	52
Figure 3.3: Dispersion curves of a 115 lb A.R.E.M.A. section rail; high frequency resolution is needed to correctly interpret the solutions in proximity of two modes	65
Figure 3.4: Wavenumber as a function of frequency for damped rail. Case of double solutions.....	71
Figure 3.5: Mesh of the rail cross-section generated by ABAQUS™	72
Figure 3.6: Wavenumbers as a function of frequency.....	73

Figure 3.7: Phase velocity dispersion curves	74
Figure 3.8: Mode filtering by thresholding the phase velocity	75
Figure 3.9: Energy Velocity (Poynting vector) dispersion curves	76
Figure 3.10: Imaginary part of the wavenumber (mode attenuation).....	76
Figure 3.11: Modeshape of S_0 at 200 kHz.....	77
Figure 3.12: Modeshape of A_0 at 200 kHz	78
Figure 3.13: Modeshape of S_1 at 200 kHz.....	79
Figure 3.14: Modeshape of A_1 at 200 kHz	80
Figure 3.15: Modeshape of S_2 at 200 kHz.....	81
Figure 3.16: Modeshape of A_2 at 200 kHz	82
Figure 3.17: Modeshape of S_3 at 200 kHz.....	83
Figure 3.18: Modeshape of A_3 at 200 kHz	84
Figure 3.19: Time and frequency domain of the load function	91
Figure 3.20: Wavenumbers as a function of frequency in undamped rail.....	93
Figure 3.21: Phase velocity dispersion curves in undamped rail	94
Figure 3.22: Group velocity dispersion curves in undamped rail	95
Figure 3.23: Frequency domain displacement of the central node of the top of the railhead subjected to laser excitation at 4" (≈ 102 mm) away from the node.....	96
Figure 3.24: Time domain displacement of the central node of the top of the railhead subjected to laser excitation at 4" away from the node	97
Figure 3.25: Laser-generated signal acquired by the air-coupled sensor	98
Figure 3.26: Raw simulated(red) and 40 kHz high-passed(blue) simulated response	98

Figure 3.27: Raw and 500 kHz low-passed experimental signals.....	99
Figure 3.28: Comparison of the measured response with the results of the SAFE simulation	100
Figure 3.29: Simulated and experimental spectral responses.....	100
Figure 3.30: Phase velocity dispersion curves filtering.....	101
Figure 3.31: Symmetric excitation pattern on the meshed rail section	110
Figure 3.32: Nonsymmetric excitation pattern on the meshed rail section	111
Figure 3.33: Mode excitability curves for a case of symmetric excitation of the rail head	112
Figure 3.34: Mode excitability curves for a case of nonsymmetric excitation of the rail head	113
Figure 3.35: S_0 , A_0 , A_1 and S_2 mode excitability curves for a symmetric excitation of the rail head	114
Figure 3.36: S_0 , A_0 , A_1 and S_2 mode excitability curves for a nonsymmetric excitation of the rail head	114
Figure 3.37: Nonsymmetric load as sum of symmetric and antisymmetric contributions	115
Figure 3.38: S_1 , A_2 , S_3 and A_3 mode excitability curves for a symmetric excitation of the rail head	117
Figure 3.39: S_1 , A_2 , S_3 and A_3 mode excitability curves for a nonsymmetric excitation of the rail head	117
Figure 3.40: In-plane cross-sectional strain energy distribution for a symmetric excitation at 4" (≈ 102 mm) from the source.....	118
Figure 3.41: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at 4" (≈ 102 mm) from the source.....	118
Figure 3.42: Maximum of the signal envelope acquired at different distances from the source	119

Figure 3.43: In-plane cross-sectional strain energy distribution for a symmetric excitation ($2'' \approx 51$ mm) from the source)	120
Figure 3.44: In-plane cross-sectional strain energy distribution for a symmetric excitation at $6'' (\approx 152\text{mm})$ from the source.....	121
Figure 3.45: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at $2'' \approx 51$ mm) from the source	121
Figure 3.46: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at $6'' (\approx 152$ mm) from the source.....	122
Figure 3.47: Cross-sectional in-plane strain energy of rail response, filtered with a 300 kHz, 3-order Butterworth highpass	124
Figure 3.48: Cross-sectional in-plane strain energy of rail response, filtered with a 100-300 kHz, 3-order Butterworth bandpass	125
Figure 4.1: The Gabor wavelet with σ and $G_s = 5.34$. (a) Time domain, real (continuous line) and imaginary (dotted line) components. (b) Magnitude of Fourier spectrum	133
Figure 4.2: Wavelet spectra resulting from scaling of the mother-wavelet in the time domain	136
Figure 4.3: Replacement of an infinite wavelet set by one scaling function.....	137
Figure 4.4: (a) Discrete Wavelet filter-bank decomposition; (b) signal reconstruction from wavelet coefficients; (c) reconstruction of original signal	140
Figure 4.5: Example of discrete wavelet processing. (a) Unprocessed signal; (b), (c) wavelet coefficient vector of cD3, cD4, and cD5 (d), (e) thresholded wavelet coefficient vectors in (b) and (c); (f) and (g) reconstructed direct and echo signals	142
Figure 4.6: Detection scheme for coherent acquisition of guided waves reflected by a flaw in the rail head	146
Figure 4.7: Test setup: laser excitation and transducer spacing in proximity of the 5% H.A. reduction surface defect.....	148
Figure 4.8: Probe spacing and laser-beam position relative to the defect	148

Figure 4.9: Waterfall plot of ultrasonic signals acquired in proximity of the 8 mm deep surface defect	149
Figure 4.10: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 8 mm deep surface defect	150
Figure 4.11: CWT of the first A-scan acquired in proximity of the 8 mm deep surface defect	151
Figure 4.12: Waterfall plot of the 12 A-scans acquired in proximity of the 23% H.A. reduction internal defect	152
Figure 4.13: A-scans acquired in proximity of the third internal defect, high-passed at 100 kHz	152
Figure 4.14: Zoom in, of the waveforms reflected from the 23% H.A. reduction internal defect	153
Figure 4.15: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 23% H.A . reduction internal defect	154
Figure 4.16: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 23% H.A . reduction internal defect, after the application of the spatial averaging technique (N=5, spacing 0.5'' \approx 13 mm)	155
Figure 4.17: CWT of the first A-scan using the Gabor motherwavelet	156
Figure 4.18: CWT of the first A-scan using the Complex Morlet 1-1 motherwavelet	156
Figure 4.19: Front and 3-D view of the rail head with the 23% H.A. reduction internal flaw	157
Figure 4.20: Defect detection schemes examined. "Reflection mode" with a single air-coupled sensor (top); "transmission mode" with a pair of air-coupled sensors (bottom).....	159
Figure 4.21: Cross-sectional head area reduction as a function of the defect depth for the centre crack and for the corner crack	161

Figure 4.22: Feature-based Damage Index as a function of the center crack depth, “reflection mode”; (a) wavelet coefficients variance; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum Hilbert transform of the reconstructed signal.....	163
Figure 4.23: Feature-based Damage Index as a function of the corner crack depth, “reflection mode”; (a) wavelet coefficients variance; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum Hilbert transform of the reconstructed signal.....	164
Figure 4.24: Feature-based Damage Index as a function of the center crack depth, “transmission mode”; (a) wavelet coefficients variance; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum Hilbert transform of the reconstructed signal	166
Figure 4.25: (a) Photo of the rail track mock-up at UCSD; (b) schematic of the mock-up layout; (c) defect detection scheme in ‘transmission mode’ with a pair of air-coupled transducers	175
Figure 4.26: Rail defect detection prototype deployed on the rail track mock-up at UCSD.....	177
Figure 4.27: Typical waveforms detected by the hybrid laser/air-coupled transducer system on a rail. (a), (b) damage-free rail; (c), (d) 5% HA transverse surface notch; (e), (f) joint.....	179
Figure 4.28: (a) The Coiflet wavelet of order 4; (b) the Symlet wavelet of order 6; (c) the Daubechies wavelet of order 10	181
Figure 4.29: Example of discrete wavelet processing. (a), (b) unprocessed signals detected for a transverse surface notch; (c), (d) wavelet coefficient vector of detail levels 2 and 3 (e), (f) reconstructed rear and front sensor signals	183
Figure 4.30: RMS-based Damage Index as a function of the acquisition number for the unprocessed time waveforms, the vector of the retained wavelet coefficients, and the time-domain reconstructed signal.....	184
Figure 4.31: RMS-based Damage Index as a function of the acquisition number calculated for four different retained wavelet coefficient vectors.....	185

Figure 4.32: RMS-based Damage Index as a function of the acquisition number for the vector of the retained wavelet coefficients, computed from three different mother wavelets [J=Joint, S=Surface defect, ID=Internal defect].....	185
Figure 4.33: Typical waveforms filtered with the 300-1,100 kHz pass-band 3rd order Butterworth. (a), (b) damage-free rail; (c), (d) transverse surface notch; (e), (f) joint.....	187
Figure 4.34: Typical waveforms filtered with the 40-300 kHz pass-band 3rd order Butterworth. (a), (b) damage-free rail; (c), (d) transverse surface notch; (e), (f) joint.....	189
Figure 4.35: Damage Index as a function of the acquisition number, based on the normalized RMS of the signal amplitudes. Sensors deployed on the Center Head of the rail and signals filtered with a pass-band 40-300 kHz (low frequency range) 3rd order Butterworth.....	190
Figure 4.36: Damage Index as a function of the acquisition number, based on the normalized RMS of the signal amplitudes. Sensors deployed on the Center Head of the rail and signals filtered with a pass-band 300-1,100 kHz 3rd order Butterworth	190
Figure 4.37: Damage Index as a function of the acquisition number, based on the temporal signal coherence of signals filtered with (a) 40-1,100 kHz and (b) 40-300 kHz pass-band 3rd order Butterworth. Sensors deployed on the Center Head of the rail	191
Figure 5.1: Continuum® Surelite SLI-20 Q-Switched laser (power supply unit).....	195
Figure 5.2: Continuum® Surelite SLI-20 Q-Switched laser (laser head)	195
Figure 5.3: Continuum® Surelite SLI-20 Q-Switched laser (control panel)	196
Figure 5.4: Layout of the optical element sequence.....	197
Figure 5.5: MICROACOUSTIC® broadband air-coupled Transducer (BAT).....	198
Figure 5.6: Defect detection schemes examined. (a) “Reflection mode” with a single air-coupled sensor; (b) “reflection mode” with a pair of air-coupled sensors; (c) “transmission mode” with a pair of air-coupled sensors.....	200

Figure 5.29: Layout of the acting/sensing components of the rail inspection prototype.....	224
Figure 5.30: Prototype hardware mounted on the cart. Laser head and sensor holders are installed on the right beam of the main frame	225
Figure 5.31: Hardware layout of the hybrid laser/air-coupled transducer Long-Range method	226
Figure 5.32: Producer-Consumer scheme, allowing the Prototype for independent execution timing.....	228
Figure 5.33: Client-Server architecture of the Prototype link	228
Figure 5.34: Snapshot of the Vi LaserControl that controls the laser and allows opening the calibration task and the testing task.....	229
Figure 5.35: Snapshot of the user interface when the calibration task (VI 4Calibration) is activated.....	230
Figure 5.36: Snapshot of the user interface when the testing task (VI visual&report) is activated	232
Figure 5.37: Snapshot of the user interface when the report is shown during the testing task (VI 4visual&report) is activated.	234
Figure 5.38: 3D view of the mechanical model of the upgraded cart by ENSCO.	235
Figure 5.39: Rail flaw detection prototype installed on the cart during second field test.	236
Figure 5.40: Improved sensor arrangement adopted in second field test to exploit the bi-directionality of the laser ultrasound generation.....	237
Figure 5.41: Sensing devices: air/coupled transducers at the gage side, center head and field side	238
Figure 5.42: Details of the three-sensor holder	238
Figure 5.43: Layout of the block diagram of the bandpass filtering subVI.	240
Figure 5.44: Snapshot of the Calibration Session showing the various settings for the digital filtering of the ultrasonic measurements.....	240

Figure 5.45: Internal defect mapping (2.25 MHz ultrasonic transducer with 70° wedge).....	241
Figure 5.46: Test site near Gettysburg, PA.	243
Figure 5.47: Rail sections with internal defects plugged in the railroad.	245
Figure 5.48: Particular of the surface transverse cuts.....	245
Figure 5.49: Details of the oblique surface cuts added later.	246
Figure 5.50: Laser head and sensors on Ensco’s cart during first field tests.....	246
Figure 5.51: Detail of the air-coupled sensors.....	247
Figure 5.52: Laser/air-coupled sensor layout for the field test. Dimension in inches. Drawing not to scale.....	248
Figure 5.53: The inspection prototype towed by the FRA Hy-railer managed by ENSCO during the first field tests.....	248
Figure 5.54: Results of Test 01 (2006).....	252
Figure 5.55: Results of Test 02 (2006).....	253
Figure 5.56: Sensor layout in presence of a defect (a) between sensors#1 and #2, and (b) between sensors #2 and #3.....	254
Figure 5.57: Results of a sample test, performed at the speed of 5 mph (2007)	261
Figure 5.58: Results of a sample test, performed at the speed of 10 mph (2007)	262

LIST OF TABLES

Table 2.1: FRA Office of Safety and Analysis track failure data for years 1992-2002	7
Table 2.2: FRA Office of Safety and Analysis track failure data for year 2001	7
Table 2.3: Advantages and disadvantages of different rail inspection methods	31
Table 3.1: Modeshapes and strain energy distributions of S0, A0 and S1, computed at 200 kHz.....	105
Table 3.2: Modeshapes and strain energy distributions of A1, S2 and A2, computed at 200 kHz.....	106
Table 3.3: Modeshapes and strain energy distributions of S3 and A3, computed at 200 kHz.....	107
Table 3.4: In-plane strain energy distribution and defect sensitivity of the eight lowest order modes in rails	108
Table 3.5: Correspondence between frequency and penetration depth (wavelength) into the rail head for different values of the spectrum	123
Table 4.1 Experimental parameters for the three tests performed (Ld = distance laser source – crack; L1 = distance laser source-air-coupled sensor)	161
Table 4.2: List of features considered in the unsupervised learning defect detection.	168
Table 4.3: Outliers detected from the different D.I. features and for each defect type when the ultrasonic signals were corrupted with the low-level noise.....	169
Table 4.4: Outliers detected from the different D.I. features and for each defect type when the ultrasonic signals were corrupted with the high-level noise.	169
Table 4.5: Outliers detected from the multivariate analysis for three different D.I. feature combinations and for each defect size when the ultrasonic signals were corrupted with low noise.....	170

Table 4.6: Outliers detected from the multivariate analysis for three different D.I. feature combinations and for each defect size when the ultrasonic signals were corrupted with high noise.	170
Table 4.7: Different sources of variabilities affecting the D.I. computation.....	172
Table 4.8. Mock-up layout	176
Table 5.1: Continuum™ Surelite SLI-20 specifications	194
Table 5.2: MICROACOUSTIC® broadband air-coupled transducer specifications	197
Table 5.3: MICROACOUSTIC® charge-sensitive amplifier specifications	198
Table 5.4: NI PXI-1031 specifications	203
Table 5.5: NI PXI-8186 specifications	204
Table 5.6: NI PXI-6115 specifications	204
Table 5.7: NI BNC-2090 specifications	205
Table 5.8: Big Sky Laser® CFR-400 specifications	223
Table 5.9: Test site layout	242
Table 5.10: Summary of test conditions.....	249
Table 5.11: Summary of Gettysburg first field test results	257
Table 5.12: Summary of second field test results.....	263

ACKNOWLEDGEMENTS

I am forever grateful for the opportunity to perform this research under the precious guidance of my advisor, Prof. Francesco Lanza di Scalea. His selfless support and academic passion was critical for the success of my research. I would also like to thank my old time colleagues of the NDE/SHM Laboratory of UCSD, Prof. Piervincenzo Rizzo, Prof. Alessandro Marzani, Dr. Howard Matt, and especially Dr. Ivan Bartoli for their invaluable contributions, humor and friendship. A special thanks goes to Prof. Rizzo and Dr. Salvatore Salamone for the fruitful collaboration in producing the results shown in chapter 4 of this dissertation. I am also particularly grateful to Prof. Marzani for helping in the development of part of chapter 3 of this work.

A very special thanks goes to Kimberly Ravey, for the immense help and support that she gave me during the last preparation stage of this dissertation.

Thank you also to all my closest friends of the Structural Engineering Department of UCSD, Maurizio Gobbato, Marco Faggella, Prof. Michele Barbato, Dr. Babak Moaveni, Robert Phillips, Ankit Srivastava, Andre Barbosa, Elisa Sorri, Giuseppina Vitale, Daniel He, Ozgur Ozelik, Carmen Amaddeo, and to anybody who shared with me the moments of the graduate student life; a special thanks goes to my longtime roommate Gregory Andreev.

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 DTFR53-02-G-00011. The work was sponsored by FRA's Office of

Research and Development. The author is grateful to Mahmood Fateh, FRA's Technical Representative, for his technical guidance and valuable comments during the different phases of the research, and to Gary Carr, formerly Senior Engineer of Ensco Inc. and currently FRA's Chief of Track Research Division, for providing technical and logistic support during the prototype assembling and field testing. The in-kind support of San Diego Trolley, Inc. which donated parts of the rail specimens used for testing purposes to the NDE & SHM Laboratory at UCSD, is acknowledged. The research was also partially funded by the UCSD/Los Alamos Cooperative Agreement on Research and Education (CARE), for which the author is thoroughly appreciative. Thanks are extended to Dr. Gyuhae Park and Prof. Charles Farrar for the established fruitful collaboration.

Chapter 3, in part, will be submitted for publication in *Journal of Acoustical Society of America*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Marzani, A.. The current running title of this paper is "Forced solution of guided waves in rails". The dissertation author will be the primary investigator and author of this paper.

Part of chapter 4 of this dissertation was printed in *Insight –Non –Destructive Testing & Condition Monitoring*, Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I., Fateh, M., Viola, E. and Pascale, G., 2006. "Non-Contact Ultrasonic Inspection of Rails and Signal Processing for Automatic Defect Detection and Classification". The dissertation author was co-author of this paper. A segment in chapter 4 has been published in *Insight - Non-Destructive Testing & Condition Monitoring*, Rizzo, P., Bartoli, I., Cammarata, M. and Coccia, S., 2007. "Digital signal processing for rail

monitoring by means of ultrasonic guided waves.” The dissertation author was co-author of this paper.

Chapter 5, in part, has been published in *Transportation Research Record, Journal of the Transportation Research Board*, Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I. and Fateh, M., 2006. “Laser-Air-Coupled Hybrid Detection in Rail Tracks: Status of FRA Prototype Development at UC San Diego”. The dissertation author was co-author of this paper. An additional part in chapter 5 was presented at *World Forum on Smart Materials and Smart Structures Technology*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P., Fateh, M. and Yang, T.-L., 2007. “Non-contact Ultrasonic Rail flaw Detection”. The dissertation author was the main contributor of this paper. A section in chapter 5 was presented at *IWSHM - The 6th International Workshop Structural Health Monitoring*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P. and Fateh, M., 2007. “Non-contact Ultrasonic Rail Defect Detection System. Prototype development and field testing”. The dissertation author was the main contributor of this paper. An additional segment in chapter 5 is in *Technical Report for the University of San Diego, California*, Rizzo, P., Coccia, S., and Lanza di Scalea, F., 2005. “On line High-Speed Rail Defect Detection, Final Report-Phase III”. The dissertation author was co-author of this paper. In part, chapter 5 was in *Technical Report for the Federal Railroad Administration*, Coccia, S., Bartoli, I., Lanza di Scalea, F. and Rizzo, P., 2007. “Non-contact Rail Defect Detection: First and Second Field Tests”. The dissertation author was the main contributor of this paper.

VITA

2003	Bachelor of Science, University of Bologna, Italy
2004 – 2006	Research Assistant, University of California, San Diego
2006	Master of Science, University of California, San Diego
2006 – 2007	Research Assistant, University of California, San Diego
2007	Doctor of Philosophy, University of California, San Diego

JOURNAL PAPERS

1. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I., Fateh, M., Viola, E. and Pascale, G., “Non-Contact Ultrasonic Inspection of Rails and Signal Processing for Automatic Defect Detection and Classification,” *Insight - Non-Destructive Testing & Condition Monitoring*, Vol. 47, No. 6, pp. 346-353 (2005).
2. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I. and Fateh, M. “Laser-Air-Coupled Hybrid Detection in Rail Tracks: Status of FRA Prototype Development at UC San Diego,” *Transportation Research Record, Journal of the Transportation Research Board*, No. 1943, pp. 57-64 (2006).
3. Park, G., Farrar, C.R., Lanza di Scalea, F., Coccia, S. “Performance Assessment and Validation of Piezoelectric Active Sensors in Structural Health Monitoring,” *Smart Materials and Structures*, Vol. 15, No. 6, pp. 1673-1683 (2006).
4. Rizzo, P., Bartoli, I., Cammarata M. and Coccia S. “Digital signal processing for rail monitoring by means of ultrasonic guided waves,” *Insight - Non-Destructive Testing & Condition Monitoring*, Vol. 49, No. 6, pp. 327-332 (2007).
5. Lanza di Scalea, F., Matt, H., Bartoli, I., Coccia, S., Park, G. and Farrar, C. “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, 373-388 (2007).
6. Rizzo, P., Coccia, S., Bartoli, I. and Lanza di Scalea, F. “Non-Contact Rail Monitoring by Ultrasonic Guided Waves,” in Chapter 10 (D. Inman, ed.) of *Encyclopedia of Structural Health Monitoring*, C. Boller, F-K. Chang and Y. Fujino, eds., John Wiley & Sons, Chichester, UK, in press (2007).

BOOKS AND CHAPTERS

1. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli I. and Fateh M. "On-Line High-Speed Rail Defect Detection–Phase III", chapter 4 of Technology in Rail Transport Management, Prabha Shastri Ranade ed.

PATENTS

1. Lanza di Scalea, F., Rizzo, P., Bartoli, I., Coccia, S., "Enhancement in Reliability Of Defect Detection And Quantification In Non-Contact Ultrasonic Testing With Application To Rail Inspections," Invention Disclosure UCSD docket no. SD 2005 235.
2. Lanza di Scalea, F., Rizzo, P., Bartoli, I., Coccia, S., "Metodo ed apparecchiatura per l'ispezione in corsa dei binari ferroviari mediante una tecnica ultrasonica senza contatto," patent application no. 2005-000042, Italy.

CONFERENCE PROCEEDINGS

1. Rizzo, P., Bartoli, I., Coccia, S., Lanza di Scalea, F. and Fateh, M. "Enhancement in Guided Wave Structural Health Monitoring by the Discrete Wavelet Processing: Application to Multi-wire Strands and Rail Tracks," Proceedings of the 5th International Workshop on Structural Health Monitoring: Advancements and Challenges for Implementation, Ed. F.-K. Chang, Stanford, CA, September 12-14, pp. 1259-1266 (2005).
2. Rizzo, P., Bartoli, I., Lanza di Scalea, F., Coccia, S., and Fateh, M. "High-speed Defect Detection in Rails by Non-contact Guided Ultrasonic Testing," Proceedings of SPIE's 10th Annual International Symposium on Nondestructive Evaluation Health Monitoring and Diagnostics – Health Monitoring and Smart Nondestructive Evaluation of Structural and Biological Systems IV, Ed. Tribikram Kundu, Proc. of SPIE Vol. 5768, pp. 274-284 (2005).
3. Matt, H., Lanza di Scalea, F., Marzani, A., Coccia, S., Oliver, J. A., Kosmatka, J. B., Rizzo, P., and Restivo, G. "A Guided-Wave System for Monitoring the Wing Skin-to-Spar Bond in Unmanned Aerial Vehicles", Proceedings of SPIE's 12th Annual International Symposium on Smart Structures and Materials - Sensors and Smart Structures Technologies for Civil, Mechanical and Aerospace Systems, Ed. Masayoshi Tomizuka, Proc. of SPIE Vol. 5765, pp. 758-768 (2005).
4. Matt, H., Bartoli, I., Coccia, S., Lanza di Scalea, F., Kosmatka, J., Park, G. and Farrar, C. "Built-in guided-wave diagnostic system for the wing skin-to-spar bond in unmanned aerial vehicles," IWSHM - The 5th International Workshop Structural Health Monitoring, Stanford, CA, September 12-14 (2005).

5. Rizzo, P., Matt, H., Coccia, S., Lanza di Scalea, F., Bartoli, I., Marzani, A., and Viola, E. "Guided ultrasonic waves for the inspection of structural components," XVI Convegno Italiano di Meccanica Computazionale (XVI Italian Congress of Computational Mechanics), GIMC 2006, Bologna, Italy, June 26-28, (2006).
6. Matt, H., Bartoli, I., Coccia, S., Lanza di Scalea, F., Kosmatka, J., Park, G. and Farrar, C. "Guided Wave Interrogation of Composite Bonded Joints Using MFC Sensors," Proceedings of the SEM IMAC 24th Conference and Exposition on Structural Dynamics, St. Louis, MO, (2006).
7. Rizzo, P., Lanza di Scalea, F., Matt, H., Bartoli, I., Coccia, S., Salamone, S., Sorrivi, E., Marzani, A., Viola, E., Park, G., Farrar, C., Di Paola, M. and Pirrotta, A. "Health Monitoring of Civil and Aerospace Structural Components by Guided Ultrasonic Waves," 4th World Conference on Structural Control and Monitoring, San Diego, CA 11-13 July, (2006).
8. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I. and Fateh, M. "Development of a Non-contact Defect Detection System for Railroad Tracks for the US Federal Railroad Administration," Third European Workshop on Structural Health Monitoring, July, 5th-7th, 2006 Granada, Spain (2006).
9. Rizzo, P., Coccia, S., Bartoli, I., Lanza di Scalea, F. and Fateh, M. "Unsupervised learning algorithm for high-speed defect detection in rails by laser/air-coupled non-contact ultrasonic testing," Proceedings of SPIE's 13th Annual International Symposium on Smart Structures and Materials – Sensors and Smart Structures Technologies for Civil, Mechanical and Aerospace Systems, Proc. of SPIE Vol. 61742G, (2006).
10. Matt, H., Bartoli, I., Coccia, S., Lanza di Scalea, F., Oliver, J., Kosmatka, J., Park, G. and Farrar, C. "Ultrasonic Guided Wave Monitoring of Composite Bonded Joints Using Macro Fiber Composite Transducers," Proceedings of SPIE's 13th Annual International Symposium on Smart Structures and Materials – Sensors and Smart Structures Technologies for Civil, Mechanical and Aerospace Systems, Proc. of SPIE Vol. 61742D, (2006).
11. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I. and Fateh, M. "Laser/Air-Coupled Hybrid Non-Contact System for Defect Detection in Rail Tracks: Status of FRA Prototype Development at UC San Diego," Proceeding of the 85th Annual Conference of the Transportation Research Board, (2006).
12. Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P., Fateh, M. and Yang, T.-L. "Non-contact Ultrasonic Rail flaw Detection," World Forum on Smart Materials and Smart Structures Technology, May 22-27, 2007, Chongqing & Nanjing, China (2007).
13. Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P. and Fateh, M. "Non-contact Ultrasonic Rail Defect Detection System. Prototype development and field testing," IWSHM - The 6th International Workshop Structural Health Monitoring, Stanford, CA, September 10-13 (2007)

14. Rizzo, P., Coccia, S., Lanza di Scalea, F., Bartoli, I. and Fateh, M. "Non-contact rail flaw detection system: first field test," Proceedings of SPIE's 14th Annual International Symposium on Smart Structures and Materials – Sensors and Smart Structures Technologies for Civil, Mechanical and Aerospace Systems, Proc. of SPIE Vol. 65293W, (2007).

TECHNICAL REPORTS

1. Rizzo, P., Coccia, S., and Lanza di Scalea, F. "On line High-Speed Rail Defect Detection, Final Report-Phase III", Technical Report No. SSRP-05/04, University of California, San Diego, (2005).
2. Rizzo, P., Coccia, S., Bartoli, I. and Lanza di Scalea, F. "On-line High-speed Rail Defect Detection – Phase IV: Prototype Assembling and Field Testing," Technical Report No. SSRP-06/17 to the Federal Railroad Administration, University of California, San Diego, (2006).
3. Coccia, S., Bartoli, I., Lanza di Scalea, F. and Rizzo, P. "Non-contact Rail Defect Detection: First and Second Field Tests," Technical Report No. SSRP-07/15 to the Federal Railroad Administration, University of California, San Diego, (2007).

ABSTRACT OF THE DISSERTATION

Ultrasonic Guided Waves for Structural Health Monitoring and Application to Rail
Inspection Prototype for the Federal Railroad Administration

by

Stefano Coccia

Doctor of Philosophy in Structural Engineering

University of California, San Diego, 2007

Professor Francesco Lanza di Scalea, Chair

Recent train accidents, associated direct and indirect costs, as well as safety concerns, have reaffirmed the need for developing rail defect detection systems more effective than those used today. Current methods for detecting internal flaws in rails rely primarily on ultrasonic pulse-echo technology operated in a water-filled wheel or sled. Presence and loss of echoes along each tested direction are analyzed in parallel to map surface and internal cracks in the rail. While this technology has served the industry well, several inherent weaknesses exist. Contact heads and cross-sectional inspection limit the speed of the measurement; liquid couplant is required to maintain efficient wave/echo transmission through the contact patch. Despite the acoustic couplant, significant transmission loss results from the pulse/echo passing through the

contact patch twice. More importantly, ultrasonic beams launched vertically from the top of the rail head can miss internal defects located under horizontal shelling; this was the case, in the disastrous train derailment at Superior, WI in 1991. As a proposal to address these issues, the use of ultrasonic guide waves appears promising. One objective of this work is extending the fundamental knowledge of the guided wave propagation in rails by predicting modal and forced solutions in the high frequency (<500 kHz) range. The selection of guided wave features sensitive to the presence of the different type of defects is essential for a successful defect detection performance. Another accomplishment of this work is the development of a rail defect detection prototype based on a laser/air-coupled ultrasonic technique. The prototype has been successfully tested twice in the field for the Federal Railroad Administration and it shows promise for implementation in rail inspection cars.

1 Introduction

1.1 Health monitoring of railroad tracks. Introductory discussion and motivation for research

Safety statistics data from the U.S. Federal Railroad Administration [1] indicate that train accidents caused by track failures including rail, joint bars and anchoring resulted in 700 derailments and \$441 million in direct costs during the 1992-2002 decade. The primary cause of these accidents is the ‘transverse defect’ type that was found responsible for 541 derailments and \$91 million in cost during the same period. Transverse defects are cracks developing in a direction perpendicular to the rail running direction, and include transverse fissure initiating in a location internal to the rail head, and detail fractures, initiating at the head surface as Rolling Contact Fatigue defects.

Rail failures may collaterally cause the spill of hazardous materials. In June 1992, for example, a derailment in Superior, WI (USA), forced the evacuation of an entire town due to hazard concerns. In 2000, hazardous materials were transported in 725 trains that were involved in railroad accidents: in those trains 75 cars released hazardous material [2].

The most common methods of rail inspection are magnetic induction and contact ultrasonic testing [3, 4]. The first method is affected by environmental magnetic noise and it requires a small lift-off distance for the sensors in order to produce adequate sensitivity [5, 6]. Ultrasonic testing is conventionally performed from the top of the rail head in a pulse-echo configuration, with the use of water-filled wheel probe. Such an approach suffers from a limited inspection speed and from other drawbacks associated with the requirement for contact between the rail and the inspection wheel. More

importantly, horizontal surface cracks such as shelling and head checks can prevent the ultrasonic beams from reaching the internal defects, resulting in false negative readings. The problem of surface shelling was highlighted in the above mentioned derailment in Superior, caused by the presence of a transverse crack missed during a previous inspection. Unfortunately, rail safety concerns can only become more serious given the unavoidable aging of the transportation infrastructure and the increasing rail tonnage. The need to develop more reliable defect detection systems for rails has produced promising results in recent years based on the use of guided ultrasonic waves [7, 8, 9, 10, 11].

As a proposal to address this pressing issue, this dissertation presents a promising method that uses ultrasonic guided waves excited by a pulsed laser and detected by an array of air-coupled sensors [12, 13, 14]. A prototype based on this technology has been packaged, installed, and tested in the field twice by the author and his research colleagues of the UCSD NDE & SHM Laboratory, with the help of the Federal Railroad Administration.

1.2 Outline of the numerical and experimental approach of the dissertation

The emphasis of this dissertation is placed upon the use of ultrasonic guided waves for performing health monitoring of railroad structures. Specifically, the dissertation extends the current state of knowledge of guided waves rail inspection both from a fundamental wave propagating viewpoint and from an experimental and prototype development viewpoint.

Chapter 2 gives insights on the state of the art of rail inspection. 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. Features and limitations of the current inspection systems are described, and the guided wave solution is presented as a promising technique with potential advantages compared to the traditional method.

Chapter 3 presents the use of a semi-analytical finite element method to model the propagation of guided waves in rails. Part of the novelty of this research consists of obtaining the response of the rail to different configurations of broadband laser excitation. The important results achieved in the modeling part of the dissertation, are critical to the selection of guided wave features (propagating modes) which are sensitive to defects located in specific portion of the rail head cross-section. The numerical predictions were validated by an extensive set of experimental results obtained at the UCSD rail test site.

Chapter 4 gives an overview of different signal processing techniques including joint time-frequency transforms, to extract meaningful features from the ultrasonic signals. The efficiency of different signal conditioning tools, such as discrete wavelet transform and digital filtering, is evaluated.

An additional unique part of the research is the development of a prototype based on the laser/air-coupled ultrasonic technique, which is presented in chapter 5. The description of the hardware design and software implementation of the prototype is presented in detail. Along with the different development phases of the prototype, results of two field tests performed in March 2006 and in April 2007 for the Federal Railroad Administration are also presented.

The dissertation concludes with a brief discussion on the key results of this research, and research topics requiring further study.

2 State of the art review

2.1 Description of structural defect in rails

One way of classifying rail defects can be based on their origin [15]. There are three families of defects: rail manufacturing defects, improper use or handling rail defects, and rail wear and fatigue. A good review of rail defects is given in the Sperry Rail Defect Manual [16].

Rail manufacturing defects are generally inclusions of nonmetallic origin or wrong local mixings of the rail steel components that, under operative loads, generate local concentration of stresses, which can trigger the rail failure process [17]. Defects deriving from improper use or handling of rails are generally due to spinning of train wheels on rails or sudden train brakes. The last class of defects is due to wearing mechanisms of the rolling surface and/or to fatigue. Not all rail defects are critical. A critical defect is a rail defect that will affect the safety of train operations. Noncritical defects are defects that occur in the rail but do not affect the structural integrity of the rail or the safety of the trains operating over the defect. Different regulations define the operating restrictions for each defect type; as an example Figure 2.1 reports the allowed remedial actions for the different cases of rail defects set by the U.S. Department of the Army.

Defect Type	Operating Restrictions Until Repairs Are Completed (Maximum Operating Speed)	Remedial Actions		
		Replace Entire Defective Rail	Crop Defect ²	Apply Joint Bars (Fully Bolted)
Bolt hole crack	10 mph	Allowed	Allowed	--
Broken base	5 mph	Allowed	Allowed	Not allowed
Corrosion - greater than 1/4 inch	10 mph	REQUIRED	Not allowed	Not allowed
Complete break - clean and square	CLOSE TO TRAFFIC	Preferred	--	Allowed ¹
Complete break - rough or angled	CLOSE TO TRAFFIC	Preferred	Allowed	Not allowed ⁴
Crushed head	10 mph	Preferred	Allowed	Not allowed
Defective weld	10 mph	Preferred	Preferred	Not allowed
End batter - greater than 1/4 inch	10 mph	Allowed	Allowed	--
Fissure - compound ⁵	CLOSE TO TRAFFIC	Preferred ¹	Allowed	Not allowed ⁴
Fissure - transverse ⁵	- Size less than 40% ⁶	5 mph	Allowed ¹	Allowed
	- Size greater than 40% ⁶	CLOSE TO TRAFFIC	Preferred ¹	Allowed
Fracture - detail ³	- Size less than 40% ⁶	5 mph	Allowed ¹	Allowed
	- Size greater than 40% ⁶	CLOSE TO TRAFFIC	Preferred ¹	Allowed
Fracture - engine burn ³	- Size less than 40% ⁶	5 mph	Allowed ¹	Allowed
	- Size greater than 40% ⁶	CLOSE TO TRAFFIC	Preferred ¹	Not allowed ⁴
Head/web separation	CLOSE TO TRAFFIC	REQUIRED	Not allowed	Not allowed
Piped rail	10 mph	REQUIRED	Not allowed	--
Running surface damage (depth greater than 1/4 inch)	10 mph	Allowed	Allowed	Not allowed
Short rail (rail less than 13 feet long)	NO RESTRICTION	REQUIRED	--	Not allowed
Split rail - horizontal	5 mph	REQUIRED	Not allowed	Not allowed
Split head - vertical	CLOSE TO TRAFFIC	REQUIRED	Not allowed	Not allowed
Split web	5 mph	REQUIRED	Not allowed	Not allowed
Torch cut - rail ends	10 mph	Preferred	Allowed	Not allowed
	bolt holes	5 mph	Preferred	Allowed
Wear on 90 lb or larger rail				
Side wear greater than 1/2 inch	10 mph	REQUIRED ⁷	Not allowed	Not allowed
Vertical wear greater than 1/2 inch	10 mph	REQUIRED	Not allowed	Not allowed
Wear on rail less than 90 lb				
Side wear greater than 3/8 inch	10 mph	REQUIRED	Not allowed	Not allowed
Vertical wear greater than 3/8 inch	10 mph	REQUIRED	Not allowed	Not allowed

Notes:

¹If two or more of these defects are found in any individual rail, that rail shall be replaced.²Rails may be cropped by cutting the rail with a rail saw or other appropriate cutting tool, at least 6 inches either side of the defect.³Not allowed if results in a rail length of less than 13 feet (See "Short Rail" below).⁴May be allowed as an emergency measure until defect is removed provided operations are restricted to 5 mph and an inspector is present.⁵If broken through or cracked out, rules for rough or angled complete break apply.⁶Defect size estimated from internal rail flaw testing. If size is unknown, assume greater than 40%.⁷90 lb or larger rail with wear on one side only may be transposed.

Figure 2.1: Rail defects, operational restrictions and remedial actions for the U.S. Department of Army [Railroad Track Standards, <http://www.usace.army.mil>]

The Federal Railroad Administration Office of Safety Analysis produces a report each year that list the costs of track failures occurred in the U.S., associated with the responsible defects. Reports from the FRA Office of Safety and Analysis can be found at the URL <http://safetydata.fra.dot.gov/officeofsafety/>. Table 2.1 lists six relevant rail defects, their prevalence as a percent of the total track failures due to all the defects listed in the original report for the years between and including 1992 and 2002 [1, 2], and the

cost associated with the track failures due to that particular type of defect. The second column lists the defects prevalence in terms of rank.

Table 2.1: FRA Office of Safety and Analysis track failure data for years 1992-2002

Type of Defect*	% of Total Defects	Reported Damage (\$)
Transverse Fissure	5.5 (3 rd highest)	91,448,042
Broken Rail Base	4.6 (4 th highest)	49,362,600
Vertical Split Head	3.6 (6 th highest)	45,922,310
Head and web Separation	3.5 (7 th highest)	38,820,132
Detail Fracture	2.7 (10 th highest)	71,392,239
Horizontal Split Head	0.7 (35 th highest)	5,643,239
Piped Rail	0.1 (43 rd highest)	717,093
Total	20.7	303,305,655

* Notes: Track failures due to rail, joint bar, and rail anchoring defects. Sixty three types of defects considered. Relevant defects are listed in Table 2.1.

For example between 1992 and 2002 transverse fissures were the cause of 5.5 % of all rail failures. This was the 3rd most prevalent defect listed. Table 2.2 contains similar data for year 2001.

Table 2.2: FRA Office of Safety and Analysis track failure data for year 2001

Type of Defect*	% of Total Defects	Reported Damage (\$)
Transverse Fissure	2.5 (3 rd highest)	17,342,722
Broken Rail Base	1.1 (7 th highest)	4,269,293
Vertical Split Head	1.7 (4 th highest)	6,757,293
Head and web Separation	1.0 (8 th highest)	6,691,432
Detail Fracture	1.4 (5 th highest)	10,772,281
Horizontal Split Head	0.4 (14 th highest)	2,453,153
Piped Rail	0.5 (35 th highest)	448,858
Total	8.6	48,735,032

* Notes: Track failures due to rail, joint bar, and rail anchoring defects. Fifty five types of defects considered. Relevant defects are listed in Table 2.2.

The costs of rail failures include train delay, cost of inspection, costs of remedial and pre-emptive treatments, loss of business confidence and customer support.

Among the many stresses that operate in a rail, bending and shear stresses arise principally from the gross vehicle load and they can be significantly increased by dynamic effects. As well as bending and shear stresses, the rail is also subjected to thermal and residual stresses. The following section will give insight to the different types of defects, with a brief description of their mechanical aspects and sketches picturing their appearance in the rail.

2.1.1 Rail manufacturing defects

It should be noted that the defects described in this section could also be originated by wearing mechanisms of the rolling surface and/or to fatigue, even if their occurrence is mainly due to the presence of original manufacturing defects. A further division within this class of rail defects can be based on the direction of propagation of the flaws under operative loads.

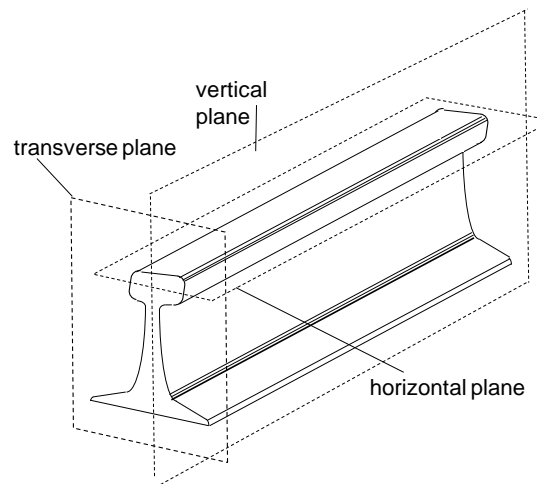


Figure 2.2: Relative position of planes through a rail

Two types of damages are classified [18, 19] in this group: *transverse* (extending primarily in the rail transverse plane) and *longitudinal* defects (extending primarily in the rail vertical plane) (Figure 2.2). Any progressive fracture occurring in the rail head having a transverse separation, however slight is defined as a transverse defect. The exact type of transverse defect cannot be determined until after the rail is broken for examination. It might stay hidden until it reaches an outer surface.

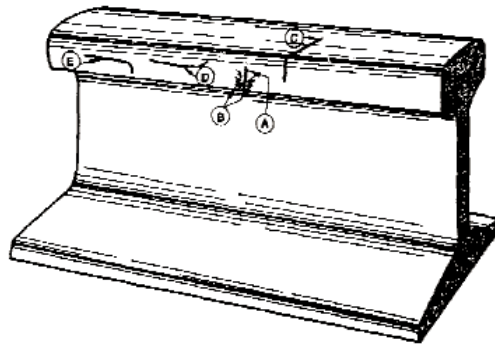


Figure 2.3: Appearance of different cases of transverse defects

A transverse defect (TD) may be recognized by one or more of the following characteristics (Figure 2.3):

- (a) A hairline crack on the side of the head at right angles to the running surface, at the fillet under the head, and occasionally on the running surface.
- (b) Bleeding at the crack.
- (c) A hairline crack at the gage corner of the rail head. On turned rail, this condition may occur at the field corner. Numerous small gage cracks or head checks are often present but should not cause suspicion unless a single crack extends much farther down the side and/or cross the running surface.

(d) A horizontal hairline crack in the side of the rail head turning upward or downward at one or both ends usually accompanied by bleeding. Under such conditions a flat spot will generally be present on the running surface.

(e) A hairline crack extending downward at right angles from a horizontal crack caused by shelling of the upper gage corner of the rail head.

TDs include *transverse/compound fissures* and *detail fractures*; pictures of different type of TDs are in Figure 2.4, where the coupling of a TD with a case of surface *shelling* is depicted.

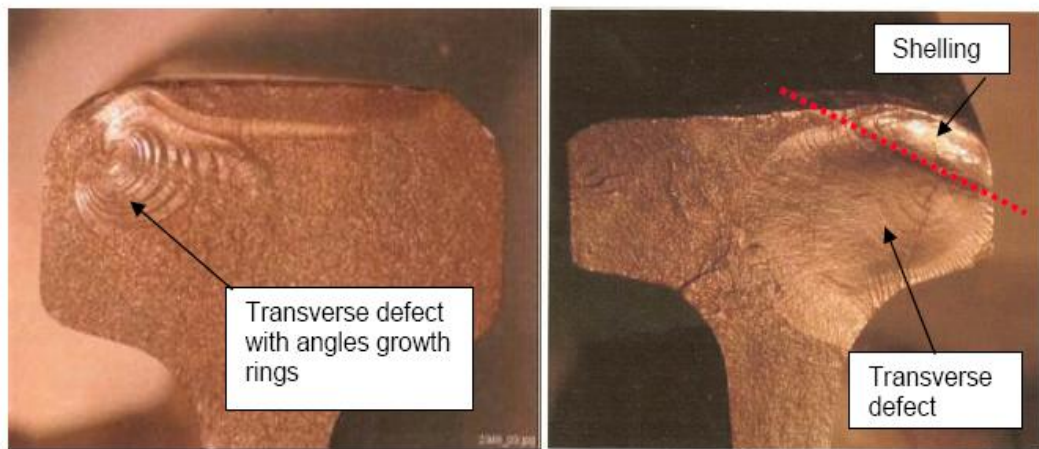


Figure 2.4: Transverse defects. Case of coupling with shelling

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. The shelling phenomenon is usually caused by wearing mechanisms of the rolling surface, and it will be described in depth later in the chapter. Among longitudinal defects, the *vertical* and the *horizontal split heads* are the most potentially dangerous. The horizontal split head (HSH) is a

progressive longitudinal fracture in the rail head parallel to the running surface. An example of HSH is depicted in Figure 2.5.

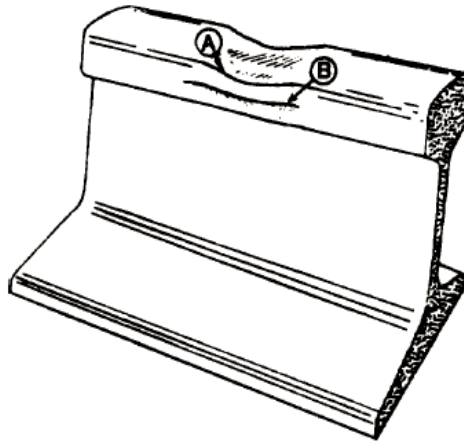


Figure 2.5: General appearance of an horizontal split head

Before cracking out (a), a moderate size horizontal split head will appear as a flat spot on the running surface often accompanied by a slight widening or dropping of the rail head. The flat spot will be visible as a dark spot on the bright running surface. After cracking out (b), the horizontal split head will appear as a hairline crack in either side or both sides of the rail head. A *vertical split head* (VSH) is a progressive longitudinal fracture in the head of the rail perpendicular to the running surface. It can be recognized on the track for the presence of one or more of the following (Figure 2.6):

- (a) A dark streak on the running surface.
- (b) Widening of the head for the length of the split. The cracked side of the head may show signs of sagging.
- (c) Sagging of the head causing a rust streak to appear on the fillet under the head.
- (d) A hairline crack near the middle of the rail head.

(e) In advanced stages, a bleeding crack is apparent on the rail surface and in the fillet under the head.

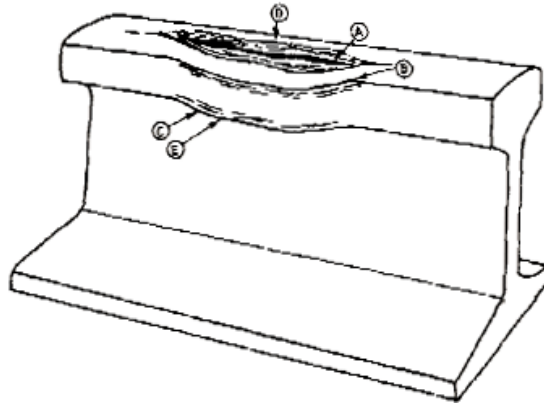


Figure 2.6: General appearance of a vertical split head

Another defect belonging to this group, is known as *head/web separation*, which is a progressive fracture separating the head and web of the rail. Figure 2.7 sketches a case of head/web separation. In earlier stages, wavy lines can appear along the fillet under the head (a); as the condition develops, a small crack will appear along the fillet on either side progressing longitudinally with slight irregular turns upward and downward (b). In advanced stages, bleeding cracks will extend downward from the longitudinal separation through the web and may extend through the base (c).

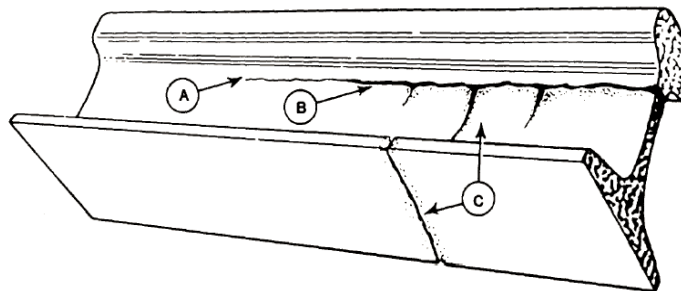


Figure 2.7: Case of a head/web separation

The *piped rail*, whose sketch is depicted in Figure 2.8, is another flaw most likely due to defective manufacturing.

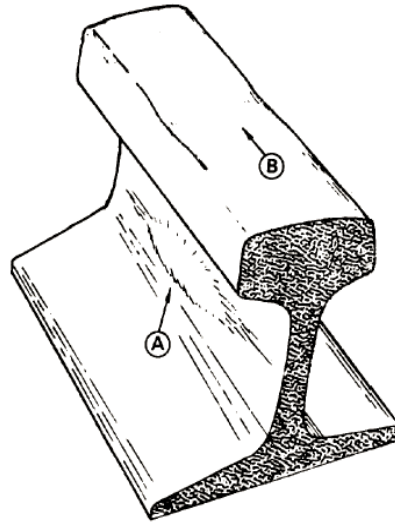


Figure 2.8: General appearance of a piped rail

The piped rail is a progressive longitudinal fracture in the web of the rail with a vertical separation or seam, forming a cavity in the advanced state of development. A bulging of the web (a) and a slight sinking of the rail head (b) above the pipe can be present.

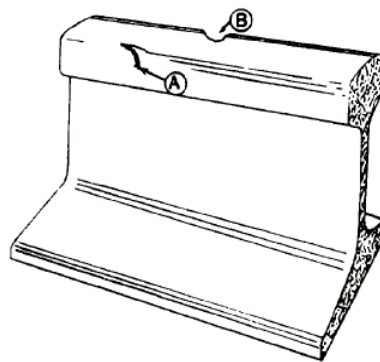


Figure 2.9: Sketch of cases of mill defects

Last groups of defects belonging to this category includes the *mill defects*, which appears in the track, as deformations (a), cavities (b) or inclusion in the head (Figure 2.9) and *split webs*. A split web is a progressive fracture through the web in a longitudinal or transverse direction, or both (Figure 2.10).



Figure 2.10: General appearance of a split web

2.1.2 Rail defects caused by improper use or bad maintenance

Examples of damage due to improper use of railroad tracks, are the *engine burns* and *engine burn fractures*. Engine burns are rail that have been scarred on the running surface by the friction of slipping locomotive wheels. An engine burn is not a critical defect, however an engine burn may lead to an engine burn fracture. Round or oval rough spots or holes on the tread of the running surface can characterize engine burns (Figure 2.11). An engine burn fracture is a progressive fracture in the rail head starting from a point where engine wheels have slipped and burned the rail (Figure 2.12).

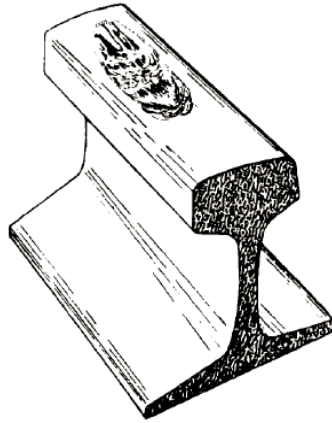


Figure 2.11: Sketch of an engine burn

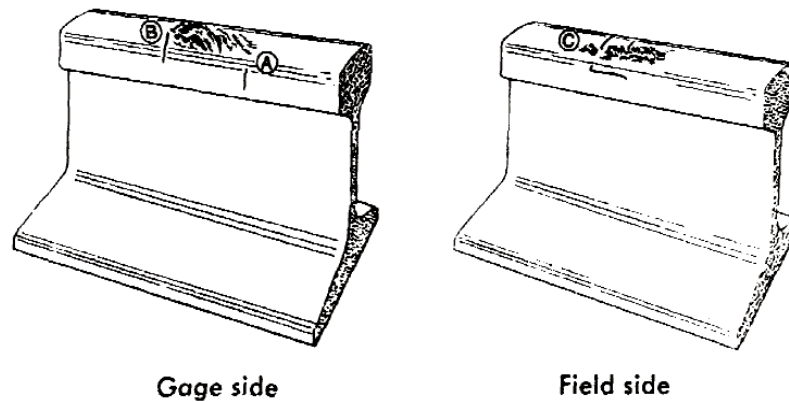


Figure 2.12: General appearance of engine burn fractures

Another type of defect of this category is the *crushed head*, which is the flattening of several inches of the rail head (Figure 2.13 (a)), usually accompanied by a crushing down of the metal but with no signs of cracking in the fillet under the head. A crushed head usually shows also small cracks in a depression on the running surface (b), and in advanced stages, it can present a bleeding crack at the fillet under the head (c).

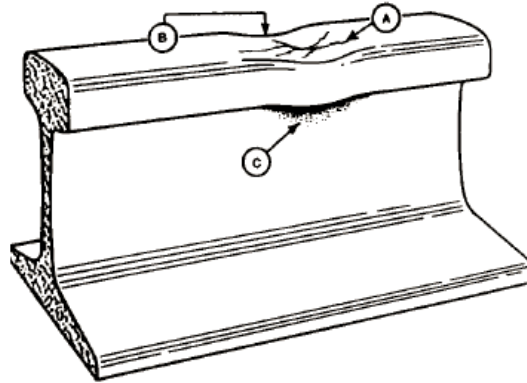


Figure 2.13: Sketch of a crushed head

Examples of defect caused by bad maintenance are *defective welds* and *torch cut rails*. The former usually present a progressive transverse separation within an area where two rails have been joined by welding or a rupture at a weld where improper fusion has occurred; the latter instead, describes any rail that is cut or otherwise modified using an acetylene torch or other open flame (Figure 2.14).

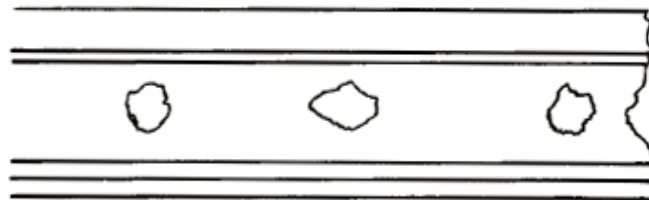


Figure 2.14: Sketch of a case of torched cut rail

2.1.3 Rail defects due to wearing mechanisms of the rolling surfaces and to fatigue (Rolling Contact Fatigue – RCF defects)

Cracking can be found in the head of all types of track, but is predominantly found on highly canted curves (Figure 2.15) where stresses develop due to extra pressure and wear of the wheel on the “high” rail (Figure 2.16).

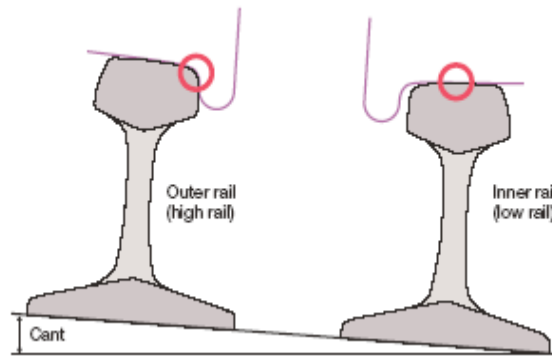


Figure 2.15: Wheel contact on a curve with the “high” rail and the “low” rail [20]

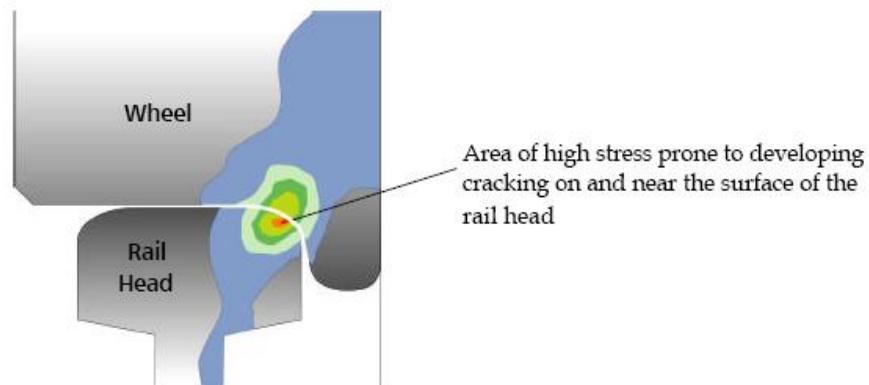


Figure 2.16: Contact stresses on tight curved track

The causes of this flaw typology can be divided in four classes: rolling contact fatigue (RCF) damage, structural failures of the rail, surface wearing damages and adverse environmental conditions.

RCF defects were ranked the most severe in the U.S. in the years between 1990 and 1995, while data from Europe state that their cost accounts for 300 million euros per year on average, in the European Union. RCF damages are much more severe from the point of view of the structure integrity because they are able to initiate the formation of

hidden transverse defects. As an example, the October 2000 disastrous derailment at Hatfield in the UK, that caused four deaths and over thirty injuries, is an outstanding example of the potential harm of RCF damage, or GCC (Gauge Corner Crack), as it was initially named. In the United States, the famous derailment near Superior, in Wisconsin that caused hazardous material spillage and the evacuation of 40,000 people was more than likely caused by RCF damage. A report from the National Transportation Safety Board states that the rail was inspected 6 weeks before the disaster and the ultrasonic inspection car operator recognized that the rail contained shelling but did not consider the conditions severe enough to warrant an exception report. After conducting additional ultrasonic inspections with handheld equipment, the operator considered the rail to be free from internal defects based on his evaluation of the test and his experience.

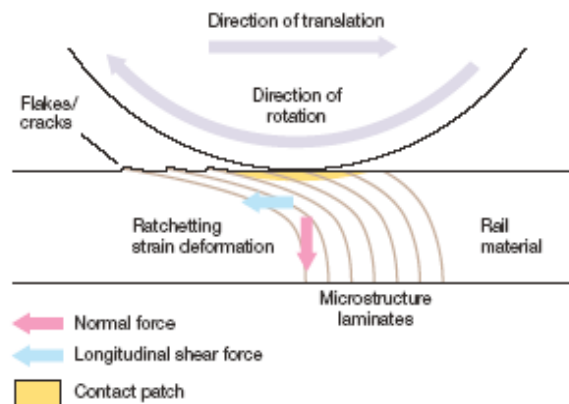


Figure 2.17: Contact patch pressure and forces causing surface elastic deformation of the rail [20]

An RCF crack initiates on or very close to the rail rolling surface; the crack, that initially could be microscopic in size, propagates through the heavily deformed (and then orthotropic) layers of steel at a shallow angle of about 10° to the rail running surface,

until it reaches a depth where the steel retains its original isotropic properties. At this point 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 detected [15, 20, 21, 22]. A sketch of the basic process of the initiation and development of RCF damage is illustrated in Figure 2.17. Its occurrence is increasing on high speed passenger lines, mixed and heavy haul railways and can lead to expensive rail grinding in the attempt to remove it, premature removal of the rails and complete rail failure. The rolling contact fatigue damage on rails can be headchecks, squats and spalling. RCF damages incidence can be reduced by lubricating rolling surfaces. Although, fluid entrapment in the metal is one of the most common causes for speeding up the surface initiated crack growth [15].

A type of structural failure of the rail is the *broken base*. This type is classified as any break in the base of the rail. The broken base defect generally appears as a half-moon crack break in the rail base (Figure 2.18).

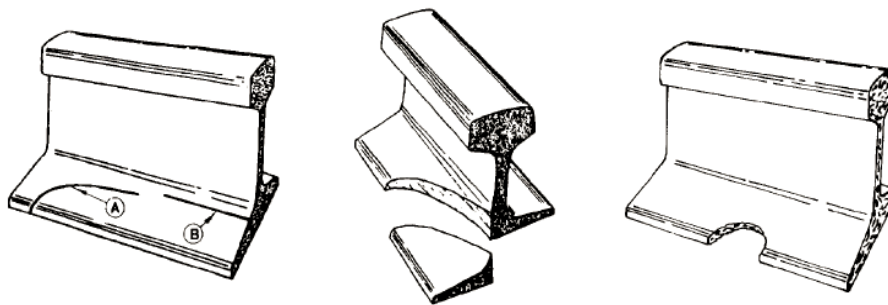


Figure 2.18: Example of the broken base defect

The *surface bent rail* is another type of structural failure of the rail. This defect is a permanent downward bending of the rail ends due to long-term passage of traffic over the track with loose or poorly supported joints. This is not a critical defect and cannot be corrected without replacing the rail; it appears as a downward bending of the rail head near the rail ends giving the appearance of low joints. When a track with a surface bent rail is raised and tamped, the rail ends soon return to a lower elevation. In more serious cases the vertical curve in the rail head is still visible after surfacing.

Another type of failure of the rail is the *bolt hole crack*, which is a progressive fracture originating at a bolt hole. Bolt-hole cracks account for about the 50% of the rail defects in joined tracks [23]. This type of defect is not visible until a bolt or joint bar has been removed or unless the defect has progressed beyond the bar. A bolt hole crack may be recognized by a hair line crack extending from the bolt hole (Figure 2.19). These cracks originate on the closest bolt-hole surface to the rail end and propagate with an angle $\pm 45^\circ$ from the vertical until the web-railhead junction. Fretting fatigue of the bolt shank against the hole surface is believed to be the principal cause of this typology of crack.

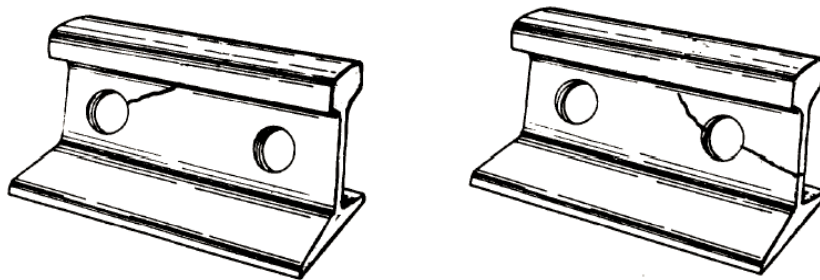


Figure 2.19: Sketches of bolt hole cracks

The extreme structural failure of the rail is the *complete break* also known as *broken rail*, which is a complete transverse separation of the head, web, and base of the rail. This defect may appear as a hairline crack running completely around the rail, usually accompanied by bleeding or a separation of the rail at the break with one or both of the broken ends battered down (Figure 2.20).

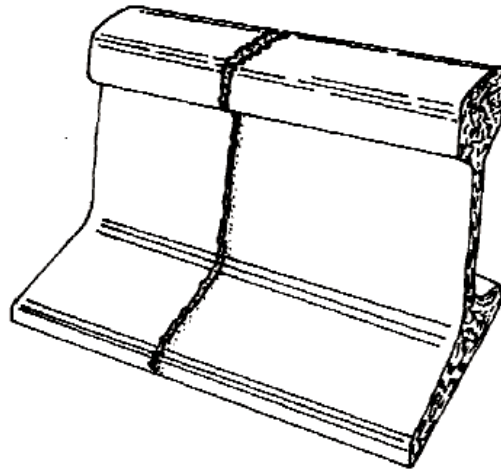


Figure 2.20: Complete rail break (broken rail)

Among the defects caused by surface wear damage, there is *rail corrugation*. Rail corrugation is an event strictly correlated to the wearing of the railhead, generated by a wavelength-fixing mechanism related to train speed, distance between the sleepers [24] friction [25] and so on. Six type of corrugation can be identified [26]: heavy haul corrugation, light rail corrugation, corrugation on resiliently booted sleepers, contact fatigue corrugation in curves, rutting and roaring rails or short-pitch corrugation. The corrugation itself does not compromise rolling safety, but has an adverse effect on track elements and rolling stocks by increasing noise emissions, loading and fatigue [27]. In the rail corrugation usually appears as a repeated wavelike pattern on the running surface

of the rail (Figure 2.21). Corrugations develop over a long period of time. A number of factors contribute to the development of corrugations with the actual cause dependent on the track and operating conditions. Corrugations are not critical defects.

Another defect due to surface damage is the *end batter*, which is defined as damage caused by the wheels striking the rail ends; it appears as damage to or a depression in the top surface of the rail head at the ends of the rail (Figure 2.22)

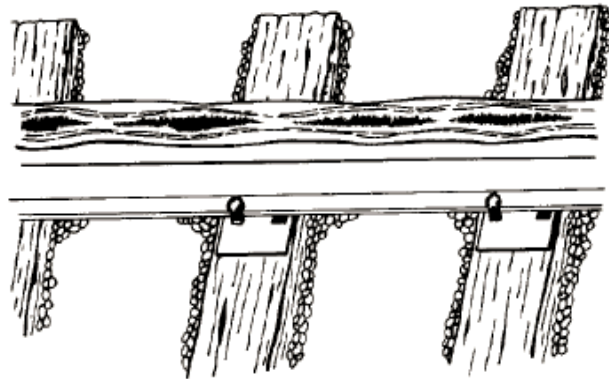


Figure 2.21: Sketch of the corrugation phenomenon

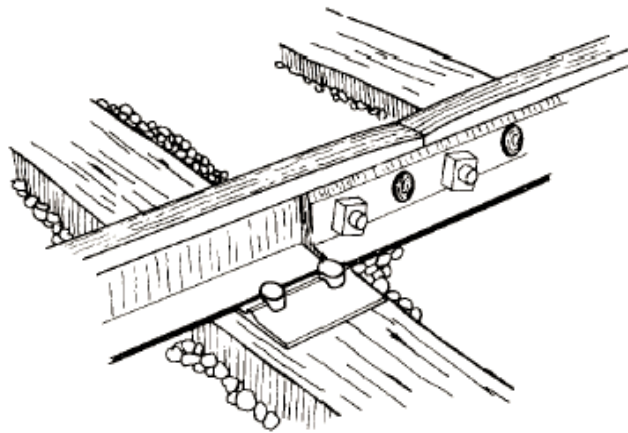


Figure 2.22: General appearance of the end batter defect

Flaking is also due to surface damage. Flaking is a progressive horizontal separation on the running surface near the gage corner often accompanied by scaling or chipping; it is not to be confused with shelling as it occurs only on the running surface near the gage corner and is not as deep as shelling. Flaking is not considered to be a critical defect. This type of defect can be recognized by the following: shallow depressions with irregular edges, and horizontal hairline cracks along the running surface occurring near the gage corner (Figure 2.23).

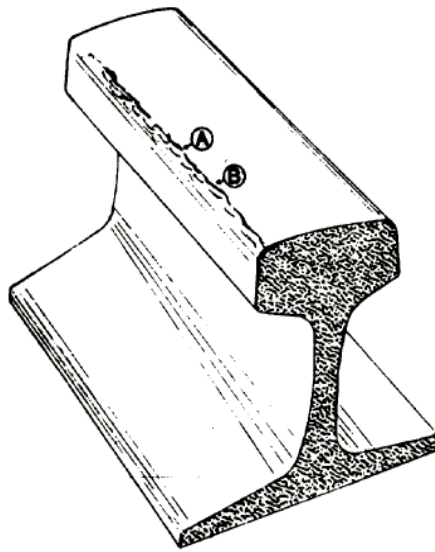


Figure 2.23: Sketch of the flaking phenomenon

Similar to flaking is the case of *slivers*. A sliver is the separation of a thin, tapered mass of metal from the surface of the head, web, or base of a rail; sliver is not a critical defect. Another type of defect caused by surface damage is the *flowed rail*, which is a rolling out of the tread metal beyond the field corner with no breaking down of the underside of the head. This defect is not considered to be critical; it appears on the track in the following ways: (a) a surface metal on the head flowed toward the field side giving

a creased appearance on the running surface near the field corner, (b) a protruding lip extending along the length of the rail or, in the advanced stage, as (c) a jagged nonuniform flow that may hand down or separate from the rail head.

Rail wear is the loss of material from the running surface and 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 2.24).

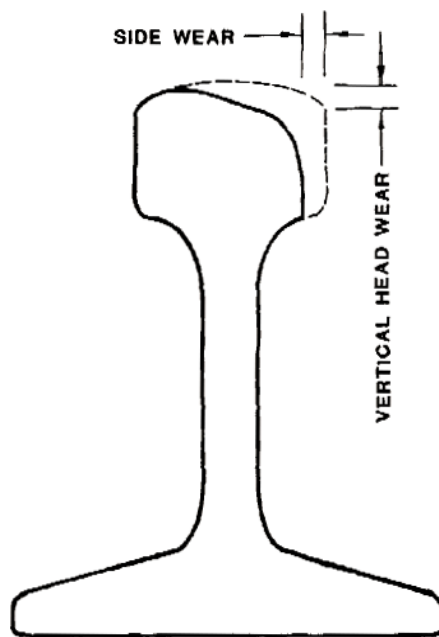


Figure 2.24: General appearance of rail wear

Already mentioned in section 2.1.1 is the phenomenon of *shelling*. Shelling is a progressive horizontal separation which may crack out at any level on the gage side but generally at the gage corner. It extends longitudinally not as a true horizontal or vertical crack, but an angle related to the amount of rail wear. Shelling is not a critical defect, but it can assume a relevant importance in the case it is masking a deeper and more threatening transverse defect. A sketch of a case of shelling is reported in Figure 2.25.

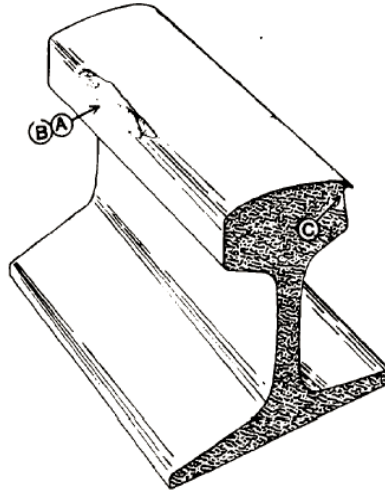


Figure 2.25: Example of shelling

Corrosion is a well known damage that can come from adverse environmental conditions; it is the decaying or corroding of the metal in the web or base of the rail. It can appear as pits or cavities in the upper base of the web of the rail. In advanced stages, a significant loss of material can be experienced. Finally, it should be mentioned that water from rain, snow or dew can become trapped in some of the described defects, in the rail along with oil and diesel. When a wheel runs over a track with entrapped fluid in a crack, a very high localized pressure at the crack tip can cause the crack to grow (Figure 2.26).

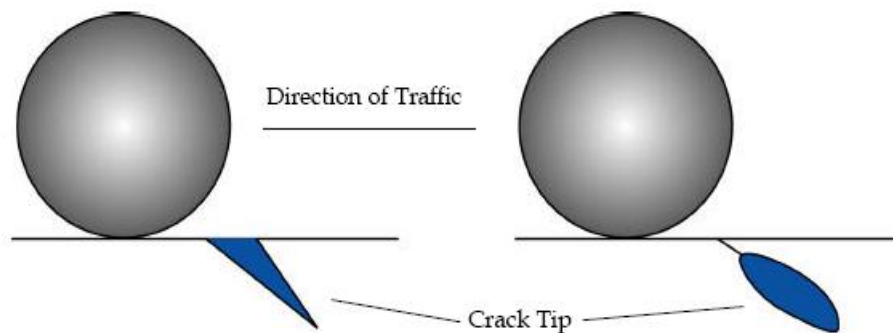


Figure 2.26: Trapped fluids causing cracking to worsen

2.2 Current rail inspection methods

2.2.1 Historical background

The first type of inspection ever used was the visual inspection. A sand mirror inspector was capable of inspecting one mile of rail per day. Unfortunately, many external and internal defects are overlooked in this way. Nondestructive testing had its beginning in rail applications in 1927, when Dr. Elmer Sperry noticed an increase in the number of disastrous train derailments. Dr. Sperry developed the first rail inspection car (Car SRS #101) to detect transverse fissures in railroad tracks, which was the first car built for commercial use, to inspect railroad tracks in Ohio and Indiana. Car SRS #102 detected a large transverse fissure in the head of a rail; much to the annoyance of the railroad, the rail was taken out of service and the following day was tested by Sperry's chief research engineer. The rail was tested and broke at the location of the transverse fissure. After the convincing test of the SRS #102, Sperry expanded his services and increased his fleet of 10 cars by the end of 1930. The first inspection car employed an induction method. A heavy current was induced through the rail to be tested. Search coils were moved through the resultant magnetic field to find perturbations in this field due to the presence of defects. By 1959, Sperry developed the first ultrasonic test car for the New York City Transit Authority. By the 1950s, track conditions significantly improved when continuously welded track replaced joint bars. This resulted in a shift of defect types found in rails, away from joint defects. Over the next 10 to 20 years the average age of the rail naturally increased and so did the average load of the rail cars. Then, in the 1980s, with increased use of inspection cars through North America, the general trend in

detection was lower defect rates. Fifteen to twenty years ago typical rail life was 800 million gross tons, while today 1.5 billion gross tons is not unusual.

Today, interest in repair and maintenance has changed, along with the characteristics of railroad industry. There is great pressure on operations to gain efficiencies for greater financial returns with increased traffic. These pressures have resulted in heavier axle loads and higher train speeds. Due to higher train speeds, there are shorter work windows for conducting inspections. Federal regulations require immediate remediation of detected defects, regardless of the type, size, or quantity of traffic that the rail carries.

Ultrasonic inspection technology is the predominant rail inspection technology used in North America. Most of the major contractors rely exclusively on ultrasonic rail testing cars, however some companies employ larger rail bound units that have both ultrasonic and magnetic induction technologies. Currently, the ultrasonic rail testing cars are capable of an average speed of 30 mph in a nonstop testing mode. However, in practice the vehicle stops frequently to hand verify indications from the rail testing. With stopping and verification, rail inspection vehicles sometimes can average only 6 to 8 mph in practice.

2.2.2 Current inspection systems

The different NDT inspection methods currently used on railways around the world can be mainly classified in the following categories:

- Visual inspection
- Traditional ultrasonic inspection

- Contact induction technique inspection
- Non-contact induction technique (eddy current) inspection

The visual inspection technique is widely used, but produces the poorest results of all the methods. It is now becoming widely accepted that even surface cracking often cannot be seen by the naked eye. The visual inspection methods can be manual or automated, through the development of automated visual inspection/detection systems. The automated visual inspection/detection systems use state of the art computer vision and pattern recognition techniques to assess the condition of the rail surface.

Ultrasonic techniques scan railhead through ultrasonic beams and detect the return of reflected or scattered energy using ultrasonic transducers [28]. The ultrasonic system is traditionally a pulse-echo method where standard ultrasonic piezo-electric transducers are mounted into fluid filled wheels or sleds. To transmit the ultrasonic wave from the wheel to the rail, a coupling media is needed. The coupling fluid used is usually oil- or water- based. The transducer orientation within the wheel is fixed. Inside each wheel, three to six transducers are mounted, usually with two wheels per rail. A total of as many as 12 transducers per rail can be hosted in a wheel or a sled. The most common configuration (Figure 2.27) uses transducer orientations to generate ultrasonic beams propagating at normal incidence (0°), and 70° from the normal to the rail surface. The 0° degree probe targets horizontal cracks that tend to grow in a 20° direction from the transverse plane and it is usually mounted in a forward and reverse position in the wheel. A 37° or 45° probe is also often used to target other discontinuities, including bolt hole cracks and defective welds. In addition, a *side looking transducer* may be located in each wheel, generating beams in the transverse plane of the track, to target vertical

discontinuities. The side looking transducer is typically oriented at 45° in respect to the vertical plane.

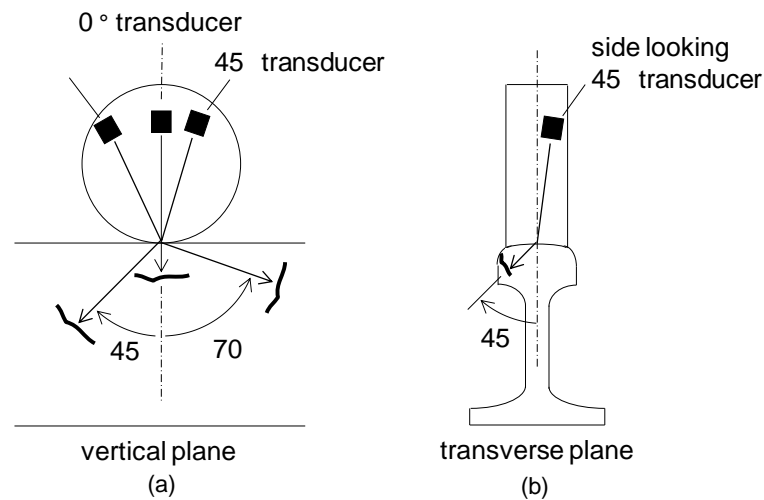


Figure 2.27: Ultrasonic probe wheel transducer arrangements

The amplitude of the reflections and their arrival times indicate the presence, the location and the severity of the damage [3]. Although, ultrasonic testing is capable of inspecting the whole railhead [29] it has several drawbacks such as:

- Limited ultrasonic inspection car speeds [15]. These are generally much slower than the line speed compelling the inspection operation to be carried out outside the commercial track periods, in order to avoid disruption at the normal train timetables. Typical operational speeds are between 25 and 45 mph; unfortunately, these speeds are theoretical, because if a damage is detected the inspection crew, generally, checks immediately the nature and the severity of the damage, reducing the inspection speed to an average of less than 10 mph.

- Shallow crack shadowing [15, 19] Small shallow cracks can shadow much more severe cracks by reflecting the ultrasonic beams, preventing so the detection of deeper defects.

- False defect detection. Current data reports that from 70% to 80% of the defect detection reveals to be false during the hand test verification. This contributes to a further slowing down of the inspection operations.

The induction technique, exploits electromagnetic phenomena. Basically, the rail becomes part of a circuit, where an ampere generator through brushes, in contact with the rail surface, injects high currents, which in presence of a defect are distorted. This distortion of the current flux in the rails generates discrepancies, in the associated magnetic field, detected by a special designed group of sensors placed on the inspection car [17]. The contact brushes required by this technique are often considered cumbersome.

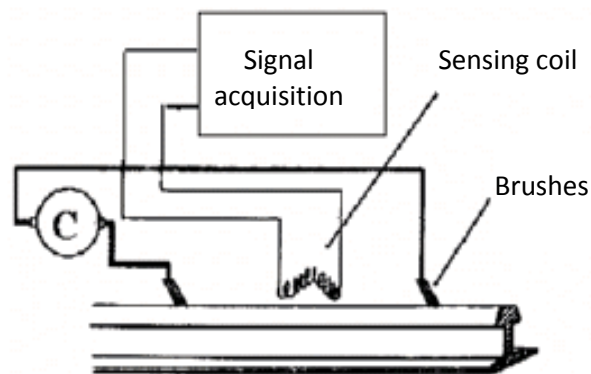


Figure 2.28: Scheme of the contact induction technique

Recently, a non-contact induction technique based on eddy currents was developed [5]. This method does not employ brushes to close the circuit with the rail, but

uses, instead, a magnetic core to induce eddy currents in the rail, which losses are correlated to the damage presence. However, although this technology allows using bogies of a standard railway car or coach at the standard speed of the line, only damages on the surface, or close to it, can be monitored.

Table 2.3: Advantages and disadvantages of different rail inspection methods

Inspection technique	Advantages	Disadvantages
Visual inspection	Easy to perform. Increased speed in case of automated visual inspection.	Not reliable. Low testing speed
Traditional UT	Well established. Reliable for detecting many internal and surface defects.	Unable to detect transverse defects under shelling. Limited testing speed. Contact required.
Contact induction	Reliable for surface defect detection.	Unreliable for internal deep defect detection. Limited testing speed. Contact required. Sensitivity to electromagnetic noise. Sensitivity to presence of joint, bar, holes.
Eddy current	Reliable for surface defect detection. Non-contact deployment. High testing speed.	Unreliable for internal deep defect detection. Small sensor lift-off. Sensitivity to electromagnetic noise. Sensitivity to presence of joint, bar, holes.
Guided waves*	Potentially reliable for surface and internal defect detection; Potentially sensitive to transverse defects covered by shelling. High testing speed. Non contact deployment.	Signal processing technique required to increase signal SNR. Lower frequencies may reduce sensitivity to certain small defects. Reliable deployment needs to be fully demonstrated in the field.

* Next section will describe the use of guided waves for rail inspection

Other disadvantages of the induction methods are the required small lift-off of the core from the rail head surface, and the pronounced sensitivity to electromagnetic noise. Rail inspection systems based on induction methods are also sensitive to the presence of joints, bars and holes in the rail, resulting in false detections that need to be purged from the results.

Other methods for rail inspection are currently being investigated, such as the *magnetic particle* and the *acoustic emission* techniques [30], but both are in the early stages of research.

Current data indicate that 70 to 80% of rail detector car indications prove false after hand-test verifications. Significant time is therefore, consumed hand-verifying false indications. Another significant concern is that current detector cars systems fail to detect detail fractures or other defects masked by shelled rail, details fractures, under spall, or even dirty rail. Thus, the frequency of detecting false indications has to be reduced while also improving the ability to detect certain types of defects.

In the last decade, the ultrasonic guided wave method appears to be a promising technique, which has been and currently is a topic of interest among NDT research groups around the world. A more detailed description of the use of ultrasonic guided waves for structural health monitoring of rails is presented in the next section. Table 2.3 summarizes the features of the different techniques described in the chapter, along with their advantages and disadvantages.

2.3 Ultrasonic guided waves for defect detection in rails

Guided waves result from the constructive interference of bulk longitudinal and shear waves propagating within a confined geometry. An illustration of the guided wave phenomenon, thought of as the superposition of bulk longitudinal and shear waves, is shown in Figure 2.29. To satisfy the boundary conditions at each interface, mode conversion into both longitudinal and shear waves occurs due to an incoming bulk wave. With enough propagation distance, the large number of mode converted bulk waves result in bulk wave resonances, otherwise known as guided waves.

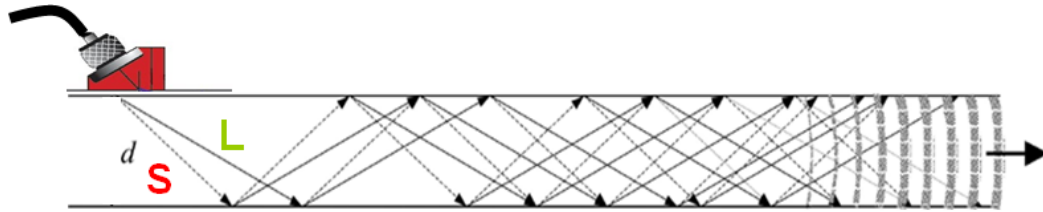


Figure 2.29: Guided waves propagating within a confined geometry.

Most structural elements such as plates, rods, pipelines, wing skin, or rails are natural waveguides, thereby lending themselves to ultrasonic guided wave based SHM methods. The guided wave method can be an effective diagnostic tool due to its capability of short- and long- range inspection. For the former, sensitivity is of key importance and the effect of dispersion is relatively unimportant when the monitored distances are small.

The ability to inspect hidden structures under water, coatings, insulations and concrete with great sensitivity, is another advantage of guided waves. From the economical point of view, the high cost effectiveness of guided wave-based techniques, is

related to large coverage and potential high testing speeds, in case of non-contact probing. The difficulties associated with the use of guided waves for SHM include their dispersive nature (wave velocity is frequency dependent) and the existence of multiple modes propagating simultaneously. The effect of dispersion is that the energy in a wavepacket propagates at different speeds depending on its frequency content. This manifests itself as a spreading of the wavepacket in space and time as it propagates through a structure, and it appears as an increase in the duration of the wavepacket in time and a decrease in its amplitude, which is undesirable in long range guided wave testing and consequently reduces the resolution and the sensitivity of the testing system.

A remarkable feature of guided wave analysis is the flexibility in selecting sensitive mode-frequency combinations. In view of dispersion, in a large structure test, it is essential to excite a mode in a nondispersive region since otherwise the shape of the wave packet will change as it propagates along the structure. Low dispersion is most conveniently achieved by testing at frequency-thickness products close to the maxima in the group velocity dispersion curves [31, 32, 33]. In addition, before choosing a particular mode, it is essential to check that it is sensitive to the defect types of interest. The wave structures give a good indication of the likely sensitivity to defects at different positions through the waveguide's cross-section.

Guided wave excitation methods are mainly classified into two classes: contact ones and non-contact ones. For the former, the probe contacts the surfaces of tested structure. Contact excitation can be conducted using relatively inexpensive, unobtrusive, piezoelectric ceramics such as PZT ($\text{Pb}(\text{Zr-Ti})\text{O}_3$) [33] or piezoelectric composites like

the MFC (Macro Fiber Composite) [34] or AFC (Active Fiber Composite) [35] transducers.

More sophisticated transducer systems for generating guided waves are laser spatial arrays [36] and comb transducers [37]. Comb transducers are made of bulk waves transducers (usually longitudinal waves) that are pasted on comb structures. The transducer element size and spacing dimensions are arbitrarily chosen to set a desired wavelength, and therefore excite points along a straight line on the phase velocity dispersion curves diagram (Figure 2.30). For this reason comb transducers are very useful to select individual modes in excitation.

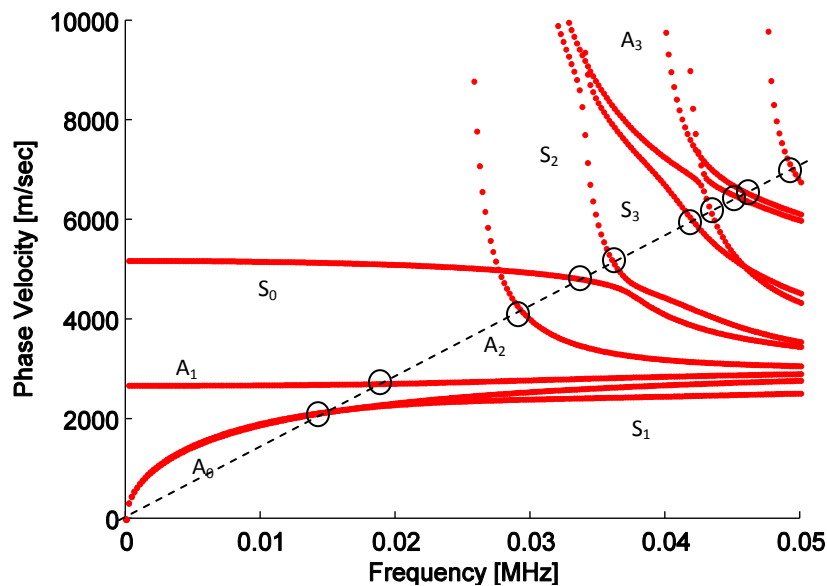


Figure 2.30: Comb transducer excitation of guided waves in a typical 115-lb A.R.E.M.A. rail

The use of wedges in contact ultrasonics or particular angles in air-coupled ultrasonics, allows an even greater selectivity of the excited guided modes. Owing to Snell's law, fixed excitation angles set the wave phase velocity in the medium, which allows to set a fixed dispersion point by selecting a specific frequency (Figure 2.31).

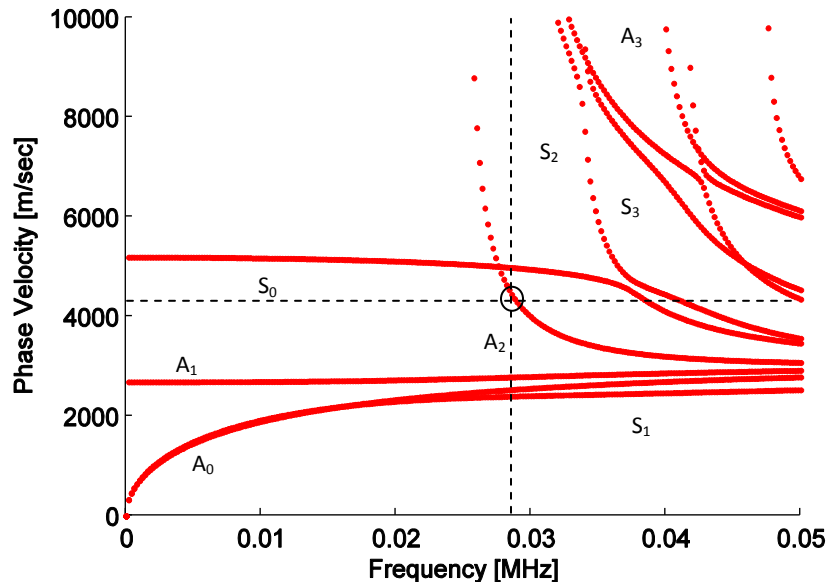


Figure 2.31: Angle beam ultrasonic for guided waves excitation in a typical 115-lb A.R.E.M.A. rail

Although the contact methods have advantage of simple operation, strict surfaces and couplant are needed, and this limits their application ranges. For example it is unsuitable for the inspection of the rough surface, curve surface, complex structure or structures that prohibit using couplant.

Ultrasonic guided waves can also be produced by a variety of different techniques, such as electromagnetic acoustic transducers (EMATs) [38] and magnetostrictive type sensor [39]. Based on the excitation, the non-contact methods are mainly divided into three classes: angle beam transducers, laser and EMAT excitation. These two techniques do not contact the surface and hence do not require couplant. Sensing of guided waves can be achieved by similar piezoelectric transducers in addition to fiber optic sensors and piezoelectric films such as PVDF (polyvinylidene fluoride). Top of the art capacitive and piezoelectric air-coupled transducers allow for a non-contact sensing approach, very useful in circumstances which the contact sensor-specimen is not

possible, or not favorable to the achievement of good system performance, such as high testing speed. The most common 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 [40, 41], frequency [42, 43] and joint-time frequency domain [44-47] of reflected [48], diffracted [49] and transmitted waves [50].

2.3.1 Advantages of guided waves for rail inspection

Guided wave based damage detection techniques of rails exploits reflection coefficient spectra [51] and wave mode conversion signatures for various defects [9, 52]. In the research work conducted by the author, the first approach to the rail inspection with the use by ultrasonic guided waves consisted of a low-frequency (\sim kHz) method probing the rail along its longitudinal, running direction (“long-range method”), shown in Figure 2.32. Laboratory and field testing of 115-lb A.R.E.M.A. rails were performed to

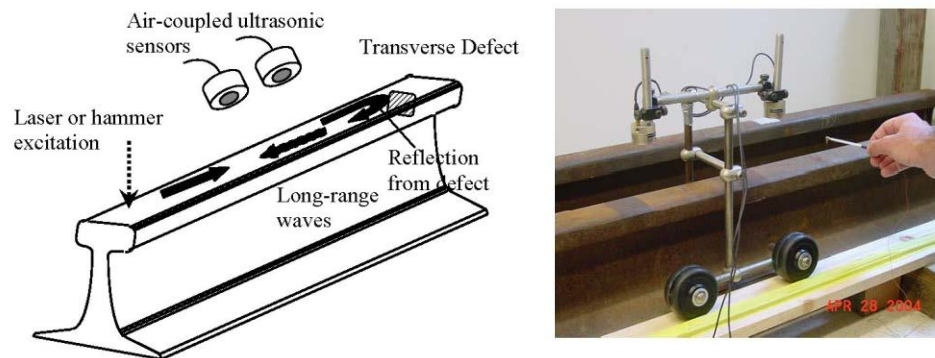


Figure 2.32: Rail defect detection: “long-range method”

characterize speed and attenuation of three primary guided-wave modes, namely the longitudinal, the lateral and the vertical modes. The interaction of these modes with rail defects were then examined in terms of wave reflection coefficients as a function of the

wave frequency. The method proved successful in detecting transverse head defects of various sizes and orientations including transverse surface cracks as shallow as 1 mm from the head surface. These tests were carried out by using a combination of impulse hammers, pulsed lasers and air-coupled ultrasonic sensors. The primary advantages of this method are 1) increased reliability of detection of transverse-type defects, even in the presence of shallow, horizontal surface cracks that often mask critical internal defects in conventional inspections, 2) potential for extremely high testing speeds, since guided waves propagate at the speed of sound in steel ($\approx 5,000$ mph), 3) extended rail coverage in a single test, since several feet of rail can be inspected at once, and 4) ease of field deployment since sensors can be positioned above the rail head and as far as 3.5" (≈ 89 mm) from the rail surface thus exceeding the recommended clearance envelope of 2" for rail inspection systems.

2.3.2 Laser generation of ultrasonic guided waves

Laser generation of ultrasound was first demonstrated by White (1963). Since then, lasers have been used to generate ultrasound in solids, liquids, and gases for a number of applications. An excellent reference for the laser ultrasonic science is given by Scruby and Drain [53]. The dominant mechanisms involved in laser generation of ultrasound in a solid are easy to outline. A pulsed laser beam impinges on a material and is partially absorbed by it. The optical power that is absorbed by the material is converted to heat, leading to rapid localized temperature increase. This results in rapid thermal expansion of a local region, which leads to generation of ultrasound into the medium. If the optical power is kept sufficiently low enough that the material does not melt and

ablate, the generation regime is called thermoelastic (see Figure 2.33 (a)). In this case the excitation can be represented by two in-plane forces opposite to one another. If the optical power is high enough to lead to melting of the material and plasma formation, once again ultrasound is generated, but in this case via momentum transfer due to material ejection (see Figure 2.33(b)). In this ablative regime the excitation can be represented by a force normal to the illuminated surface.

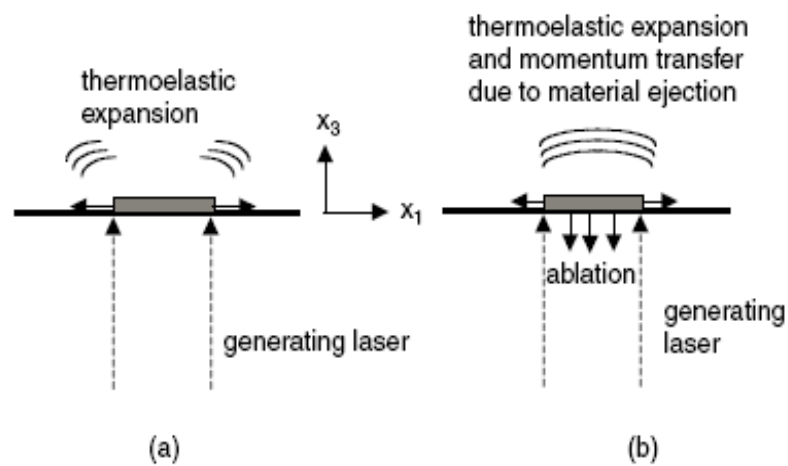


Figure 2.33: Laser generation of ultrasound in (a) thermoelastic regime and (b) ablative regime.

The ablative regime of generation is typically less preferred for nondestructive characterization of materials. However, it is useful in some process monitoring applications, especially since it produces strong bulk wave generation normal to the surface. In some cases where a strong ultrasonic signal is needed but ablation is unacceptable, a sacrificial layer (typically a coating or a fluid) is used either unconstrained on the surface of the test medium or constrained between the medium and an optically transparent plate. The sacrificial layer is then ablated by the laser, again leading to strong ultrasound generation in the medium due to momentum transfer. It is

important to characterize the ultrasound generated by laser heating of a material in order to determine the amplitude, frequency content, and directivity of the ultrasound generated. If the material ablates, the ultrasound that results from momentum transfer can be modeled as arising from a normal impulsive force applied to the surface.

The prototype described in this dissertation will use laser generation of guided waves in the slightly ablation regime. The ablated material, however, consists of foreign material usually present on the rail surface, such as dust, dirt or moisture, and it does not cause appreciable damage to the surface of the rail. It should also be noted that the small amount of material that is removed at each laser ablation is well below that removed by conventional rail grinding techniques.

2.4 Modeling guided wave propagation in rails

Various numerical and experimental studies have examined the propagation of ultrasonic guided waves in railroad tracks. This topic is of interest in the area of rail defect detection by long-range ultrasonic inspection as well as in the context noise generated by a passing train. The classical theories of Euler–Bernoulli and Timoshenko for rail vibrations are only accurate at frequencies below 0.5 and 1.5 kHz, respectively. At higher frequencies these theories cannot account for the significant cross-sectional deformations of the rail [54, 55] Numerical approaches have been proposed to predict the acoustic and ultrasonic modal properties of rails as dispersion curves. In Ref. [54], the finite element method was used to model transient vibrations in rails with beam and plate elements at frequencies as high as 6 kHz. A double Timoshenko beam model that allowed relative displacement between the head and the foot of the rail was used in Ref.

[55] to model vertical waves, again below 6 kHz. In the area of rail defect detection by long-range ultrasonic waves, however, frequencies in the range 10–50 kHz have shown the best promise [9, 57]. In this case, complications arise from the multimode and dispersive character of high frequency guided waves in rails that sometimes makes the mode identification challenging. For example, at 50 kHz, about 20 vibrational modes theoretically propagate in typical rails. In references [58, 59], the complexity of finite element modeling of waves in acoustic waveguides including rails was relaxed by proposing a discretization in the rail cross-sectional plane only. The three-dimensional problem was thus successfully treated as a bi-dimensional one with significant savings in computational efforts. Among the different solutions obtained with the use of numerical methods, Nigro [60] used the Ritz method for bars with rectangular cross-section while Fraser [61, 62] obtained dispersion curves for bars with rectangular and elliptical cross-section using the method of collocation. Elliptical and truncated elliptical cross-sections were investigated by Nagaya [63], still using a collocation method. The bi-dimensional method was employed in ref. [64] where a cyclic symmetry condition was imposed in the wave propagation direction. With the cyclic symmetry approximation, Gavric's bi-dimensional method could be implemented in standard finite element programs.

Semi Analytical Finite Element (SAFE) methods, also referred to in the literature as spectral or waveguide finite element methods, have emerged for modeling the guided wave propagation numerically as an alternative to “exact” methods based on the superposition of bulk waves (SPBW), that include the popular matrix-based method [65]. The main advantage of SAFE methods consists in the fact that the only bi-dimensional discretization of the cross-section is needed, with considerable computational savings

compared to a 3-D model of the entire waveguide. Recent studies [66] allowed the inclusion of damping in the model of the waveguide. The described techniques are very efficient to calculate the wave modal solutions in terms of dispersion curves and cross-sectional mode shapes. However, predicting the interaction of the waves with structural defects generally requires a three-dimensional model. Sophisticated numerical methods have been used to predict reflections of ultrasonic waves from rail defects, including three-dimensional time marching models [9, 52] and adaptive mesh refinement models [67]. The reflection coefficients of guided waves from defects are the basis of defect detection in rails by long-range inspection. Defects that develop transversely to the rail running direction are notoriously the most dangerous in rails, as seen earlier in the chapter. A recent paper [10] demonstrated the use of a commercial finite element code (ABAQUS) to model the reflection of vertical bending waves from transverse defects in a rail that is subject to impulsive excitation. The examined defects were transverse-type flaws in the rail head of different sizes and orientations. Reflection coefficients were predicted in the 20÷45 kHz frequency range and good agreement was shown with a limited set of experimental data. Although the implementation of the SAFE algorithm is straightforward and the solutions can be found in a stable manner for several types of problems, when the complexity of the waveguide requires a relative fine mesh to obtain the high frequency spectrum of the solutions, the algorithm can still become computationally demanding. The improvements to the SAFE method proposed in this dissertation will be given in chapter three.

3 Modeling ultrasonic wave propagation in waveguides of arbitrary cross-section

3.1 Introduction

This chapter discusses the ultrasonic wave propagation in arbitrary cross-section waveguides. As an introduction to this phenomenon, the analytical solution for the simple case of wave propagation in an isotropic single plate is presented. The geometrical complexity of the rail cross-section requires a different approach to be used in order to obtain the solution. A semi-analytical finite element method is described as an alternative yet still robust approach to get the solution of the problem. The free-vibration solution (unforced) and the forced solution to a laser excitation are shown for the case of an undamped rail. The effects of different load patterns and frequency range of the propagating wave are discussed. Experimental results are provided corroborating the numerical predictions.

3.2 Guided Waves in an unbounded medium

Prior to deriving the guided wave solutions for the case of a bounded medium such as a plate, it is important to understand wave propagation in an unbounded medium. According to Newton's second law, the equations of motion of a three-dimensional body (neglecting body forces) are described in eq. (3.1), [68] where ρ is the density of the material and u_x, u_y and u_z corresponds to the displacement along the x, y and z Cartesian coordinates.

$$\begin{aligned}
\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} &= \rho \frac{\partial^2 u_x}{\partial t^2} \\
\frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yz}}{\partial z} &= \rho \frac{\partial^2 u_y}{\partial t^2} \\
\frac{\partial \sigma_z}{\partial z} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{xz}}{\partial x} &= \rho \frac{\partial^2 u_z}{\partial t^2}
\end{aligned} \tag{3.1}$$

These equations hold regardless of the stress-strain behavior of the medium. The generalized form of Hooke's law relating stress and strain in elastic material is:

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{Bmatrix} \tag{3.2}$$

where C_{ij} represent the elastic constants of the material. For a material which is anisotropic, 21 independent elastic constants exist. 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 relation for an isotropic medium can thereby be represented in tensor form as:

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij} \tag{3.3}$$

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. (3.1) leads to Navier's equation of motion for an isotropic, elastic medium:

$$(\lambda + \mu)\vec{\nabla}(\vec{\nabla} \cdot \vec{u}) + \mu\nabla^2 \vec{u} = \rho \frac{\partial^2 \vec{u}}{\partial t^2} \quad (3.4)$$

In Navier's equation, \vec{u} is the displacement vector, and $\vec{\nabla}$ represents the divergence vector and the operator $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Based upon 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:

$$\vec{u} = \vec{\nabla}\phi + \vec{\nabla} \times \vec{\psi} \quad \text{with} \quad \vec{\nabla} \cdot \vec{\psi} = 0 \quad (3.5)$$

Substituting the displacement vector into Navier's equation results in the following expression:

$$\vec{\nabla} \left[(\lambda + 2\mu)\nabla^2 \phi - \rho \frac{\partial^2 \phi}{\partial t^2} \right] + \vec{\nabla} \times \left[\mu\nabla^2 \vec{\psi} - \rho \frac{\partial^2 \vec{\psi}}{\partial t^2} \right] = 0 \quad (3.6)$$

This equation is satisfied, if both the scalar and vector potential vanish, resulting in two decoupled wave equations:

$$c_1^2 \nabla^2 \phi = \frac{\partial^2 \phi}{\partial t^2}, \quad c_2^2 \nabla^2 \vec{\psi} = \frac{\partial^2 \vec{\psi}}{\partial t^2}, \quad (3.7)$$

where:

$$c_1 = \left(\frac{\lambda + 2\mu}{\rho} \right)^{1/2}, \quad c_2 = \left(\frac{\mu}{\rho} \right)^{1/2} \quad (3.8)$$

Therefore, two types of bulk waves exist; namely dilatation waves traveling with a constant speed of c_1 and rotational waves traveling at a constant speed of c_2 .

3.3 Guided waves in a bounded medium. Case of a single isotropic plate

Now let us consider wave propagation in a bounded isotropic elastic plate with thickness $2h$. Displacement u_y and the y axis corresponds to the through thickness direction and the displacement u_x and the x axes corresponds to the direction of propagation, as shown in Figure 3.1.

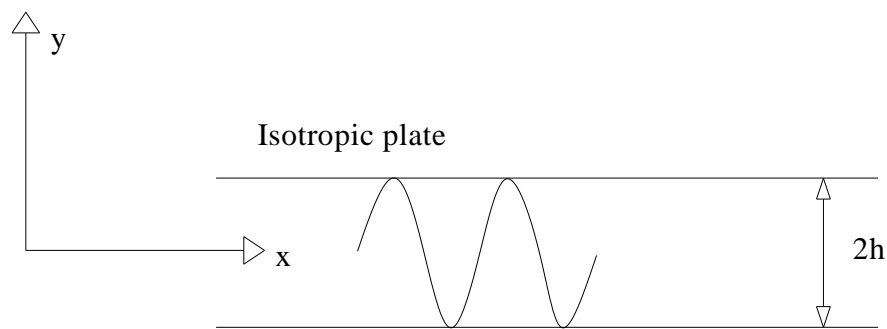


Figure 3.1: Guided wave propagation within a thin isotropic plate

The coordinate $y = 0$ is taken to be at the mid-plane of the plate. It is assumed that plane strain conditions exist such that the displacement $u_z = 0$ and $\frac{\partial u_z}{\partial z} = 0$. From eq. (3.5), the displacement is expressed as [69]

$$\begin{aligned} u_x &= \frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y} \\ u_y &= \frac{\partial \phi}{\partial y} - \frac{\partial \psi}{\partial x} \end{aligned} \quad (3.9)$$

For simplicity, the subscript z has been omitted from ψ in eq. (3.9). Substituting the expressions for u_x and u_y into Navier's equation, we obtain the two partial differential wave equations for which plane strain is two-dimensional:

$$\begin{aligned} \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} &= \frac{1}{c_L^2} \frac{\partial^2 \phi}{\partial t^2} \\ \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} &= \frac{1}{c_T^2} \frac{\partial^2 \psi}{\partial t^2} \end{aligned} \quad (3.10)$$

The general solution for ϕ and ψ are expressed as:

$$\begin{aligned} \phi &= \Phi(y) \exp[i(\xi x - \omega t)] \\ \psi &= \Psi(y) \exp[i(\xi x - \omega t)] \end{aligned} \quad (3.11)$$

where $\Phi(y)$ and $\Psi(y)$ is an expression that represents standing waves in the y direction and the $\exp[i(\xi x - \omega t)]$ term represents a propagating wave in the x direction. The terms ξ and ω represent the spatial distribution and temporal frequency of the wave, otherwise known as wavenumber and angular frequency respectively. Now, substituting eq. (3.11)

into the two partial differential wave equations, they reduce to two ordinary differential equations such that the solution, with p and q imaginary, leads to two complex exponential expressions for $\Phi(y)$ and $\Psi(y)$, which can be simplified as:

$$\begin{aligned}\Phi(y) &= A_1 \sin(py) + A_2 \cos(py) \\ \Psi(y) &= B_1 \sin(qy) + B_2 \cos(qy)\end{aligned}\tag{3.12}$$

where A_1, A_2, B_1 and B_2 are the wave amplitudes. The terms p and q are represented as:

$$\begin{aligned}p &= \left(\frac{\omega^2}{c_L^2} - \xi^2 \right)^{1/2} \\ q &= \left(\frac{\omega^2}{c_T^2} - \xi^2 \right)^{1/2}\end{aligned}\tag{3.13}$$

Using the wave equation solutions in eqs. (3.11), the displacements and stresses can be evaluated from eq. (3.3) and eq. (3.9) respectively as:

$$\begin{aligned}u_x &= [ik\Phi + \frac{d\Psi}{dy}] \exp[i(\xi x - \omega t)] \\ u_y &= [\frac{d\Phi}{dy} - i\xi\Psi] \exp[i(\xi x - \omega t)] \\ \sigma_{xy} &= [\mu(2i\xi \frac{d\Phi}{dy} + \xi^2\Psi + \frac{d^2\Psi}{dy^2})] \exp[i(\xi x - \omega t)] \\ \sigma_{yy} &= [\lambda(-\xi^2\Phi + \frac{d^2\Phi}{dy^2}) + 2\mu(\frac{d^2\Phi}{dy^2} - i\xi \frac{d\Psi}{dy})] \exp[i(\xi x - \omega t)]\end{aligned}\tag{3.14}$$

where $\Phi(y)$ and $\Psi(y)$ are defined in eq. (3.12). From eq. (3.14) it can be seen that the displacement components can be written in terms of elementary functions. For the

displacement in the x direction, the motion is symmetric (antisymmetric) with respect to the mid-plane, if u_x contains cosines (sines). The displacement in the y direction is symmetric (antisymmetric) if u_y contains sines (cosines) [69]. The modes of wave propagation can thus be split into symmetric and antisymmetric solutions. Therefore for symmetric and antisymmetric modes:

$$\begin{aligned} \Phi(h) &= A_2 \cos(py) \\ \Psi(h) &= B_1 \sin(qy) \end{aligned} \quad \textbf{Symmetric} \quad (3.15)$$

$$\begin{aligned} \Phi(h) &= A_1 \sin(py) \\ \Psi(h) &= B_2 \cos(qy) \end{aligned} \quad \textbf{Antisymmetric} \quad (3.16)$$

Where the expression relating frequency ω to the wavenumber ξ is obtained by utilization of the boundary conditions. For free boundaries, we thus impose at $y = \pm h$ that $\sigma_{xy} = \sigma_{yy} = 0$. For the symmetric modes, application of these boundary conditions yields a system of two homogeneous equations for the constants A_2 and B_1 . Similarly, for the antisymmetric modes we obtain two homogeneous equations for the constants A_1 and B_2 . Since the systems are homogeneous, we take the determinant of each system to obtain the simplified expressions which relate ω and ξ . These expressions are shown as:

$$\frac{\tan(qh)}{\tan(ph)} = -\frac{4\xi^2 pq}{(q^2 - \xi^2)^2} \quad \textbf{Symmetric modes} \quad (3.17)$$

$$\frac{\tan(qh)}{\tan(ph)} = -\frac{(q^2 - \xi^2)^2}{4\xi^2 pq} \quad \textbf{Antisymmetric modes} \quad (3.18)$$

These are the commonly found expressions for the Rayleigh-Lamb frequency equations.

An infinite number of eigensolutions exist for eqs. (3.14) and (3.15), thereby an infinite number of guided wave modes exist. Each eigenvalue corresponds to a particular angular frequency and mode of propagation, namely symmetric or antisymmetric. At low frequencies, only two propagating modes exist corresponding to the fundamental symmetric mode, S_0 and antisymmetric mode, A_0 . For each eigenvalue, a corresponding set of eigencoefficients also exist: A_2, B_1 and A_1, B_2 for the symmetric and antisymmetric case respectively. These coefficients can be used in eq. (3.14) to evaluate the Lamb mode shapes across the plate depth. The Lamb wave speeds are also a function of the frequency and are given as $c_p = \frac{\omega}{\xi}$ where c_p denotes the phase velocity for a particular mode. As a result, Lamb waves are dispersive. The speed at which the guided wave packet (or envelope) travels is known as the group velocity and is evaluated as $c_{gr} = \frac{d\omega}{d\xi}$. The guided wave dispersion solutions for a particular system are most commonly represented through phase velocity-frequency and group velocity-frequency plots.

3.4 Semi Analytical Finite Element Method

The exact analytical analysis of wave propagation characteristics is only possible for wave guides with simple cross-sections such as cylinders, ellipses, and rectangles with certain ratios of width to depth [70, 71, 72]. In order to obtain dispersion curves for an arbitrary cross-section like the rail over a wide range of frequencies and wave numbers, it is necessary the use of numerical methods. Semi-Analytical Finite Element

(SAFE) methods have emerged for modeling the guided wave propagation numerically as an alternative to the “exact” methods based on the superposition of bulk waves (SPBW), that include the popular matrix-based methods [65]. Motivations for the numerical methods include the necessity for modeling waveguides with arbitrary cross-section for which exact solutions do not generally exist. The complexity of the problem consists in the existence of multiple modes, in the frequency-dependent velocities (dispersion), and in the frequency-dependent attenuation. For example, the knowledge of the wave velocity is important for mode identification. Similarly, the knowledge of those mode-frequency combinations propagating with minimum attenuation losses helps maximizing the inspection coverage. The general SAFE approach for extracting dispersive solutions uses a finite element discretization of the cross-section of the waveguide alone. The displacements along the wave propagation direction are conveniently described in an analytical fashion as harmonic exponential functions. Thus only a bi-dimensional discretization of the cross-section is needed, with considerable computational savings compared to a three-dimensional discretization of the entire waveguide. In addition, since polynomial approximation of the displacement field along the waveguide is avoided, the method is applicable to predicting waves with very short wavelengths, where a traditional three-dimensional approximation may fail. A SAFE method for waveguides of arbitrary cross-section was demonstrated for the first time in 1973 [73, 74]. In these works dispersive solutions were obtained for the propagative modes only (i.e. real wavenumbers only). The same technique was used a decade later [75] to calculate both propagative modes and nonpropagative, evanescent modes (complex wavenumbers) for anisotropic cylinders. While the evanescent modes do not transport any energy along the

structure, they are important from a theoretical viewpoint to satisfy the boundary conditions. More recently, SAFE methods confined to obtaining the propagative solutions were applied to thin-walled waveguides [58], and railroad tracks [59]. Similar FE methods were proposed for example by Gry [76], Hayashi et al. [77, 78] and Damljanovic and Weaver [79]. A recent work extended the SAFE method for modeling dispersive solutions in waveguides of arbitrary cross-sections by accounting for material damping [76]. When accounting for damping, the exact energy velocity, rather than the conventional group velocity, is calculated along with the frequency-dependent attenuation of the modes.

3.4.1 SAFE Mathematical Framework

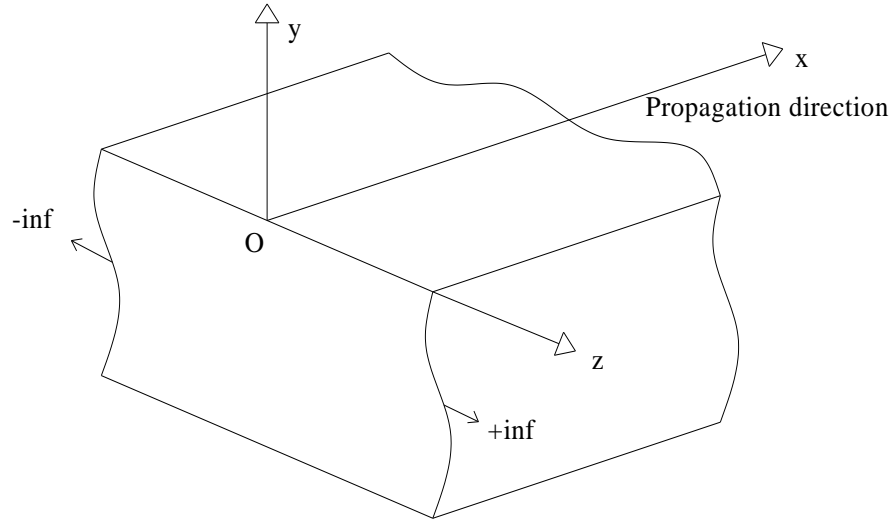


Figure 3.2: Waveguide in vacuum: infinitely wide plate.

The mathematical model is presented here for the case of a waveguide immersed in vacuum, as shown in Figure 3.2 for an infinitely wide plate; this formulation is also

applicable to an arbitrary cross-section such as the rail. The wave propagates along direction x with wavenumber ζ and frequency ω . The arbitrary cross-section lies in the y - z plane. The waveguide can generally be composed of anisotropic viscoelastic materials. The harmonic displacement, stress and strain field components at each point of the waveguide are expressed by:

$$\mathbf{u} = \begin{bmatrix} u_x & u_y & u_z \end{bmatrix}^T \quad (3.19)$$

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z & \sigma_{yz} & \sigma_{xz} & \sigma_{xy} \end{bmatrix}^T \quad (3.20)$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_x & \varepsilon_y & \varepsilon_z & \gamma_{yz} & \gamma_{xz} & \gamma_{xy} \end{bmatrix}^T \quad (3.21)$$

The constitutive relations at a point are given by $\boldsymbol{\sigma} = \tilde{\mathbf{C}} \boldsymbol{\varepsilon}$, where $\tilde{\mathbf{C}}$ is generally complex.

The compatibility equations can be written in matrix form as:

$$\boldsymbol{\varepsilon} = \left[\mathbf{L}_x \frac{\partial}{\partial x} + \mathbf{L}_y \frac{\partial}{\partial y} + \mathbf{L}_z \frac{\partial}{\partial z} \right] \mathbf{u} \quad (3.22)$$

where:

$$\mathbf{L}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{L}_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{L}_z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.23)$$

The equations of motion for the cross-section are formulated by inserting the kinetic and potential energies into Hamilton's equation. Although the nonconservative form of Hamilton's principle should be used to account for dissipation, the following analysis adopts a simplified approach that assumes a conservative waveguide; the resulting imaginary cross-sectional strain energy distribution is used to estimate the power dissipated by the section via imaginary wavenumbers. The time-averaged variation of the Hamiltonian of the waveguide, which vanishes at all material points, is:

$$\delta H = \int_{t_1}^{t_2} \delta (\Phi - K) dt = 0 \quad (3.24)$$

where Φ is the strain energy and K is the kinetic energy. The strain energy is given by:

$$\Phi = \frac{1}{2} \int_V \boldsymbol{\varepsilon}^T \tilde{\mathbf{C}} \boldsymbol{\varepsilon} dV \quad (3.25)$$

where the upper script T means a transpose vector and V is the volume. The kinetic energy is given by:

$$K = \frac{1}{2} \int_V \dot{\mathbf{u}}^T \rho \dot{\mathbf{u}} dV \quad (3.26)$$

where ρ is the mass density and the dot represents a time derivative. By integrating by parts the kinetic term, eq. (3.24) can be written as:

$$\int_{t_1}^{t_2} \left[\int_V \delta \boldsymbol{\varepsilon}^T \tilde{\mathbf{C}} \boldsymbol{\varepsilon} dV + \int_V \delta \mathbf{u}^T \rho \ddot{\mathbf{u}} dV \right] dt = 0 \quad (3.27)$$

The displacement field is assumed harmonic along the propagation direction, x , and spatial functions are used to describe its amplitude in the cross-sectional plane $y - z$:

$$\mathbf{u}(x, y, z, t) = \begin{bmatrix} u_x(x, y, z, t) \\ u_y(x, y, z, t) \\ u_z(x, y, z, t) \end{bmatrix} = \begin{bmatrix} u_x(y, z) \\ u_y(y, z) \\ u_z(y, z) \end{bmatrix} e^{i(\xi x - \omega t)} \quad (3.28)$$

where $i = \text{sqrt}(-1)$ is the imaginary unit. The waveguide's cross-sectional domain Ω can be represented by a system of isoparametric finite elements with domain Ω_e . Commercial Finite Element software can be used to obtain the discretization of Ω . The discretized version of the displacement expressions in eq. (3.28) over the element domain can be written in terms of the shape functions, $N_k(y, z)$, and the nodal unknown displacements, (U_{xk}, U_{yk}, U_{zk}) in the x , y and z directions:

$$\mathbf{u}^{(e)}(x, y, z, t) = \begin{bmatrix} \sum_{k=1}^n N_k(y, z) U_{xk} \\ \sum_{k=1}^n N_k(y, z) U_{yk} \\ \sum_{k=1}^n N_k(y, z) U_{zk} \end{bmatrix}^{(e)} e^{i(\xi x - \omega t)} = \mathbf{N}(y, z) \mathbf{q}^{(e)} e^{i(\xi x - \omega t)} \quad (3.29)$$

where:

$$\mathbf{N}(y, z) = \begin{bmatrix} N_1 & & N_2 & & \ddots & & N_n \\ & N_1 & & N_2 & & \ddots & \\ & & N_1 & & N_2 & & \\ & & & N_1 & & N_2 & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & N_n \end{bmatrix} \quad (3.30)$$

$$\mathbf{q}^{(e)} = [U_{x1} \ U_{y1} \ U_{z1} \ U_{x2} \ U_{y2} \ U_{z2} \ \dots \ \dots \ \dots \ U_{xn} \ U_{yn} \ U_{zn}]^T \quad (3.31)$$

and n denotes the number of nodes per element. The strain vector in the element can be represented as a function of the nodal displacements:

$$\boldsymbol{\varepsilon}^{(e)} = \left[\mathbf{L}_x \frac{\partial}{\partial x} + \mathbf{L}_y \frac{\partial}{\partial y} + \mathbf{L}_z \frac{\partial}{\partial z} \right] \mathbf{N}(y, z) \mathbf{q}^{(e)} e^{i(\xi x - \omega t)} = \mathbf{B}_1 + i\xi \mathbf{B}_2 \quad \mathbf{q}^{(e)} e^{i(\xi x - \omega t)} \quad (3.32)$$

where \mathbf{L} is given by eq. (3.23), $\mathbf{B}_1 = \mathbf{L}_y \mathbf{N}_{,y} + \mathbf{L}_z \mathbf{N}_{,z}$, $\mathbf{B}_2 = \mathbf{L}_x \mathbf{N}$, and $\mathbf{N}_{,y}$ and $\mathbf{N}_{,z}$ are the derivatives of the shape function matrix with respect to the y and z directions, respectively. Indicating by n_{el} the total number of cross-sectional elements, the discrete form of the Hamilton formulation of eq. (3.27) becomes:

$$\int_{t_1}^{t_2} \left\{ \bigcup_{e=1}^{n_{el}} \left[\int_{V_e} \boldsymbol{\varepsilon}^{(e)T} \tilde{\mathbf{C}}_e \boldsymbol{\varepsilon}^{(e)} dV_e + \int_{V_e} \delta \mathbf{u}^{(e)T} \rho_e \ddot{\mathbf{u}}^{(e)} dV_e \right] \right\} dt = 0 \quad (3.33)$$

where $\tilde{\mathbf{C}}_e$ and ρ_e are the element's complex stiffness matrix and density, respectively.

The substitution of eq. (3.32) into the strain energy term in eq. (3.33), followed by algebraic manipulations, yields:

$$\begin{aligned} & \int_{V_e} \delta \boldsymbol{\varepsilon}^{(e)T} \tilde{\mathbf{C}}_e \boldsymbol{\varepsilon}^{(e)} dV_e = \\ & \int_{\Omega_e} \int_x \delta \mathbf{q}^{(e)T} \left[\mathbf{B}_1^T - i\xi \mathbf{B}_2^T \right] \left[e^{i(\xi x - \omega t)} \right]^* \tilde{\mathbf{C}}_e \left[\mathbf{B}_1 + i\xi \mathbf{B}_2 \right] \mathbf{q}^{(e)} e^{i(\xi x - \omega t)} dx d\Omega_e = \\ & \int_{\Omega_e} \delta \left[\mathbf{q}^{(e)T} \left[\mathbf{B}_1^T - i\xi \mathbf{B}_2^T \right] \right] \tilde{\mathbf{C}}_e \left[\mathbf{B}_1 + i\xi \mathbf{B}_2 \right] \mathbf{q}^{(e)} d\Omega_e = \\ & \delta \mathbf{q}^{(e)T} \int_{\Omega_e} \left[\mathbf{B}_1^T \tilde{\mathbf{C}}_e \mathbf{B}_1 - i\xi \mathbf{B}_2^T \tilde{\mathbf{C}}_e \mathbf{B}_1 + i\xi \mathbf{B}_1^T \tilde{\mathbf{C}}_e \mathbf{B}_2 + \xi^2 \mathbf{B}_2^T \tilde{\mathbf{C}}_e \mathbf{B}_2 \right] d\Omega_e \mathbf{q}^{(e)} \end{aligned} \quad (3.34)$$

where $()^*$ indicates complex conjugate and $i^T = -i$. Thus the element stiffness matrix can be calculated by integrating over the cross-sectional domain Ω_e only, since the integration over x reduces to a unity factor due to the complex conjugate terms $e^{\pm i(\xi x - \omega t)}$. For viscoelastic materials, the strain energy defined by eq. (3.34) consists of a real component, describing the time-averaged elastic energy in the section, and an imaginary component, related to the time-averaged power dissipated by the section. As for the element kinetic energy contribution in eq. (3.33), by using the displacement expressions of eq. (3.29) and simplifying the harmonic terms $e^{\pm i(\xi x - \omega t)}$, the following can be written:

$$\int_{V_e} \delta \mathbf{u}^{(e)T} \rho_e \ddot{\mathbf{u}}^{(e)} dV_e = \int_{\Omega_e} \int_x \delta \mathbf{u}^{(e)T} \rho_e \ddot{\mathbf{u}}^{(e)} dx d\Omega_e = -\omega^2 \delta \mathbf{q}^{(e)T} \int_{\Omega_e} \mathbf{N}^T \rho_e \mathbf{N} d\Omega_e \mathbf{q}^{(e)} \quad (3.35)$$

Substituting eqs. (3.34) and (3.35) into eq. (3.33) yields:

$$\int_{t_1}^{t_2} \left\{ \bigcup_{e=1}^{n_{el}} \delta \mathbf{q}^{(e)T} \left[\mathbf{k}_1^{(e)} + i\xi \mathbf{k}_2^{(e)} + \xi^2 \mathbf{k}_3^{(e)} - \omega^2 \mathbf{m}^{(e)} \right] \mathbf{q}^{(e)} \right\} dt = 0 \quad (3.36)$$

where:

$$\begin{aligned} \mathbf{k}_1^{(e)} &= \int_{\Omega_e} \left[\mathbf{B}_1^T \tilde{\mathbf{C}}_e \mathbf{B}_1 \right] d\Omega_e \\ \mathbf{k}_2^{(e)} &= \int_{\Omega_e} \left[\mathbf{B}_1^T \tilde{\mathbf{C}}_e \mathbf{B}_2 - \mathbf{B}_2^T \tilde{\mathbf{C}}_e \mathbf{B}_1 \right] d\Omega_e \\ \mathbf{k}_3^{(e)} &= \int_{\Omega_e} \left[\mathbf{B}_2^T \tilde{\mathbf{C}}_e \mathbf{B}_2 \right] d\Omega_e \\ \mathbf{m}^{(e)} &= \int_{\Omega_e} \mathbf{N}^T \rho_e \mathbf{N} d\Omega_e \end{aligned} \quad (3.37)$$

Applying to eq. (3.36) standard finite element assembling procedures yields:

$$\int_{t_1}^{t_2} \delta \mathbf{U}^T \left[\mathbf{K}_1 + i\zeta \mathbf{K}_2 + \zeta^2 \mathbf{K}_3 - \omega^2 \mathbf{M} \right] \mathbf{U} dt = 0 \quad (3.38)$$

where \mathbf{U} is the global vector of unknown nodal displacements, and:

$$\mathbf{K}_1 = \bigcup_{e=1}^{n_{el}} \mathbf{k}_1^{(e)}, \quad \mathbf{K}_2 = \bigcup_{e=1}^{n_{el}} \mathbf{k}_2^{(e)}, \quad \mathbf{K}_3 = \bigcup_{e=1}^{n_{el}} \mathbf{k}_3^{(e)}, \quad \mathbf{M} = \bigcup_{e=1}^{n_{el}} \mathbf{m}^{(e)} \quad (3.39)$$

Due to the arbitrariness of $\delta \mathbf{U}$, the following homogeneous general wave equation is finally obtained:

$$\left[\mathbf{K}_1 + i\zeta \mathbf{K}_2 + \zeta^2 \mathbf{K}_3 - \omega^2 \mathbf{M} \right]_M \mathbf{U} = 0 \quad (3.40)$$

where the subscript M is the number of total degrees of freedom of the system. Nontrivial solutions can be found by solving a twin-parameter generalized eigenproblem in ζ and ω . The frequency ω is a real positive quantity. The wavenumber ζ can be either real or complex and can have both positive and negative signs. A classic technique to solve the eigenvalue problem $\zeta(\omega)$ consists of recasting eq. (3.40) to a first-order eigensystem by doubling its algebraic size:

$$\mathbf{A} - \zeta \mathbf{B}_{2M} \mathbf{Q} = \mathbf{0} \quad (3.41)$$

Where \mathbf{A} , \mathbf{B} and \mathbf{Q} are defined in eq. (3.42).

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{K}_1 - \omega^2 \mathbf{M} \\ \mathbf{K}_1 - \omega^2 \mathbf{M} & \hat{\mathbf{K}}_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{K}_1 - \omega^2 \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_3 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{U} \\ \xi \mathbf{U} \end{bmatrix} \quad (3.42)$$

Solving eq. (3.41) at each frequency ω , $2M$ eigenvalues ξ_m and, consequently, $2M$ eigenvectors are obtained.

The eigenvectors are the M forward and the corresponding M backward modes. In general the eigenvalues are pairs of complex conjugate numbers $(\pm \xi_{Re} \pm i\xi_{Im})$, representing propagative waves exponentially decaying in the $\pm x$ -directions. The phase velocity can be then evaluated by $c_p = \omega / \xi_{Re}$ and the attenuation, in Nepers per meter, by ξ_{Im} .

3.4.2 Modeling the attenuation in the rail

Although damping was not considered for the rail model examined later, the basic SAFE formulation for a damped wave guide is discussed here following the work in [76]. Linear viscoelasticity can be modeled by allowing complex components in the material's stiffness matrix

$$\tilde{\mathbf{C}} = \mathbf{C}' - i\mathbf{C}'' \quad (3.43)$$

where \mathbf{C}' contains the storage moduli and \mathbf{C}'' contains the loss moduli. The matrix $\tilde{\mathbf{C}}$ can be expressed as a combination of the elastic stiffness tensor, \mathbf{C} , and the viscosity tensor,

$\boldsymbol{\eta}$:

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ & & C_{33} & C_{34} & C_{35} & C_{36} \\ & & & C_{44} & C_{45} & C_{46} \\ & & & & C_{55} & C_{56} \\ \text{Sym.} & & & & & C_{66} \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} \eta_{11} & \eta_{12} & \eta_{13} & \eta_{14} & \eta_{15} & \eta_{16} \\ & \eta_{22} & \eta_{23} & \eta_{24} & \eta_{25} & \eta_{26} \\ & & \eta_{33} & \eta_{34} & \eta_{35} & \eta_{36} \\ & & & \eta_{44} & \eta_{45} & \eta_{46} \\ & & & & \eta_{55} & \eta_{56} \\ \text{Sym.} & & & & & \eta_{66} \end{bmatrix} \quad (3.44)$$

In the hysteretic model [77] the complex component of the stiffness matrix is independent of frequency, thus:

$$\tilde{\mathbf{C}} = \mathbf{C}' - i\boldsymbol{\eta} \quad (3.45)$$

As a consequence, the hysteretic stiffness matrix has to be determined only once for the entire frequency range examined. The wave attenuation, defined as the loss per unit distance travelled, is commonly modelled as proportional to the frequency times the imaginary part of the stiffness matrix \mathbf{C}'' [77]. It is obvious from eq. (3.45) that the attenuation is a linear function of the frequency. Since the stiffness matrix $\tilde{\mathbf{C}}$ is now complex, and consequently, the matrices \mathbf{A} and \mathbf{B} of eq. (3.41) are now complex as well, $2M$ complex eigenvalues ζ_m and, consequently, $2M$ complex eigenvectors are obtained for each input frequency ω . The stiffness matrices \mathbf{K}_1 and \mathbf{K}_3 in eq. (3.40) are symmetric, while \mathbf{K}_2 is skew symmetric. When damped motion is being considered, the

\mathbf{K}_i are all generally complex. The mass matrix \mathbf{M} , is real symmetric and positive definite.

3.4.3 Rail as undamped waveguide

Modeling the rail as an undamped waveguide, the stress-strain relation is governed by a real stiffness matrix $\tilde{\mathbf{C}}$. In this case a $3n_{el} \times 3n_{el}$ transformation diagonal matrix \mathbf{T} is useful to eliminate the imaginary unit in eq. (3.40). The elements of \mathbf{T} corresponding to the u_y and u_z displacement components are equal to 1, while those corresponding to u_x are equal to the imaginary unit:

$$\mathbf{T} = \begin{bmatrix} i & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & i & \\ & & & & & 1 \\ & & & & & & 1 \end{bmatrix} \quad (3.46)$$

This matrix has the properties $\mathbf{T}^T = \mathbf{T}^*$ and $\mathbf{T}^* \mathbf{T} = \mathbf{T} \mathbf{T}^* = \mathbf{I}$, where \mathbf{I} is the identity matrix. The terms in eq. (3.40) are pre-multiplied by \mathbf{T}^T and post-multiplied by \mathbf{T} . This manipulation does not alter the symmetric matrices \mathbf{K}_1 , \mathbf{K}_3 and \mathbf{M} since they do not mix u_x with u_y or u_z :

$$\begin{aligned} \mathbf{T}^T \mathbf{K}_1 \mathbf{T} &= \mathbf{K}_1 \\ \mathbf{T}^T \mathbf{K}_3 \mathbf{T} &= \mathbf{K}_3 \\ \mathbf{T}^T \mathbf{M} \mathbf{T} &= \mathbf{M} \end{aligned} \quad (3.47)$$

The skew symmetric matrix \mathbf{K}_2 , instead, mixes u_x with u_y and u_z but it does not mix u_y and u_z with each other. It follows that:

$$\mathbf{T}^T \mathbf{K}_2 \mathbf{T} = -i \hat{\mathbf{K}}_2 \quad (3.48)$$

where $\hat{\mathbf{K}}_2$ is now a symmetric matrix. The introduction of the matrix \mathbf{T} is equivalent to multiplying u_x by the imaginary unit to force the quadrature with u_y and u_z as done in previous works [58, 59]. The final form of the eigenvalue problem in eq. (3.40) is:

$$\left[\mathbf{K}_1 + \xi \hat{\mathbf{K}}_2 + \xi^2 \mathbf{K}_3 - \omega^2 \mathbf{M} \right]_M \hat{\mathbf{U}} = 0 \quad (3.49)$$

where $\hat{\mathbf{U}}$ is a new nodal displacement vector. If the full complex spectrum for both propagative and evanescent modes is of interest, the unknown complex wavenumbers $\xi(\omega)$ must be obtained for a given frequency ω , solving eq. (3.49) as a second-order polynomial eigenvalue problem or the linearized eigensystem of eq. (3.40). In this case, the eigenvalues occur as pairs of real numbers ($\pm \xi_{Re}$), representing propagative waves in the $\pm x$ -directions, as pairs of complex conjugate numbers ($\pm \xi_{Re} \pm i \xi_{Im}$), representing evanescent waves decaying in the $\pm x$ -directions, or as pairs of purely imaginary numbers ($\pm i \xi_{Im}$), representing the exponentially decaying near fields in the $\pm x$ directions. The waves which are evanescent at low frequencies become propagative above the cut-off frequencies, ω_c . The cut-off frequencies can be computed by letting $\xi = 0$ in eq. (3.49) and solving the simple eigenvalue problem of eq. (3.50).

$$\left[\mathbf{K}_1 - \omega_c^2 \mathbf{M} \right]_M \hat{\mathbf{U}} = 0 \quad (3.50)$$

3.4.4 Group and energy velocity

In order to compute the group velocity by the conventional manner, the derivatives of the frequency-wavenumber dispersion relations must be calculated based on the differences of the values for adjacent points of the same mode, i.e. $c_g = \partial\omega/\partial\xi = (\omega_{k+1} - \omega_k)/(\xi_{k+1} - \xi_k)$. This implies that the accuracy of the velocity solution is very sensitive to that of the (ξ, ω) solutions. Also, in this case the (ξ, ω) solutions must be categorized for the different modes (mode tracking). Tracking the modes is not straightforward when one mode approaches another; a technique to track the same mode will be shown later in the chapter. The necessity for mode tracking, as well as the dependency on the frequency step resolution, can be avoided by calculating the group velocity directly at each (ξ, ω) solution points without any contribution from adjacent points ref. [80]:

$$c_g = \frac{\partial\omega}{\partial\xi} = \frac{\hat{\mathbf{U}}_L^T \hat{\mathbf{K}}_2 + 2\xi \hat{\mathbf{K}}_3 \hat{\mathbf{U}}_R}{2\omega \hat{\mathbf{U}}_L^T \mathbf{M} \hat{\mathbf{U}}_R} \quad (3.51)$$

From this relation the group velocity can be evaluated for each individual solution (ω, ξ) of the dispersion relations at a time independently of any adjacent solution. As reported in refs. [82, 83], the group velocity definition is not valid in damped waveguides.

The energy velocity, V_e , is the appropriate property for damped media. The definition of the energy velocity can be found in classical textbooks [68]. The expression used in the present work is given in eq. (3.52)

$$V_e = \frac{\frac{1}{\Omega} \int_{\Omega} \mathbf{P} \cdot \hat{\mathbf{x}} d\Omega}{\frac{1}{T} \int_T \left(\frac{1}{\Omega} \int_{\Omega} e_{tot} d\Omega \right) dt} \quad (3.52)$$

where $\hat{\mathbf{x}}$ is the unit vector along the wave propagation direction, $1/T \int_T (..)dt$ denotes the time average over one period T , e_{tot} is the total energy density (kinetic and potential), and \mathbf{P} represents the time averaged Poynting vector (real part only). The time averaged Poynting vector can be calculated from:

$$\mathbf{P} = -\frac{1}{2} \text{Re}(\boldsymbol{\sigma} \dot{\mathbf{u}}^*) \quad (3.53)$$

where σ is the classical 3x3 stress tensor, and $\dot{\mathbf{u}}^*$ is the complex conjugate of the particle velocity vector. The denominator in eq. (3.52) can be evaluated by introducing the expressions of the time averaged energy for the kinetic component, $\langle e_k \rangle_t$, and the potential component, $\langle e_p \rangle_t$, following the formulation in reference [81]:

$$\langle e_k \rangle_t = \frac{\omega^2}{4} \rho \mathbf{u}^T \mathbf{u} \quad (3.54)$$

$$\langle e_p \rangle_t = \frac{1}{4} \boldsymbol{\varepsilon}^T \mathbf{C} \boldsymbol{\varepsilon} \quad (3.55)$$

3.4.5 Methods to accelerate the SAFE solution process

Solving the eigensystem can be computationally consuming, especially when the system contains more than 1000 x 1000 elements. As an example, the computational time spent to solve a 2,104x2,104 matrix eigensystem, in the range of frequency 0-300

kHz with 1 kHz step, was measured at 150 hours on a Pentium IV 3 GHz workstation with 3 GB of RAM, running Matlab on Windows XP 32 bit O. S., using the traditional ‘eig’ function to solve the generalized eigenvalue problem. One way to reduce the computational time is to increase the frequency step, which unfortunately reduces, the frequency resolution of the analysis. This is not always feasible, since a poor resolution could miss some important solutions in the proximity of mode coupling or mode cut-off points; low-resolution could also lead to misinterpretation of the results, as depicted in Figure 3.3 referring to phase velocity dispersion curves of a 115 lb A.R.E.M.A. section

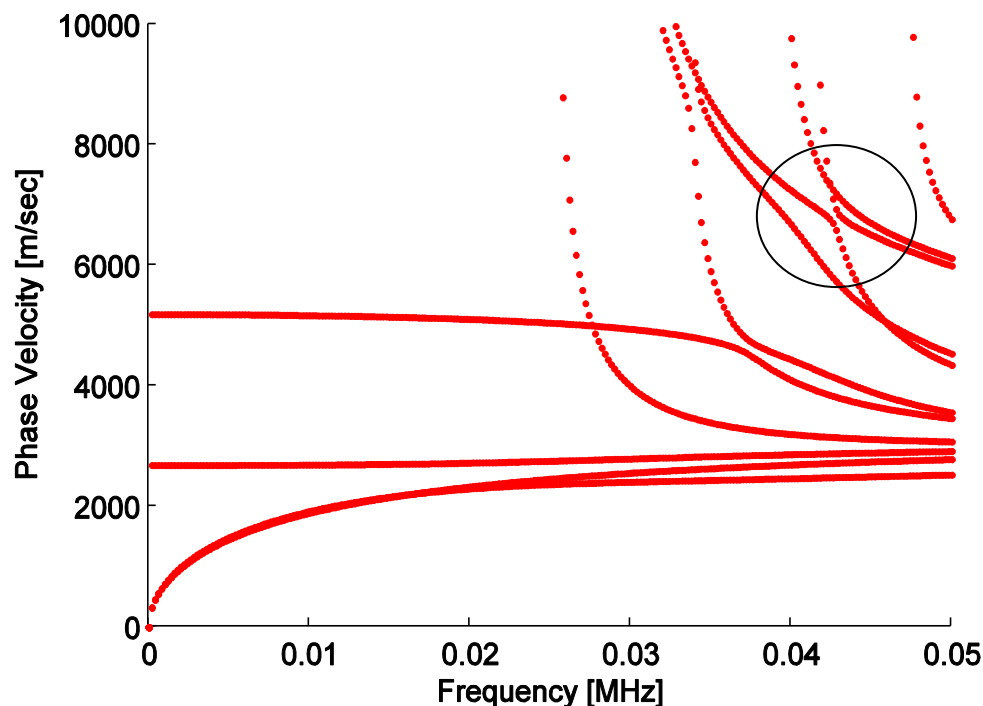


Figure 3.3: Dispersion curves of a 115 lb A.R.E.M.A. section rail; high frequency resolution is needed to correctly interpret the solutions in proximity of two modes (detail inside the circle)

rail, when a larger number of points would be needed to track the correct mode after the intersection. For undamped waveguides; the introduction of the \mathbf{T} operator is very

helpful since it eliminates any imaginary components of the elements of the matrices **A** and **B**, and the eigensystem becomes real and symmetric. When the matrices of the system are real and symmetric, the ‘eig’ function can compute the eigensolutions of the standardized problem much faster than in the complex case. If damping is considered in the waveguide and an accurate frequency resolution is needed to address the issues presented previously, the ‘eig’ function performance is very limited. At every frequency step ω_i only a few elements $n(\omega_i)$ of the eigenvalue spectrum are needed, since the remaining eigenvalues are discarded because associated to nonpropagative highly attenuated modes (e.g. end-modes). Moving from zero to the maximum frequency of interest, the number of computed eigenvalues $n(\omega)$ increases; looking at the phase velocity curves, the lowest frequency at which one mode appears in the graph, is called cut-off frequency. An algorithm able to solve the eigensystem only for a finite number of eigenvalues would accelerate the process of finding the solution. An algorithm implemented by the author is explained here. The eigensystem in eq. (3.41) is equivalent to:

$$\mathbf{A}\mathbf{Q} = \xi\mathbf{B}\mathbf{Q} \quad (3.56)$$

A shift and invert transformation is applied to the system in order to enhance the convergence to a desired portion of the eigenvalue spectrum [84]. If \mathbf{Q}, ξ is an eigenpair for **A, B** and $\sigma \neq \xi$, the shifting process consists in rewriting the eq. (3.56) in the following form:

$$\mathbf{A} - \sigma \mathbf{B}^{-1} \mathbf{B} \mathbf{Q} = \nu \mathbf{Q} \quad (3.57)$$

where:

$$\nu = \frac{1}{\xi - \sigma} \quad (3.58)$$

and σ is a first trial value for the sought eigenvalues. This transformation is effective for finding eigenvalues near σ since the n largest-in-magnitude eigenvalues of the standardized problem:

$$\mathbf{C} \mathbf{Q} = \nu \mathbf{Q}, \quad (3.59)$$

where

$$\mathbf{C} = \mathbf{A} - \sigma \mathbf{B}^{-1} \mathbf{B} \quad (3.60)$$

correspond to the n eigenvalues ξ_i of the original problem that are nearest to the shift σ in absolute value. These transformed eigenvalues of largest magnitude are precisely the eigenvalues that are easy to compute with a Krylov method [85]. Once they are found, they may be transformed back to eigenvalues of the original problem using the following relation :

$$\xi_i = \sigma + \frac{1}{\nu_i}. \quad (3.61)$$

The eigenvector \mathbf{Q}_i associated with ν_i in the transformed problem is also an eigenvector of the original problem corresponding to ξ_i . In order to implement this transformation,

the linear system of eq. (3.60) has to be solved. One feasible solution is to operate an LUP matrix factorization of $\mathbf{A} - \sigma\mathbf{B}$, where

$$\mathbf{A} - \sigma\mathbf{B} = \mathbf{L}\mathbf{U}\mathbf{P} \quad (3.62)$$

being \mathbf{L} , \mathbf{U} , \mathbf{P} a lower triangular, an upper triangular and a permutation matrix respectively. Matrix \mathbf{C} can be obtained as a solution of the system

$$\mathbf{B} = \mathbf{L}\mathbf{U}\mathbf{P}^* \mathbf{C} \quad (3.63)$$

computed by Gaussian elimination. In the described process, the application of the Krylov method is iterative, and the convergence to the solution is depending on the spectral distribution of the eigenvalues. If the sought eigenvalues are clustered, the performance of this approach is very elevated. The steps of the described algorithm, following the shift and invert spectral transformation, are known in literature as the Arnoldi Iterations. Software has already been implemented to perform this iterative process, and it is known under the acronym of ARPACK (Arnoldi Package) [86]; ARPACK is a collection of Fortran77 subroutines designed to solve large scale eigenvalue problems. Matlab routines are also available, such as the 'eigs' and 'eigs2' functions, which have been used to compute the results shown later on in this chapter. Tests of the application of the described process showed substantially smaller computational time in comparison to traditional methods; the solution of the 2,104 x 2,104 matrix eigensystem, presented at the beginning of this section, was obtained in just about 13 hours, approximately one twelfth of the 150 required with the 'eig' routine.

3.4.6 Mode tracking algorithm

The solution of the eigenvalue problem is the dispersion curves diagram. The dispersion curves are sets of frequency-wavenumber points, rather than a family of curves in which the wavenumber and the phase velocity are functions of frequency. The results of step ω_j are mathematically independent from the results of step ω_{j+1} , even if considerations can be made for two adjacent frequency values, as it will be shown in this section.

In some cases it could be helpful to focus on just one particular mode of propagation; hence it would be necessary to extract the wavenumber of this particular mode at each frequency. This can be a nontrivial task when the solutions are very close to each other. It is therefore important to develop a mode tracking algorithm, which can isolate the individual modes in the dispersion curves. Considering the orthogonality of the modes, a technique was developed to perform mode tracking [82]. Mode orthogonality can be expressed in the following way:

$$\mathbf{\Phi}_r^T(\omega) \cdot \mathbf{B} \cdot \mathbf{\Phi}_s^T(\omega) = \begin{cases} 0 & r \neq s \\ \neq 0 & r = s \end{cases} \quad (3.64)$$

where $\mathbf{\Phi}_r(\omega)$ and $\mathbf{\Phi}_s(\omega)$ are eigenvectors of the eigensystem in eq.(3.41) and matrix \mathbf{B} is expressed by eq.(3.42). It is reasonable to assume that if the frequency resolution of the problem is high, the mode shapes will not vary significantly between adjacent frequency steps. Hence, the B-orthogonality can also be expressed between modes related to two adjacent frequency steps:

$$\mathbf{\Phi}_r^T(\omega_k) \cdot \mathbf{B} \cdot \mathbf{\Phi}_s^T(\omega_{k+1}) \simeq \begin{cases} 0 & r \neq s \\ \neq 0 & r = s \end{cases}. \quad (3.65)$$

The creation of an orthogonality matrix between the modes of adjacent frequency steps is then useful to obtain a tracking of the computed modes. Defining the orthogonality matrix

$$\Theta = \mathbf{\Phi}^T(\omega_k) \cdot \mathbf{B} \cdot \mathbf{\Phi}^T(\omega_{k+1}) \quad (3.66)$$

for each value of ω along the spectrum, its principal diagonal can give information about the position of the different modes; if the wavenumber versus frequency curves have not crossed in the considered frequency interval, the largest terms in the matrix Θ will be the diagonal ones. The presence of an off-diagonal term that is larger than the corresponding diagonal term indicates that the curves have crossed. As an example, if the terms $\Theta_{i,i+1}$ and $\Theta_{i+1,i}$ are larger than the terms $\Theta_{i,i}$ and $\Theta_{i+1,i+1}$ this indicates that the i^{th} mode at frequency step k became the i^{th+1} mode at frequency step $k+1$ and that the i^{th+1} mode at frequency step k became the i^{th} mode at frequency step $k+1$. Applying the introduced algorithm, a correct mode tracking can be performed. The implementation of the algorithm requires some numerical tolerance in order to satisfy eq.(3.65). In some cases, a mode may have two wavenumbers at certain frequency values, as shown in Figure 3.4, related to the analysis performed on the damped rail. In that case, the first portion of the curve with a negative slope indicates a negative energy velocity, implying components that propagate in the negative x -direction. The ordering algorithm has to check for this occurrence, and needs to exclude points with negative energy velocity.

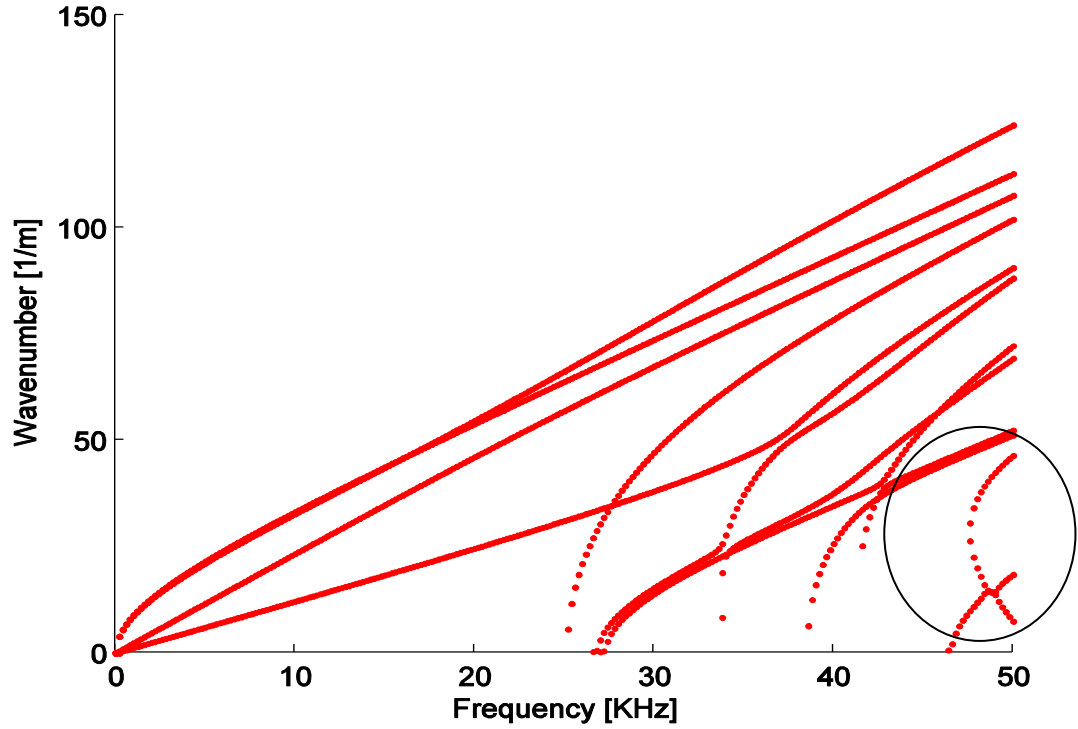


Figure 3.4: Wavenumber as a function of frequency for damped rail. Case of double solutions

3.5 Unforced solution of guided waves in rails

This section will present the results obtained analyzing the guided wave propagation in a railroad track using the SAFE method. The solutions of the system eq. (3.41) are the eigenvalues and eigenvectors, computed for the entire frequency spectrum of interest. For a given frequency step ω , the $n(\omega)$ obtained eigenvalues are the wavenumbers, and the eigenvectors are the associated vibrational modeshapes of the rail cross-section. The rail considered is a typical 115-lb A.R.E.M.A. section, modeled as an isotropic material with hysteretic damping, and having the following properties: $\rho=7932$ kg/m³, $c_L = 5960$ m/s, $c_T = 3260$ m/s, $\kappa_L = 0.003$ Np/wavelength and $\kappa_T = 0.008$

$N_p/\text{wavelength}$. The rail cross-section has a complex geometry with one vertical axis of symmetry. The mesh, shown in Figure 3.5 and generated by the FEM software package ABAQUS™, used 351 nodes for 632 triangular elements with linear interpolation displacement functions. The rule of 5 or more nodes per wavelength at the maximum frequency, has been followed along the cross the section. It should be noted that only the rail head and the upper part of the web have been modeled; the mesh is also finer in the rail head, while is coarser towards the bottom of the web.

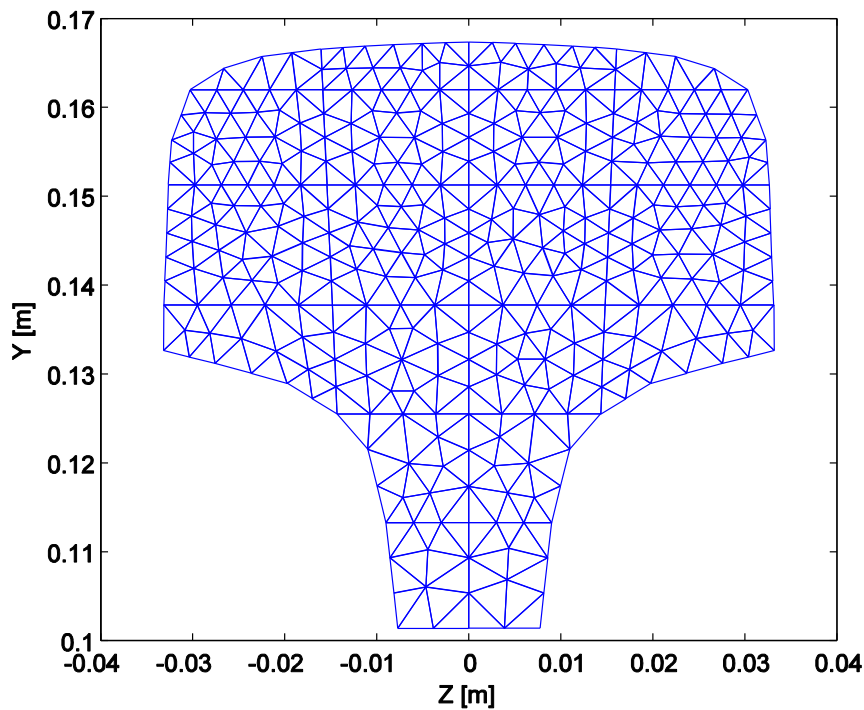


Figure 3.5: Mesh of the rail cross-section generated by ABAQUS™

This is under the assumption that the loading considered in this context will induce low or no energy in the foot. The results are shown for a frequency range up to 50 kHz. Wavenumbers as function of the frequency are depicted in Figure 3.6, while phase velocity dispersion curves are plotted in Figure 3.7. Looking at the phase velocity

curves, it can be deduced that the maximum number of eigenvalues at the frequency of 50 kHz is 11, not accounting for the modes with negative phase velocity, which are not reported in the graph.

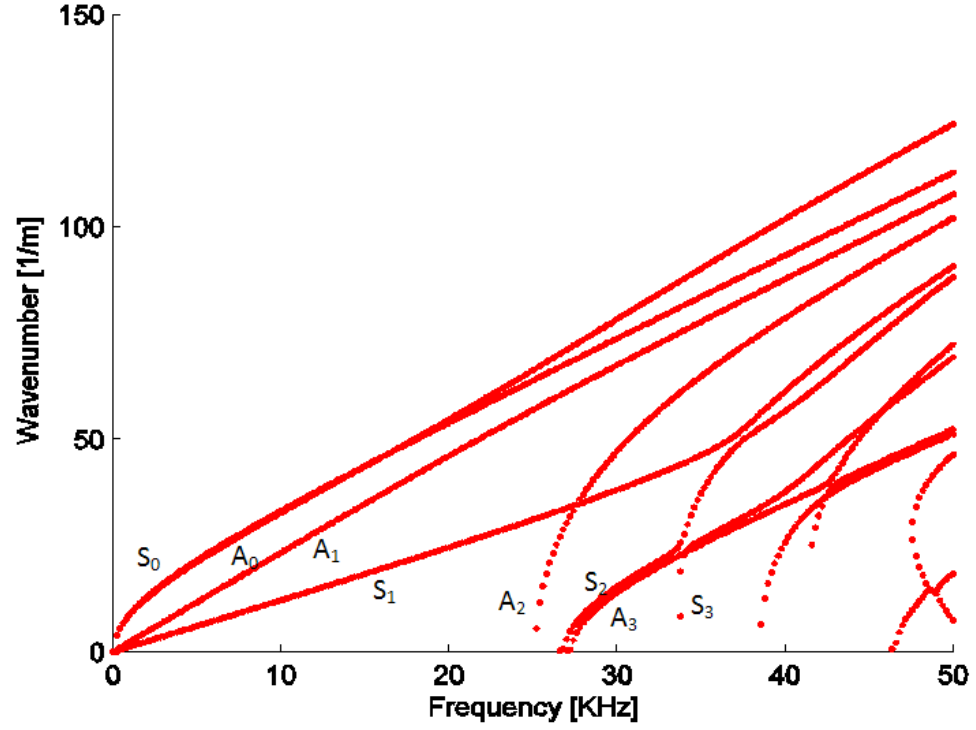


Figure 3.6: Wavenumbers as a function of frequency

The number of branches present in the wavenumber plots are instead 12; this can be explained by the fact that plotting only the modes with phase velocities smaller than 10,000 m/s is equivalent to filtering out branches that are located below the separation line shown in Figure 3.8, whose equation is:

$$\xi = \frac{\xi_c}{\omega} \omega \quad (3.67)$$

where $\bar{\omega} = 2\pi 50,000$ and $\bar{\xi} = \frac{\bar{\omega}}{10,000} = 10\pi \approx 31.4 \left[\frac{1}{m} \right]$

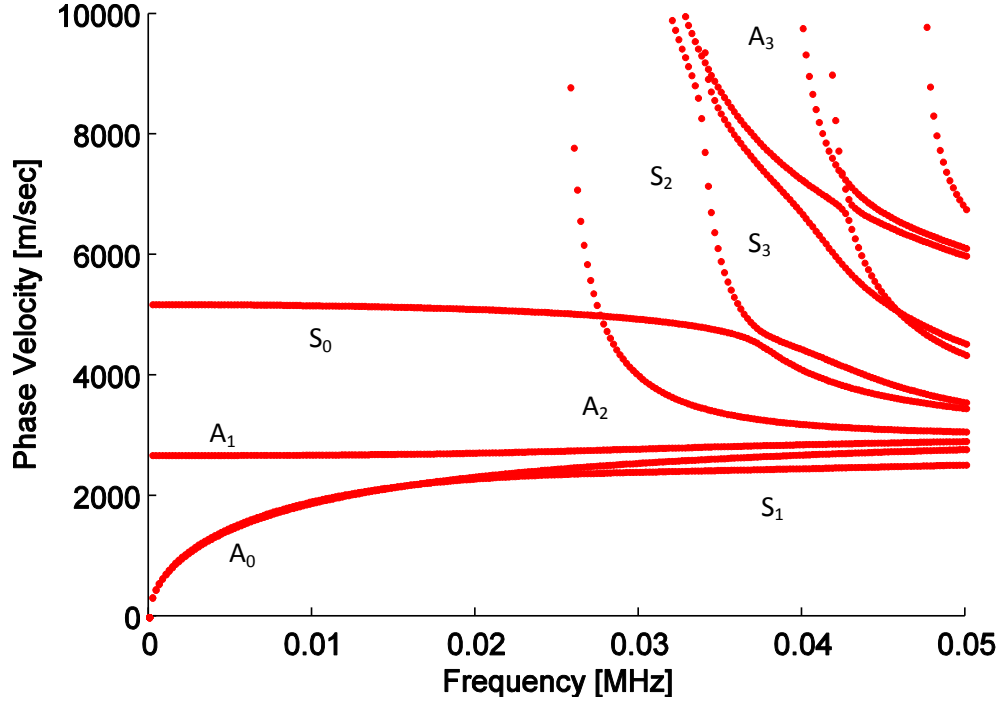


Figure 3.7: Phase velocity dispersion curves

The lowest order modes have been named and labeled according to their symmetry (S_n) or antisymmetry (A_n) in respect to the y axis. In this case one branch (mode) is located below the separation line, and that is the one excluded from the phase velocity plot. The plots of the energy velocities (Poynting vector) and attenuation are also shown in Figure 3.9 and Figure 3.10. The plot of the energy velocities shows that modes S_0 , S_1 , A_0 and A_1 have small dispersion on the frequency range considered. This makes these modes good candidates for defect probing, since they carry energy traveling with almost constant velocity.

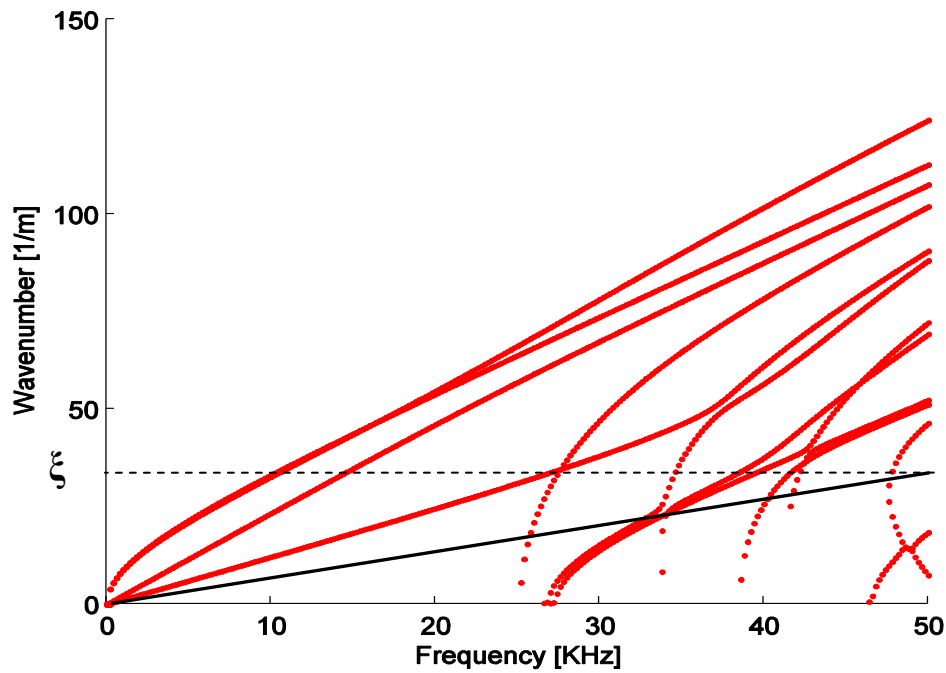


Figure 3.8: Mode filtering by thresholding the phase velocity

Later in the chapter it will be shown that these modes stay nondispersive up higher frequencies. Higher order modes velocity curves become stable at higher values of the spectrum. From the attenuation curves it is evident that the lowest order symmetric and antisymmetrical modes S_0 , S_1 , A_0 and A_1 , show low attenuation along the spectrum. At 200 kHz, the eight lowest order modeshapes have been obtained as eigenvectors of the system in eq. (3.41) and are shown in Figure 3.11 through Figure 3.18. Next section will present the forced solution of the rail subjected to an impulsive excitation.

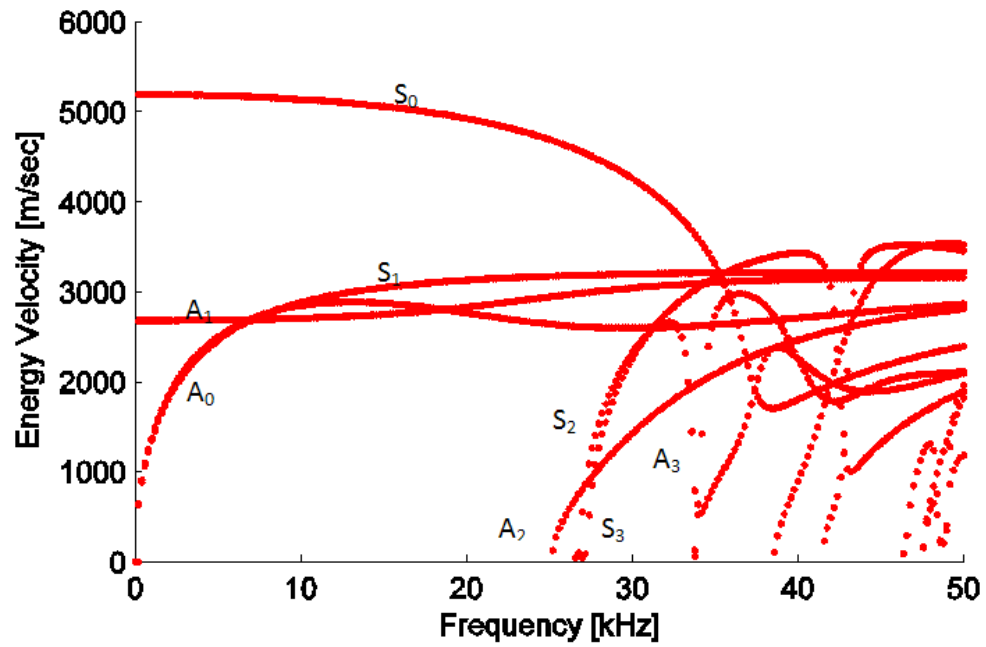


Figure 3.9: Energy Velocity (Poynting vector) dispersion curves

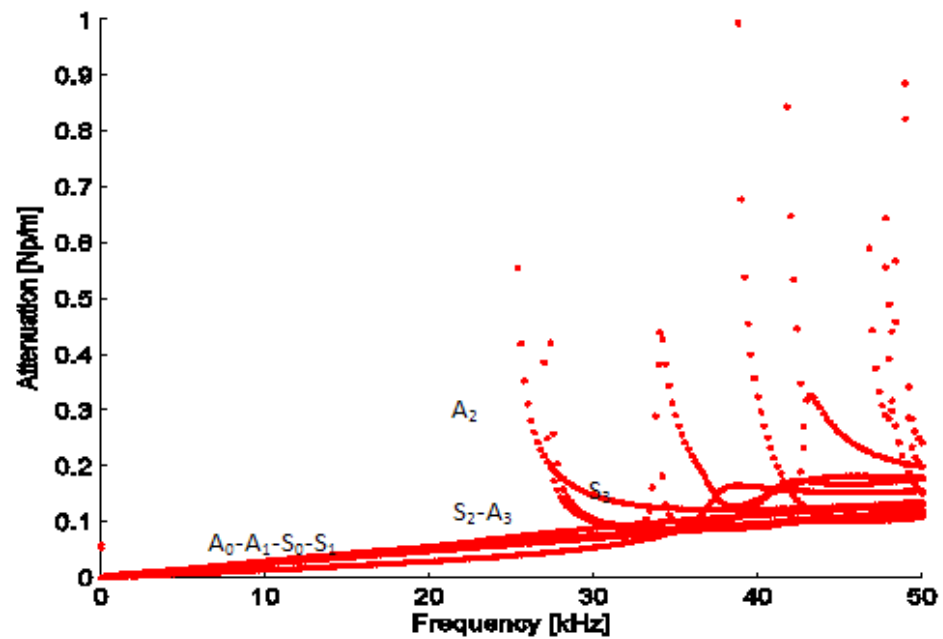


Figure 3.10: Imaginary part of the wavenumber (mode attenuation)

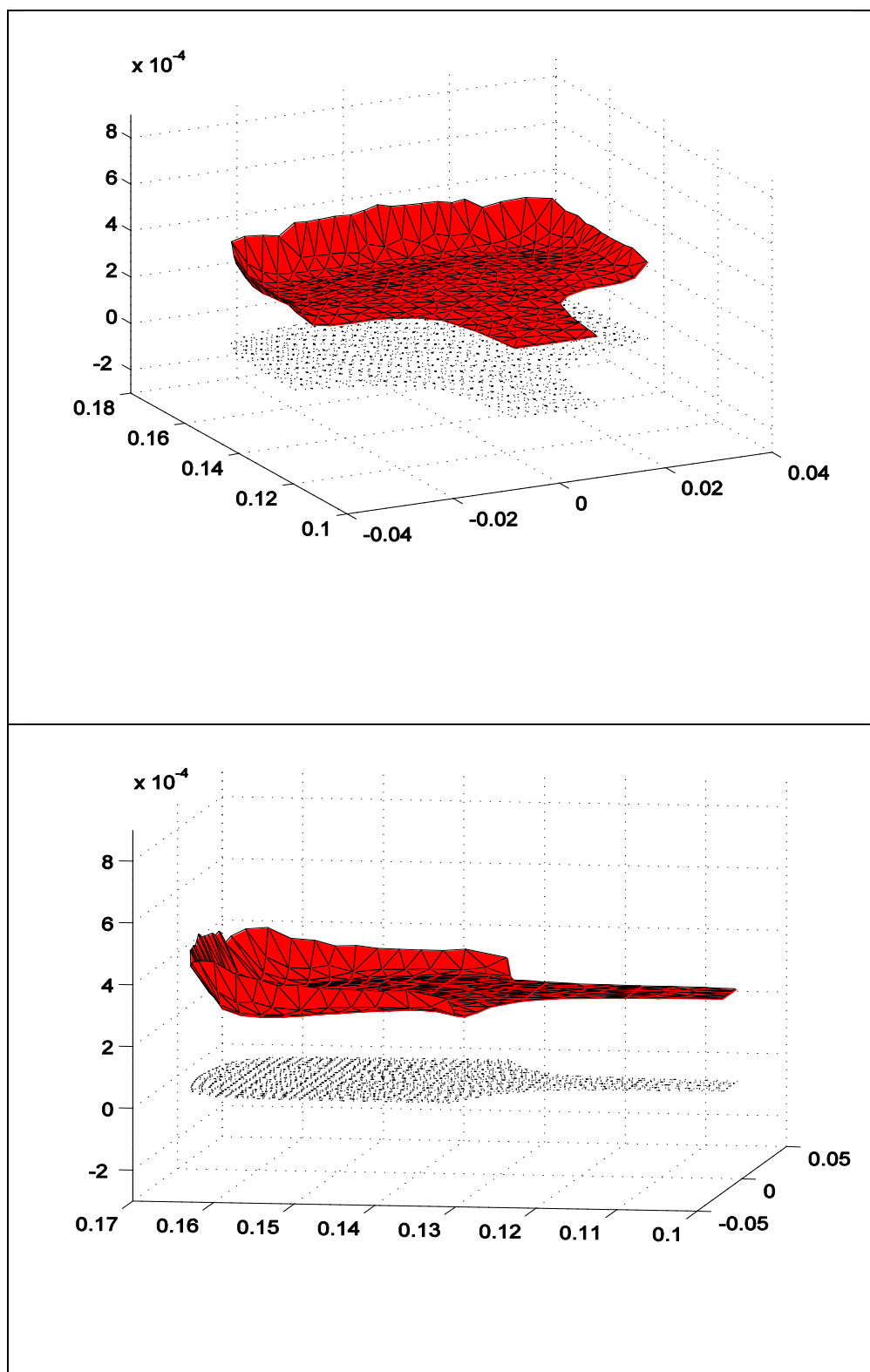


Figure 3.11: Modeshape of S_0 at 200 kHz

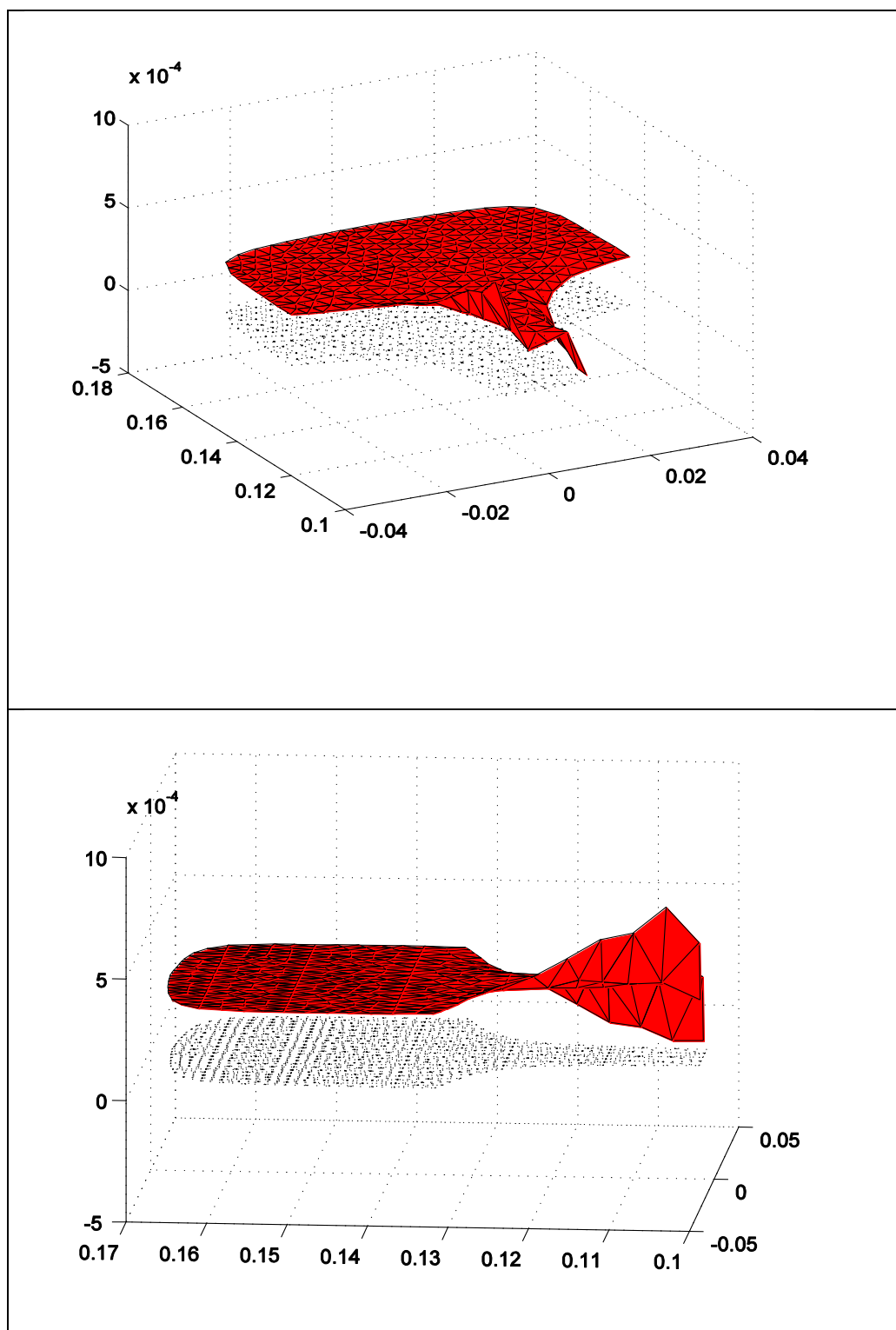


Figure 3.12: Modeshape of A_0 at 200 kHz

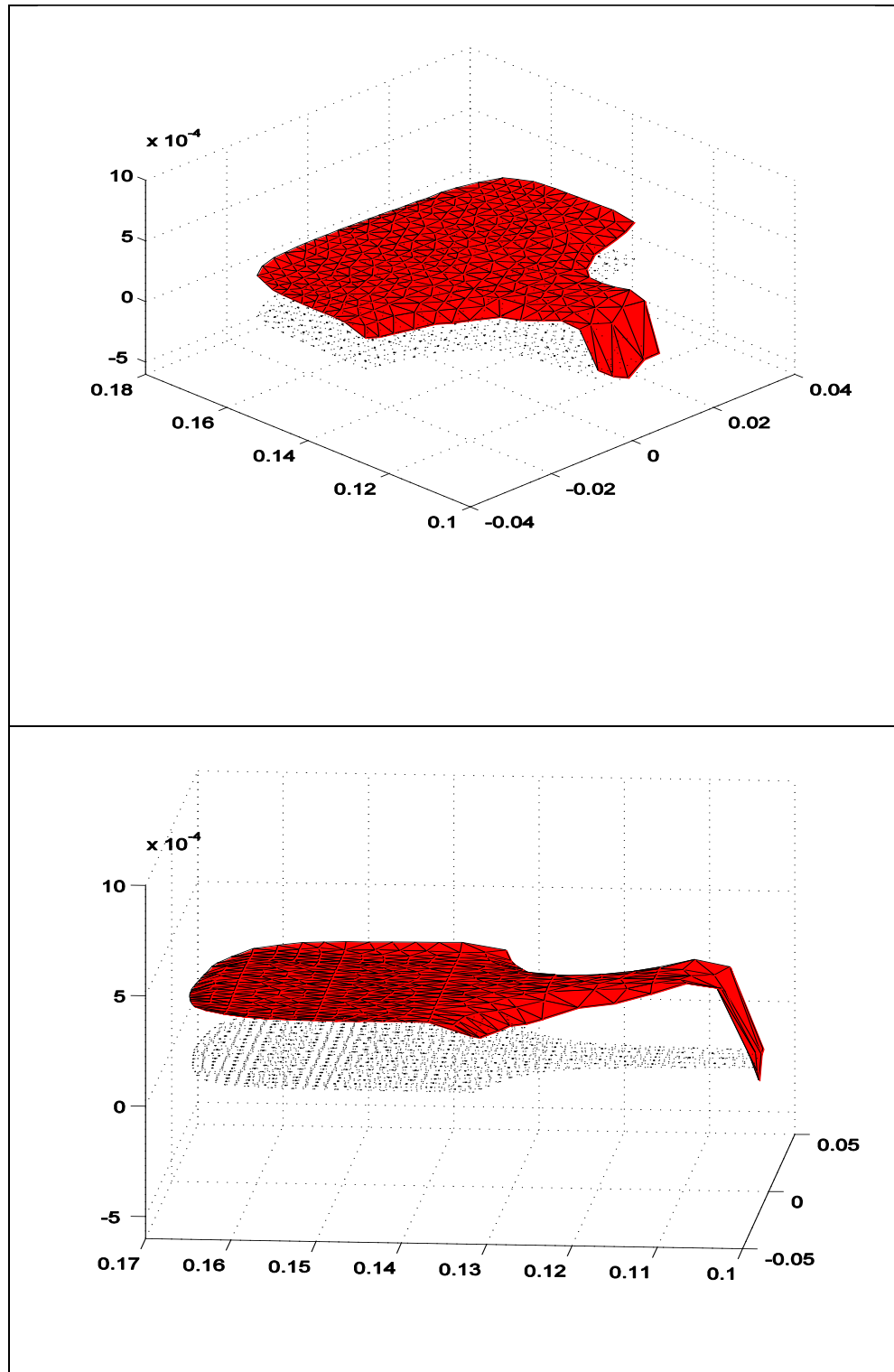


Figure 3.13: Modeshape of S_1 at 200 kHz

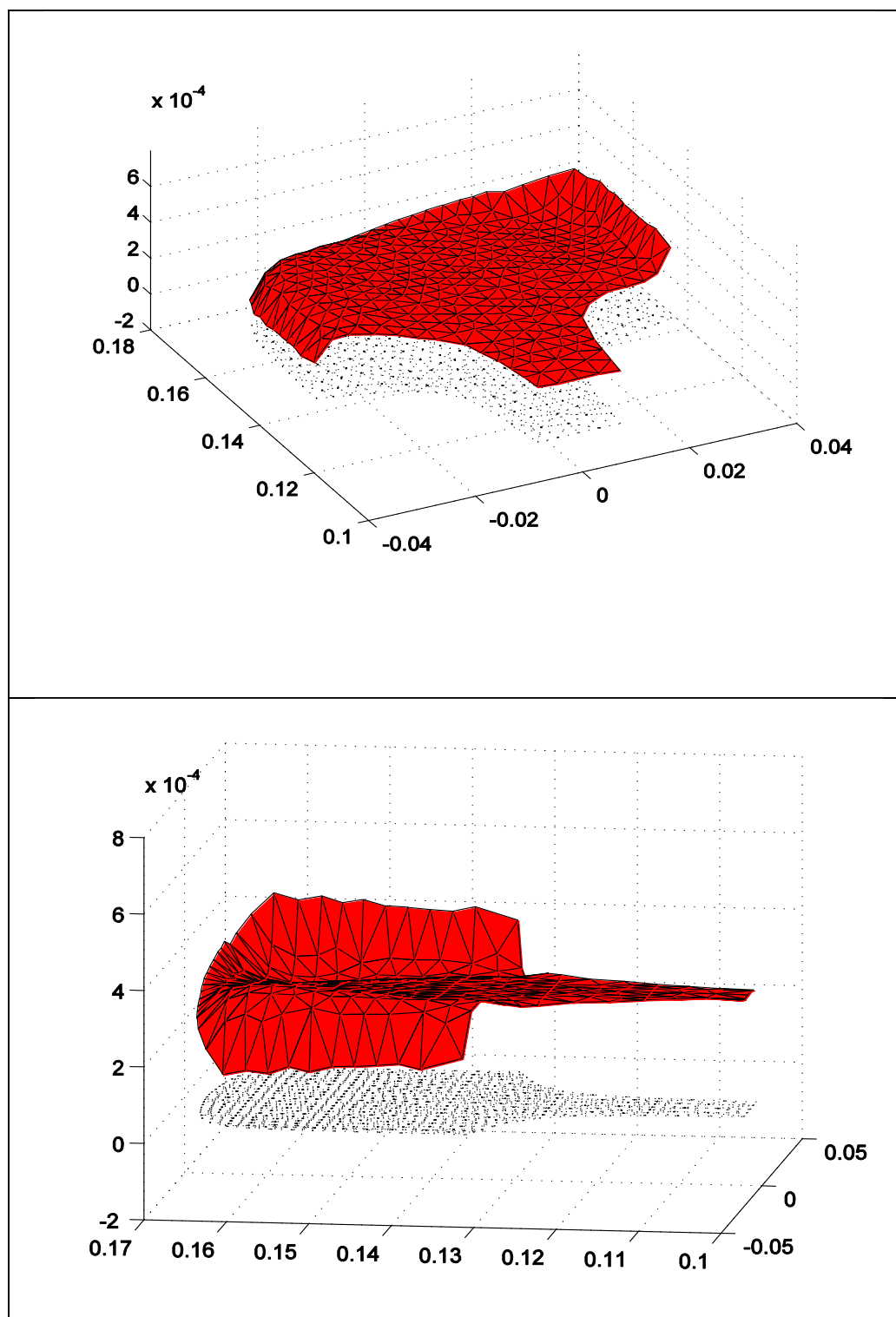


Figure 3.14: Modeshape of A_1 at 200 kHz

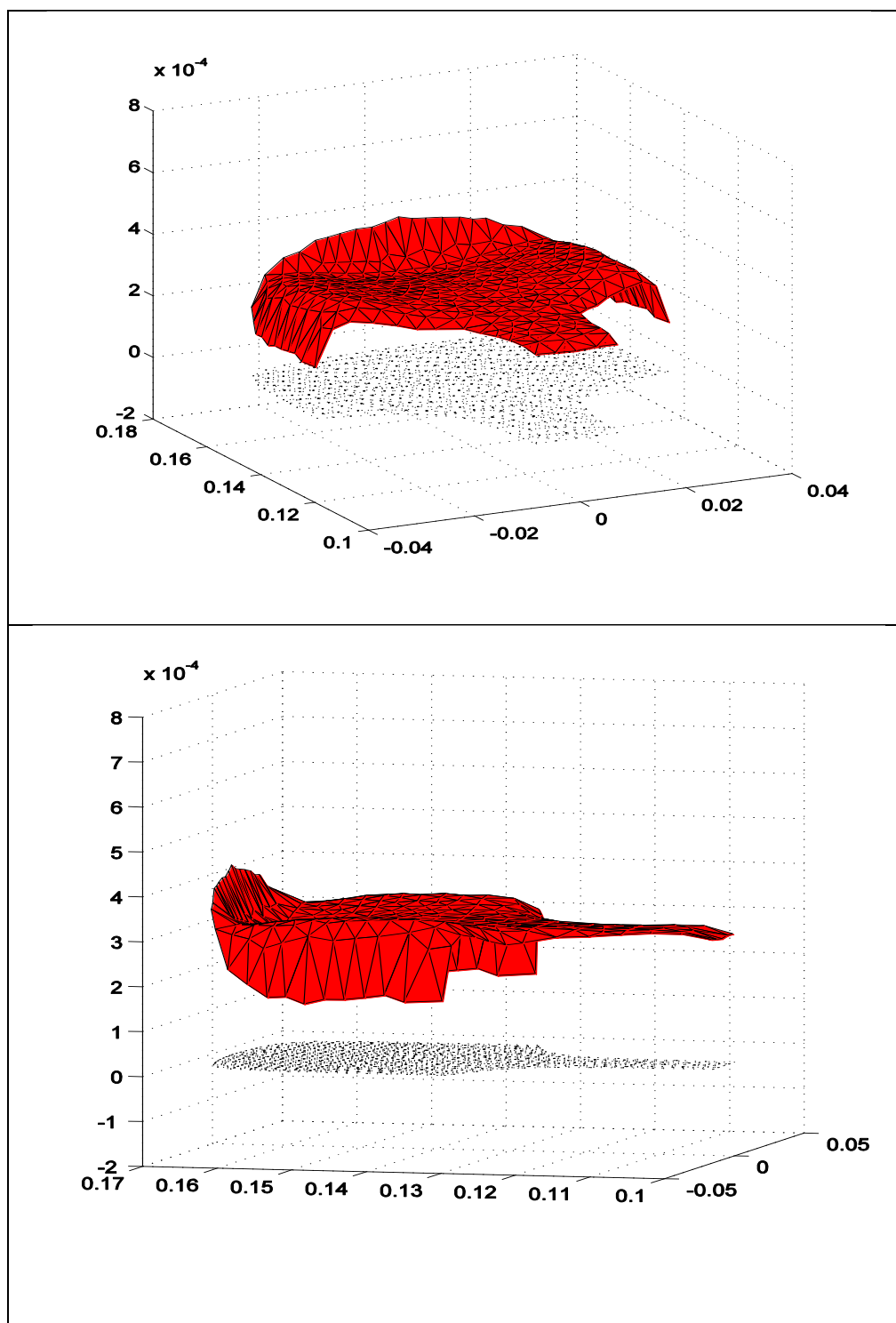


Figure 3.15: Modeshape of S_2 at 200 kHz

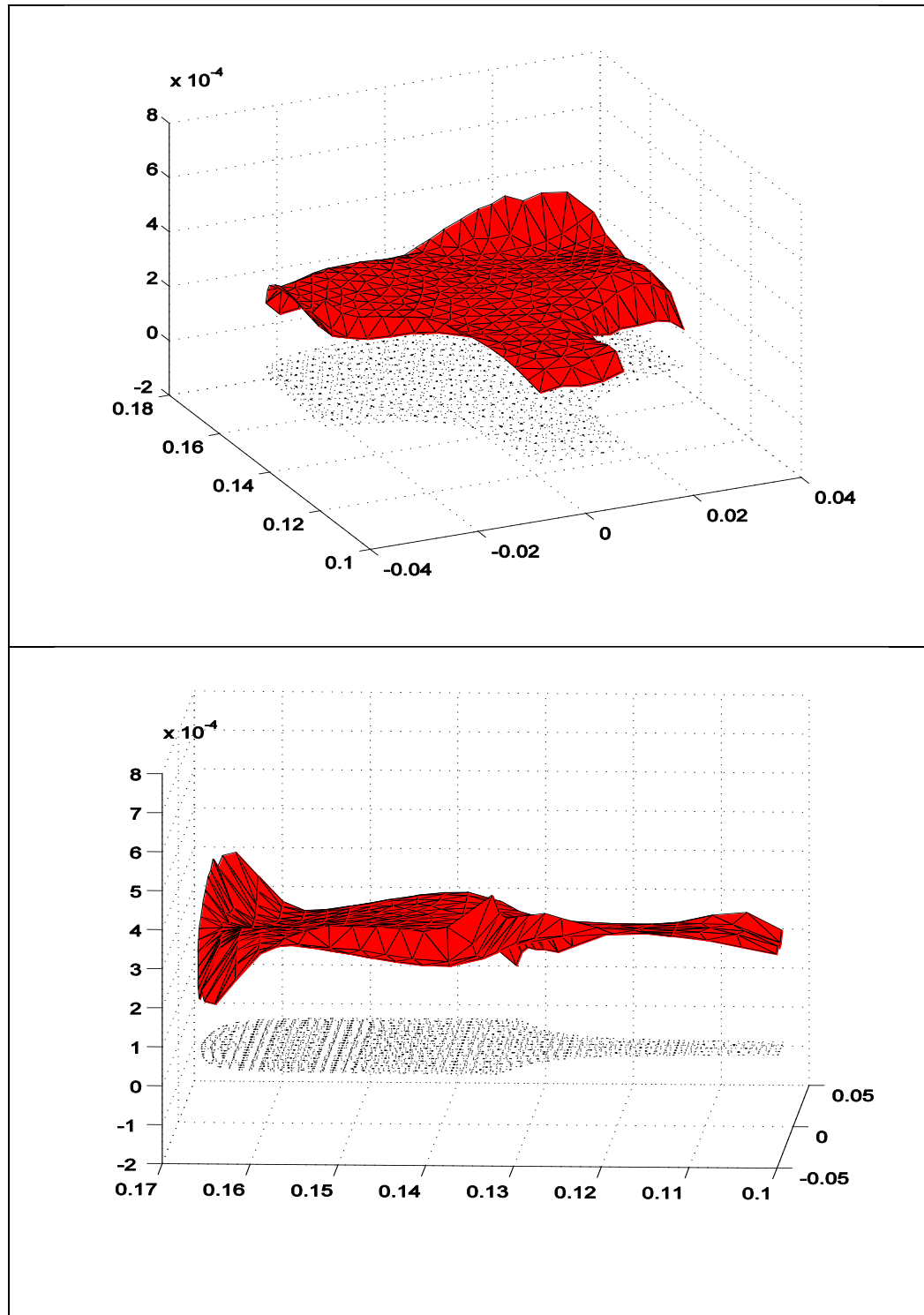


Figure 3.16: Modeshape of A_2 at 200 kHz

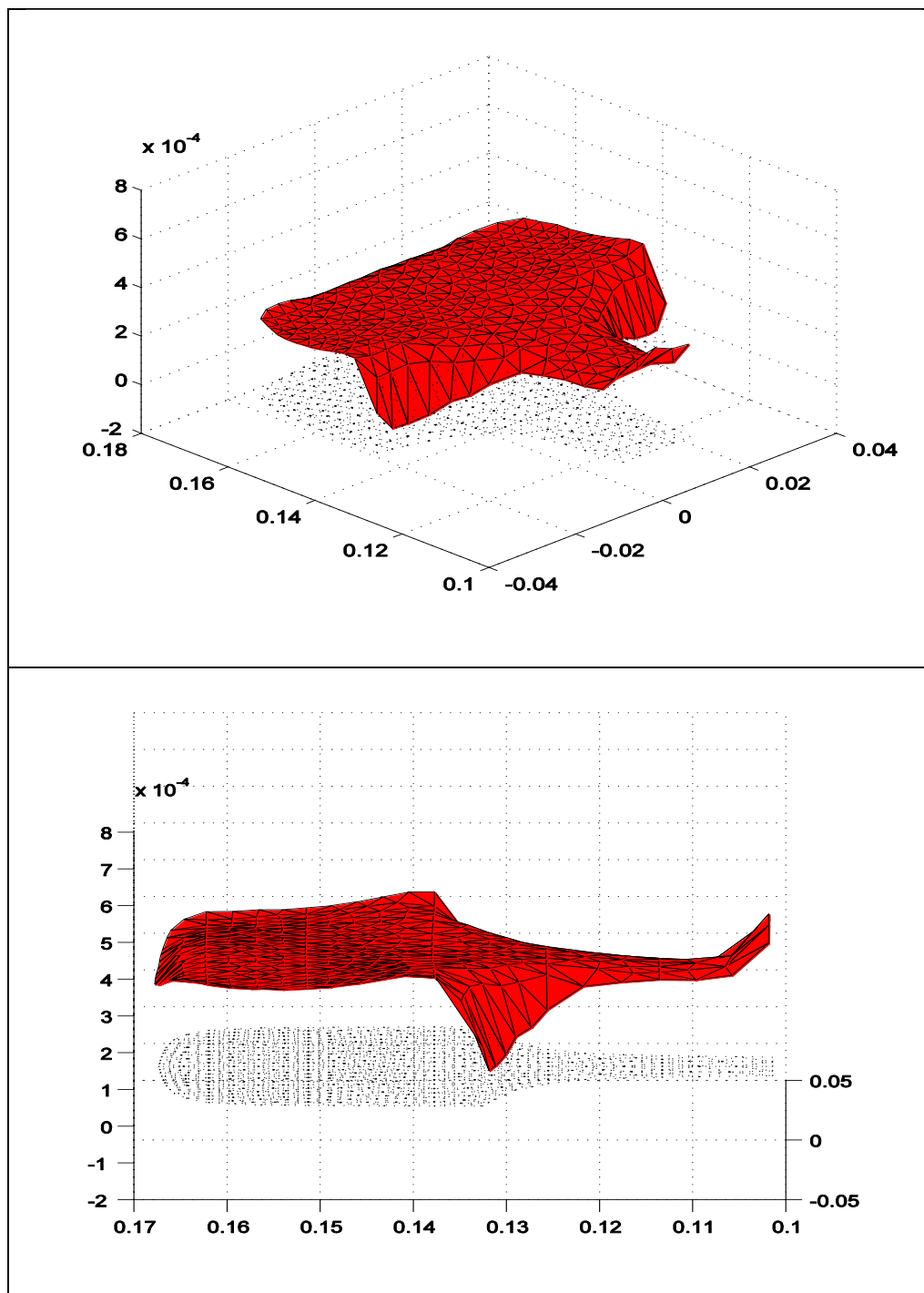


Figure 3.17: Modeshape of S_3 at 200 kHz

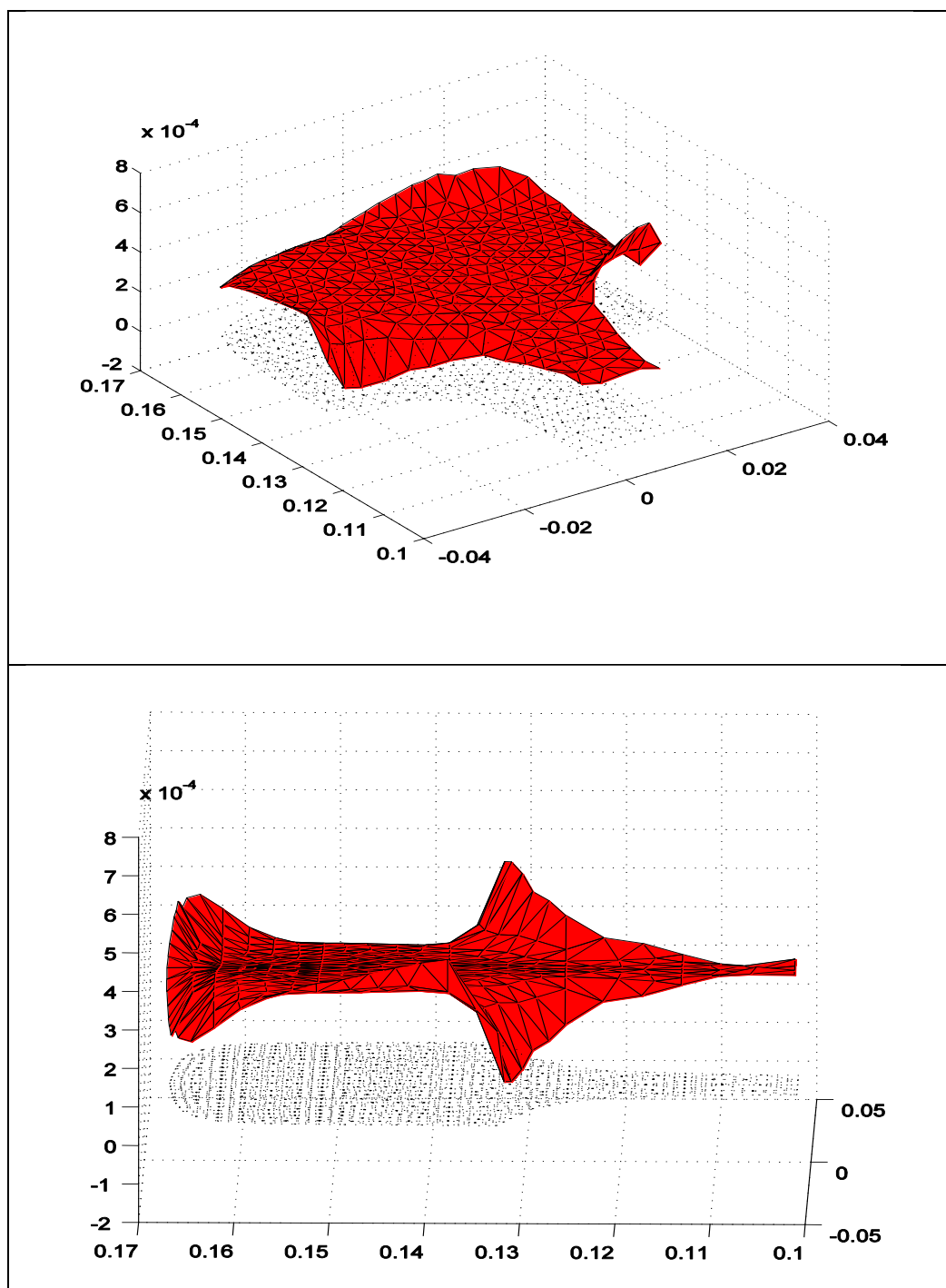


Figure 3.18: Modeshape of A_3 at 200 kHz

3.6 Forced solution for guided waves in rails

This section will discuss the response of the rail to an arbitrary load acting on the section at $x = x_s$. The contribution of the potential of the harmonic external load has to be now taken in account in the Hamilton's principle. The expression of such potential associated to the external traction \mathbf{t} is equal to:

$$\int_{\Omega} \delta \mathbf{u}^T \mathbf{t} d\Omega \quad (3.68)$$

where the traction \mathbf{t} acting on the domain Ω is expressed as:

$$\mathbf{t} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T \quad (3.69)$$

The new contribution of potential has now to be added in eq. (3.27), and the following governing equations are then obtained:

$$\int_{\Omega} \delta \mathbf{u}^T \mathbf{t} d\Omega = \int_V \delta \boldsymbol{\varepsilon}^T \tilde{\mathbf{C}} \boldsymbol{\varepsilon} dV + \int_V \delta \mathbf{u}^T \rho \ddot{\mathbf{u}} dV \quad (3.70)$$

The same notation used in the previous section is here maintained for displacement \mathbf{u} , stress $\boldsymbol{\sigma}$ and strain $\boldsymbol{\varepsilon}$ expressions. The displacement vector at an arbitrary point of the e -th element can be still written as a function of the nodal displacements $\mathbf{q}^{(e)}$

$$\mathbf{u}(x, y, z, t) = \mathbf{N}(y, z) \mathbf{q}^{(e)} e^{-i\omega t} \quad (3.71)$$

The nodal displacements can be rewritten in terms of their Fourier transform expressions

$\bar{\mathbf{q}}^{(e)}$, calculated with respect to the wavenumber [77]:

$$\mathbf{q}^{(e)} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \bar{\mathbf{q}}^{(e)} e^{i\xi x} d\xi \quad (3.72)$$

The displacement vector at an arbitrary point in the e -th element, considering only a certain wavenumber ξ is

$$\mathbf{u} = \frac{1}{2\pi} \mathbf{N}(y, z) \bar{\mathbf{q}}^{(e)} e^{i(\xi x - \omega t)}. \quad (3.73)$$

In the e -th element, the external traction vector can be expressed by the interpolation function:

$$\mathbf{t}(x, y, z, t) = \mathbf{N}(y, z) \mathbf{T}^{(e)} e^{-i\omega t} \quad (3.74)$$

Hence, \mathbf{t} can be expressed by

$$\mathbf{t} = \frac{1}{2\pi} \mathbf{N}(y, z) \bar{\mathbf{T}}^{(e)} e^{i(\xi x - \omega t)} \quad (3.75)$$

being $\bar{\mathbf{T}}^{(e)}$ the Fourier transform of $\mathbf{T}^{(e)}$ with respect to the wavenumber.

Substituting the displacement and the traction vectors, along with the stress and strain expressions and rearranging, this gives:

$$\bar{\mathbf{f}}^{(e)} = \left[\mathbf{K}_1^{(e)} + i\xi \mathbf{K}_2^{(e)} + \xi^2 \mathbf{K}_3^{(e)} - \omega^2 \mathbf{M}^{(e)} \right]_M \bar{\mathbf{U}} \quad (3.76)$$

where

$$\bar{\mathbf{f}}^{(e)} = \int_{\Omega'} \mathbf{N}^T \mathbf{N} \bar{\mathbf{T}}^{(e)} d\Omega' \quad (3.77)$$

and $\bar{\mathbf{U}}$ is the special Fourier transform of the global vector of unknown nodal displacements. Applying standard assembling procedures the following final expression can be obtained:

$$\bar{\mathbf{f}} = \left[\mathbf{K}_1 + i\xi \mathbf{K}_2 + \xi^2 \mathbf{K}_3 - \omega^2 \mathbf{M} \right]_M \bar{\mathbf{U}} \quad (3.78)$$

This system can be reordered in the same fashion of eq. (3.41) obtaining the following eigensystem with a first order wavenumber ξ :

$$\mathbf{A} - \xi \mathbf{B}_{2M} \mathbf{Q} = \bar{\mathbf{p}} \quad (3.79)$$

where:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{K}_1 - \omega^2 \mathbf{M} \\ \mathbf{K}_1 - \omega^2 \mathbf{M} & \hat{\mathbf{K}}_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{K}_1 - \omega^2 \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_3 \end{bmatrix}, \quad (3.80)$$

$$\mathbf{Q} = \begin{bmatrix} \bar{\mathbf{U}} \\ \xi \bar{\mathbf{U}} \end{bmatrix}, \quad \bar{\mathbf{p}} = \begin{bmatrix} 0 \\ \bar{\mathbf{f}} \end{bmatrix}$$

The solution of eq. (3.79) can be expressed as a linear combination of the right eigenvector $\Phi_m^{\mathbf{R}}$:

$$\mathbf{Q} = \sum_{m=1}^{2M} Q_m \Phi_m^{\mathbf{R}} \quad (3.81)$$

The coefficients Q_m can be calculated substituting last expression in the eigensystem equation, obtaining:

$$(\mathbf{A} - \xi \mathbf{B}) \sum_{m=1}^{2M} Q_m \mathbf{\Phi}_m^{\mathbf{R}} = \bar{\mathbf{p}} \quad (3.82)$$

Pre-multiplying the last equation by the left eigenvector $\mathbf{\Phi}_n^{\mathbf{L}}$

$$\mathbf{\Phi}_n^{\mathbf{L}} (\mathbf{A} - \xi \mathbf{B}) \sum_{m=1}^{2M} Q_m \mathbf{\Phi}_m^{\mathbf{R}} = \mathbf{\Phi}_n^{\mathbf{L}} \bar{\mathbf{p}} \quad (3.83)$$

and applying the bi-orthogonality conditions:

$$\mathbf{\Phi}_n^{\mathbf{L}} \mathbf{A} \mathbf{\Phi}_m^{\mathbf{R}} = \begin{cases} 0 & n \neq m \\ \mathbf{\Phi}_m^{\mathbf{L}} \mathbf{A} \mathbf{\Phi}_m^{\mathbf{R}} = \xi \mathbf{\Phi}_m^{\mathbf{L}} \mathbf{B} \mathbf{\Phi}_m^{\mathbf{R}} & n = m \end{cases} \quad (3.84)$$

$$\mathbf{\Phi}_n^{\mathbf{L}} - \xi \mathbf{B} \mathbf{\Phi}_m^{\mathbf{R}} = \begin{cases} 0 & n \neq m \\ -\xi \mathbf{\Phi}_m^{\mathbf{L}} \mathbf{B} \mathbf{\Phi}_m^{\mathbf{R}} & n = m \end{cases}$$

the following expression is obtained

$$\mathbf{\Phi}_m^{\mathbf{L}} (\mathbf{A} - \xi \mathbf{B}) Q_m \mathbf{\Phi}_m^{\mathbf{R}} = \mathbf{\Phi}_m^{\mathbf{L}} \bar{\mathbf{p}} \quad (3.85)$$

Substituting eq. (3.84) in the last expression, the sought coefficients are obtained as:

$$Q_m = - \frac{\mathbf{\Phi}_m^{\mathbf{L}} \bar{\mathbf{p}}}{(\xi - \xi_m) B_m} \quad (3.86)$$

where

$$B_m = \Phi_m^L \mathbf{B} \Phi_m^R. \quad (3.87)$$

The nodal displacement $M \times 1$ vector $\bar{\mathbf{U}}$ can be obtained by the following expression:

$$\bar{\mathbf{U}} = \sum_{m=1}^{2M} -\frac{\Phi_m^L \bar{\mathbf{p}}}{(\xi - \xi_m) B_m} \Phi_m^{\text{Rup}} \quad (3.88)$$

where Φ_m^{Rup} is the upper part of the m -th right eigenvector Φ_m^R . If point loading on the surface at $x = x_s$ is considered, the forces can be represented by the Dirac delta expression $\delta = \delta(x - x_s)$.

Being $\bar{\mathbf{p}}$ the Fourier transform with respect to ξ of the point load, it can be expressed in the following way:

$$\bar{\mathbf{p}} = \int_{-\infty}^{+\infty} \hat{\mathbf{p}} \delta(x - x_s) e^{-i\xi x} dx = \hat{\mathbf{p}} e^{-i\xi x_s} \quad (3.89)$$

where $\hat{\mathbf{p}}$ contains the spectral amplitudes of the nodal loads applied in $x = x_s$.

To obtain the time domain displacement vector \mathbf{U} , the inverse Fourier transform of eq. (3.88) must be computed. Substituting the expression of $\bar{\mathbf{p}}$ in the inverse Fourier transform of $\bar{\mathbf{U}}$, the final expression for the displacement vector \mathbf{U} is

$$\mathbf{U} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \bar{\mathbf{U}} e^{i\xi x} d\xi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{m=1}^{2M} -\frac{\Phi_m^L \hat{\mathbf{p}}}{(\xi - \xi_m) B_m} \Phi_m^{\text{Rup}} e^{-i\xi x_s} e^{i\xi x} d\xi \quad (3.90)$$

Computing the integral in eq. (3.90) by using the residue theorem, the following expression is obtained:

$$\mathbf{U} = \frac{1}{2\pi} \sum_{m=1}^{2M} -\frac{\Phi_m^L \hat{\mathbf{p}}}{B_m} \Phi_m^{\text{Rup}} e^{i[\xi_m(x-x_s)]} \quad (3.91)$$

It should be cited that except for the $\frac{1}{2\pi}$ term, the expression of this equation was first proposed by Hayashi et al. in [77]. For $x > x_s$ the last equation can be written extending the sum only to the M forward propagating modes:

$$\mathbf{U} = \frac{1}{2\pi} \sum_{m=1}^M -\alpha_m \Phi_m^{\text{Rup}} e^{i[\xi_m(x-x_s)]} \quad (3.92)$$

where the modal participation factors are:

$$\alpha_m = -\frac{\Phi_m^L \hat{\mathbf{p}}}{B_m} \quad (3.93)$$

If $x < x_s$ the included modes needs to be replaced by the M backward propagating ones. It should be noted that the vector $\hat{\mathbf{p}} = \hat{\mathbf{p}}(x_s, y, z, \omega)$ contains the spectral amplitudes of the applied nodal loads, and the calculated vector $\mathbf{U} = \mathbf{U}(x, y, z, \omega)$ contains the displacement spectral amplitudes of the cross-sectional nodes at the coordinate x .

In order to obtain the time domain displacement of those nodes, an inverse Fourier Transform from the frequency to the time domain is needed; said transformation is expressed in the following equation:

$$\mathbf{U}(x, y, z, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathbf{U}(x, y, z, \omega) e^{i\omega t} d\omega \quad (3.94)$$

3.7 Response of the rail to a laser excitation

The forced response theory of the previous chapter has been applied to the 115-lb A.R.E.M.A. rail studied in section 3.5 for the unforced solution. The rail cross-section mesh was the same as that used for the unforced solution (Figure 3.5), 351 nodes for 632 triangular elements with linear interpolation displacement functions.

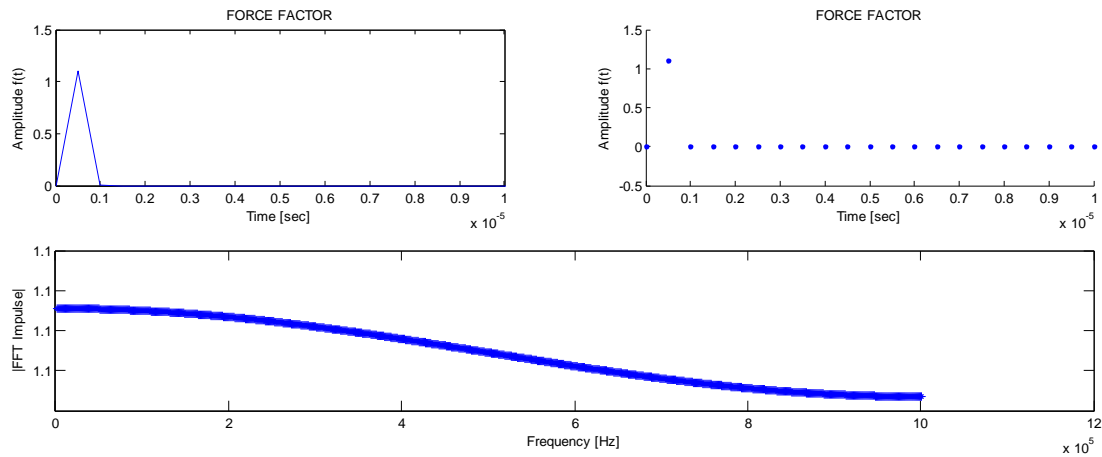


Figure 3.19: Time and frequency domain of the load function

The acting load was created to simulate the effect of a laser generator, able to provide a very short pulse; as depicted in Figure 3.19. The simulated 1 μ sec pulse is able

to provide energy up to the frequency of 1 MHz, which was consistent with the experimental data for the available laser.

In order to obtain a good approximation of the rail response to this type of excitation, the SAFE model should be run for the entire sub-spectral range up to 1 MHz, the maximum frequency at which the simulated excitation is still providing energy. Since the computational effort for the execution of the entire process over such a wide spectral range would be extremely demanding, some approximations were made. A first assumption is not accounting for the damping in the rail, which along with the introduction of the T operator, will result in a real and symmetric eigensystem to be solved at each frequency step. Another approximation is increasing the frequency step to 1 kHz and limiting the sub-spectral range to 500 kHz, beyond which, very low energy is generated. The plots of the wavenumbers, phase velocity and group velocity are shown in Figure 3.20 through Figure 3.22.

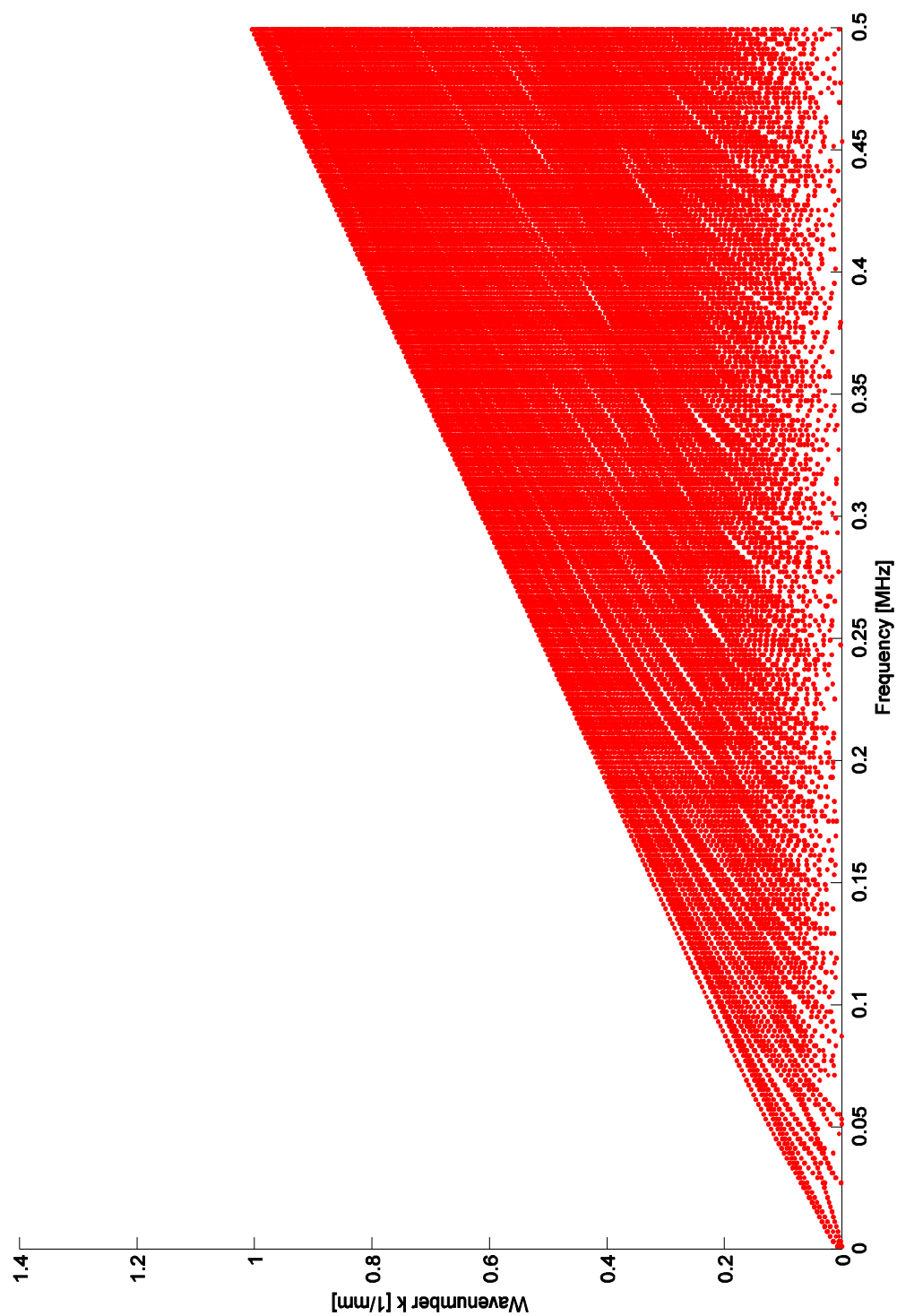


Figure 3.20: Wavenumbers as a function of frequency in undamped rail.

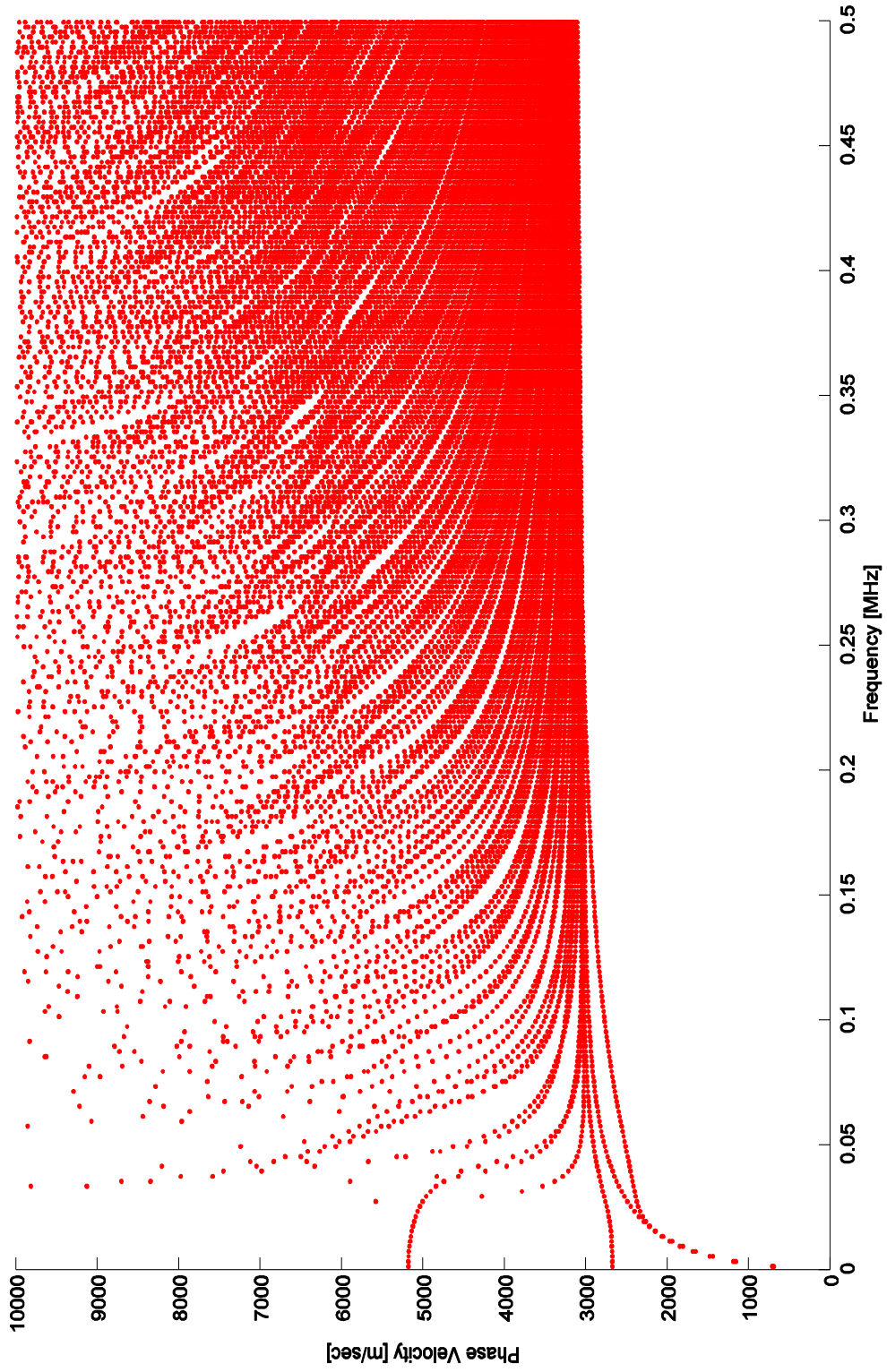


Figure 3.21: Phase velocity dispersion curves in undamped rail

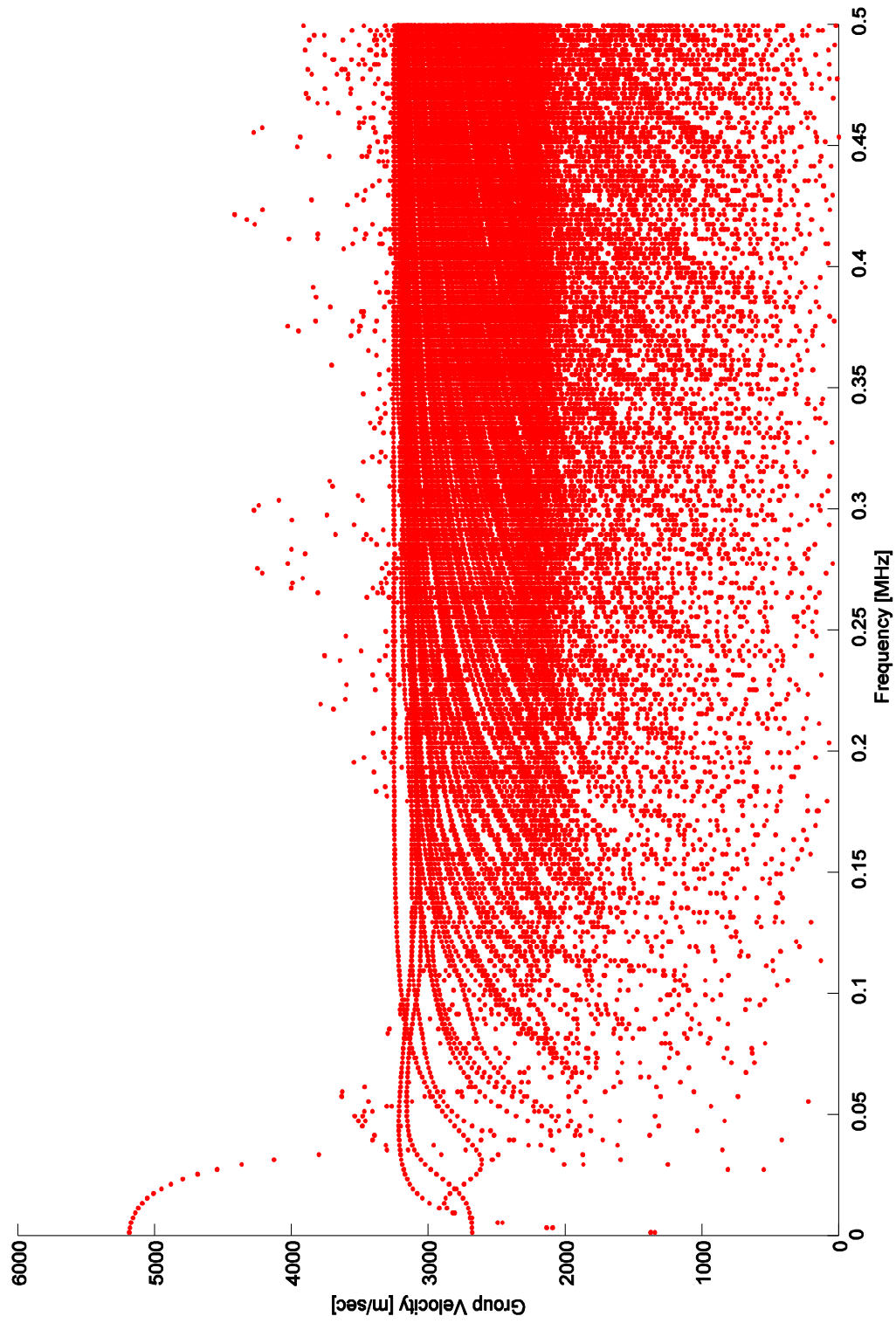


Figure 3.22: Group velocity dispersion curves in undamped rail

The plotted curves show the complexity of the propagating modes between 0 and 500 kHz. At the highest frequency, 152 forward propagating modes are found limiting the mode phase velocity to 10,000 m/s.

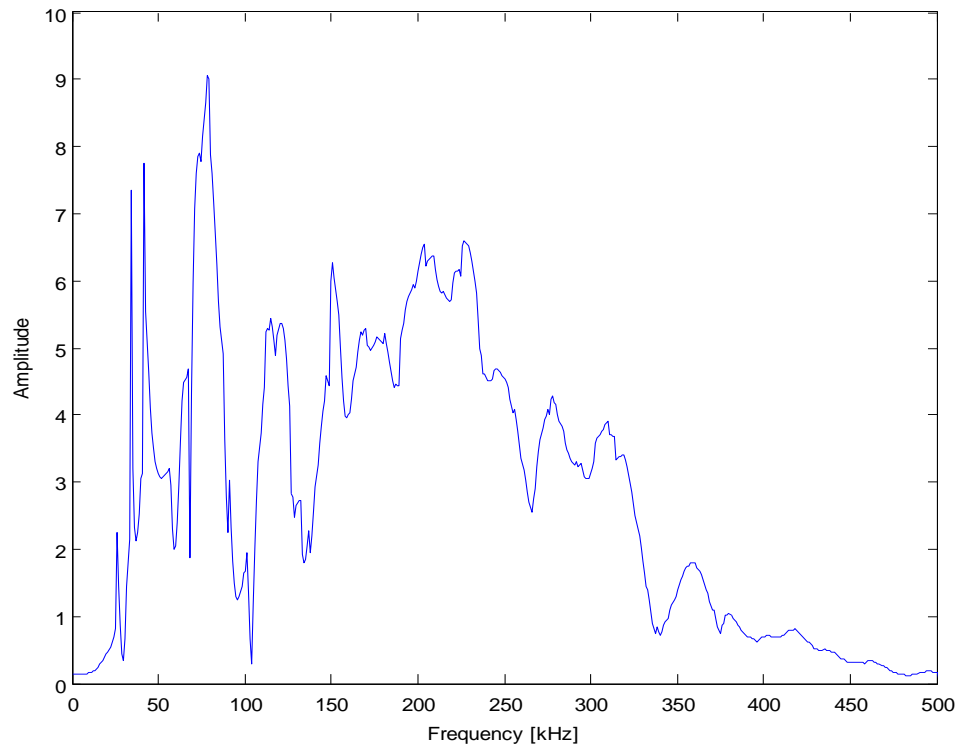


Figure 3.23: Frequency domain displacement of the central node of the top of the railhead subjected to laser excitation simulated at 4" (≈ 102 mm) away from the node

Including all the concurrent modes, the frequency and time domain nodal displacements of any section at arbitrary distance x along the propagation direction, can be obtained. The frequency domain displacement of the central node on the top of the rail head is shown in Figure 3.23. The analyzed node was on the cross-section located at $x_s = 4''$ (≈ 102 mm). The time domain displacement of the same node is shown in Figure 3.24.

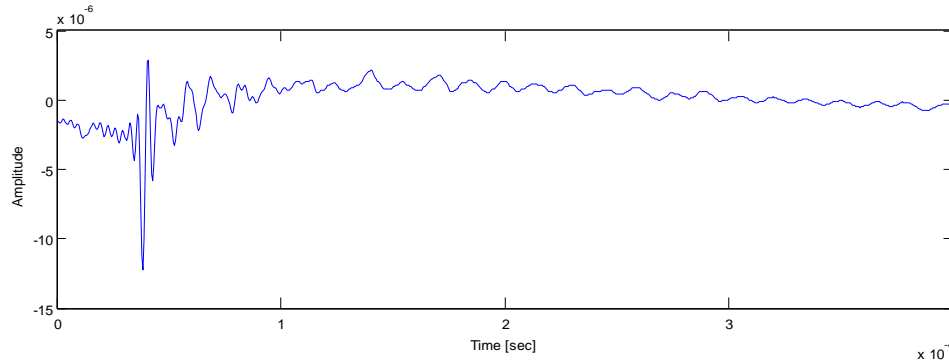


Figure 3.24: Time domain displacement of the central node of the top of the railhead subjected to laser excitation simulated at 4'' (≈ 102 mm) away from the node

A validation of the model and the overall procedure was performed by checking the match between the calculated response and experimental results. A pre-amplified air-coupled transducer was deployed at 4'' (≈ 102 mm) from the excitation source, which was provided by a pulsed laser beam in the shape of a full-head wide line; the air-coupled lift-off was about 1.5'' (≈ 38 mm). The transducer was deployed at an angle $\theta = 6.3^\circ$, following Snell's law of refraction (see detail in eq. (4.25)). The laser generator was a Q-Switched Big Sky Laser CFR-400 capable of 300 mJ pulse of ~ 10 ns duration, on 30 Hz max repetition rate. The frequency spectrum of the air-coupled transducer was very broad, ranging between 40 kHz to 2.25 MHz. A detailed description of the characteristics of the laser generator and transducer can be found on chapter five. The signal acquired by the air-coupled sensor is shown in Figure 3.25, while the raw and filtered simulated responses are shown in Figure 3.26. The delay in arrival time of the measured signal, with respect to the simulated one, is due to the path that the ultrasonic wave had to travel within the air, in order to reach the transducer, from the top of the rail head. In fact, being that distance equal to the 1.5'' (≈ 38 mm) transducer lift-off, and

assuming the speed of the sound in air equal to 340 m/sec, the delay corresponds to about 1 μ sec.

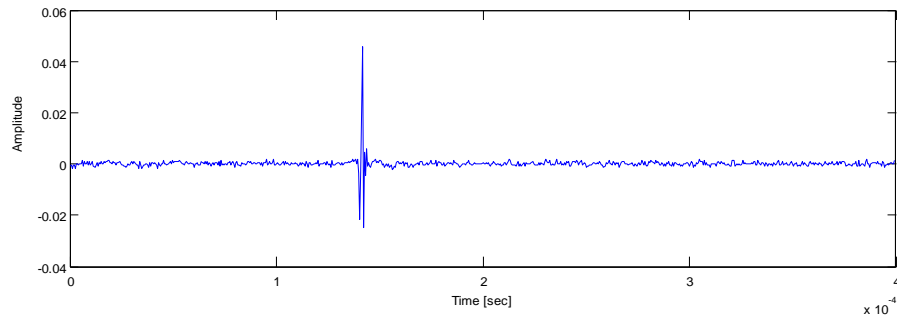


Figure 3.25: Laser-generated signal, acquired by the air-coupled sensor 4” away from the excitation source

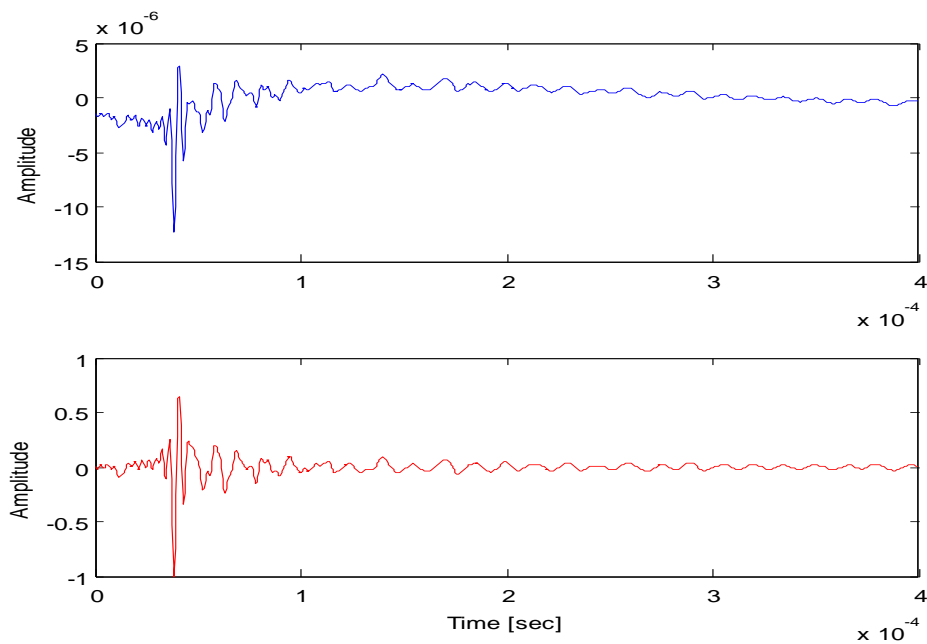


Figure 3.26: Raw simulated(red) and 40 kHz high-passed(blue) simulated response

Since the simulated response was calculated in the spectral range of 0 to 500 kHz and the experimental signal was high-passed at 40 kHz by the sensor pre-amplifier, a filtering process is needed to prepare the signals for a test of matching. A third order Butterworth

filter was used to perform the signal conditioning. Figure 3.27 contains the measured and the 500 kHz low-passed experimental signals.

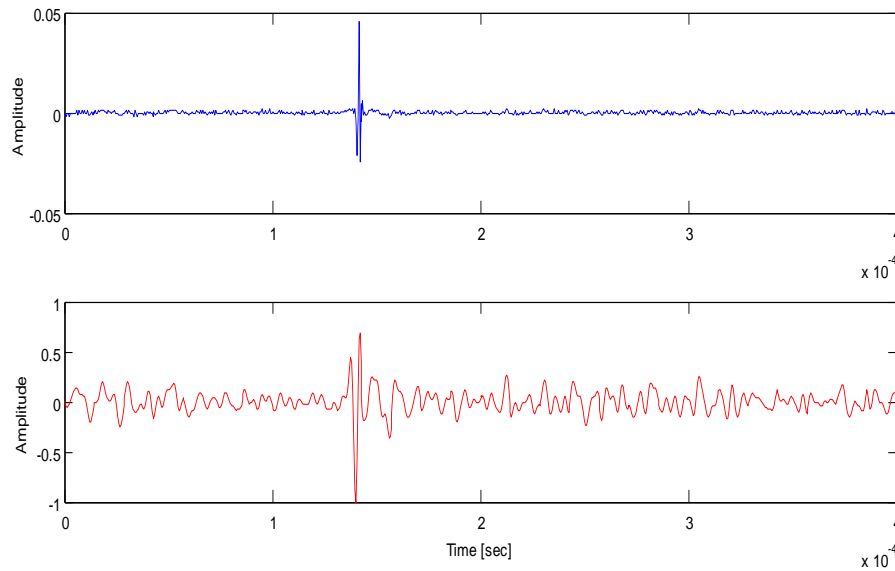


Figure 3.27: Raw and 500 kHz low-passed experimental signals

In both graphs, the signals in red, or filtered signals, have been normalized to their maximum absolute amplitude. The predicted and experimental filtered normalized signals are compared in Figure 3.28, after appropriately delaying the experimental response, in order to account for the time traveled in air. Clearly, the response obtained by the SAFE model shows a very good match with the experimental results; the first arrival is almost perfectly reproduced, while the scattered peaks which follow are caused by other phenomena such as multiple reflections along the rail. The mismatch of parts of the signals can also be due to the sensor vertical orientation, whose effect is band pass filtering the phase velocities at values close to the Rayleigh velocity. Figure 3.29 shows the comparison between the simulated and experimental responses in the frequency domain.

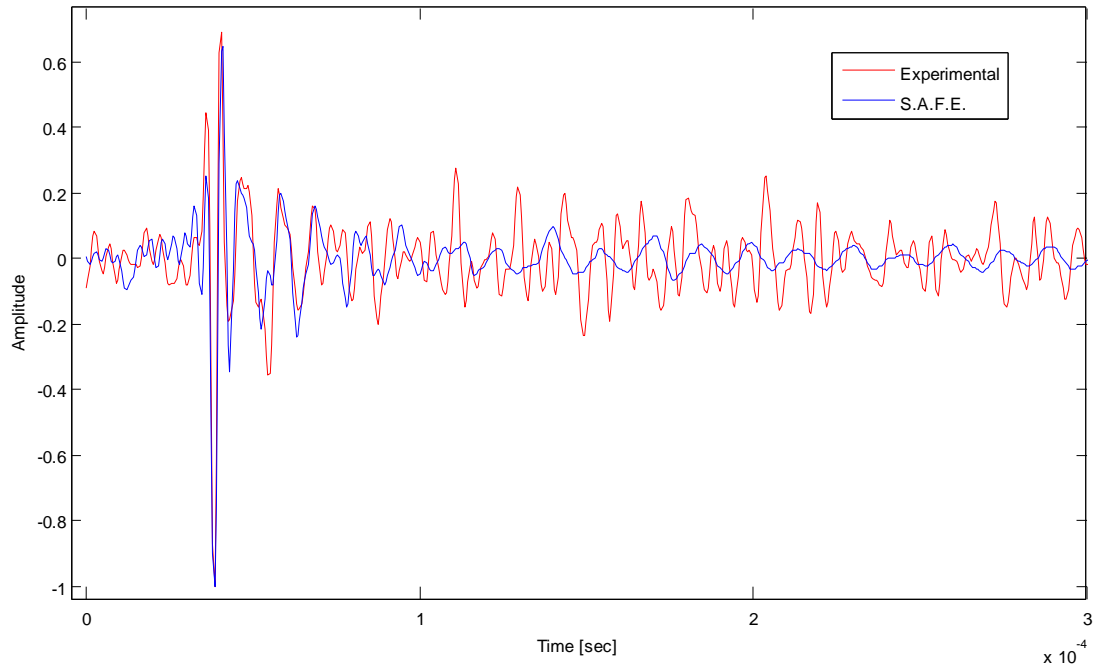


Figure 3.28: Comparison of the measured response with the results of the S.A.F.E. simulation

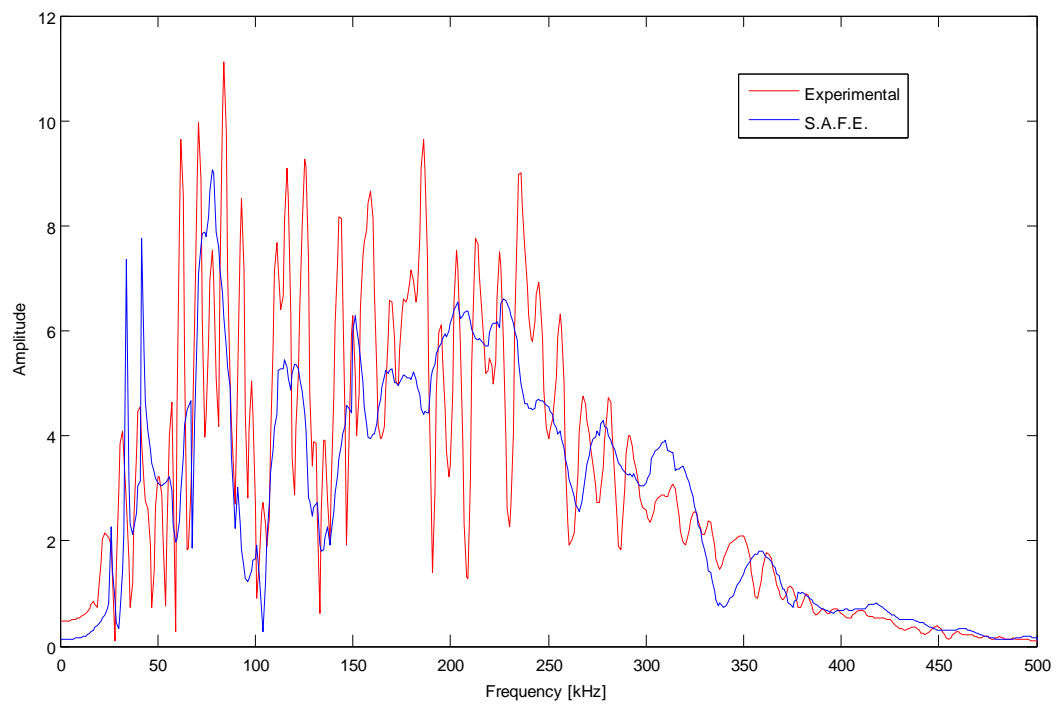


Figure 3.29: Simulated and experimental spectral responses

The frequency content of the simulated response is very similar to the experimental one; the peaks around 80 and 110 kHz are reproduced, along with the valleys at about 100 and 130 kHz. The simulated response contains two unrecognized peaks around 30 and 40 kHz, circled in Figure 3.29; that can be explained considering that those are the cut-off frequencies of two low order propagating modes [82]. In the undamped SAFE model, it is known that modes may have a resonance-type behavior in proximity of the cut-off frequency, showing an amount of energy which is erroneously larger than the real one. This should not be considered as a failure of the semianalytical model, but it is showing that the SAFE response is sensitive to the unreal modeling of the rail waveguide as undamped. There are a couple of ways to overcome this problem. The first possibility is to include damping in the model, which automatically reduces the amplification phenomena. Another method is to filter out the modes with large phase velocities, as previously mentioned in the chapter. This filtering is shown in Figure 3.30 where

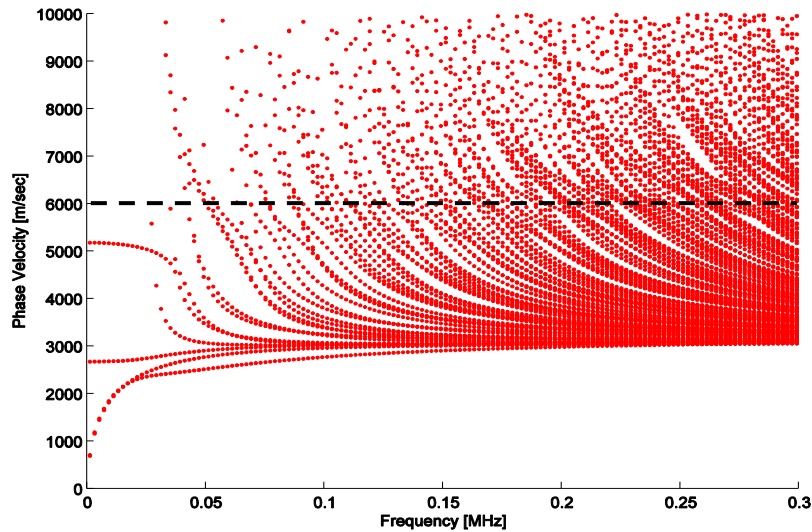


Figure 3.30: Phase velocity dispersion curves filtering.

solutions with phase velocity greater than 6,000 *m/sec* are discarded. Following a single mode backward up to its cut-off frequency, it is clear that the curve is approaching a vertical tangent. Eliminating the points where the tangent becomes pseudo-vertical helps solving the resonance type problem. Theoretically, since each mode approaches its cut-off frequency with a different slope along the spectrum, a different value of limiting phase velocity should be chosen for each mode. Unfortunately this process removes some of the propagation components which in reality would be in the response itself. It should be noted that in order to limit the phase velocity per mode, the application of the mode tracking algorithm illustrated in section 3.4.6 is necessary.

3.7.1 **Distribution of in-plane strain energy in the rail cross-section**

Critical to this study is the knowledge of how the energy provided by the excitation is distributed across the waveguide. A propagating guided wave is a probe, which is sensitive to the occurrence of discontinuities along his path. Depending on the cross-sectional distribution of the wave energy, the inspection can be targeted to different areas. Since the most critical defects are usually located in the rail head, being able to excite the modes which show most of the energy in that area greatly increases the sensitivity of the detection. This process can be performed through different steps; initially the mode-tracking algorithm needs to be applied in order to isolate each of the propagating modes. The forced response process can be run multiple times, limiting only one mode per run, in the computation. Plotting the cross-sectional strain energy will then provide a description of the different rail defect-sensitive areas for each of the considered

modes. The expression of the cross-sectional strain energy of the e -th element is given by

$$e_p^{(e)} = \int_{V_e} \boldsymbol{\varepsilon}^{(e)T} \tilde{\mathbf{C}}_e \boldsymbol{\varepsilon}^{(e)} dV_e \quad (3.95)$$

Linear interpolation functions imply constant strain within any element of the bi-dimensional mesh. Limiting the strain energy calculation to the cross-sectional plane (y - z), the expression of the e -th element in-plane strain energy is the following:

$$e_{p,yz}^{(e)} = \frac{1}{4} \hat{\boldsymbol{\varepsilon}}^{(e)T} \hat{\mathbf{C}}_e \hat{\boldsymbol{\varepsilon}}^{(e)} \quad (3.96)$$

where $\hat{\boldsymbol{\varepsilon}}^{(e)}$ and $\hat{\mathbf{C}}_e$ represent respectively the expressions of element strain and stiffness matrix, reduced to the in-plane coordinates y and z , and are expressed as:

$$\hat{\boldsymbol{\varepsilon}}^{(e)} = \begin{bmatrix} \varepsilon_y \\ \varepsilon_z \\ \gamma_{yz} \end{bmatrix}^{(e)} \quad (3.97)$$

$$\hat{\mathbf{C}}_e = \frac{E}{1 + \nu (1 - 2\nu)} \begin{bmatrix} 1 - \nu & 0 & 0 \\ 0 & 1 - \nu & 0 \\ 0 & 0 & 1 - 2\nu \end{bmatrix} \quad (3.98)$$

where E and ν respectively represent the Young Modulus and Poisson coefficient of steel.

The reduced strain can then be expressed by:

$$\hat{\mathbf{\epsilon}}^{(e)} = \begin{bmatrix} \epsilon_y \\ \epsilon_z \\ \gamma_{yz} \end{bmatrix}^{(e)} = \frac{1}{2A^{(e)}} \begin{bmatrix} z_k - z_j & 0 & z_i - z_k & 0 & z_j - z_i & 0 \\ 0 & y_j - y_k & 0 & y_k - y_i & 0 & y_i - y_j \\ y_j - y_k & z_k - z_j & y_k - y_i & z_i - z_k & y_i - y_j & z_j - z_i \end{bmatrix}^{(e)} \begin{bmatrix} u_{y,i} \\ u_{z,i} \\ u_{y,j} \\ u_{z,j} \\ u_{y,k} \\ v_{z,k} \end{bmatrix}^{(e)} \quad (3.99)$$

being $(y, z)_{(i,j,k)}$ and $(u_y, u_z)_{(i,j,k)}$ respectively the in-plane coordinates and displacements of the i -th, j -th and k -th nodes of the e -th triangular element. The term $A^{(e)}$ in the reduced expression of the element strain corresponds to the e -th triangular element area, that can be calculated in the matricial form

$$A^{(e)} = \det \begin{bmatrix} y_i & y_j & y_k \\ z_k & z_k & z_k \\ 1 & 1 & 1 \end{bmatrix}^{(e)}. \quad (3.100)$$

Table 3.1, Table 3.2 and Table 3.3 report the eight lowest order propagating modes, along with their modeshapes and in-plane cross-sectional strain energy distributions, calculated at 200 kHz.

Table 3.1: Modeshapes and strain energy distributions of S_0 , A_0 and S_1 , computed at 200 kHz

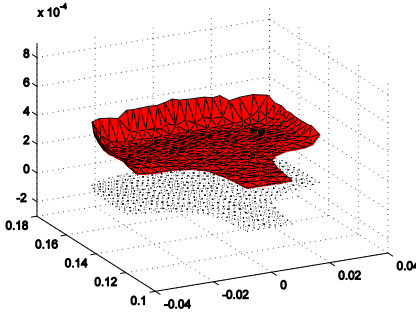
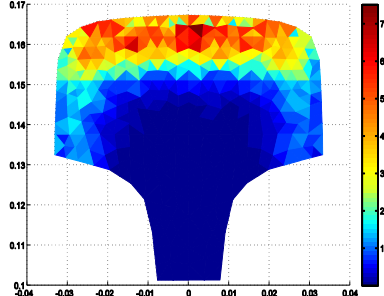
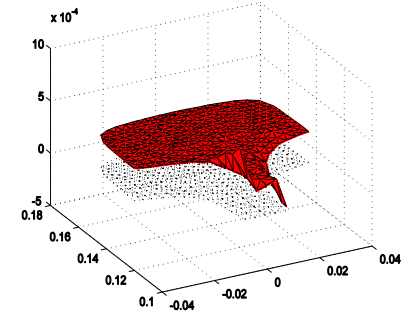
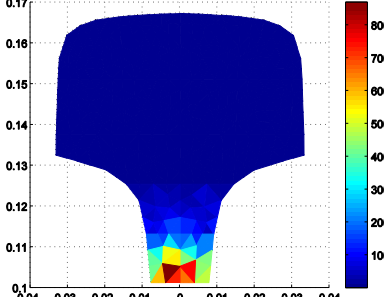
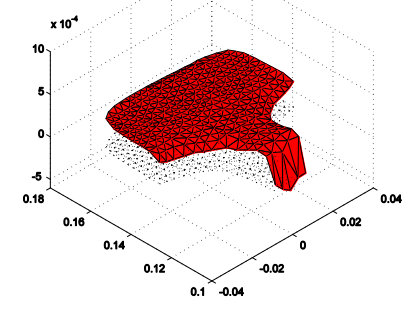
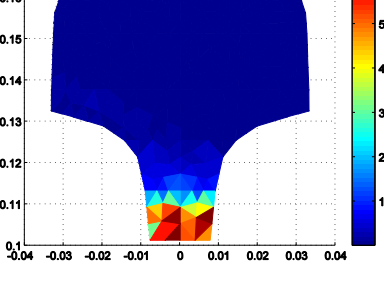
Mode	Modeshape @ 200 kHz	In-plane strain energy density [a.u.]
S_0		
A_0		
S_1		

Table 3.2: Modeshapes and strain energy distributions of A_1 , S_2 and A_2 , computed at 200 kHz

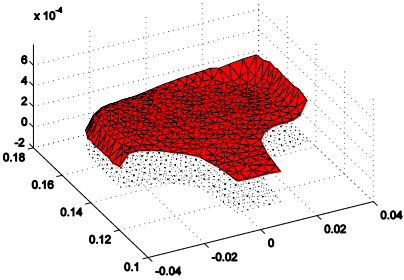
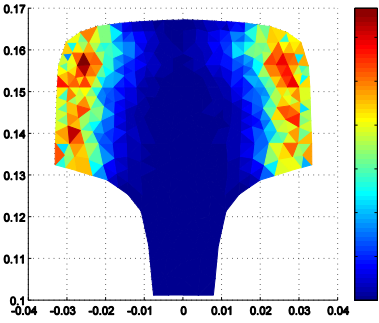
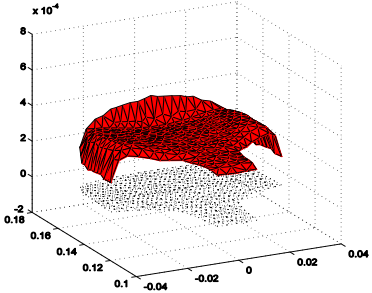
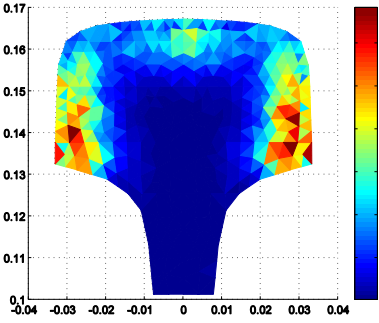
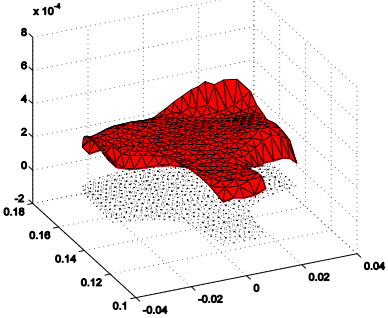
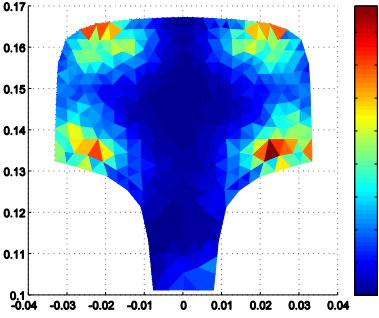
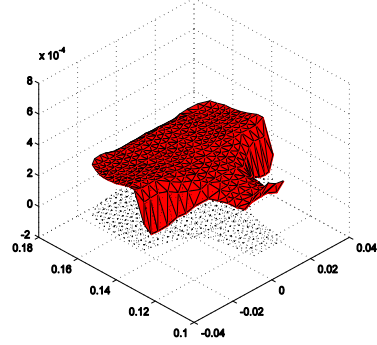
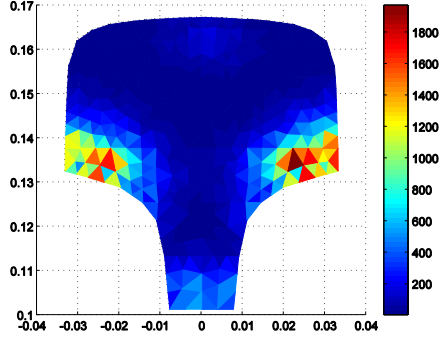
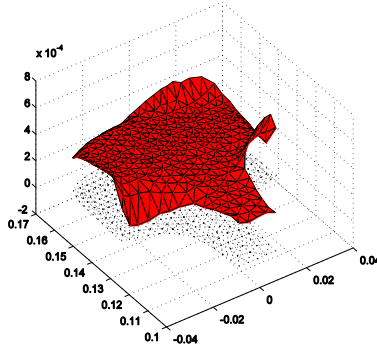
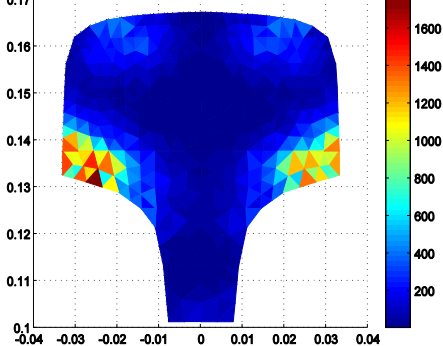
Mode	Modeshape @ 200 kHz	In-plane strain energy density [a.u.]
A_1		
S_2		
A_2		

Table 3.3: Modeshapes and strain energy distributions of S_3 and A_3 , computed at 200 kHz

Mode	Modeshape @ 200 kHz	In-plane strain energy density [a.u.]
S_3		
A_3		

The energy plots show that mode S_0 is a candidate for the detection of surface cracks, chips or shallow internal defects located anywhere within the full width of the rail. Modes S_2 and A_2 are good probes for surface cracks, chips or shallow internal defects located respectively in the center head and in the gage/field sides. Modes A_1 and S_2 develop most of the strain energy along the depth of the gage and field sides of the rail section; they are candidates for the detection of detail fractures and transverse fissures. Higher order modes S_3 and A_3 develop energy at the bottom of the gage and field sides of the head, being sensitive to deep internal defects at the earliest stage of their growth.

Table 3.4: In-plane strain energy distribution and defect sensitivity of the eight lowest order modes in rails

Mode	Area of the rail cross-section with highest strain-energy	Type of detectable defects
S_0	Top 15 mm of rail head	Surface cracks, chips, shallow internal defects anywhere along the full width of the rail head
A_0	Bottom 10 mm of meshed web	Not usable for head defect detection
S_1	Bottom 15 mm of meshed web	Not usable for head defect detection.
A_1	Areas at gage and field sides of the head, 15 mm wide	Surface cracks, chips, shallow internal defects, deep internal defects such as transverse fissures or detailed fractures at the gage and field side of the head.
S_2	Middle and bottom part of the gage and field sides of the head, plus a localized area at the top of the center head	Deep internal defects such as transverse fissures or detailed fractures at the gage and field side of the head.
A_2	Areas at the four corners of the head side wings	Surface cracks, chips at the gage and field side of the head
S_3	Areas at the bottom corners of the gage and field sides of the head	Deep internal defects such as transverse fissures or detailed fractures at the bottom of the gage and field side of the head.
A_3	Areas at the bottom corners of the gage and field sides of the head	Deeper internal defects such as transverse fissures or detailed fractures at the bottom of the gage and field side of the head.

Modes A_0 and S_1 are unsuitable for rail head inspection, since they carry most of the energy in the rail web; as it will be shown later, those modes are barely excited by the laser excitation on the top of the head. Table 3.4 contains the summary of the analysis of

the strain energy distribution for each mode in regard to its defect sensitivity. Among the good candidates for surface defect detection, S_0 is almost nondispersive and propagates with low attenuation. Among the good candidates for internal defect detection, mode A_1 shows low attenuation and limited dispersion. This leads to the conclusion that an antisymmetric excitation would be needed in order to inspect the rail head gage and field sides for internal defects, while a symmetric excitation would be optimal to detect surface flaws. Experimental results will be shown in chapter four to validate this concept. As an intermediate step, the following section will show how the excitation pattern influences the generated propagating modes.

3.7.2 Influence of the excitation pattern on the propagating mode

A symmetric and a nonsymmetric excitation patterns were applied on the modeled rail. The nodal amplitudes of the load were adjusted so that the same amount of energy was transferred to the rail in both cases. Figure 3.31 shows the symmetric pattern, while the nonsymmetric one is depicted in Figure 3.32. In order to understand how a different pattern excitation influences the response of the rail, the values of the modal participation factors $\alpha_m(f)$ have been plotted along the frequency spectrum up to 500 kHz. The excitability curves, which are plots of the modal participation factors as a function of frequency, are shown in Figure 3.33 and Figure 3.34 for the case of symmetric and nonsymmetric excitation respectively.

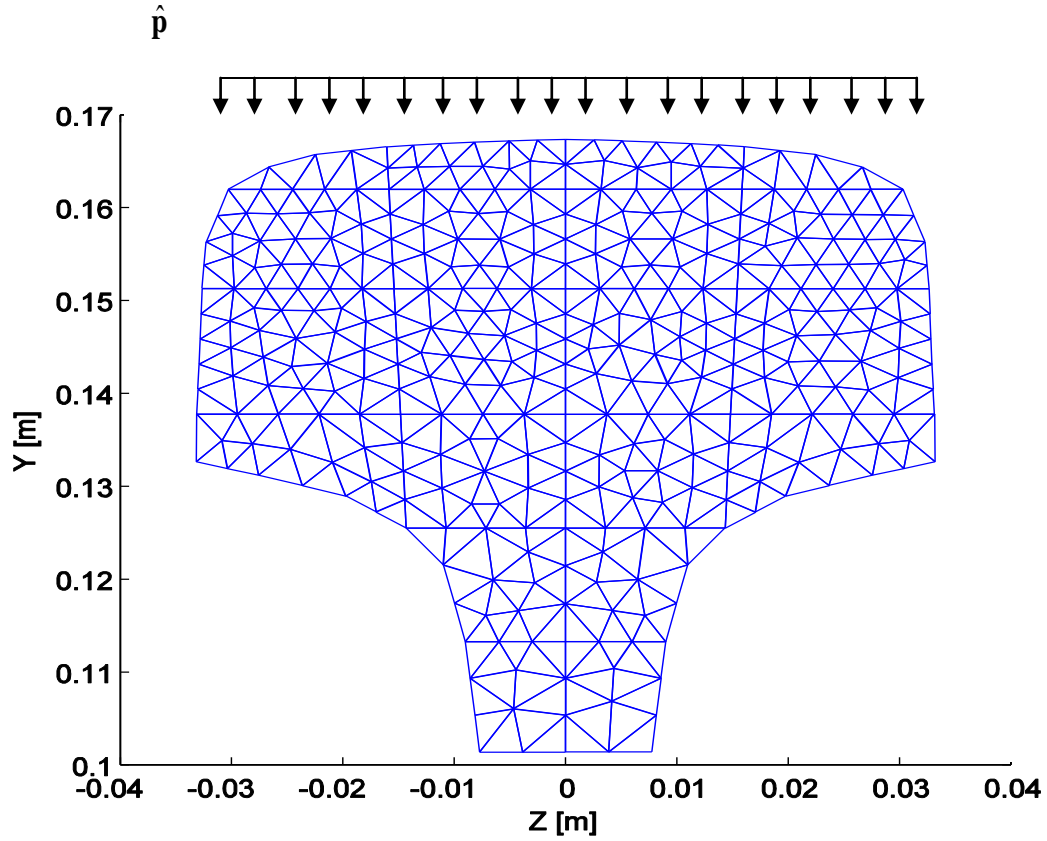


Figure 3.31: Symmetric excitation pattern on the meshed rail section

With increasing value of frequency, the complexity of the scenario is evident; 152 forward propagating modes are identified at 500 kHz. In both the excitability plots, it is clear that a few number of lowest order modes are dominant along the majority of the spectrum. An accurate analysis of the eight lowest order mode excitability is useful to appreciate the influence of the loading pattern on the response. Figure 3.35 and Figure 3.36 give a closer look at the excitability of S_0 , A_0 , A_1 and S_2 modes for the two load cases.

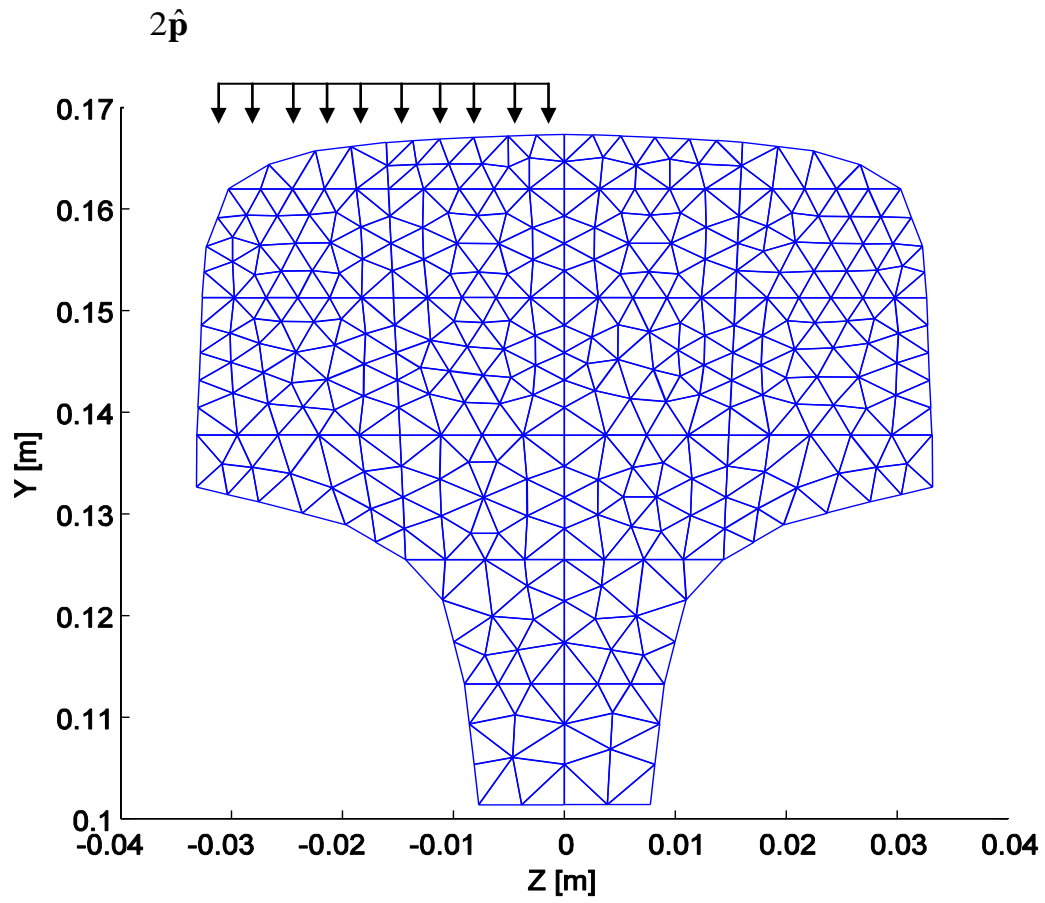


Figure 3.32: Nonsymmetric excitation pattern on the meshed rail section

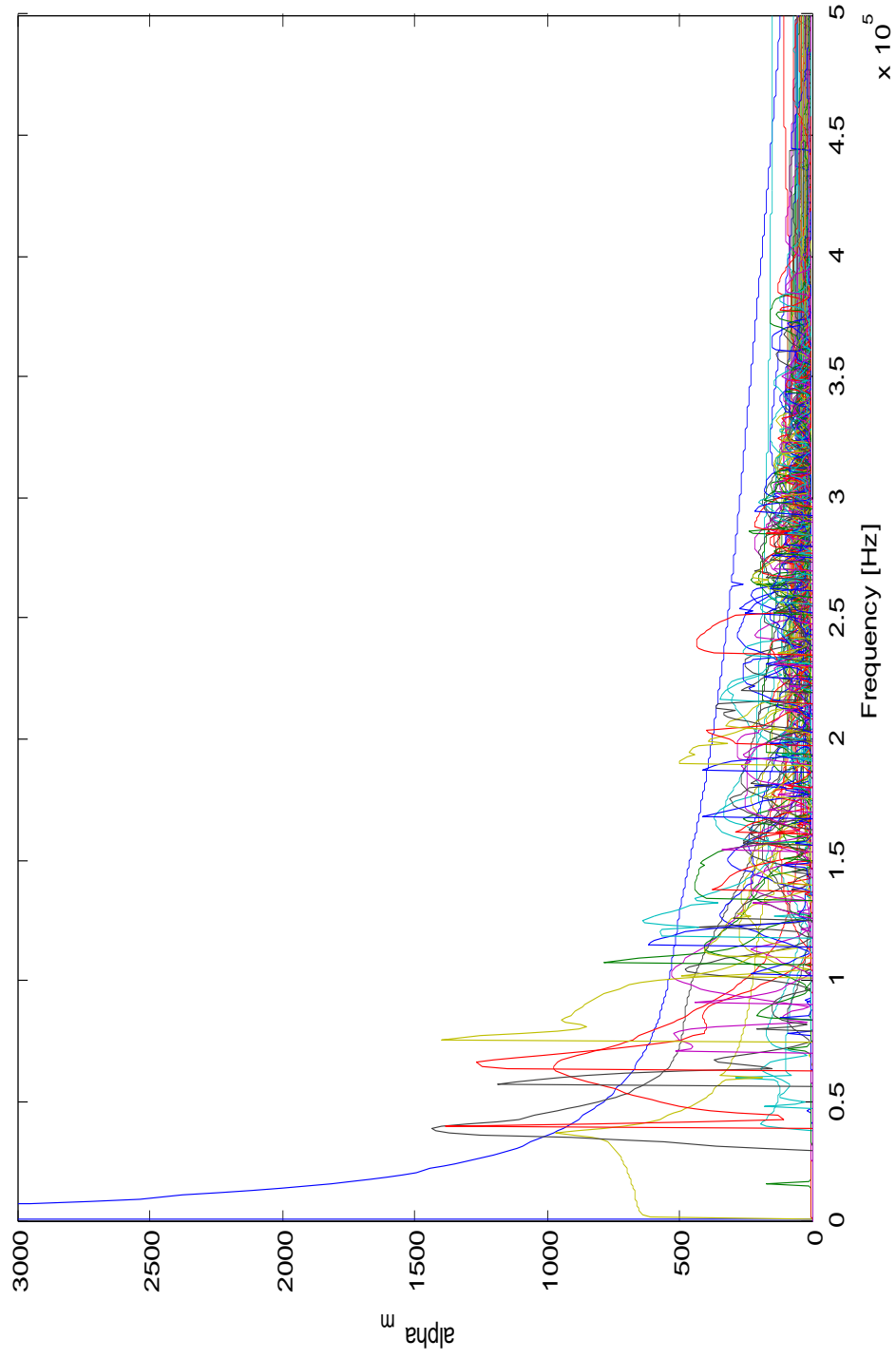


Figure 3.33: Mode excitability curves for a case of symmetric excitation of the rail head

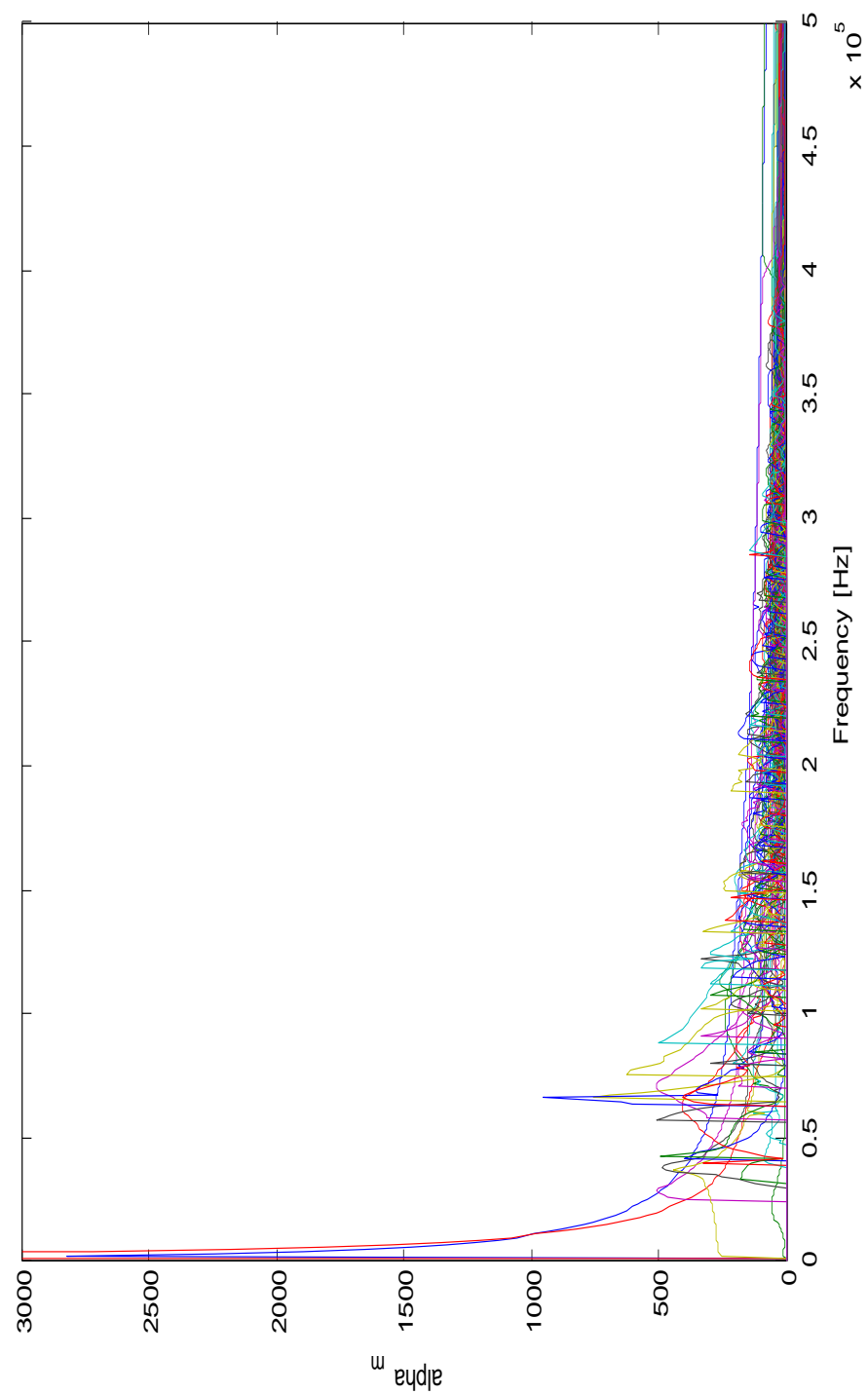


Figure 3.34: Mode excitability curves for a case of nonsymmetric excitation of the rail head

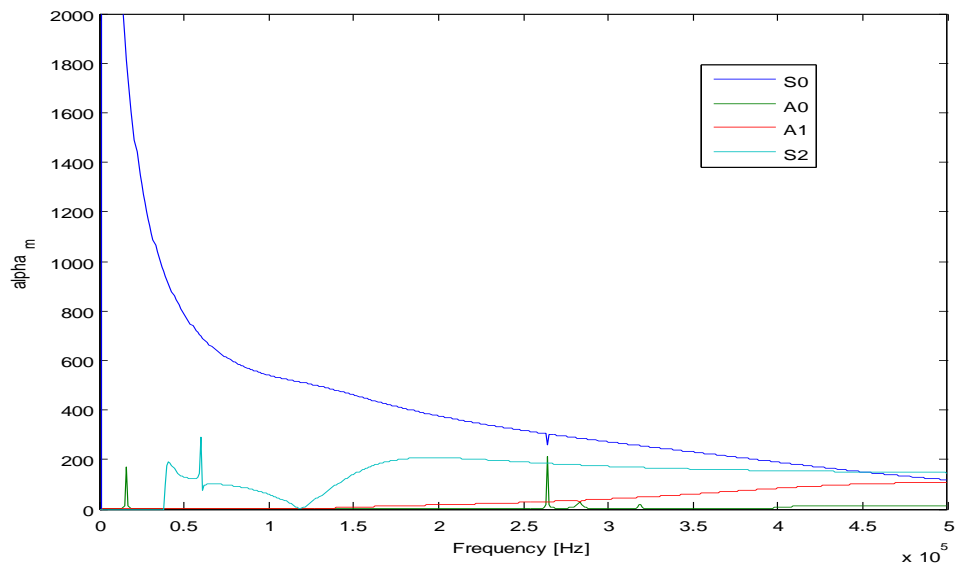


Figure 3.35: S_0 , A_0 , A_1 and S_2 mode excitability curves for a symmetric excitation of the rail head

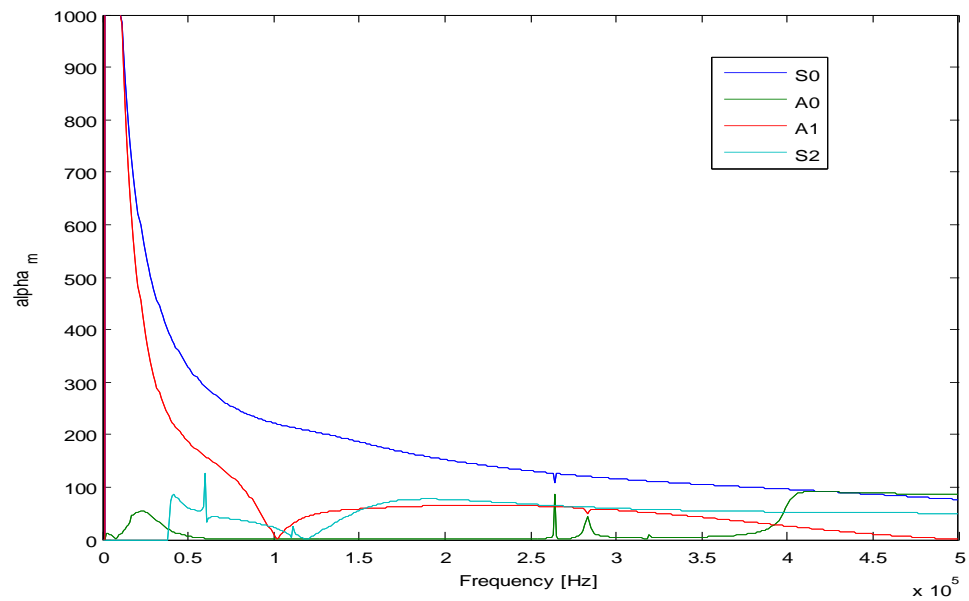


Figure 3.36: S_0 , A_0 , A_1 and S_2 mode excitability curves for a nonsymmetric excitation of the rail head

It should be noted that in the discussion that follows, the nonsymmetric load is considered as the sum of a symmetric and an antisymmetric contribution, as depicted in Figure 3.37.

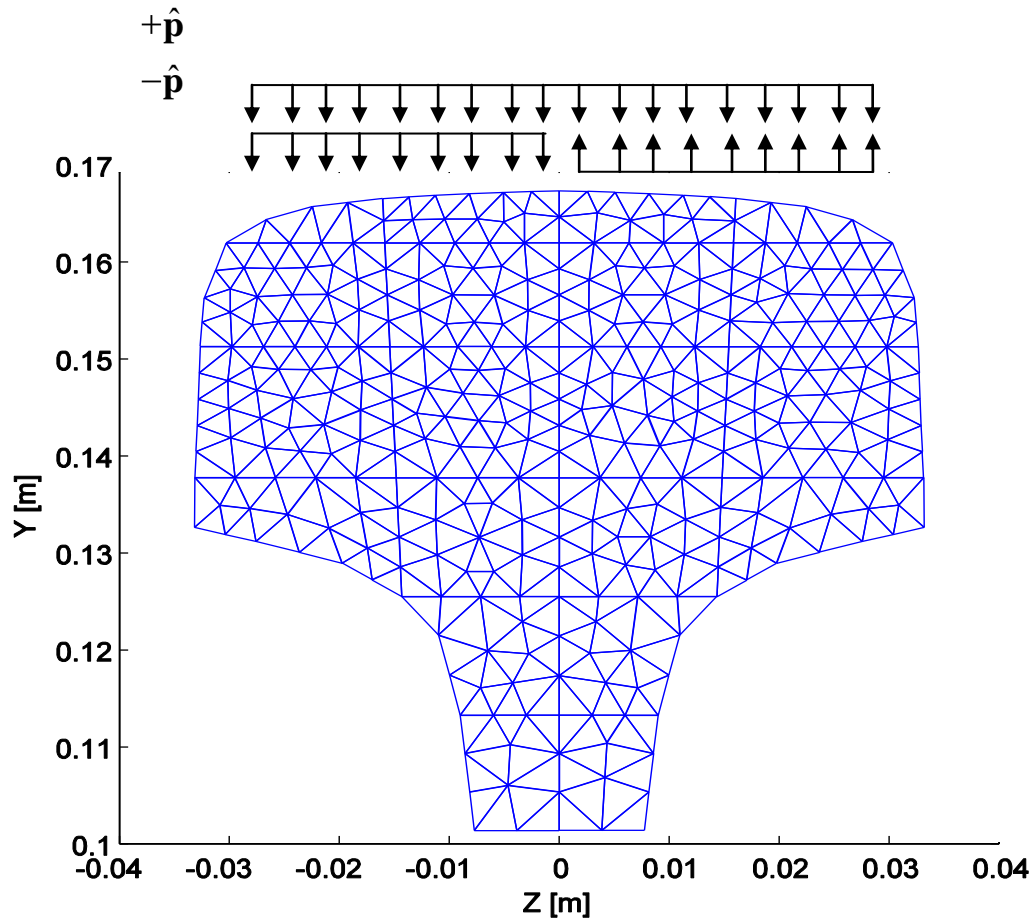


Figure 3.37: Nonsymmetric load as sum of symmetric and antisymmetric contributions

In both symmetric and nonsymmetric loading cases, mode S_0 is the most excited along the spectrum; this is due to the fact that its modeshape shows maximum displacements in the area where the load has been applied. In the nonsymmetric load case, mode A_1 is excited as well. It is clear from the plots in Figure 3.38 and Figure 3.39, that in the symmetric cases the excitability is almost only confined to symmetric modes; in the

nonsymmetric cases, instead, symmetric and antisymmetric modes are simultaneously excited. Mode A_0 is never excited since it transfers energy mainly through the rail web, while mode S_2 shows medium excitability for the symmetric case; its modal participation factors lower than S_0 since it shows energy in critical areas of the rail, but not in close proximity to the loaded modes. Figure 3.38 and Figure 3.39 show the excitability plots of S_1 , A_2 , S_3 and A_3 modes for the two different excitation patterns; in general, these modes show lower excitability than S_0 , A_0 , S_2 , and A_1 , which remain dominant in the response. Modes S_1 and S_3 are excited in the symmetric loading case, while A_0 and A_1 are triggered only in the nonsymmetric. Two cases of mode-coupling between S_1 and A_2 are occurring in proximity of 140 and 215 kHz, as highlighted in Figure 3.39. Mode coupling is a known phenomenon that can be caused by damping effects in a single layer plate [82] or solely by geometrical effects in undamped multilayered structures [88]; it is also known to be dependent on the material properties of the media where the propagation occurs [11]. The phenomenon has not yet been studied intensively for the case of arbitrary cross-sections, such as the rail, although, it is believed that its occurrence is due to the complex shape of the rail cross-section. Including all the propagating modes, the response has been obtained for the two excitation patterns, and the related cross-sectional in-plane strain energy distributions have been plotted in Figure 3.40 and Figure 3.41.

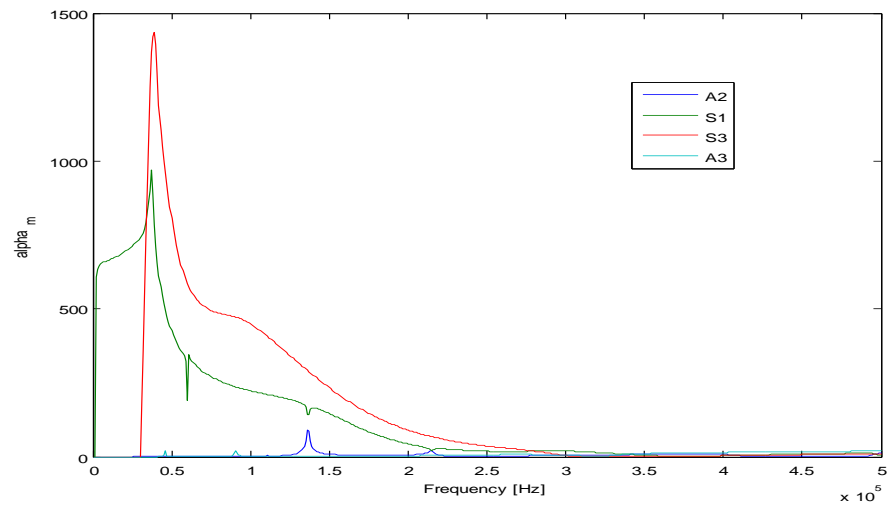


Figure 3.38: S_1 , A_2 , S_3 and A_3 mode excitability curves for a symmetric excitation of the rail head

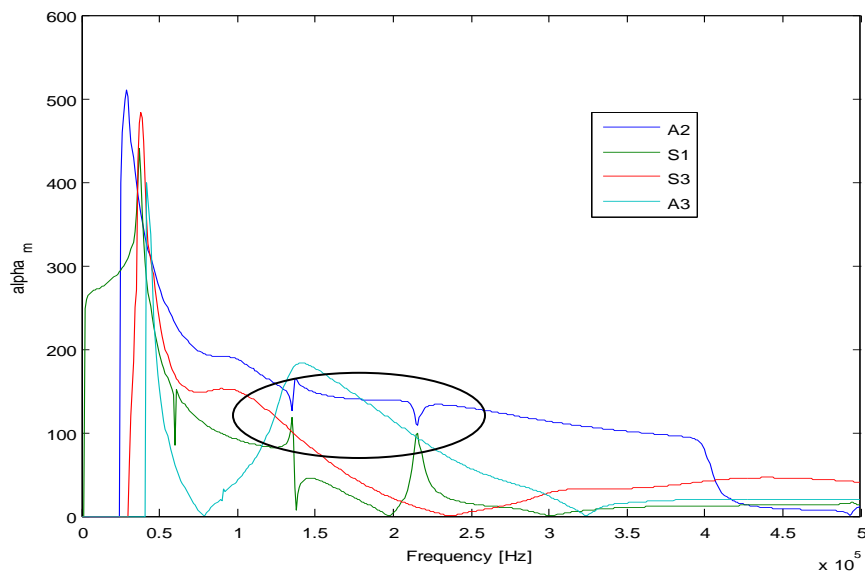


Figure 3.39: S_1 , A_2 , S_3 and A_3 mode excitability curves for a nonsymmetric excitation of the rail head

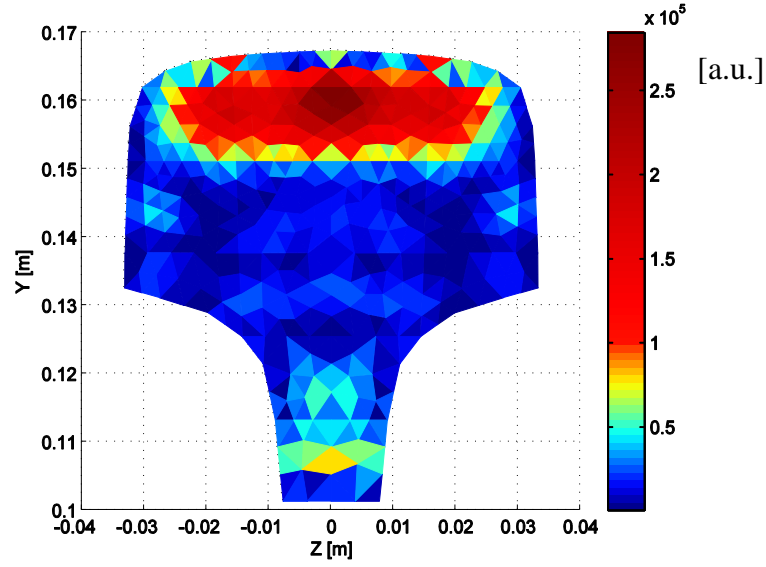


Figure 3.40: In-plane cross-sectional strain energy distribution for a symmetric excitation at 4'' (≈ 102 mm) from the source

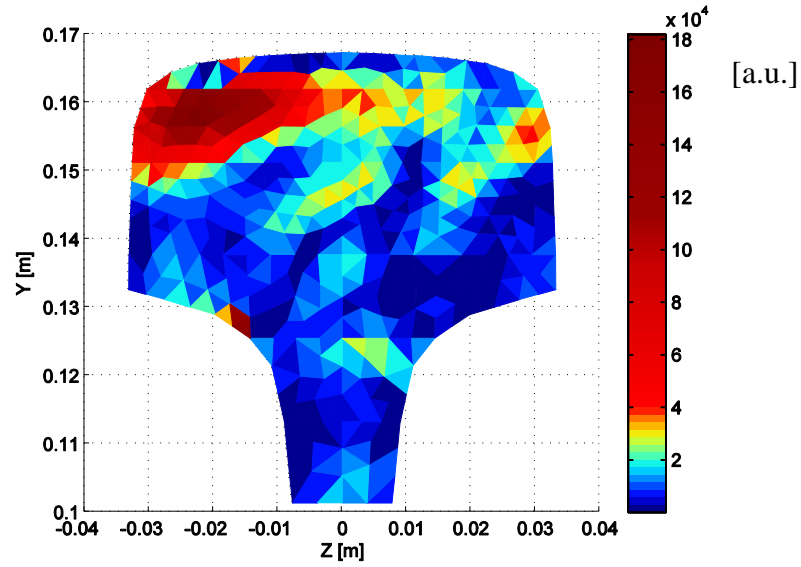


Figure 3.41: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at 4'' (≈ 102 mm) from the source

It is evident that the energy distribution on the rail section reflects the symmetry of the excitation pattern. This assertion will be validated by experimental results shown

later on in chapter four. Calculating the response at different distances from the excitation, modulation among modes occurs be experienced when many of them are excited. Hesse and Cawley [11] studied this phenomenon plotting the maximum of the signal envelope, measured at different distances from a narrowband excitation source; while a gradually decaying amplitude was expected, the resulting amplitude showed an irregular pattern, caused by constructive and destructive interference of multiple modes [11].

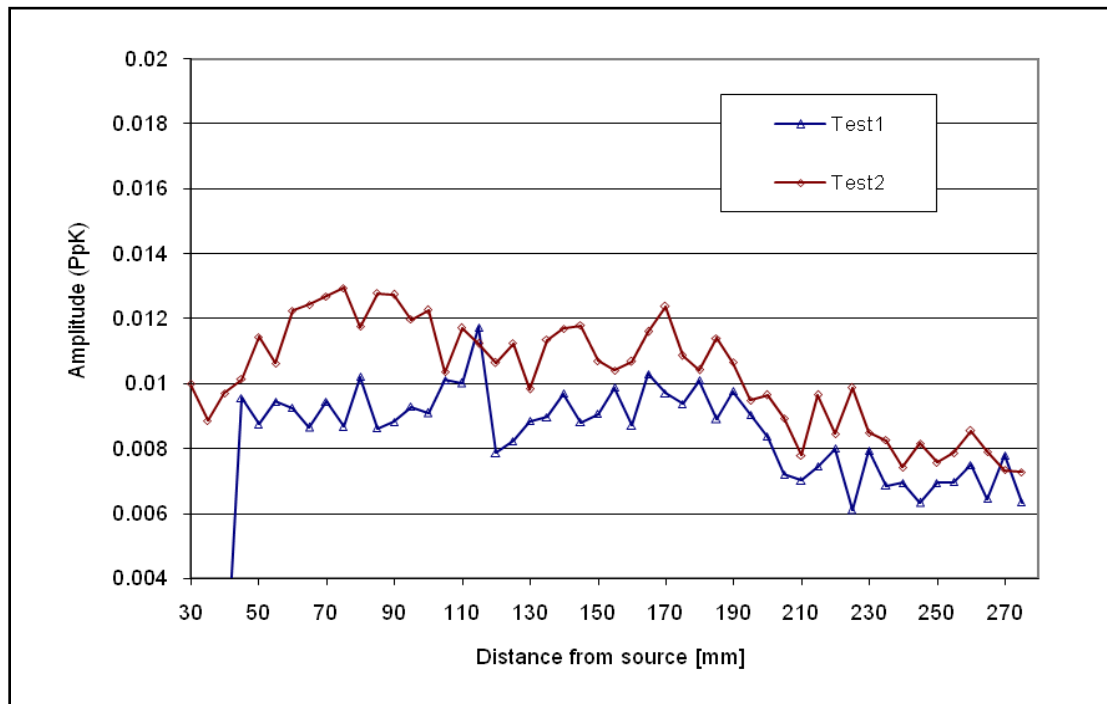


Figure 3.42: Maximum of the signal envelope acquired at different distances from the source

Similar plots were obtained using laser excitation on the test rail at UCSD. The results of the UCSD test are shown in Figure 3.42, where the irregular pattern can still be observed, although the variability of the measure amplitude is very limited. This is due to the broad spectrum of the excitation, where higher frequencies contribute to an earlier

formation of a constant wavefront. In order to generalize the results obtained regarding the influence of the excitation pattern on the rail response, this response was calculated at different distances from the source (100 ± 50 mm), for both the symmetric and nonsymmetric load cases. The distribution of the in-plane strain energy across the cross-section shows to be slightly different, but it still follows the symmetry (or nonsymmetry) of the load as shown in Figure 3.43 through Figure 3.46

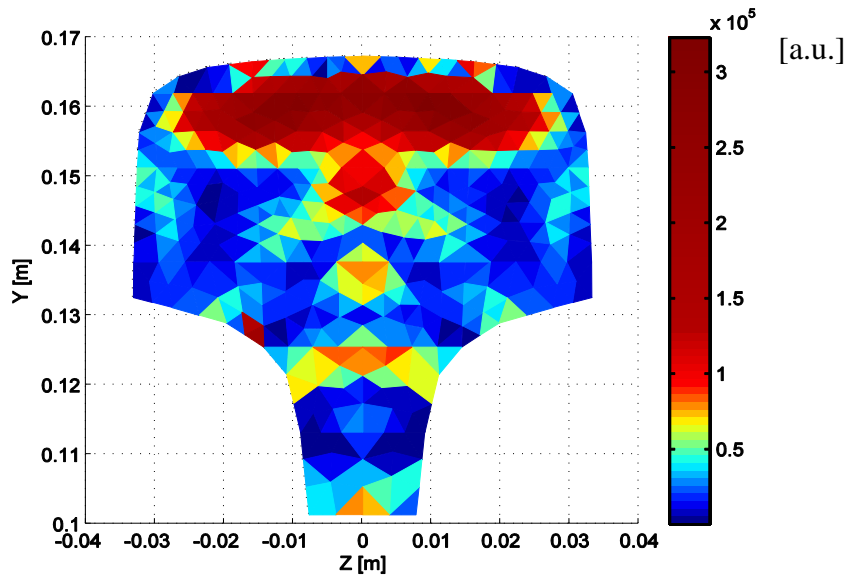


Figure 3.43: In-plane cross-sectional strain energy distribution for a symmetric excitation ($2'' \approx 51$ mm) from the source)

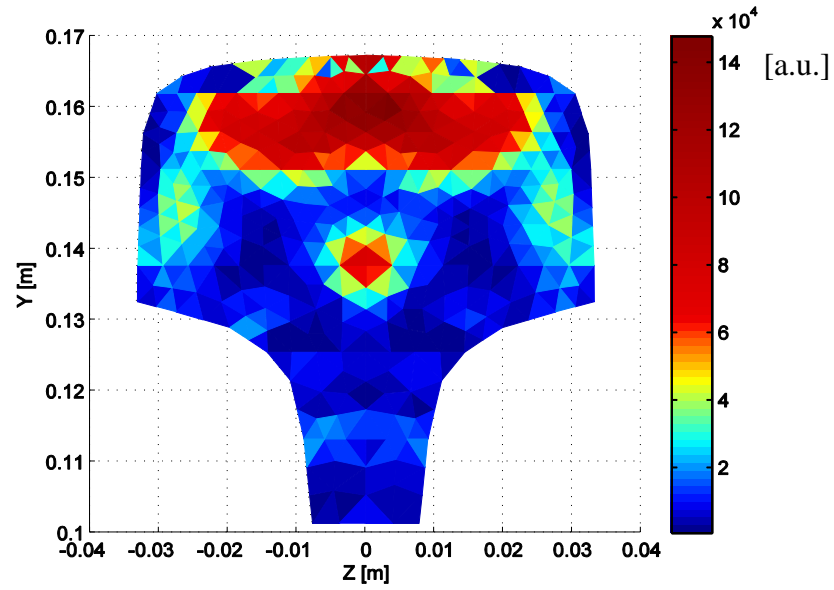


Figure 3.44: In-plane cross-sectional strain energy distribution for a symmetric excitation at 6" (≈ 152 mm) from the source

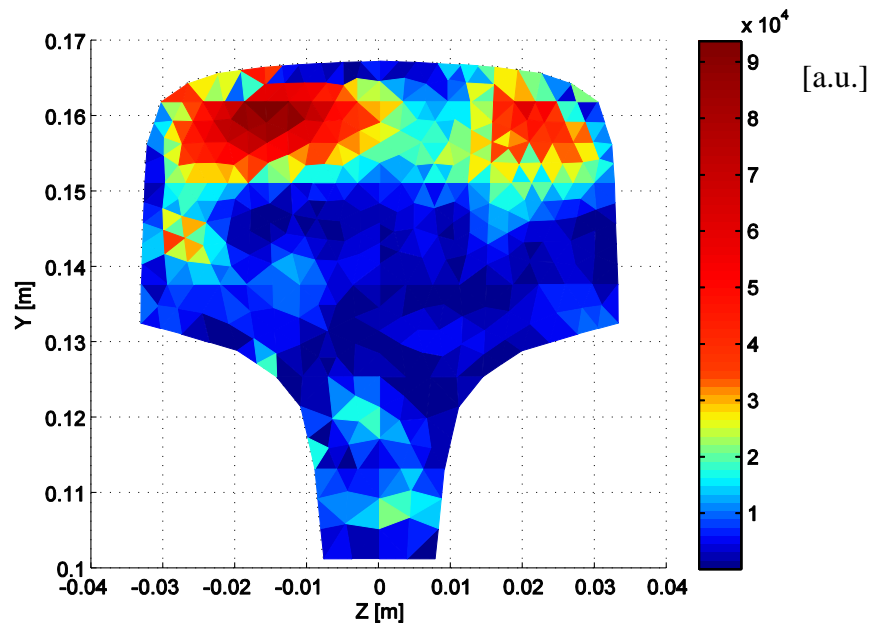


Figure 3.45: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at 2" (≈ 51 mm) from the source

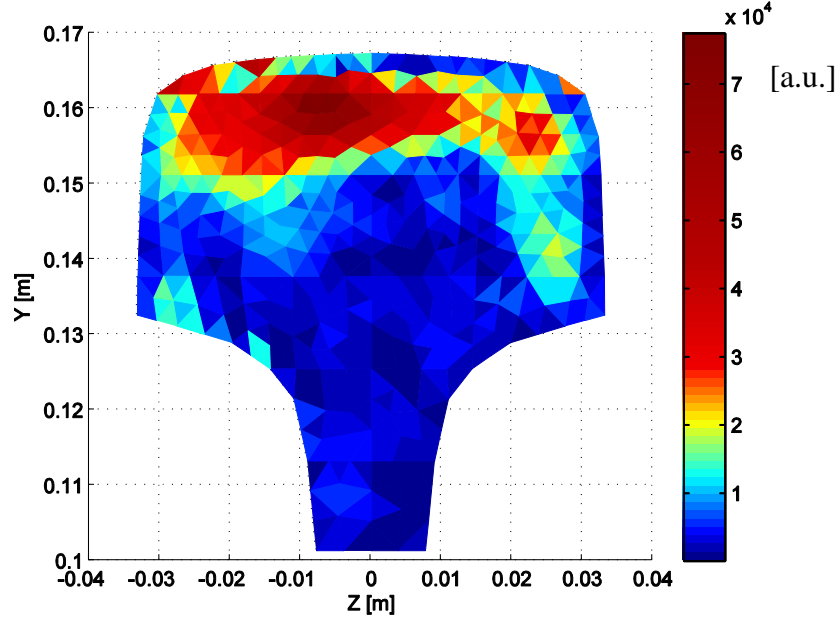


Figure 3.46: In-plane cross-sectional strain energy distribution for a nonsymmetric excitation at 6'' (≈ 152 mm) from the source

3.7.3 Effect of the frequency on the penetration depth of the guided wave

In this section, the penetration depth of the propagating guide wave is discussed.

A mode j propagating with phase velocity

$$c_p^j = \frac{\omega_k}{\xi_j} \quad (3.101)$$

at the circular frequency ω_k , will develop a penetration depth on the order of its wavelength λ_j [69]. Substituting the values of circular frequency and wavenumber in eq.

(3.101), as expressed in eq. (3.102)

$$c_p^j = \frac{\omega_k}{\xi_j} = \frac{2\pi f_k}{\frac{2\pi}{\lambda_j}} = f_k \lambda_j \quad (3.102)$$

the mode wavelength can be expressed as the ratio between the phase velocity and the frequency:

$$\lambda_j = \frac{c_p^j}{f_k}. \quad (3.103)$$

From the dispersion curves of Figure 3.21, it can be deduced that lower-order modes, which contribute to the majority of the strain energy, tend to the Rayleigh wave velocity at higher frequencies. Hence it is reasonable to assume that the penetration depth of these modes is proportional to their wavelength, or inversely proportional to their frequency. Table 3.5 summarizes the correspondence between frequency and penetration depth (wavelength) for different values along the spectrum.

Table 3.5: Correspondence between frequency and penetration depth (wavelength) into the rail head for different values of the spectrum

Frequency [kHz]	Penetration depth (wavelength) [mm]
60	48
100	29
200	14.5
300	≈10
500	≈6
700	≈4

The value of 60 kHz has been included in Table 3.5 since modes S_0 and A_1 phase velocities are already stable at the value of the Rayleigh wave velocity. The semianalytical model solution proved the above considerations; this can be shown plotting the cross-sectional in-plane strain energy for the case of symmetric pattern excitation for different frequency ranges. Figure 3.47 contains the in-plane strain energy of the rail response, after being filtered through a 3-order Butterworth highpass set at the cut-off frequency of 300 kHz. As expected the energy developed in the rail is confined to a surface depth of about 10 mm; these results are also well confirmed in literature [11]. Figure 3.48 instead, contains the reduced strain energy after being filtered through a 3-order Butterworth bandpass, between 100 and 300 kHz.

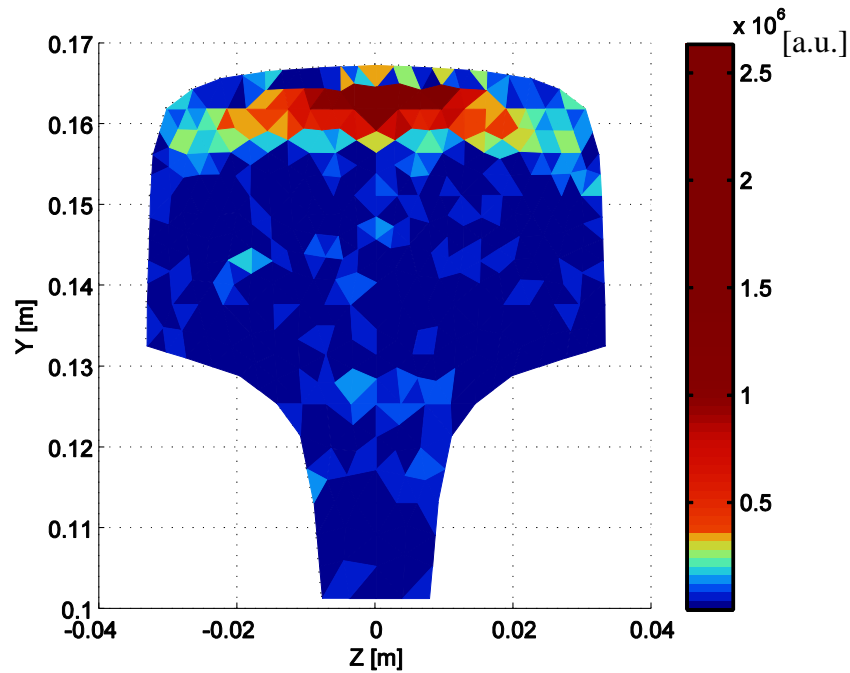


Figure 3.47: Cross-sectional in-plane strain energy of rail response, filtered with a 300 kHz, 3-order Butterworth highpass

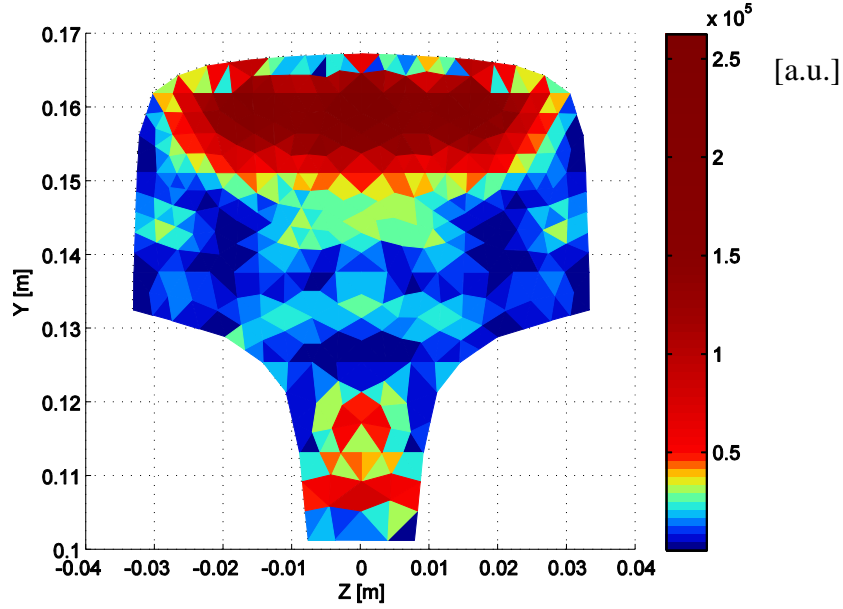


Figure 3.48: Cross-sectional in-plane strain energy of rail response, filtered with a 100-300 kHz, 3-order Butterworth bandpass

It can be seen that at the lower frequency, the excitation can induce energy up to a depth of about 25 mm into the rail head. It is important mentioning that in both cases no energy appears within the top 3 mm of the rail cross-section; that is because the mesh used is not fine enough to have the spatial resolution required to match the short wavelength of frequencies around 700 kHz, which can penetrate to that depth.

3.8 Conclusions

A semi analytical approach was shown to be successful for modeling the ultrasonic wave propagation in rails. The unforced and forced solutions of the rail response were obtained with the SAFE method and some were validated by experimental results. An algorithm which uses the Arnoldi Iterative solution process was presented and its efficiency was validated by computational time test results. Analytical and

experimental results showed how the load pattern influences the rail response. Understanding cross-sectional energy distribution of the propagating modes is crucial to the design of a detection system targeted to defects of interest. A symmetric excitation pattern on the top of the rail develops strain energy in a symmetric fashion along the cross-section, with maxima close to the top of the head; a nonsymmetric pattern, instead, induces energy at the sides of the head. For the symmetric load case, high frequency waves generate energy close to the top of the head, while lower frequency waves induce energy in deeper areas of the head. Among the lower order modes, S_0 and A_1 are shown effective for surface defect detection, while A_1 and S_2 are more suitable for the detection of internal flaws. Material damping was included in the model for the unforced response solution, and it was neglected in the forced response solution. Neglecting material damping caused the occurrence of resonance-like phenomena, which were corrected by filtering out modes at cut-off frequencies, after applying a mode tracking algorithm, which employs mode orthogonality.

Chapter 3, in part, will be submitted for publication in *Journal of Acoustical Society of America*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Marzani, A.. The current running title of this paper is “Forced solution of guided waves in rails”. The dissertation author will be the primary investigator and author of this paper.

4 Defect-sensitive feature extraction from ultrasonic waves

4.1 Introduction

The drawback of any noncontact ultrasonic testing system is a reduced signal-to-noise ratio of the defect detection procedure, when compared to conventional contact testing. The use of signal processing is necessary, yet sometime not sufficient to overcome this problem. Different techniques are currently available and are good candidates to be employed for the signal conditioning of ultrasonic waves. As mentioned before, the main objective of a signal processing technique to be adopted in the defect detection prototype is the capability of increasing the signal-to-noise ratio and the ability of compressing data and retaining only useful information from the acquisition. Dealing with a large amount of data acquired by multi-channel systems and looking at high frequency ranges such as the spectrum of the ultrasonic region contribute to make the issue not a trivial one. The choice of a signal processing technique is also strictly connected with the implementation of an efficient feature extraction procedure; the output of the former is fed as the input to the latter. Robustness, low computational speed, low false positive and false negative rates are some of the distinctive properties which characterize the good performance of a signal processing-feature extraction procedure. This chapter introduces the theoretical background of the different signal processing approaches which have been tested for the best performance of the defect detection prototype for rail tracks based on noncontact (laser generation/air-coupled detection) ultrasonic testing. Experimental results obtained varying multiple sensitive

parameters are shown in the last section of this chapter, as a validation of the different methods.

4.2 Continuous and discrete wavelet analysis

This section will give a brief introduction of the main signal processing tools currently known in literature, which were considered for the rail inspection prototype implementation. The information contained within an ultrasonic signature is both present in the time and in the frequency domain. Time of Flight (TOF) is an important signal feature, which can be used to recognize the position of a peak corresponding to an echo reflected from a defect (“reflection mode”) or to a first arrival in the “transmission mode” case [90]. The frequency content of a signal is critical for the identification of the propagating modes and, as seen in chapter three, it is strictly related to the area of the inspected rail cross-section. The Fourier Transform $X(\omega)$ of a signal $x(t)$ is defined as:

$$X(\omega) = \int x(t) e^{-i\omega t} dt \quad (4.1)$$

where t and ω are the time and frequency variables, respectively. It defines the spectrum of $x(t)$ which consists of components at all frequencies over the range for which it is nonzero. For stationary signals, Fourier analysis is extremely useful because the only signal's frequency content is of great importance. However, Fourier analysis has a serious drawback. In transforming to the frequency domain, time information is lost. Instead, the ultrasonic signals to be processed by the prototype, contain numerous nonstationary characteristics, such as drift, abrupt changes, beginning and end of events. These characteristics are the most important part of the signal and Fourier analysis is not suited to detecting them. The complex propagation scenario described in Figure 3.19 of

the previous chapter, creates an inevitable need to retain the time information, for performing efficient mode recognition. To correct this deficiency, Dennis Gabor (1946) adapted the Fourier transform to analyze only a small section of the signal at a time. It is called the Short-Time Fourier Transform (STFT) [91, 92]. The STFT is obtained from the usual Fourier Transform by multiplying the time signal $x(t)$ by an appropriate sliding time window $w(t)$.

The location of the sliding window adds a time dimension and gets a time-varying frequency analysis. Thus, instead of the usual Fourier Transform expression, a time frequency expansion can be expressed in the form:

$$X(t, \omega) = \int x(\tau) w(\tau - t) e^{-i\omega\tau} d\tau \quad (4.2)$$

where $w(t)$ is the time window applied to the signal. It provides some information about time and frequencies that a signal event occurs. However, this information can only be obtained with limited precision, which is determined by the size of the window. The STFT has also strict limitations on the time-frequency resolution. If greater frequency resolution is required, it can be achieved only at the expense of temporal resolution. Many signals require a more flexible approach where the window size can be varied to determine more accurately either time or frequency. This adaptive approach can be obtained using the Wavelet Transform, in its continuous and discrete expressions.

4.2.1 Theory behind wavelet analysis

The Continuous Wavelet Transform (CWT) has been extensively [93, 94, 95, 96, 97] used to process ultrasonic signals in time and frequency domains.

The CWT of a signal $x(t)$ with respect to an analyzing or parent, or mother) wavelet $\psi(t)$ is defined as:

$$Wx(u, s) = \int_{-\infty}^{+\infty} x(t) \psi_{u,s}^*(t) dt \quad (4.3)$$

where $*$ denotes a complex conjugate and $\psi_{u,s}(t)$ is obtained by scaling the mother wavelet $\psi(t)$ by the scaling parameter s and translating it by the translation parameter u , as expressed by:

$$\psi_{u,s} = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (4.4)$$

The mother wavelet should have a certain number of properties. The importance is satisfying square integrability, and being concentrated in the time and frequency as much as possible.

The admissibility condition is expressed by:

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (4.5)$$

where $\Psi(\omega)$ is the Fourier transform of $\psi(t)$. The weak condition for the existence of the integral in eq. (4.5) is

$$\Psi(\omega)|_{\omega=0} \quad \text{which is equivalent to} \quad \int_{-\infty}^{+\infty} \psi(t) dt = 0. \quad \text{This means that wavelets must behave}$$

like a band-pass like spectrum. A zero at the zero frequency also means that the average value of the wavelet in the time domain must be zero, and therefore the function has to be oscillatory.

An analytical wavelet can be created by modulating the frequency of a real and symmetric window $g(t)$ as follows:

$$\psi(t) = g(t) \exp(i\eta t) \quad (4.6)$$

whose Fourier transform is

$$\Psi(\omega) = G(\omega - \eta) \quad (4.7)$$

with a central frequency of η , being $G(\omega)$ the Fourier transform of $g(t)$. According to Heisenberg uncertainty principle, a signal cannot be fully resolved both in time and in frequency simultaneously. The time-frequency resolution of the signal depends on the time-frequency resolution of the analyzing wavelet $\psi_{u,s}(t)$ [98]. Each time-frequency window $\psi_{u,s}(t)$ can be represented by a Heisenberg box centered on the frequency axis at

$\frac{\eta}{s}$, and on the time axis at u , being η the central frequency of $\Psi(\omega)$. The sides of the

Heisenberg box are of lengths $\frac{\sigma_\omega}{s}$ and $s\sigma_t$, where σ_ω is the spread of the wavelet

window in frequency and σ_t is the spread of the wavelet window in time. The

uncertainty principle implies that the area of the Heisenberg box must be lower bounded

by 0.5, that is $\sigma_t \sigma_\omega \geq 0.5$. The Gabor wavelet transform provides the best balance

between time resolution and frequency resolution since it uses the smallest possible

Heisenberg uncertainty box ($\sigma_t \sigma_\omega = 0.5$) [98]. In the case of the Gabor wavelet, $g(t)$ is a

Gaussian window defined as

$$g(t) = \frac{1}{\sqrt[4]{(\sigma^2 \pi)}} \exp\left(\frac{-t^2}{2\sigma^2}\right) = \frac{1}{\sqrt[4]{\pi}} \sqrt{\frac{\eta}{G_s}} \exp\left(\frac{-(\eta / G_s)^2 t^2}{2}\right) \quad (4.8)$$

where η is the Gabor center frequency, σ is the standard deviation of the Gaussian window, and $G_s = \sigma\eta$ is known as the Gabor shaping factor [99]. The choice of η and G_s influences the time–frequency resolution of the analysis. As for the G_s parameter, while the *product* of the time resolution and the frequency resolution (area of the Heisenberg box) is kept at the best possible value of 0.5 with the Gabor wavelet, the choice of the factor G_s (and therefore σ , for fixed η) selects the relative resolution achievable in time and in frequency (shape of the Heisenberg box). For example, decreasing G_s decreases the achievable time resolution while increasing the frequency resolution. The optimum value of G_s must be chosen on the basis of the particular application. Although theoretically any value of $G_s \gg 1$ will satisfy the admissibility condition equation (4.5), it has been found in practice that G_s cannot have a value lower than 3.5 [95]. Most authors use $G_s = 5.336$ for dispersive wave propagation studies [94, 96, 100, 101]. The energy density spectrum of a WT, defined as $P_w f(u, s) = |Wf(u, s)|^2$ and commonly referred to as a scalogram, indicates the energy of signal $f(t)$ in the Heisenberg box of each wavelet window $\psi_{u,s}(t)$ around time $t = u$ and angular frequency $\omega = \eta/s$. The scalogram is often normalized by the scaling parameter. In this case it is referred to as the normalized scalogram $P_w f(u, s)/s$. Besides providing the time–frequency information of the signal components, the scalogram retains the signal energy content. It is therefore possible to extract both the dispersion curves and the frequency-dependent attenuation. The morlet wavelet is a particular case of the Gabor wavelet, obtained setting $G_s = 5$. Figure 4.1 shows an example of a Gabor wavelet, with central frequency $\eta = 2\pi$ and Gabor shaping factor $G_s = 5.34$, plotted in both frequency and time domain.

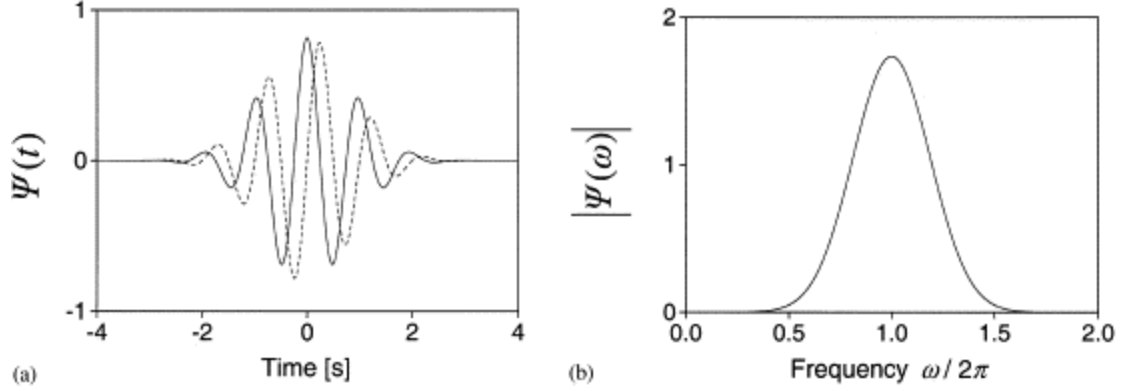


Figure 4.1: The Gabor wavelet with $\eta = 2\pi$ and $G_s = 5.34$. (a) Time domain, real (continuous line) and imaginary (dotted line) components. (b) Magnitude of Fourier spectrum

4.2.2 Discrete Wavelet Transform (DWT) analysis for signal decomposition

The implementation of the CWT is not always feasible for applications that require real-time processing, such as the rail defect detection prototype, due to the high computational load of its algorithm. The CWT is calculated by continuously shifting a continuously scalable function over a signal and calculating the correlation between the two. Some considerations must be given to accelerate the solution process, while still maintaining the advantages that an adaptive time-frequency approach has to offer. The obtained wavelet coefficients are highly redundant, since the scaled functions are not an orthogonal basis. Even resolving the redundancy problem of the CWT, the analyzing wavelets would still be infinite [102]. Last, for most functions, the wavelet transforms have no analytical solutions and can be calculated only numerically. The removal of the redundancy can be achieved through the introduction of discrete wavelets, which are

continuous functions that can only be scaled and translated in discrete steps. This is achieved modifying the wavelet representation of eq. (4.4) into the expression [103]

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \psi\left(\frac{t - ku_0 s_0^j}{s_0^j}\right) \quad (4.9)$$

Where j, k are integers and s_0 is a fixed dilation step; the translation factor u_0 depends on the dilation step. The effect of discretizing the wavelet is that the time-scale space is now sampled at discrete intervals. Choosing $s_0 = 2$ and $u_0 = 1$, the sampling of both the frequency and time axis are dyadic. The results of a Discrete Wavelet Transform (DWT) of a continuous signal will be a series of wavelet coefficients, and it is referred as the wavelet series decomposition. As a backward process, from the wavelet series decomposition it is possible to reconstruct the original signal. The necessary and sufficient condition for a stable reconstruction is that the energy of the wavelet coefficients must lie between two positive bounds, such as

$$A\|f\|^2 \leq \sum_{j,k} \left| \langle f, \psi_{j,k} \rangle \right|^2 \leq B\|f\|^2 \quad (4.10)$$

where $\|f\|^2$ is the energy of $f(t)$, $A < 0$, $B < \infty$ and A, B are independent of $f(t)$. If eq. (4.10) is satisfied, the family of basis functions $\psi_{j,k}(t)$ with $j, k \in \mathbb{Z}$ is referred to as a frame with frame bounds A and B [103]. When $A=B$ the frame is tight, and the discrete wavelets behave exactly like an orthonormal basis. If $A \neq B$ the reconstruction is still possible, at the expense of a dual frame, in which different wavelets are used for the decomposition and reconstruction processes. In order to make the discrete wavelets orthonormal, they have to satisfy the following condition

$$\int \psi_{j,k}(t) \psi_{j,m,n}^*(t) dt = \begin{cases} 1 & j = m \text{ and } k = n \\ 0 & \text{otherwise} \end{cases} \quad (4.11)$$

that means each wavelet has to be orthogonal to its dilations and translations. An arbitrary signal can be then reconstructed by summing the orthogonal wavelet basis functions, weighted by the wavelet transform coefficients [104]:

$$f(t) = \sum_{j,k} \gamma(j,k) \psi_{j,k}(t) \quad (4.12)$$

In the representation of the signal, the orthogonality is not essential and the redundancy can actually help to reduce the sensitivity to noise [104] and to improve the *shift invariance* of the transform [105]. The shift invariance of a transform, implies that the transform of a signal and of a time-shifted version of the same signal are not simply shifted versions of each other. This is a disadvantage of the DWT, which generally is not shift invariant. Even using discrete wavelets, an infinite number of scaling and translations is needed, to calculate the wavelet transform of a signal. It is possible to reduce the number of scaled wavelets to be included in the analysis of a signal, still maintaining a meaningful result. An upper bound for the number of wavelet to retain, is given by the duration of the signal under investigation, since the translations of the wavelet cannot go pass that limit. A lower bound can be obtained looking at the frequency spectrum of the analyzing wavelet. One of the properties of the Fourier transform states that compression a signal in time is equivalent to stretching its spectrum and shifting it upwards, as expressed by:

$$F[f(st)] = \frac{1}{|s|} F\left(\frac{\omega}{s}\right) \quad (4.13)$$

Using this insight, the finite spectrum of the investigated signal can be covered with the spectra of dilated wavelets in the same way as its time domain can be covered with translated wavelets. For a good coverage of the signal spectrum, the stretched wavelet spectra should touch each other, as shown in Figure 4.2, and this can be arranged by a correct design of the analyzing wavelets.

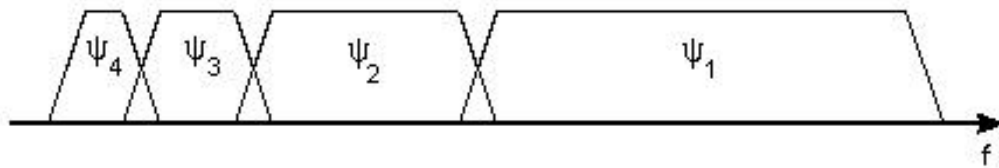


Figure 4.2: Wavelet spectra resulting from scaling of the mother-wavelet in the time domain

From this point of view, considering the wavelet as a band-pass filter, a series of dilated wavelets can be seen as a band-pass filter bank, in which the ratio between the center frequency of a wavelet spectrum and its width, is the same for all the wavelets. This ratio often referred to as the fidelity factor Q of a filter, which in this case names the above decomposition, as *constant- Q* filter bank. It should be noted that covering the time and frequency domains of the investigated signal, is only possible if the signal itself does not have infinite energy; this does not constitute a problem, since natural signals, and in particular pulsed laser generated ultrasonic signals, have finite energy. With the filter bank approach, since every time the wavelet is stretched in the time domain with a factor of 2 its bandwidth is halved, theoretically an infinite number of wavelets will still be needed in order to cover the whole spectrum of interest. The solution was found limiting the wavelet sub-spectra which would leave a small enough gap between his left boundary and zero frequency. A scaling function was introduced by Mallat [98] in order to fill this

gap; from the low-pass nature of the function, it is often referred to as the averaging or low-pass filter. The scaling function itself is composed by wavelets, and it can be expressed by:

$$\varphi(t) = \sum_{j,k} \gamma(j,k) \psi_{j,k}(t) \quad (4.14)$$

with $j \leq \bar{j}$, being \bar{j} the scale of the wavelet spectrum with lowest frequency, considered in the decomposition. The connection between wavelet and scaling function spectra is described in qualitative way, in Figure 4.3

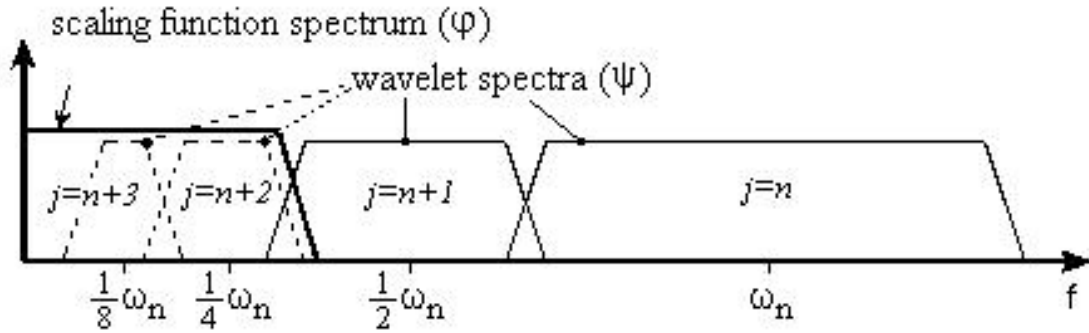


Figure 4.3: Replacement of an infinite wavelet set by one scaling function

The obtained low-pass spectrum, allows the definition of an admissibility condition for the scaling function

$$\int \varphi(t) dt = 1 \quad (4.15)$$

which shows that the 0-order moment of the function itself cannot vanish. Regarding the wavelet transform as a filter bank, then wavelet transforming a signal can be considered as passing the signal through the same filter-bank, which is often known as sub-band coding. The filter bank can be built in several ways. One way is considering many independent band-pass filters to split the spectrum into frequency bands with arbitrary

bandwidths. The drawback of this choice is that every filter would have to be designed separately, implying a very time consuming process. Instead, splitting the signal spectrum in two equal parts, a low-pass and a high-pass part, containing respectively the approximation and the details of the signal, would create an iterated filter bank, and only two filters have to be designed. The drawback of this approach is that the signal spectrum coverage is fixed, but the advantage of the ease of implementation is substantial. On top of that, implementing the wavelet transform as an iterated filter bank, eliminates the need of specifying the wavelet explicitly, which is often impossible. In order to implement the wavelet filter bank, some considerations still need to be made. Since the scaling function could be expressed in wavelets from minus infinity up to a certain scale j , if a wavelet spectrum is added to an existing scaling function spectrum, another scaling function can be obtained, with a spectrum twice as wide as the first. Through this addition, the first scaling function can be expressed in terms of the second,

$$\varphi(2^j t) = \sum_k h_{j+1}(k) \varphi(2^{j+1} t - k) \quad (4.16)$$

also known as *multi-resolution* formulation of the scaling function[105]. Since the first scaling function replaced a set of wavelets, therefore also the wavelets in this set can be expressed in terms of translated scaling functions at the next scale. The wavelet at scale j can then be expressed as

$$\psi(2^j t) = \sum_k g_{j+1}(k) \varphi(2^{j+1} t - k) \quad (4.17)$$

which is the multi-resolution relation between the scaling function and the wavelet.

Since the investigated signal $f(t)$ can be expressed in terms of dilated and translated wavelets up to a scale $j-1$, this means that $f(t)$ can also be expressed in terms of dilated and translated scaling functions at a scale j :

$$f(t) = \sum_k \lambda_j(k) \varphi(2^j t - k) \quad (4.18)$$

If the wavelet analysis is performed only up to the scale $j-1$, then wavelets need to be added in order to keep the same level of detail, and the signal $f(t)$ can be expressed as

$$f(t) = \sum_k \lambda_{j-1}(k) \varphi(2^{j-1} t - k) + \sum_k \gamma_{j-1}(k) \psi(2^{j-1} t - k) \quad (4.19)$$

If the scaling functions $\varphi_{j,k}(t)$ and the wavelets $\psi_{j,k}(t)$ are orthonormal or constitute a tight frame, then the coefficients $\lambda_{j-1}(k)$ and $\gamma_{j-1}(k)$ can be found by taking the inner products

$$\begin{aligned} \lambda_{j-1}(k) &= \langle f(t), \varphi_{j,k}(t) \rangle \\ \gamma_{j-1}(k) &= \langle f(t), \psi_{j,k}(t) \rangle \end{aligned} \quad (4.20)$$

Substituting in eq. (4.20), suitably scaled and translated versions of the scaling functions and the wavelets expressed in their multi-resolution forms of eq. (4.16) and eq. (4.17), and manipulating the results, the following important result can be obtained[105]:

$$\begin{aligned} \lambda_{j-1}(k) &= \sum_m h(m-2k) \lambda_j(m) \\ \gamma_{j-1}(k) &= \sum_m g(m-2k) \gamma_j(m) \end{aligned} \quad (4.21)$$

These equations state that both the wavelet- and scaling function coefficients at a certain scale, can be found by calculating a weighted sum of the scaling function coefficients

from the previous scale. Equations (4.21) are typical expressions of a digital filter; since the coefficients $\lambda_j(k)$ come from a low-pass part of the splitted signal spectrum, the weighting factors $h(k)$ must form a low-pass filter. Similarly, the weighting factors $g(k)$ must form a high-pass filter. The former are defined the scaling filter, while the latter are known as the wavelet filter.

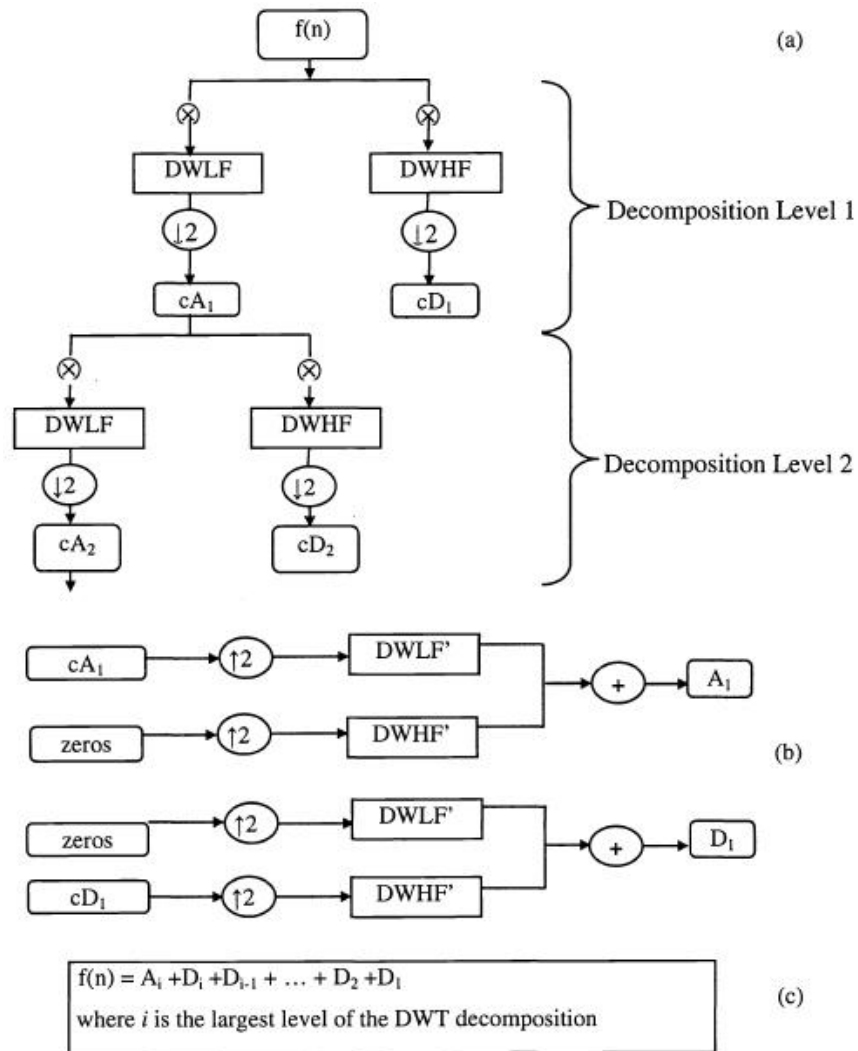


Figure 4.4: (a) Discrete Wavelet filter-bank decomposition; (b) signal reconstruction from wavelet coefficients; (c) reconstruction of original signal

The most important property of these filters is that they have both a step-size of 2 in the variable k , and this is known as the subsampling property. As a consequence of this, the convolution will include only every other $\lambda_j(k)$, with the results that the output data rate is equal to the input data rate. At every iteration step, the number of samples for the next stage is halved, so that the last iteration in theory, will end with just one sample. Usually the iteration will stop at the point where the number of samples is smaller than the length of the scaling or wavelet filters, whichever is longest. The discrete wavelet decomposition based on the iterated filter-bank algorithm is schematized in Figure 4.4, where A_n and D_n are the approximations and details at level n . The inverse reconstruction process is also shown, performed through a backward process, involving the reconstruction filters and the up-sampling stages.

4.2.3 DWT for Data De-noising and High Speed Data Manipulation

The DWT is useful for its data de-noising and data compression abilities. De-noising and compression of the original signal can be achieved if only a few wavelet coefficients representative of the signal are retained and the remaining coefficients, related to noise, are discarded [106, 107, 108]. An example of DWT processing to an ultrasonic signal is illustrated in Figure 4.5; the ultrasonic wave was generated illuminating a rail with a pulsed laser source, and it was acquired by an air-coupled sensor, positioned at some distance away from the source. The unprocessed signal recorded during the test is shown in Figure 4.5(a). Two ultrasonic signatures are visible. The first one is the direct signal propagating from the laser excitation point; the second

ultrasonic signature (echo signal) is the surface wave reflected from a 5 mm notch positioned 500 mm away from the sensor.

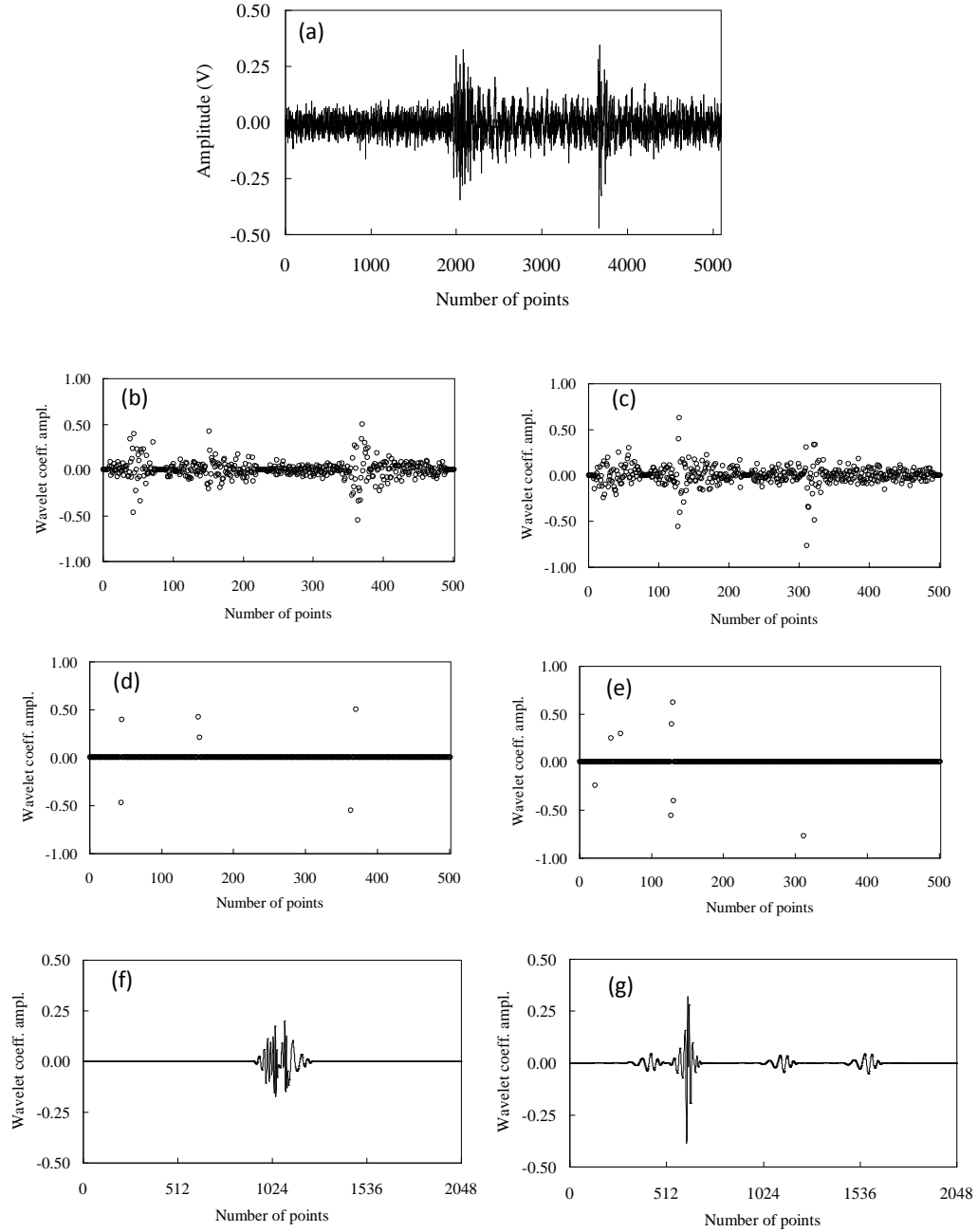


Figure 4.5: Example of discrete wavelet processing. (a) Unprocessed signal; (b), (c) wavelet coefficient vector of levels cD3, cD4, and cD5 relative to the direct and the echo signal, respectively; (d), (e) thresholded wavelet coefficient vectors in (b) and (c), respectively; (f) and (g) reconstructed direct and echo signals, respectively, using the thresholded wavelet coefficient vectors in (d) and (e).

The time interval was divided into two windows, detailed in Figure 4.5(a), containing the direct signal and the echo signal, respectively. For each window, the wavelet coefficients at detail levels 3, 4 and 5 only were considered, since these three levels contained the frequency band of interest. Figure 4.5(b) and (c) show the clustered coefficients for each of the two windows, namely the direct signal and the echo signal, respectively. Most of these coefficients have very low amplitude and thus do not carry any information on the signal of interest. These coefficients can be discarded if not exceeding a given threshold, which is related to the maximum amplitude modulus of each decomposition level. Figure 4.5(d) and (e) are the result of a threshold setting equal to 70%, 50% and 80% of the maximum coefficient modulus of levels 3, 4 and 5, respectively. Signals reconstructed from the thresholded wavelet coefficient vectors are shown in Figure 4.5(f) and (g) and they demonstrate that only a few wavelet coefficients are indeed representative of the ultrasonic signatures of interest. The small waves reconstructed in Figure 4.5(g) should be ignored compared to the predominant defect echo. In order to reconstruct the direct signal decomposed by the DWT, only 6 coefficients are needed, compared to the 2048 points of the entire first gated window of the original signal. Although the mother wavelet and the filter bank coefficients have to be stored once in the algorithm, the reached level of compression ($6/2048 \times 100 = 2.9\%$) is noticeable. In systems that require time-critical execution of data acquisition, processing and delivery of the results, the use of the DWT is highly efficient. As demonstrated, the DWT is able to give insight in the time-frequency domain of the acquired signal as the CWT does, while performing at speed that are orders of magnitude faster than the CWT; the amount of data to be handled and delivered to consuming processes is also

substantially smaller. The success of a proper DWT decomposition is highly dependent on a correct choice of the mother-wavelet. The pruning and compressing performances of the decomposition are consequently a function of the chosen analyzing wavelet. A mother wavelet that matches the shape of the signal that is being analyzed is a perfect candidate for a good DWT decomposition. It is also confirmed in recent literature [43], that the best mother wavelet is the one that generates the smallest number of nonzero wavelet coefficients when decomposing a given signal, or in other words, the one which maintains the closest shape to the signal. Other practical issues concerning the implementation of the DWT include a correct choice of the sampling rate and length of the signal to be decomposed. Due to aliasing effects, the sampling rate determines the highest wavelet scale s_{MAX} that can be meaningfully interpreted [109]; if the acquisition time step is equal to Δt_s , then s_{MAX} is given by:

$$s_{MAX} = -\log_2(2\Delta t_s) \quad (4.22)$$

Although Nyquist sampling theorem is respected, for consistent results, oversampling is necessary, and the current practice is to use sampling rate of 6-10 times the signal frequency of interest [109]. The length of the signal instead, controls the lowest wavelet scale that produces meaningful coefficients, and it is dependent on the chosen mother-wavelet. Considering the Gabor wavelet for example, given the extension of its support region, the lowest scale of interest should be greater than $\log_2(5/T_L)$ [109]. Since the DWT is not shift invariance, as stated in the previous section, the position of significant points within the signal is critical for a good decomposition. The truncation of a finite signal, originates unwanted effects in proximity of the signal ends, called end-effects. It

can be demonstrated that at least 2^{1-p} samples have to be present after the last point of interest and before the end of the signal, for a good wavelet decomposition at p scale [109].

4.3 Digital processing of ultrasonic signals

4.3.1 Digital filtering

In the implementation of the defect detection prototype, the need of looking into different frequency ranges of the acquired signals, to inspect different areas of the rail section, sometime required more freedom in the choice of the filters, than what it is allowed by the filter-bank decomposition. The calibration the D.I. computation, which implies the choice of the correct number of wavelet coefficients to be retained for each of the considered scale, can be a time demanding procedure. This procedure could be avoided or at least simplified, in particular at the time of debugging or calibration of the prototype. A solution was found in the use of digital filters.

With the drawback of not contributing to the compression of the acquired signals, the digital filters have the great advantage of giving complete freedom for setting the denoising parameters. Once a digital filter and its order are chosen, the only variables that need to be set are the cut-on and cut-off frequencies, for a case of a pass-band filter, and they can be chosen along the entire frequency spectrum up to the Nyquist frequency. Experimental tests on the prototype suggested that the use of the third order Butterworth pass-band filter was successful. Laboratory results in terms of Damage Index (D.I.) computation for known defects in the rails available at the UCSD NDE & SHM lab, will be shown later in the chapter.

4.3.2 Spatial averaging algorithm for enhancement of defect detection in “reflection mode”

In this section, a new approach for detecting defects in rails was investigated. As opposed to deploying two receiving sensors in the traditional detection scheme, a series of multiple sensors was located at one side from the laser excitation, as shown in Figure 4.6.

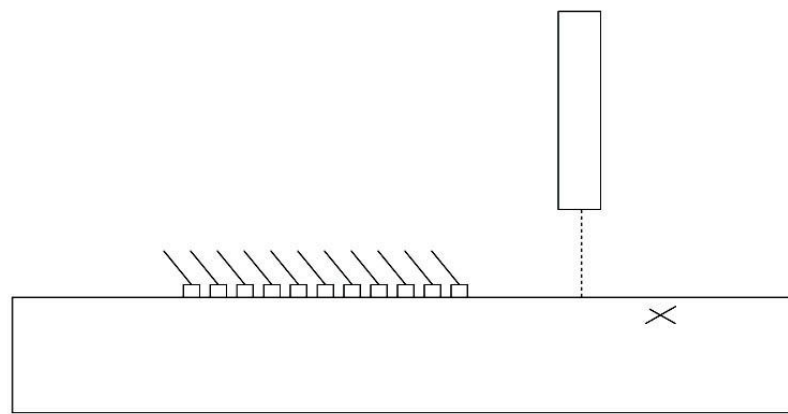


Figure 4.6: Detection scheme for coherent acquisition of guided waves reflected by a flaw in the rail head

The laser excited guided waves in the rail in the traditional fashion, and all the transducers were oriented towards the source, waiting for a reflection from a defect located downstream along the rail. The laser provided a 2" (≈ 51 mm) wide line beam across the top of the rail head, inducing a symmetrical excitation. As discussed in the previous chapter, the propagation of surface waves is confined to a superficial region of the rail head of a thickness compared with the wavelength. The broad excitation produced by the laser generates guided waves with a wide frequency spectrum, and indeed, different penetration depths. Two cases of known defects were tested in this

configuration. The first was a surface defect with a 5% H.A. reduction, 8 mm deep, while the second was an internal defect with a 23% H.A. reduction. To prove the validity of the new approach, contact sensors were used, with a better SNR than the air-coupled transducers. The Krautkramer ultrasonic contact transducer centered at 2.25 MHz was employed, with a 90 deg wedge, to receive a surface wave traveling along the rail. Gel was used in the interfaces sensor/wedge and wedge/rail, to increase the acoustic coupling of the sensing parts. As it will be demonstrated later in this section, the use of a high number of sensors improves the system sensitivity to the detection of defects that are usually difficult to catch, due to their small size, orientation or depth within the rail head [110]. A set of 12 transducers, with a constant spacing of 0.5" (≈ 13 mm) was chosen for the tests. Due to the unavailability of the 12 sensors, the test setup was reproduced by using a spatial averaging method [111]. This was done by shifting the ultrasonic signatures of adjacent positions, moving only one sensor along the probe, within the 6" (≈ 152 mm) long sensing area ($\Delta=0.5"$ (≈ 13 mm), $N=12$). A picture of the test setup is shown in Figure 4.7.

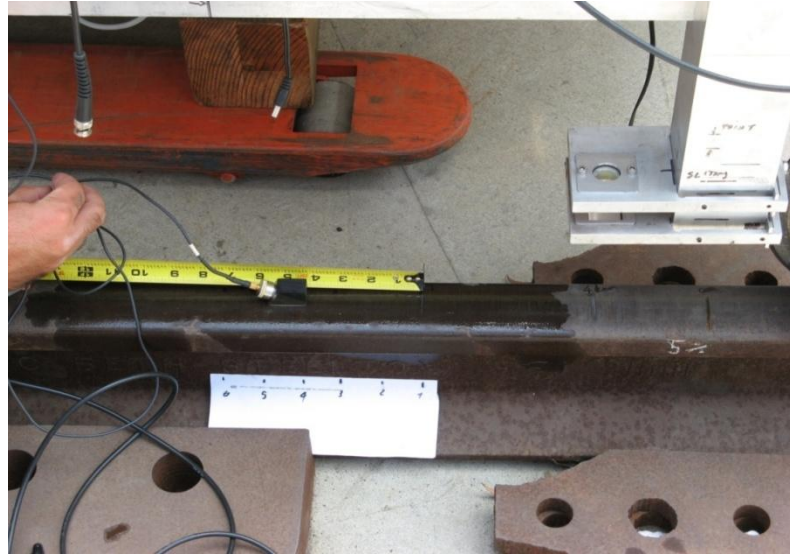


Figure 4.7: Test setup: laser excitation and transducer spacing in proximity of the 5% H.A. reduction surface defect

The first test was performed in proximity of a surface defect. The defect was custom made saw-cutting the top of the rail-head for a depth of 8 mm. The excitation line-beam was 4" (≈ 102 mm) away from the defect; the closest position of the sensor was 9" (≈ 229 mm) away from the defect, as depicted in Figure 4.8.

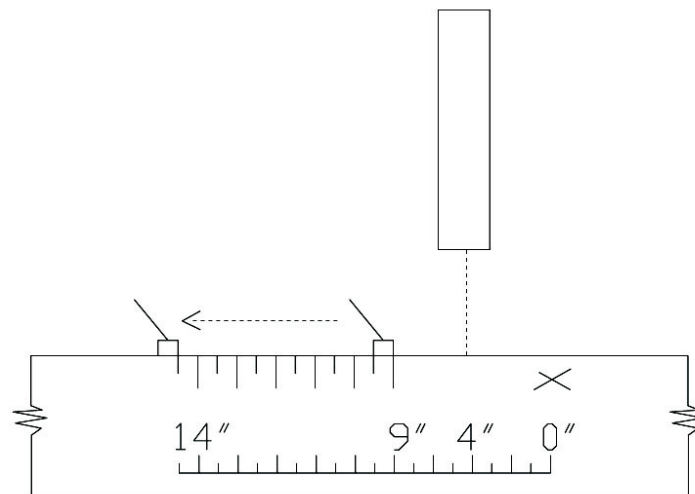


Figure 4.8: Probe spacing and laser-beam position relative to the defect

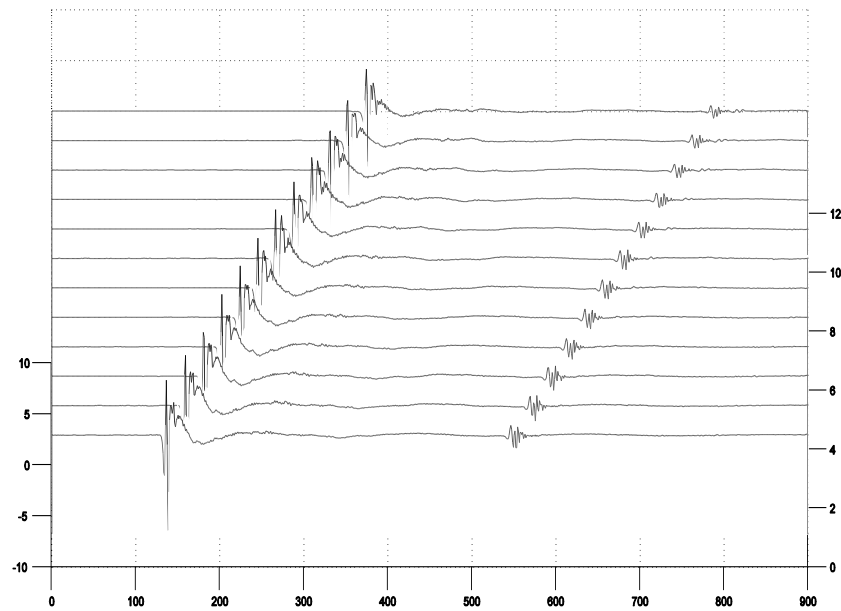


Figure 4.9: Waterfall plot of ultrasonic signals acquired in proximity of the 8 mm deep surface defect

Figure 4.9 shows a waterfall graph of the ultrasonic signatures obtained after moving the transducer over the 12 positions. Looking at the graph, the surface defect can be easily detected due to the clear presence of a reflected wave after the first arrival. The waterfall graph also shows that the energy of the reflected wave attenuates from the closest to the furthest position. A better view of the relative amplitude of the reflected waves is given in Figure 4.10, where a normalized mesh of the A-scans is depicted, excluding the first arrival from the plot (mesh starts after the first 400 points). Using a threshold-crossing method, it is possible to compute the difference in arrival times between each reflected wave and the adjacent ones. The scan-rate of the acquisition board was set at 5 MS/s; the computed average difference between the arrival of the reflected waves is 22 points, which means $4.4 \mu\text{sec}$ in terms of time.

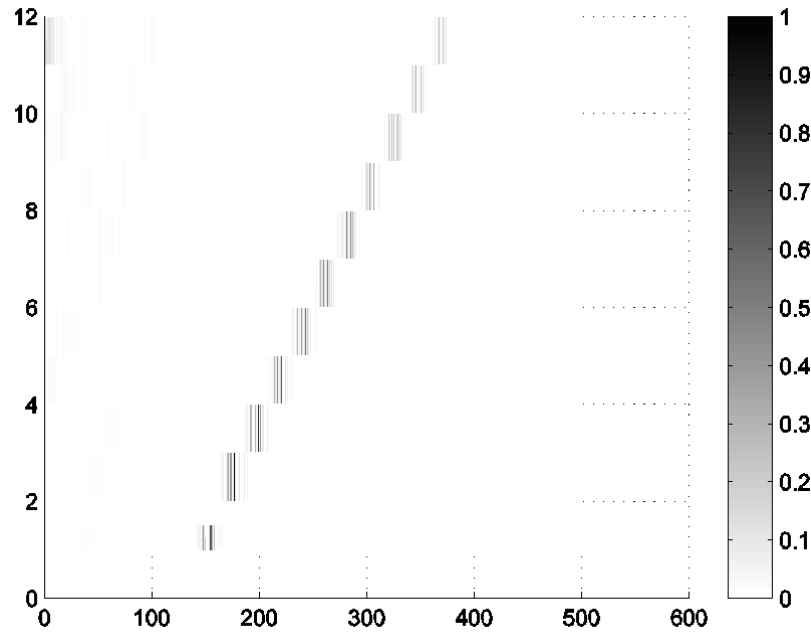


Figure 4.10: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 8 mm deep surface defect

Since the spacing of the transducer between the different acquisitions was set to $0.5''$ (≈ 13 mm) the group velocity of the reflected wave can be easily computed equal to $2,890$ mm/ μ sec. A time-scale analysis of the acquired signals, gives insight on the frequency content of the first arrival and the reflected wave as well. Figure 4.11 shows the Continuous Wavelet Transform (CWT) of the first acquired signal, obtained using the Complex Morlet 1-1 as a mother wavelet. It is clear that the reflected energy is confined to a frequency range above 350 kHz; this is consistent to the depth of the flaw, which is about 8 mm. The multiple acquisition technique was useful to provide further information about the attenuation of the propagating wave and to give accuracy on the measurement of the group velocity of the reflected wave. No spatial averaging was applied here, since it was not necessary for detecting this surface defect.

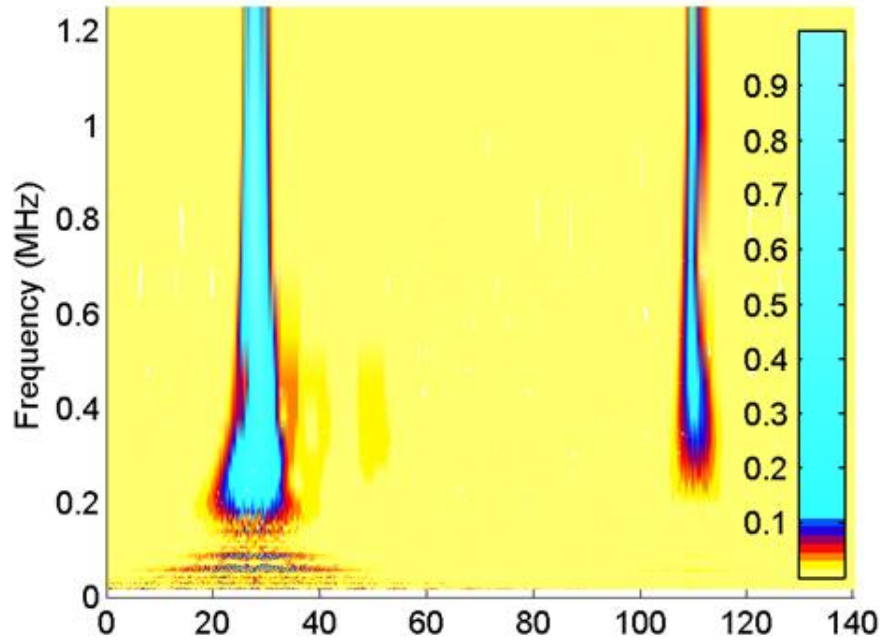


Figure 4.11: CWT of the first A-scan acquired in proximity of the 8 mm deep surface defect

The second test was performed in proximity of an internal defect equivalent to a 23 % H.A. reduction. The excitation line-beam was kept at 4" (≈ 102 mm) from the defect; the closest position of the sensor was 9" (≈ 229 mm) from the defect (Figure 4.8). The sensor was then moved over the remaining 11 positions, spaced 0.5" (≈ 13 mm) one from each other. In Figure 4.12 a waterfall of the 12 A-scans signals is depicted. As it can be deducted by looking at the saturated first arrival, the energy of the laser line-beam was here increased to generate larger guided wave amplitudes and to better detect the presence of the internal defect. It is hard to pinpoint the presence of any reflected wave just looking at Figure 4.12.

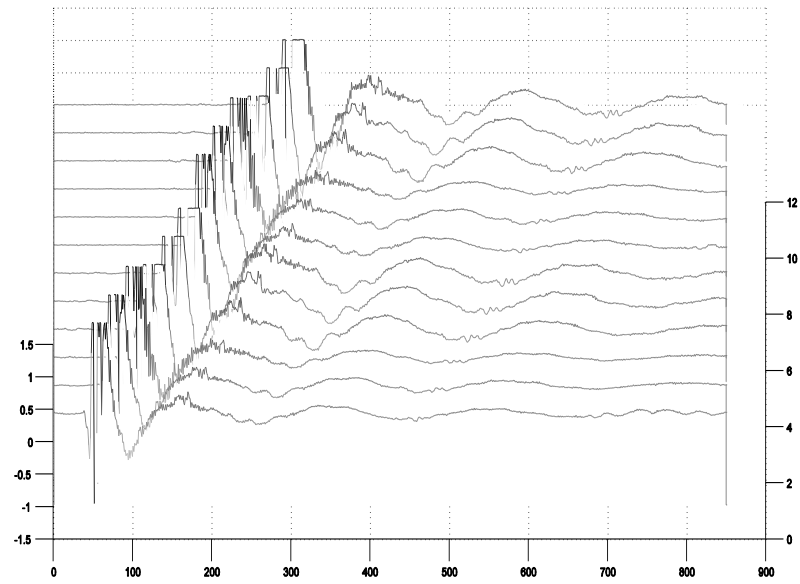


Figure 4.12: Waterfall plot of the 12 A-scans acquired in proximity of the 23% H.A. reduction internal defect

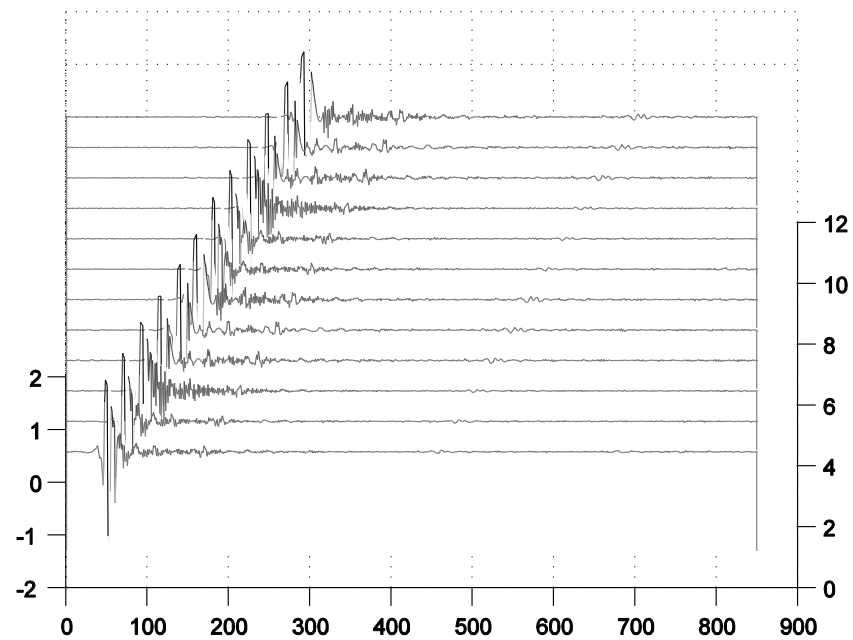


Figure 4.13: A-scans acquired in proximity of the third internal defect, high-passed at 100 kHz

An interesting plot is shown in Figure 4.13, where the A-scans were high-passed using a Butterworth filter with a cut-off frequency of 100 kHz. Eliminating of the low frequency content makes the signals much cleaner, but it is still not trivial to find out whether there is any waveform being reflected from the internal defect. A zoom on the signals excluding the first arrival (starting after the first 400 points) contains some interesting information about the presence of a small reflection which can be recognized among the incoherent noise (Figure 4.14). In Figure 4.15, a normalized mesh of the same extracted signals shows a better view of a wave coherently spaced between the different acquisitions which could be the one reflected back from the internal defect.

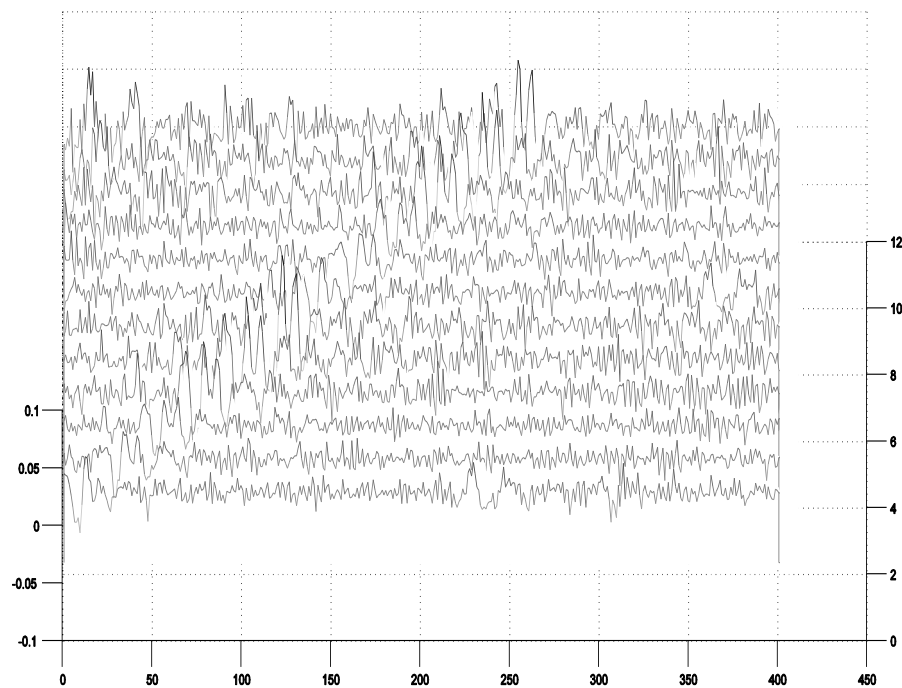


Figure 4.14: Zoom in, of the waveforms reflected from the 23% H.A. reduction internal defect

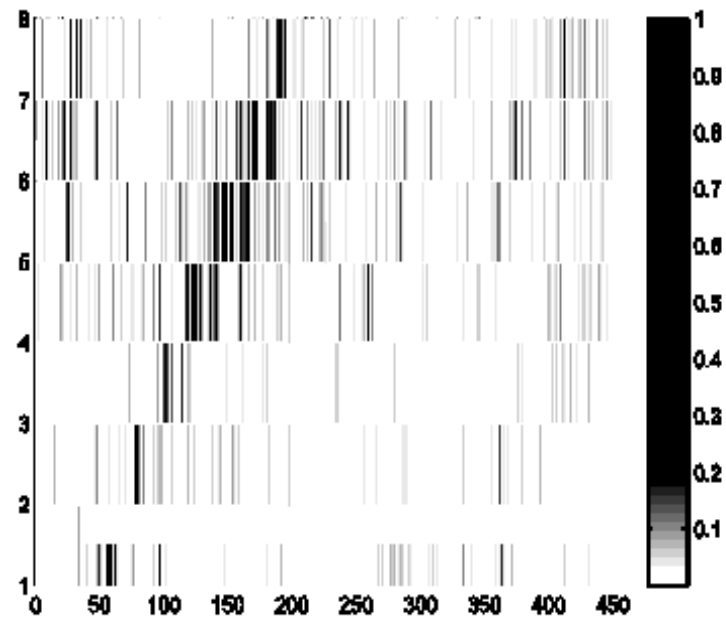


Figure 4.15: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 23% H.A. reduction internal defect

The process of spatial averaging consists of 2 operations. First, the A-scans of adjacent positions are phase-shifted such that the individually received wavefronts are synced. The delays were calculated looking for the maximum of the cross-correlation of the A-scans extracted excluding the first arrival. Considering the 0.5'' (≈ 13 mm) spacing, once again the group velocity of the reflected wave was calculated being 2,890 mm/ μ sec, as for the case of the surface defect in test #1. The second step consists in taking a fixed number N of phase-shifted A-scans for each one of the probe positions and in summing them up. This procedure allows to increase the strength of the reflected signal and to cancel out unwanted noise. This noise is due to the presence of multiple modes which have been excited by the laser. During the summation the unwanted modes are added not-in-phase, so their intensity substantially decreases relatively to the strength of the

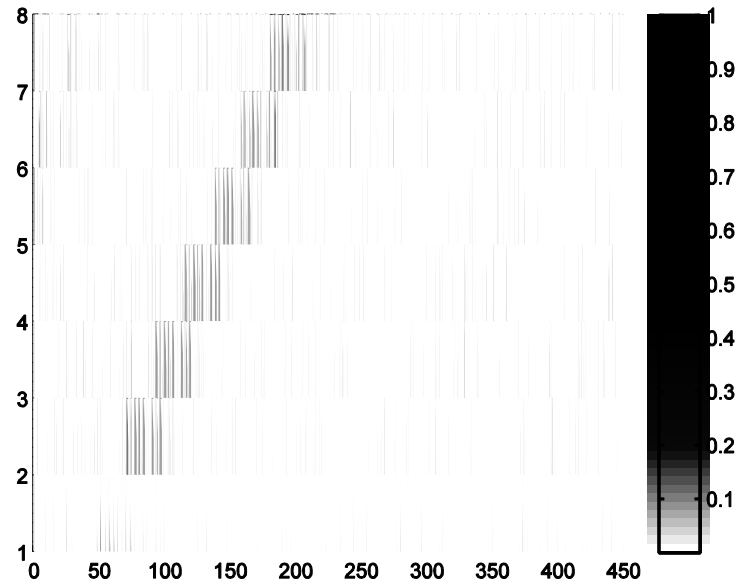


Figure 4.16: Mesh of the reflected waves extracted from the 12 A-scans acquired in proximity of the 23% H.A . reduction internal defect, after the application of the spatial averaging technique ($N=5$, spacing $0.5'' \approx 13 \text{ mm}$)

reflected wavefronts, which are added in phase. Part of the noise is also due to the reverberation of the ultrasound in the probe. The mesh of the A-scans, after spatial averaging them ($N=5$) is shown in Figure 4.16. In this case the reflected wavefront can be seen much clearly. To get a deeper insight on the problem, a time-scale analysis of the first A-scan was performed. Figure 4.17 and Figure 4.18 show the resulting scalograms, respectively adopting the Gabor and the Complex Morlet 1-1 mother-wavelets. In both the plots the reflected wave is well visible, with an arrival time of $125 \mu\text{sec}$. Considering the calculated group velocity of the reflected wave equal to $2,890 \text{ mm}/\mu\text{sec}$, and that the difference in traveled paths between the first arrival and the reflection is $8'' (\approx 203 \text{ mm})$, it is easy to identify that the reflected mode is the one indicated by the left arrow.

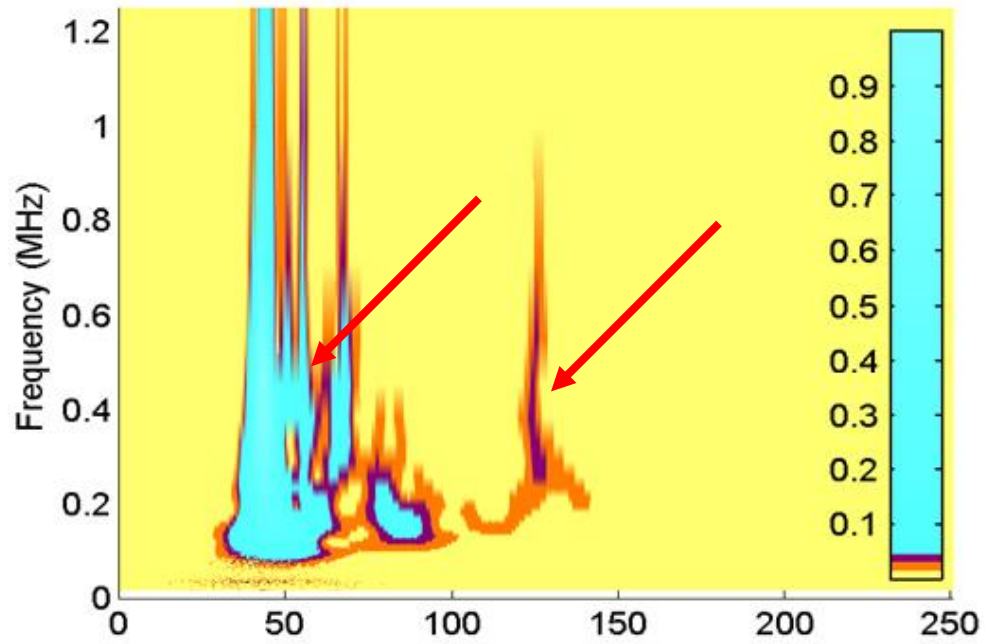


Figure 4.17: CWT of the first A-scan using the Gabor motherwavelet

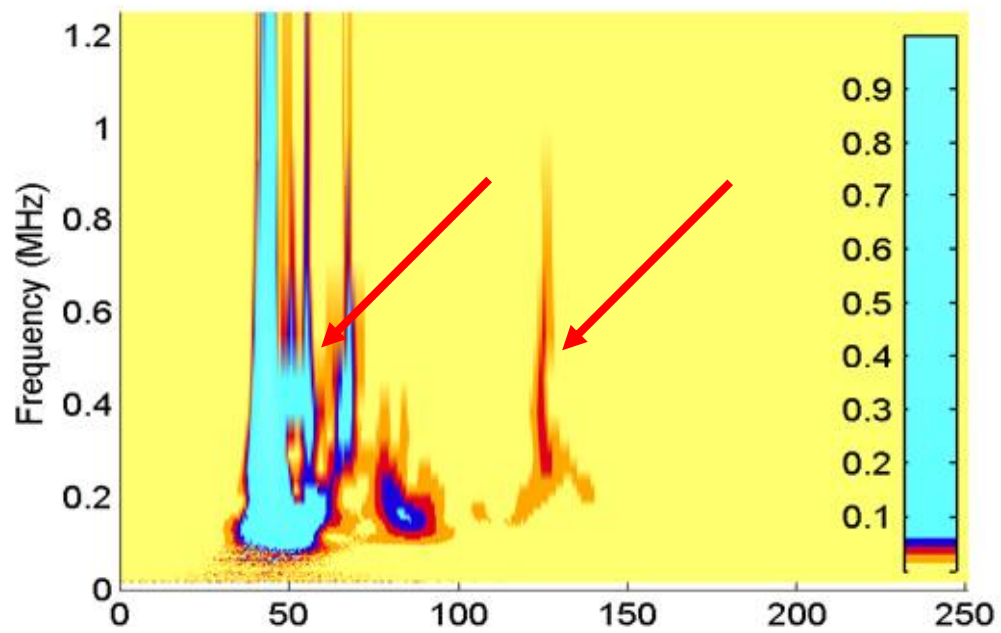


Figure 4.18: CWT of the first A-scan using the Complex Morlet 1-1 motherwavelet

The frequency of the reflection ranges between 200 kHz and 600 kHz; considering the propagating speed of the wave, the penetration depth for that frequency range is between 5 and 15 mm from the top of the rail head. In the past, the internal defect was mapped using a Krautkramer Ultrasonic Testing Machine. An approximate plot of the defect is shown in Figure 4.19, where the defect depth appears to be ranging between 5 and 20 mm. Since the defect is extending only at the gage side of the rail head, the area with high energy visible in the scalogram at about 40 μ sec can be associated to the presence of propagating modes that transmit through the part of the rail head clear from the flaw (field side) and that are not reflected. SAFE results for cross-sectional strain energy for

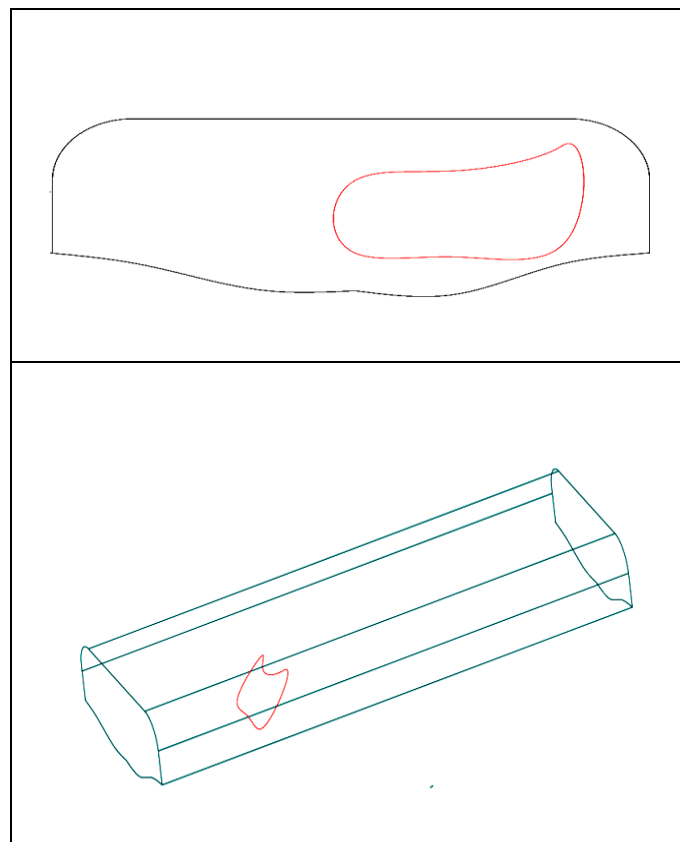


Figure 4.19: Front and 3-D view of the rail head with the 23% H.A. reduction internal flaw

the lowest order selected symmetric vibrational modes at propagation frequency of 200 kHz were shown in Table 3.1 and Table 3.2 in the previous chapter. It is clear that at this frequency range the energy of the selected modes is confined to a region with a depth up to about 10 mm from the top of the rail head. This confirms that the modes excited in the rail are surface-wave type. Also, the strain distribution confirms that these modes can be very sensitive to the defects that have been tested.

4.4 Feature extraction in ultrasonic rail inspection

4.4.1 Joint time-frequency domain feature extraction

An efficient damage index must be robust against noise and must allow for the detection, as well as for the sizing of the defect. These performances can be achieved by the appropriate selection of the signal features used for the damage index computation [112]. With reference to Figure 4.20, if the defect detection is based on reflection measurements (reflection mode), the proposed Damage Index (D.I.) uses the ratio between certain features of the echo signal $F_{reflection}$ and the same features of the direct signal, F_{direct} :

$$DI = \frac{F_{reflection}}{F_{direct}} \quad (4.23)$$

Clearly, the D.I. is expected to increase with increasing defect size or the % of head area reduction (% HA). Alternatively, if defect detection is based on transmission measurements (transmission mode), the proposed D.I. uses the ratio between certain features of the signal detected by the further sensor #2, F_{sens2} over the same features from the closer sensor #1, F_{sens1} :

$$DI = \frac{F_{sens2}}{F_{sens1}} \quad (4.24)$$

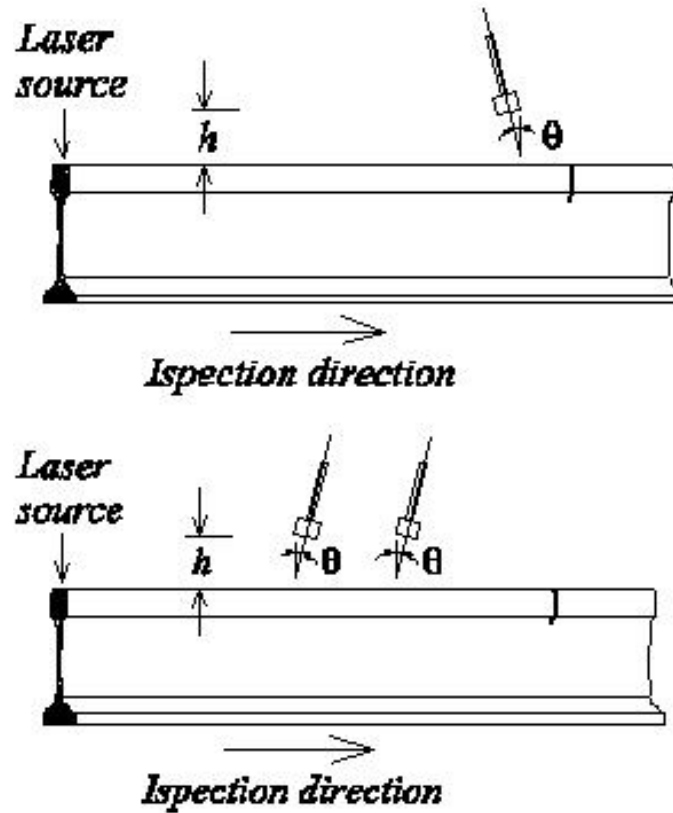


Figure 4.20: Defect detection schemes examined. “Reflection mode” with a single air-coupled sensor (top); “transmission mode” with a pair of air-coupled sensors (bottom).

Laboratory tests were made in order to analyze the performance of different features in the computation of the D.I.. The light source was a Q-switched, Nd:YAG pulsed laser operating at 1064 nm with an ~ 8 ns pulse duration. Through conventional optics the laser beam was focused on the rail head to a 20 mm long line to excite high-frequency surface waves travelling along the rail running direction. The present investigation was focused to detecting small cracks on the order of few millimetres in depth; thus only high-frequency waves were considered by using a 50 kHz high-pass filter. Micro-machined, air-coupled capacitive transducers were used for signal detection with a 76.2 mm (3") lift-off distance. As traditionally done with conventional wedge

transducers, the alignment angle of the air-coupled detectors was adjusted to maximize the sensitivity to the guided waves. The optimum air-coupled detection angle from the normal to the rail surface θ , is given by Snell's Law of refraction:

$$\theta = \arcsin\left(\frac{c_{air}}{c_p} \sin \theta_p\right) \quad (4.25)$$

where c_p is the phase velocity of the guided wave in the rail, $c_{air} = 330$ m/s is the wave velocity in air and $\theta_p = 90^\circ$ for a guided wave propagating parallel to the rail surface. Considering that high-frequency nondispersive waves, surface waves, propagate at $c_p = 3,000$ m/s, eq. (1) gives $\theta = 6.3^\circ$. The convention here is positive angle towards the defect (ie away from the generating laser light). Thus $\theta = 6.3^\circ$ is the optimum orientation to detect crack reflections. Similarly, $\theta = -6.3^\circ$ is the optimum orientation to detect the first arrival of the wave propagating from the laser source. In an effort to examine the use of both reflection and transmission measurements for crack detection, three different tests were carried out on two 2100 mm (7 ft) long, 115-lb A.R.E.M.A. rail sections donated by San Diego Trolley, Inc. The test set-ups are summarized in Table 4.1. In the first test (Test #1), a transverse crack located at the centre of the rail head top surface (centre crack) was monitored by reflection measurements with a single sensor oriented towards the notch. The same sensor orientation was used to monitor a transverse crack located in the gauge-side corner of the rail head, again by reflection measurements (Test #2). Finally, in Test # 3, a pair of air-coupled sensors oriented towards the laser source was used to monitor the centre surface crack by transmission measurements. The cracks were simulated by narrow notches that were machined at depths ranging from a minimum depth of 0.5 mm to a maximum depth of 10 mm. The relation between the defect depth

and the corresponding cross-sectional area reduction of the rail head (% HA reduction) is illustrated in Figure 4.21 for both the centre crack and the corner crack. The largest crack depth of 9 mm for the centre crack resulted in a 20% HA reduction.

Table 4.1 Experimental parameters for the three tests performed (L_d = distance laser source – crack;
 L_1 = distance laser source-air-coupled sensor)

Test #	θ	L_d (mm)	L_1 (mm)	Crack depth (mm)	Sensors used	Defect type
1	+6°	1000	500	0 to 9	1	Centre crack
2	+6°	650	350	0 to 10	1	Corner crack
3	-6°	1000	700	0 to 8.5	1 and 2	Centre crack

A number of signal features were considered to construct the D.I. for the reflection and the transmission modes. First appropriate DWT thresholds were selected to maximize the de-noising of the defect signatures.

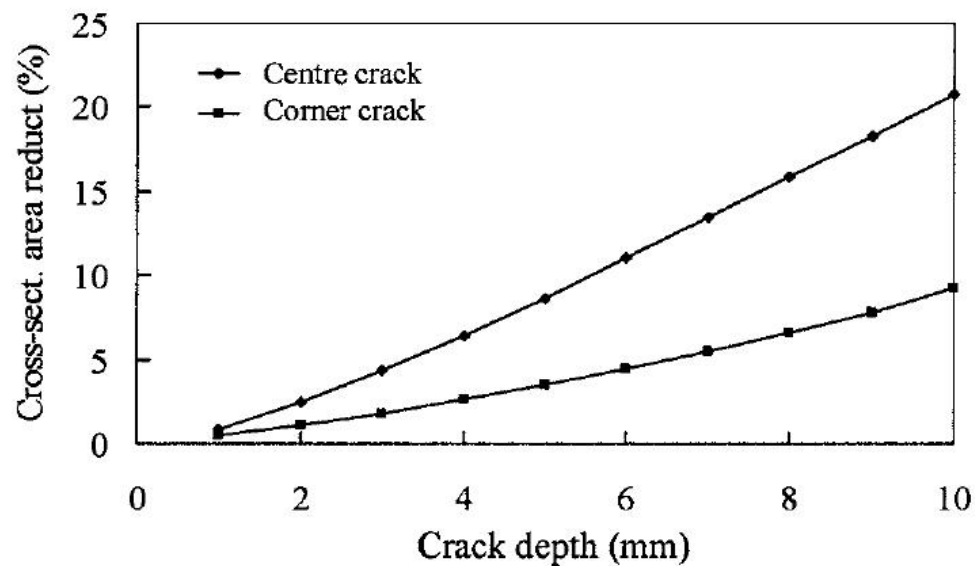


Figure 4.21: Cross-sectional head area reduction as a function of the defect depth for the centre crack and for the corner crack

The features of variance, root mean square (RMS), kurtosis, peak amplitude and peak-to-peak amplitude of the thresholded wavelet coefficient vectors were computed. The direct and echo signals were then reconstructed from the corresponding wavelet coefficient vectors and the maximum and the peak-to-peak amplitudes were calculated. Both reconstructions were then processed through the Fast-Fourier Transform (FFT) and the Hilbert Transform (HT). The following additional features were finally extracted: the peak amplitude and the area below the FFT frequency spectrum within the frequency range of 100 kHz – 900 kHz, the area below the HT, the maximum amplitude and the corresponding arrival time of the HT. All but one of the selected features were related to the size of the crack; the arrival time computed from the HT maximum amplitude was related to the location of the crack from the sensing system. Eleven different damage conditions, ranging from no defect up to a 9 mm crack depth, were monitored during Test #1 by using a single sensor optimally oriented to detect the wave reflected from the crack. It was previously demonstrated that this set-up, once aided by the DWT processing, can detect surface-breaking cracks in rails at depths as shallow as 1 mm and located 500 mm away from the sensing system [113]. The results presented here were obtained using the following threshold values for the wavelet coefficients: 60% for levels 3 and 4, and 100% for level 5. This selection was found adequate to denoise the defect signatures under the test set-up adopted. Figure 4.22 shows the D.I. results calculated in Test #1 from the variance of the wavelet coefficient vector, plot (a), from the area below the FFT frequency spectrum of the reconstructed signal, plot (b), and from the maximum amplitude of the HT of the reconstructed signal, plot (c).

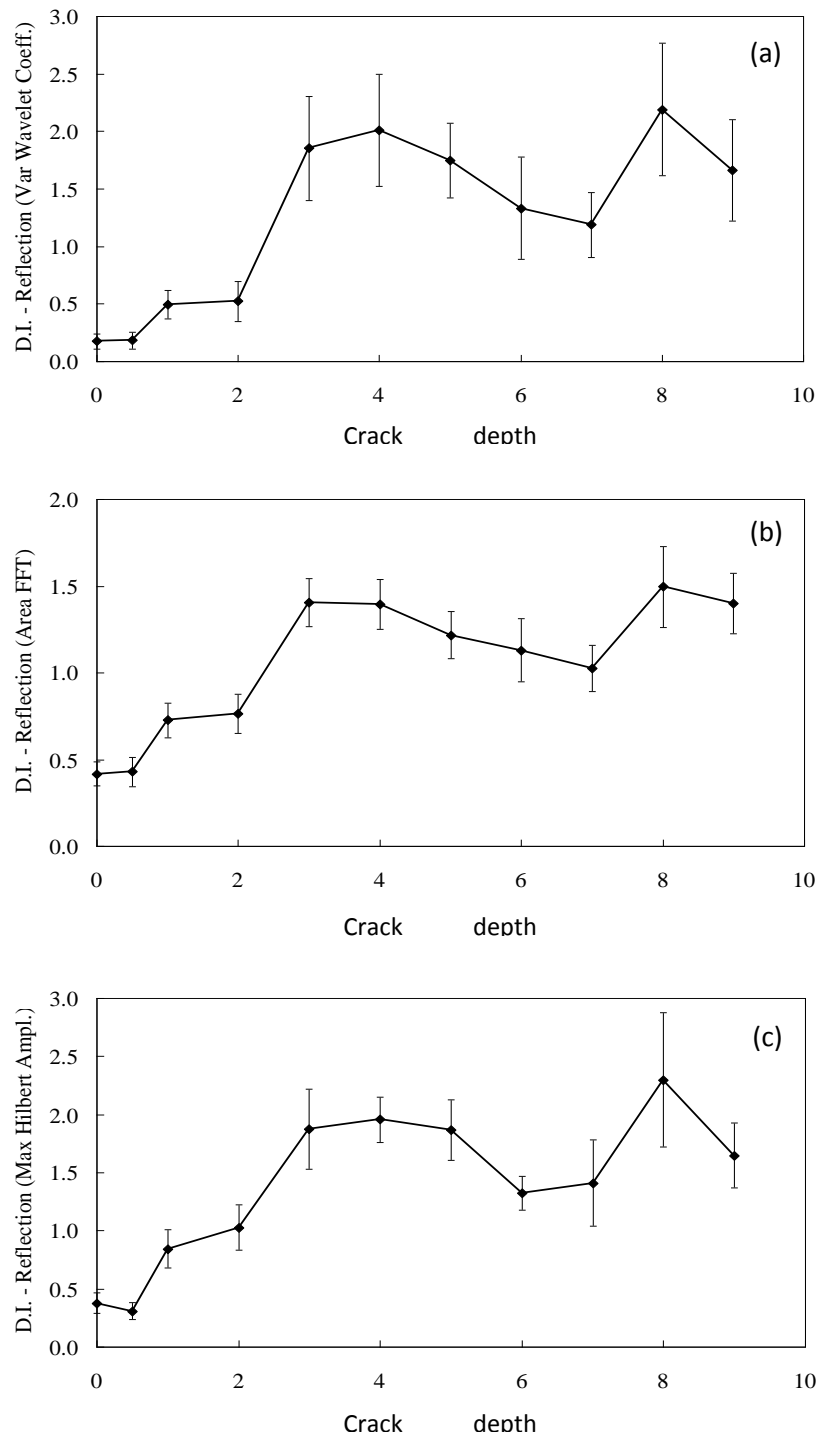


Figure 4.22: Feature-based Damage Index as a function of the center crack depth monitored in the “reflection mode”; (a) variance of the wavelet coefficients; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum amplitude of the Hilbert transform of the reconstructed signal

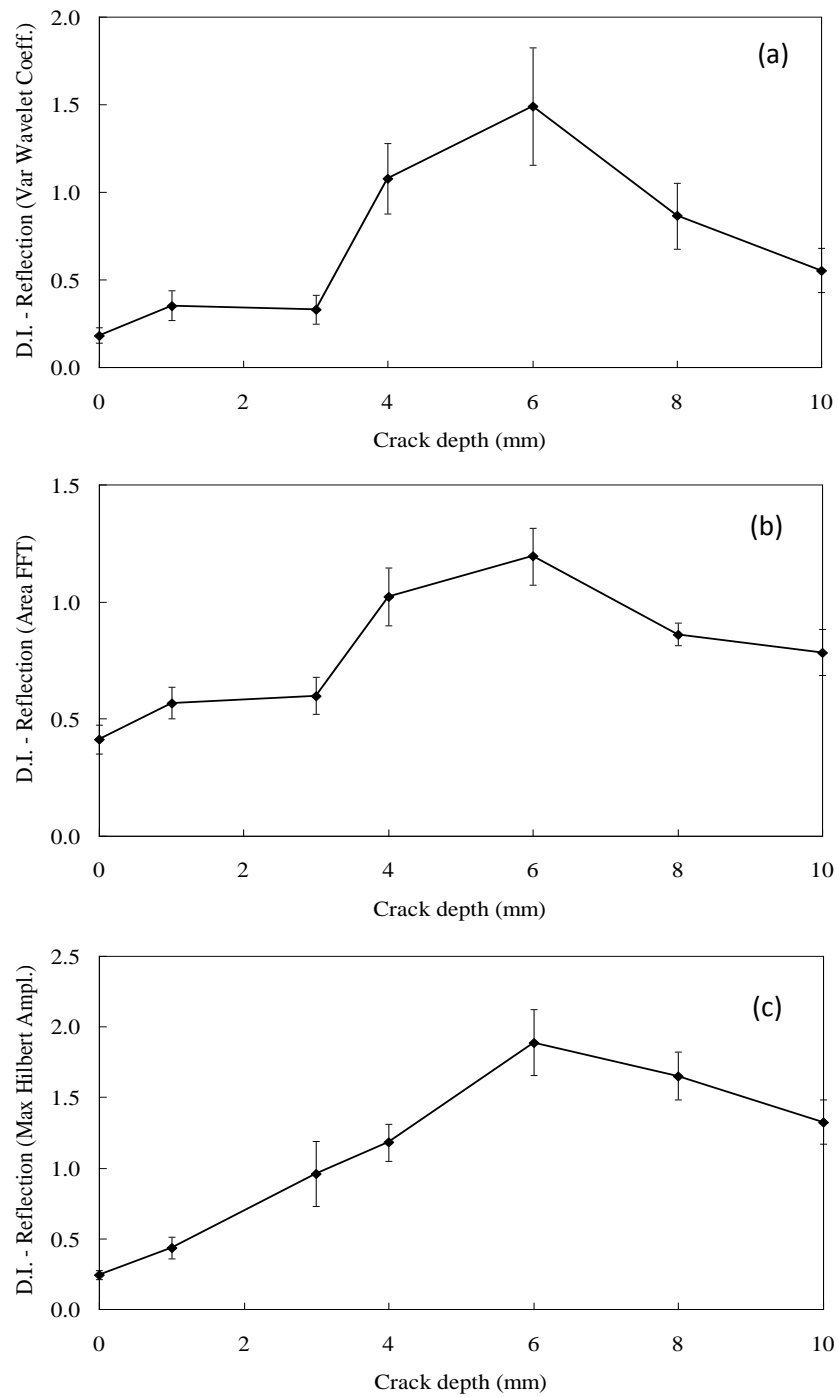


Figure 4.23: Feature-based Damage Index as a function of the corner crack depth monitored in the “reflection mode”; (a) variance of the wavelet coefficients; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum amplitude of the Hilbert transform of the reconstructed signal

The plots show the average results of ten acquisitions along with the $2\sigma_{\text{dev}}$ variations, where σ_{dev} is the standard deviation of ten measurements. It can be seen that the D.I. values monotonically increase with increasing crack depth until $s = 4$ mm, after which the values decrease and then increase again at $s = 8$ mm. It should be also noted that the D.I. values calculated for defects larger than 1 mm are clearly above the D.I. values calculated for the pristine structure and for $s = 0.5$ mm which corresponds to a head area (HA) reduction below 0.8%. Similar analyses were carried out for the oblique, corner crack (Test #2). The same feature-based indexes are presented in Figure 4.23, which illustrates seven different crack sizes between 0 and 10 mm. The general trend is similar to what found for the centre crack, with D.I. increasing from the no damage baseline and then decreasing for defect depths larger than $s = 6$ mm. Also, the D.I. values calculated for the corner crack, are generally smaller than those calculated for the centre crack at the same depth. This is an expected result since the % HA reduction is smaller for the corner cracks at the same depth, as previously illustrated in Figure 4.21. It should be noted that for the pristine cases the D.I. should be equal to zero, since the crack reflection is absent. The nonzero D.I. values reported at $s = 0$ in Figure 4.22 and Figure 4.23 are the result of the wavelet threshold that is relative to the maximum coefficient modulus. Thus, some wavelet coefficients are always retained. The nonzero values should be considered as baselines against which all D.I. from potential defect signatures are compared. In summary, the results in Figure 4.22 and Figure 4.23 are encouraging for defect detection and sizing, particularly considering that the cracks being targeted are smaller than 20% HA reduction, the distance at which they are being targeted is as large

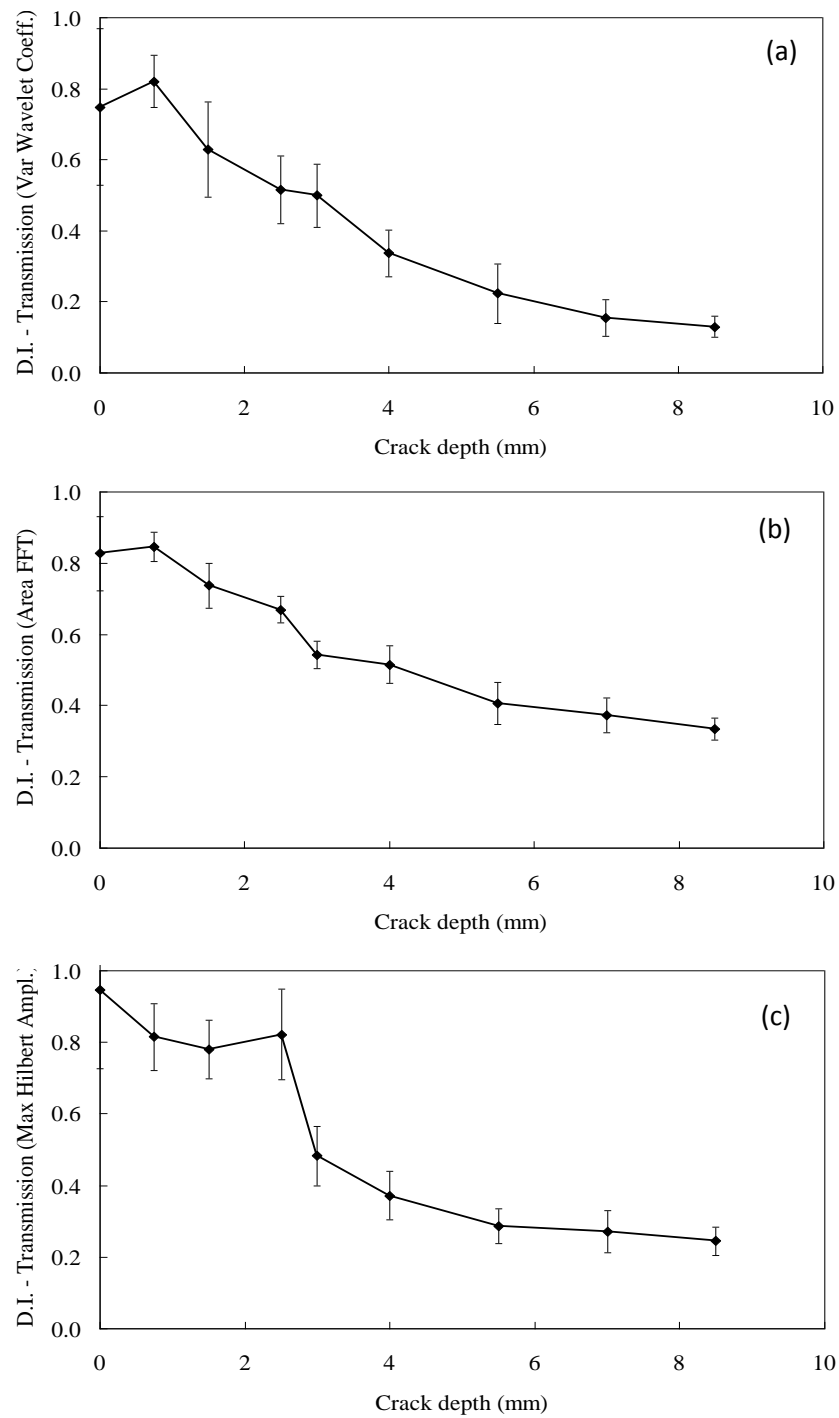


Figure 4.24: Feature-based Damage Index as a function of the center crack depth monitored in the “transmission mode”; (a) variance of the wavelet coefficients; (b) area of the frequency spectrum of the reconstructed signal; (c) maximum amplitude of the Hilbert transform of the reconstructed signal

as 500 mm from the sensing system, and the fact that air-coupled sensors are being used for noncontact signal detection at 76.2 mm (3") from the rail head. Nine different damage conditions, ranging from pristine structure to $s = 8.5$ mm (17% HA reduction) were monitored by using two sensors inclined towards the laser source in Test # 3. In this case the following thresholds were imposed on the DWT decompositions of sensor # 1 and sensor # 2: 80% for level 3, 70% for level 4 and 80% for level 5. Figure 4.24 shows the transmission D.I. calculated from eq. (4.24) using the same features as those used for the reflection D.I. in the two previous figures. Again, the mean value of ten measurements is plotted as a function of crack depth and the vertical line is equal to $2\sigma_{dev}$. The D.I. value at pristine conditions ($s = 0$) is not exactly one due to slight differences in alignment/sensitivity of the two sensors and signal attenuation. It can be seen that the desired monotonic behavior of the D.I. with respect to crack depth is more pronounced for the transmission D.I.. This suggests that the reliability of sizing the crack is higher with transmission measurements than it is from reflection measurements. This is also due to the fact that the orientation of the single sensor used to detect the defect echo was not optimized for the incoming signal and the resulting reflection D.I. may thus suffer from increased instabilities. The maximum sensitivity to the crack depth (slope of the D.I. curves) is given by the wavelet coefficient variance in Figure 4.24(a). The D.I.'s calculated with individual features suggest that the use of two sensors operated in the transmission mode provides the best determination of the damage severity. The portion of the rail that can be monitored in this case is strictly related to the distance between the two sensors; however, the method signals that a defect is present between the sensors but

it does not exactly locate its position. The reflection mode though, has the advantage to give direct information on the location of a defect ahead of the sensing system.

4.4.2 DWT-based feature extraction. A synopsis of univariate and multivariate approaches

Further testing were made on a different rail specimen, in order to investigate the performance of several DWT-based features for the D.I. computation; the features are listed in Table 4.2. The acquired signals were corrupted with two levels of noise; the SNRs of the corrupted signals were on the order of 40 dB (low noise level) and 10 dB (high noise level). An unsupervised learning algorithm was employed to rate the performance of each of the selected features, in terms of outliers detected, over a defined number of acquired test, related to defected areas of the rail. Details of the test setup and of the employed Outlier Analysis algorithm are available in [114].

Table 4.2: List of features considered in the unsupervised learning defect detection.

Feat #		Feat #	Features
1	Variance of the wavelet coefficient vector	7	Max ampl. of the DWT reconstructed time signal
2	RMS of the wavelet coefficient vector	8	Peak-to-peak ampl. of the DWT reconstructed time signal
3	Max ampl. of the wavelet coefficient vector	9	Max ampl. of the FFT of the DWT reconstructed time signal
4	Peak-to-peak ampl. of the wavelet coefficient vector	10	Area of the FFT of the DWT reconstructed time signal
5	Variance of the DWT reconstructed time signal	11	Max ampl. of the Hilbert Transform of the DWT reconstructed time signal
6	RMS of the DWT reconstructed time signal	12	Area of the Hilbert Transform of the DWT reconstructed time signal

The results of the detection in the case of low and high levels of noise corruption, are shown in Table 4.3 and Table 4.4 in terms of number of outliers for each of the selected features.

Table 4.3: Outliers detected from the different D.I. features and for each defect type when the ultrasonic signals were corrupted with the low-level noise.

<i>Def.</i>	<i>Wavelet Coefficient</i>				<i>Reconstructed Signal</i>				<i>Fourier Transf.</i>		<i>Hilbert Transf.</i>	
	<i>Var</i>	<i>Rms</i>	<i>Max</i>	<i>Ppk</i>	<i>Var</i>	<i>Rms</i>	<i>Max</i>	<i>Ppk</i>	<i>Max</i>	<i>Area</i>	<i>Max</i>	<i>Area</i>
0	10	0	0	10	0	0	10	10	2	10	10	10
1	10	32	0	0	0	32	50	30	0	20	40	26
2	0	10	0	0	0	10	0	0	0	20	0	40
3	0	230	0	20	0	230	0	0	39	242	0	179
4	75	499	0	91	87	499	228	139	225	402	160	217
5	250	580	208	530	440	580	537	498	430	560	500	413
6	329	600	549	590	590	600	570	570	572	600	570	575
7	540	600	580	600	600	600	600	600	600	600	600	595
Tot.	1204	2551	1337	1831	1717	2551	1985	1837	1866	2444	1870	2045
Tot %	27.6%	58.0%	30.4%	41.6%	39.0%	58.0%	45.1%	41.8%	42.4%	55.6%	42.5%	46.5%

It is evident that the in both cases of noise corrupted signals, the RMS of the wavelet coefficients and the RMS of the reconstructed signals are the best performing features, with the former prevailing on the latter for the low noise case. Further investigations showed that the definition of a multivariate Damage Index composed by a combination of some of the features presented in Table 4.2 was able to improve the results of the univariate definitions of the D.I..

Table 4.4: Outliers detected from the different D.I. features and for each defect type when the ultrasonic signals were corrupted with the high-level noise.

<i>Def.</i>	<i>Wavelet Coefficient</i>				<i>Reconstructed Signal</i>				<i>Fourier Transf.</i>		<i>Hilbert Transf.</i>	
	<i>Var</i>	<i>Rms</i>	<i>Max</i>	<i>Ppk</i>	<i>Var</i>	<i>Rms</i>	<i>Max</i>	<i>Ppk</i>	<i>Max</i>	<i>Area</i>	<i>Max</i>	<i>Area</i>
0	6	1	4	7	11	1	13	14	6	3	16	5
1	34	21	0	4	0	21	46	35	8	18	37	12
2	0	20	0	2	0	20	2	0	0	29	1	4
3	5	147	0	12	0	147	10	3	9	147	2	67
4	28	399	1	77	39	399	180	118	82	329	135	58
5	89	524	187	476	335	524	518	466	284	500	459	223
6	148	590	517	587	538	590	574	561	387	586	560	390
7	273	600	565	600	596	600	600	599	519	600	599	390
Tot.	577	2301	1270	1758	1508	2301	1930	1782	1289	2209	1793	1144
Tot. %	13.1%	52.3%	28.9%	40.0%	34.3%	52.3%	43.9%	40.5%	29.3%	50.2%	40.8%	26.0%

Three combination of features, named “Best” including the features (1,2,3,5,9), “Wvlt coeff” including the features (1,2,3,4) and “Worst” including the features (3,4,8,9,10,11,12) were analyzed, and the results are shown in Table 4.5 and Table 4.6.

Table 4.5: Outliers detected from the multivariate analysis for three different D.I. feature combinations and for each defect size when the ultrasonic signals were corrupted with low noise.

<i>Def</i>	<i>Best</i>		<i>Wvlt coeff</i>		<i>Worst</i>	
	<i># of outl.</i>	<i>% outl.</i>	<i># of outl.</i>	<i>% outl.</i>	<i># of outl.</i>	<i>% outl.</i>
0	40	2.5%	29	1.8%	20	1.3%
1	348	43.5%	320	40.0%	7	0.9%
2	407	67.8%	121	20.2%	0	0.0%
3	490	81.7%	243	40.5%	11	1.8%
4	577	96.2%	510	85.0%	125	20.8%
5	590	98.3%	590	98.3%	126	21.0%
6	600	100.0%	600	100.0%	491	81.8%
7	600	100.0%	600	100.0%	584	97.3%
Tot.	3612		2984		1344	
Tot. %	82.1%		67.8%		30.6%	

Table 4.6: Outliers detected from the multivariate analysis for three different D.I. feature combinations and for each defect size when the ultrasonic signals were corrupted with high noise.

<i>Def</i>	<i>Best</i>		<i>Wvlt coeff</i>		<i>Worst</i>	
	<i># of outl.</i>	<i>% outl.</i>	<i># of outl.</i>	<i>% outl.</i>	<i># of outl.</i>	<i>% outl.</i>
0	31	1.9%	25	1.6%	29	1.8%
1	291	36.4%	302	37.8%	29	3.6%
2	273	45.5%	88	14.7%	2	0.3%
3	453	75.5%	188	31.3%	15	2.5%
4	552	92.0%	335	55.8%	40	6.7%
5	587	97.8%	572	95.3%	78	13.0%
6	600	100.0%	600	100.0%	327	54.5%
7	600	100.0%	600	100.0%	498	83.0%
Tot.	3356		2685		989	
Tot. %	76.3%		61.0%		22.5%	

The advantage of using contemporary more than one defect sensitive feature substantially improved the performance of the defect detection. It should be also noted that the performance of the multivariate analysis is highly influenced by the selection of the participating features.

4.4.3 Feature extraction for D.I. computation

Different methods can be implemented for quantitatively comparing the two ultrasonic signals acquired by a sensor pair deployed on the rail in the transmission mode. An expression of D.I. based on the results of the DWT, is based on the Root Mean Square (RMS) of the wavelet transform coefficients retained after the filtering or pruning process. Given the RMS of the retained wavelet coefficients belonging to levels $j= k...m$ of the DWT decomposition of signal $x(n)$, expressed as

$$RMS\{wc[x(n)]_{j=k...m}\} = \sqrt{\frac{\sum_{n=1}^N wc[x(n)]_{j=k...m}^2}{N}} \quad (4.26)$$

where N is the total sum of the coefficients $wc[x(n)]_{j=k...m}$ throughout the considered levels $j= k...m$. The DWT-based D.I. is then expressed as

$$DI_{wc} = \max \left\{ \left| \frac{MS\{wc[x_1(n)]_{j=k...m}\}}{MS\{wc[x_2(n)]_{j=k...m}\}} \right|, \left| \frac{MS\{wc[x_2(n)]_{j=k...m}\}}{MS\{wc[x_1(n)]_{j=k...m}\}} \right| \right\} \quad (4.27)$$

where $x_1(n)$ and $x_2(n)$ are the signals acquired by the two sensors. The presence of a defect between the excitation source and either one of the transducers, will decrease the RMS of the wavelet coefficients of the signal acquired by that specific sensor. It is clear that in that case, one of the ratios expressed in eq. (4.27) will increase, while the other

will decrease. Defining the D.I. as the maximum between the two ratios of eq. (4.27) ensures that the presence of a defect will always cause an increase of the D.I..

Some operational testing parameters influencing the computation of the D.I. are not invariant throughout the test. Table 4.7 includes the main sources of variability that affect the D.I. computation in a typical noncontact acquisition, related to the laser, to the ultrasonic sensors and to the condition of the top rail surface.

Table 4.7: Different sources of variabilities affecting the D.I. computation

Laser generation	Non-contact acquisition	Rail surface condition
Variability of beam intensity	Variability of the different sensors frequency response spectrum	Variability of the ablation phenomenon effect due to local presence of material on the top surface of the rail (water, moisture, dust, dirt)
Variability of beam inclination, due to local rail misalignments	Variability on the sensor lift-off over the rail surface due to local rail misalignments	
	Variability of the sensor inclination due to local rail misalignments	

The different varying parameters described in Table 4.7, can influence the computation of the D.I., which can vary substantially even if calculated for rail areas characterized by the same damage. The direct effect of these unstable parameters is that the energy of the acquired signal is not constant. By normalizing the energy of the acquired signals, the effect of the described variations is reduced and the robustness and reliability of the computed D.I. critically improves.

This paragraph introduces the analysis of two energy-normalized D.I.'s, respectively based on RMS amplitude differencing and temporal coherence. These

features are common in acousto-ultrasonic tests where it is necessary to achieve a baseline-subtracted feature for defect detection [115]. For the two methods, parameters are computed that are a measure of the difference between the two signals. As a first case, the RMS of the amplitude of the gated signals in the time domain was analyzed. The definition of the RMS of signal $x(n)$, where $n=1 \dots N$ is given by:

$$RMS(x) = \sqrt{\frac{\sum_{n=1}^N x^2(n)}{N}} \quad (4.28)$$

The RMS quantity expressed in eq. (4.28) is also an expression of the power (time averaged energy) of the signal $x(n)$. The normalized RMS based D.I. computed on two ultrasonic signals included in the signals x_1 and x_2 is expressed

$$DI_{RMS} = \left| \frac{\sum_{gate, x_1(n)} x_1^2(n)}{N_{gate, x_1(n)}} \cdot \frac{RMS(x_2)}{RMS(x_1)} - \sqrt{\frac{\sum_{gate, x_2(n)} x_2^2(n)}{N_{gate, x_2(n)}}} \right| \quad (4.29)$$

where $gate, x_1(n)$ and $gate, x_2(n)$ are the time-windows containing the first arrival of the propagating guided wave in the signals x_1 and x_2 , and $N_{gate, x_1(n)}$, $N_{gate, x_2(n)}$ their lengths, respectively. The absolute value was considered in eq. (4.29) to generalize the method, which accounts for signal differences due to the decrease of $RMS(x_1)$ in respect to $RMS(x_2)$ or vice versa. Eq. (4.29) highlights relative differences in RMS's independently of the absolute energy of the individual signals. The normalized RMS based D.I. demonstrated to be very successful in the detection of several types of surface and internal defects during test performed in UCSD, which are discussed in the following section. While this feature showed to be sensitive to very small defects in the tested rail, its behavior was very stable in defect-free areas, proving to be robust and reliable.

Following the analysis of the time-domain amplitudes difference, another approach can be followed to identify the discrepancies between two ultrasonic signals. This approach is based on monitoring the shape of the signal, which includes both the time and frequency information. The temporal coherence is a measure of the time-dependent shape change between two signals. The temporal coherence is based upon the cross-correlation between two time invariant discrete signals $x_1(t)$ and $x_2(t)$, defined by [115]:

$$R_{12}(\tau) = E[x_1(t)x_2(t + \tau)] \quad (4.30)$$

where E denotes expectation. The time of the cross-correlation peak is typically used to estimate the time delay between the two signals if they are sufficiently similar in shape.

The normalized cross-correlation, or temporal coherence, provides an amplitude-independent measure of the similarity in shape of the two signals when the second signal is delayed by τ from the first [116]:

$$\gamma_{12}(\tau) = \frac{R_{12}(\tau)}{\sqrt{R_{11}(0)R_{22}(0)}}. \quad (4.31)$$

The expression of the D.I. based on the peak of the temporal coherence is

$$DI_{PeakCoherence} = \max[\gamma_{12}(\tau)] \quad (4.32)$$

Due to its nature, the normalized temporal coherence performs at best when applied on signals with broad frequency spectra. This behavior is due to the fact that a local change in the signal spectrum as given by a defect, will greatly affect the signal shape.

4.5 Validation of signal processing for defect detection in rails

4.5.1 Introduction

The rail defect detection prototype assembled at UCSD, described in depth in chapter five, was used at UCSD to test the performance of the different D.I.'s discussed in the previous section. A rail track mock-up, 20' (≈ 6 m) long, was built by aligning five rail sections with surface and internal defects. A photo and a sketch of the mockup are presented in Figure 4.25(a). Figure 4.26 depicts the prototype deployed on the rail mockup. Details of the site lay-out are schematized in Figure 4.25(b) and presented in Table 4.8. The most significant discontinuities, along with their locations (distance from the start of the test zone) are listed therein.

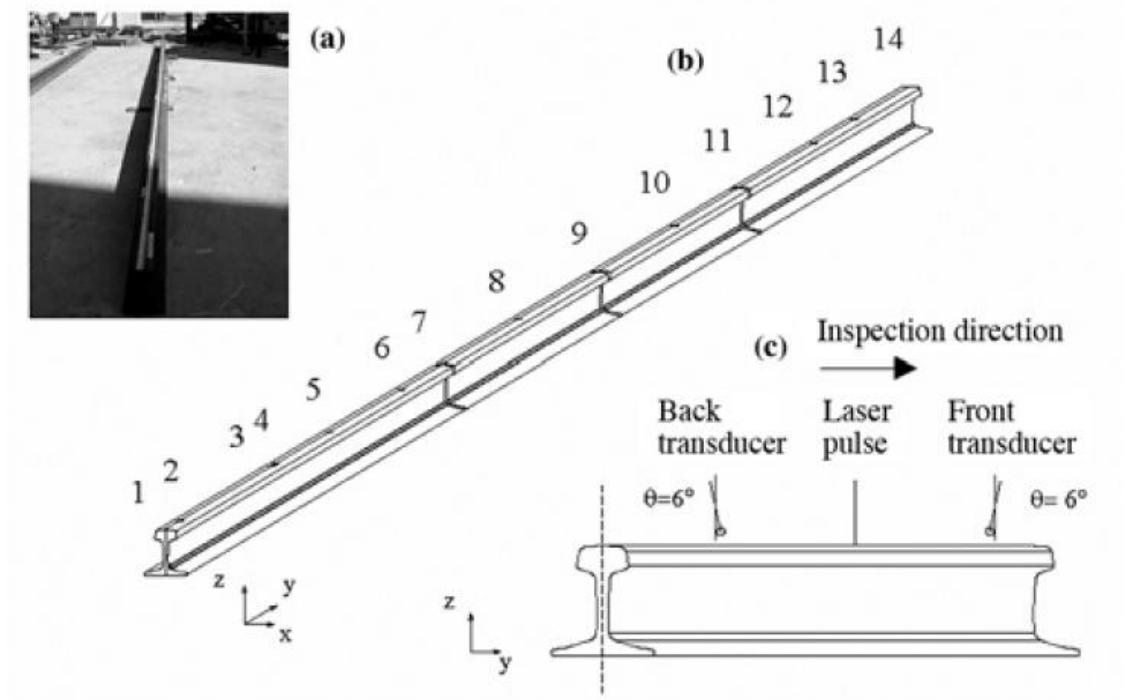


Figure 4.25: (a) Photo of the rail track mock-up at UCSD; (b) schematic of the mock-up layout; (c) defect detection scheme in 'transmission mode' with a pair of air-coupled transducers

The first joint is considered the start point (location #1). The second section contains four surface defects: three defects (namely at locations #2, #5, and #6) were transverse notches; the fourth defect (location #3-4) was a longitudinal notch over 40 mm long. The second and third section contained each an internal defect located closer to the gauge side.

Table 4.8. Mock-up layout

Location	Distance (mm)	Description	% HA reduction	Acquisition #
1	0	Joint: start of test zone		1 – 4
2	127	Surface transverse notch <i>No defect</i>	1%	1 – 6 7 – 11
3	920	Start of longitudinal surface notch		12 – 16
4	968	End of longitudinal surface notch <i>No defect</i>		12 – 16 17 – 18
5	1428	Surface transverse notch <i>No defect</i>	16%	19 – 22 23 – 26
6	2060	Surface transverse notch	5%	27 – 30
7	2432	Joint <i>No defect</i>		31 – 35 36 – 38
8	3073	Internal defect <i>No defect</i>	8%	39 – 42 43 – 47
9	3789	Joint <i>No defect</i>		48 – 52 53 – 55
10	4445	Internal defect <i>No defect</i>	23%	56 – 60 61 – 63
11	5013	Joint <i>No defect</i>		64 – 67 68 – 70
12	5657	Surface corner oblique notch (45°)	7%	71 – 75
13	6010	Surface transverse notch	20%	76 – 80

* Notes: *No defect* indicates that the acquisition was stored, passed the related defect

A hand mapping conducted by using commercial contact ultrasonic equipment estimated that the internal defects were 8% HA (location #8) and 23% HA (location #10).

Finally, the last section contained an oblique defect and a surface transverse defect accounting for 7% and 20% HA reductions, respectively. The oblique cut was machined on the gauge side a 45-degree inclination from the running direction. The relative position of the laser head and air-coupled sensors is schematized in Figure 4.25(c). The cart was manually pushed on the mock-up and an acquisition was made approximately every inch (25.4 mm). The ultrasonic signals were acquired at a 5 MHz sampling rate.



Figure 4.26: Rail defect detection prototype deployed on the rail track mock-up at UCSD

4.5.2 Parametric study of Damage Index computation using the DWT

A first set of test were made considering a DWT-based D.I., and different combinations of parameters were investigated. A preliminary study on the sensitivity of the mother wavelet selection to detect defects was carried. First, eighty acquisitions were made during one run. Column 5 of Table 4.8 shows which portion (damaged, defect free, or joint) of the rail was probed during each acquisition. Figure 4.27(a) and (b) present typical waveforms recorded from the rear and front sensors, respectively, at acquisition #17 when the laser light illuminated the rail head at 1150 mm from the start point. Ideally, the shape and the amplitude of the waveforms between the two sensors should be identical; however, slight variations on the sensor inner sensitivity, sensor alignment, and relative distance of the transducer to the laser light produced small differences in the waveforms' amplitude. Figure 4.27(c) and Figure 4.27(d) show the ultrasonic signals recorded at acquisition #27, when the laser light impinged the rail head at 2060 mm from the start of the test zone. As illustrated in Table 1 (column 5), said acquisition should be able to detect the surface defect machined at location #6 (Table 1, column 1). The fact that the amplitude of the signal detected by the front transducer (Figure 4.27(d)) is smaller than the signal amplitude from the rear transducer (Figure 4.27(c)) proves that the laser illuminated a zone between the notch and the rear sensor. This is also confirmed by noting in the time history of Figure 4.27(c) the presence of a second pulse, which is the result of the reflection from the surface defect. Finally, Figure 4.27(e) and (f) show the ultrasonic signals recorded in the acquisition #32 by the rear and front sensors, respectively. At this acquisition, the second joint was between the laser light and the front transducer. The disappearing of the ultrasonic trace is observed in Figure 4.27(f).

The signal detected by the rear sensor shows the stress wave travelling from the laser light to the sensor directly and shows also an echo from the joint.

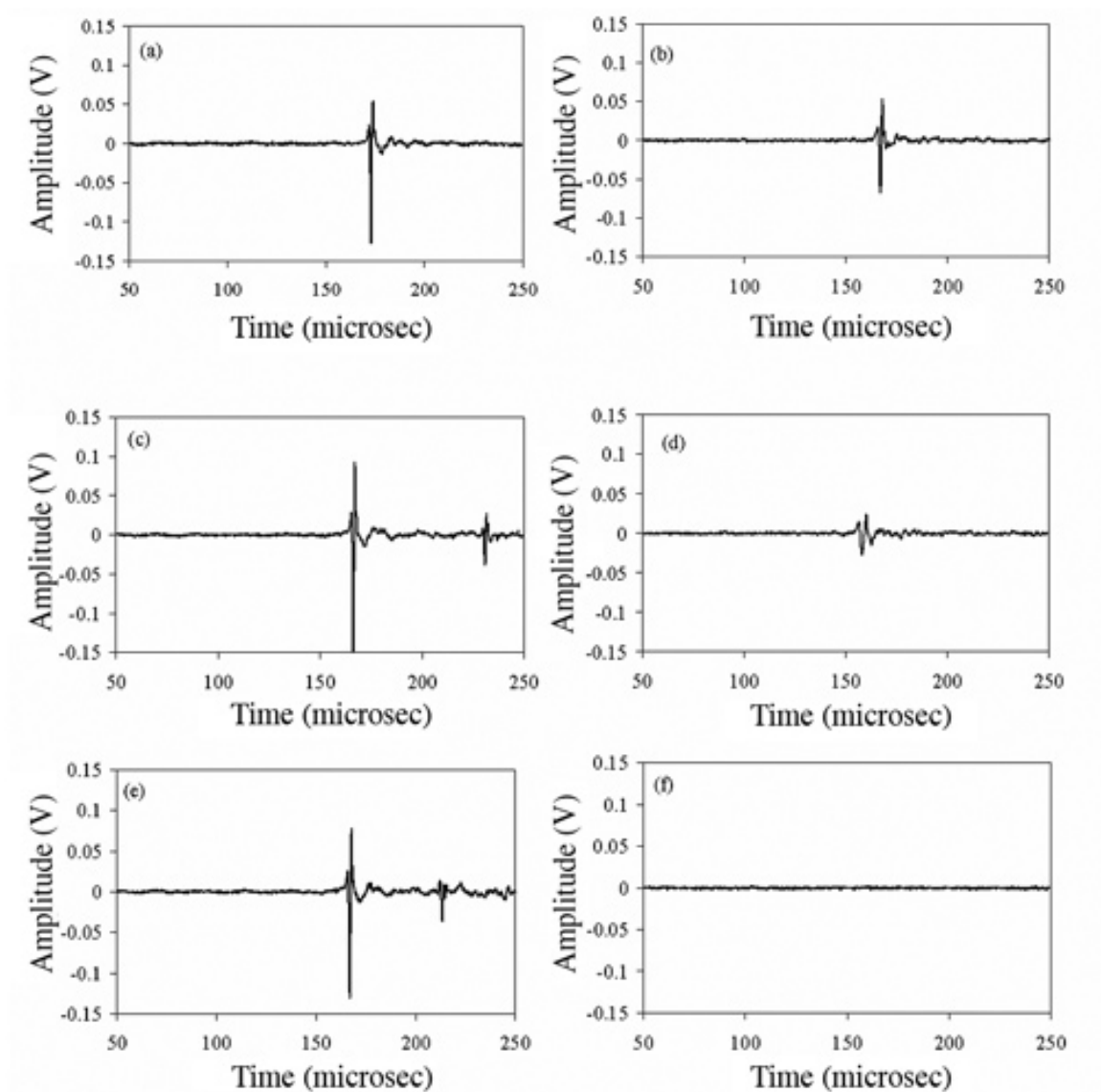


Figure 4.27: Typical waveforms detected by the hybrid laser/air-coupled transducer system on a rail.
 (a), (b) Time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a damage-free rail; (c), (d) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a 5% HA transverse surface notch; (e), (f) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a joint

For the computation of the DWT-based D.I., three mother wavelets were considered for their resemblance to the original recorded waveforms. The mother wavelets considered are (Figure 4.28) the Coiflet of order 4 (coif4), the Symlet of order 6 (Sym6), and the Daubechies of order 10 (db10), respectively. The general DWT processing applied in this study is described in Figure 4.29. The unprocessed signals recorded by the rear and front transducers are shown in Figure 4.29(a) and (b), respectively. These ultrasonic signatures are relative to acquisition #27. For the signal processing algorithm, the time interval was centered to the arrival of the signal peak amplitude and a 256-points time history was considered. Notice that the scale amplitude of the front sensor signal is set to one fifth of the rear sensor signal. For each window, the wavelet coefficients at detail levels 2 and 3 only were considered. These two levels contained the frequency band between 150 and 600 kHz, which is of interest in this study. All wavelet coefficients from other levels were set to zero. Figure 4.29(c) and (d) show the clustered coefficients for the levels 2 and 3 for each waveform, respectively. Most of these coefficients have very low amplitude and thus do not carry any information on the signal of interest. These coefficients can be discarded if not exceeding a given threshold, and set to zero in the reconstruction process. The threshold was related to the maximum amplitude modulus of each decomposition level. For level 2 only the two largest wavelet coefficient modules were selected and for level 3 the six largest modules. These eight coefficients are indicated with an empty circle in Figure 4.29(c) and (d). Signals reconstructed from the threshold wavelet coefficient vectors are shown in Figure 4.29(e) and (f) and they demonstrate that only a few wavelet coefficients are indeed representative of the

ultrasonic signatures of interest. For the transmission setup of Figure 4.25(c) the largest value between $D.I._{front/rear}$ and $D.I._{rear/front}$ was considered as the defect indicator.

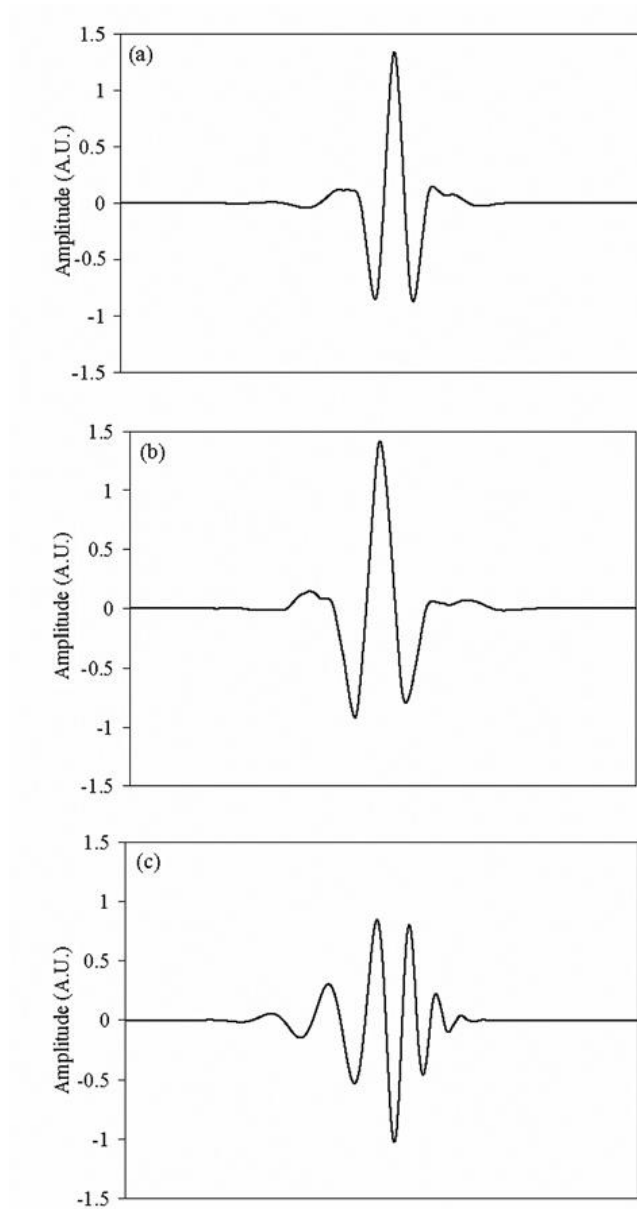


Figure 4.28: (a) The Coiflet wavelet of order 4; (b) the Symlet wavelet of order 6; (c) the Daubechies wavelet of order 10

The feature of the root mean square (RMS) was considered next. Figure 4.30(a) compares the RMS-based D.I. computed for the original unprocessed signal, the vector of the retained wavelet coefficients, and the reconstructed signals. The D.I. is plotted as a function of the acquisition number, which corresponds to eighty different locations of the rail head. Acquisition #1 corresponds to the start of the test zone. The desirable properties of the D.I. are stable low values when no defect is present, and spikes, whenever the inspection is over a defect. Clearly, the higher is the spike, the higher is the defect sensitivity of the system to the presence of a defect. The dotted line overlapping the D.I. plots, is a threshold whose value was arbitrarily chosen to highlight the areas related to the different discontinuities of the rail. The detection of the joints around acquisitions #3, #32, #48, and #64 are obvious and will not be further discussed. To improve the readability of the results, a detail of plots in Figure 4.30(a) is given in Figure 4.30(b). It can be observed that the RMS-based D.I. computed for the vector of the wavelet coefficients and for the reconstructed signal are more sensitive to the presence of defects compared to the same feature calculated for the unprocessed signal. By looking at the spikes at acquisition #20, for example, it can be seen that the D.I. computed over the vector of the wavelet coefficients is two times higher than the unprocessed feature. At acquisition #27 and #76, the improvement is even more dramatic. According to the fifth column of Table 1, the system was able to detect the following defects: the longitudinal surface notch expected at acquisitions #12-16; the 16% HA and 5% HA surface transverse notches expected at acquisitions #19-22 and #27-30, respectively; the 16% HA and the surface transverse notch expected at acquisition #76-80, respectively. The first transverse notch is probably too small and too close to the first joint to be discriminated

from the D.I. associated with the joint. The internal defects, expected at acquisitions #39-42 and #56-60, as well as the 45° oblique notch located at the gauge side respectively, are not clearly visible.

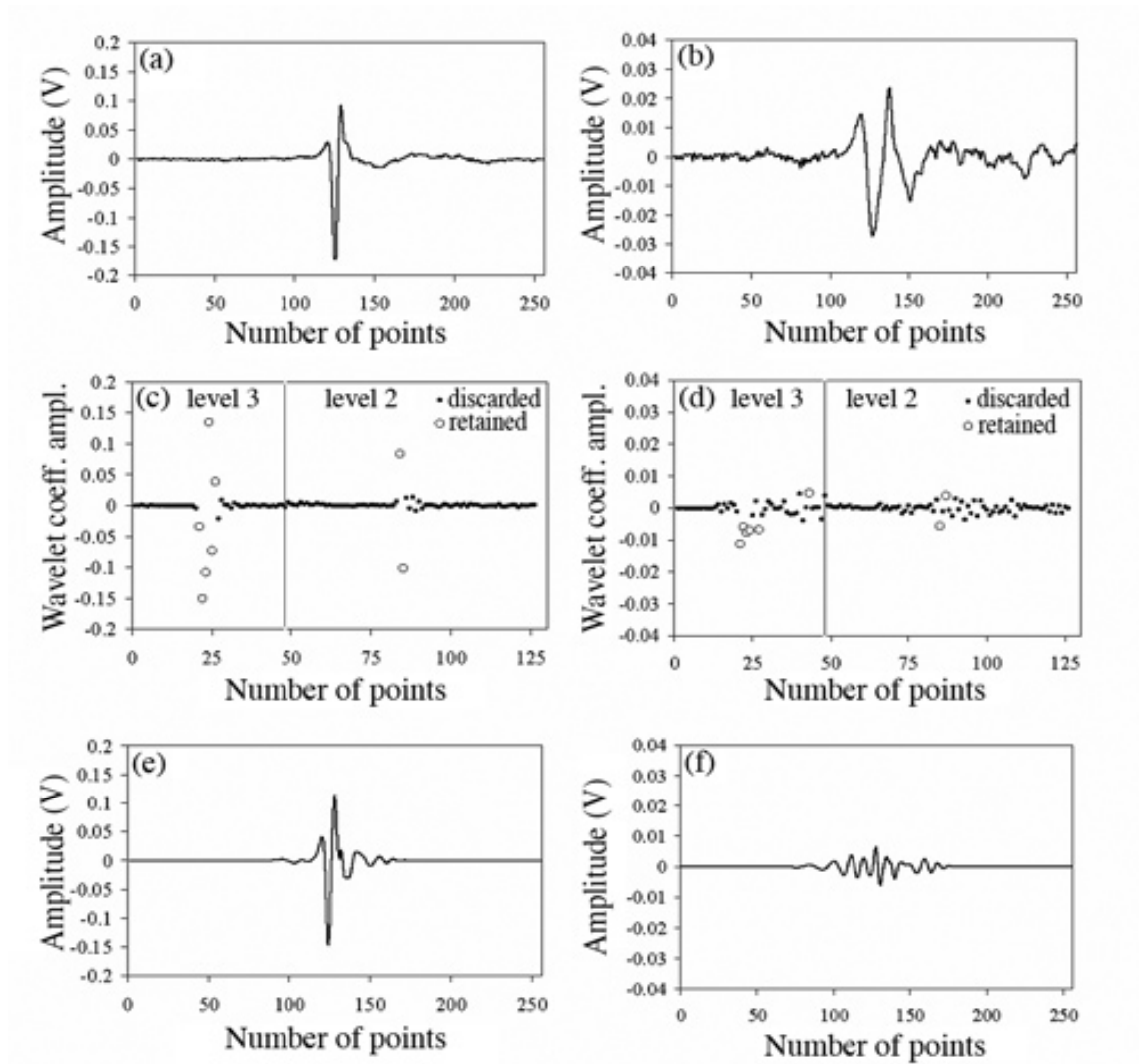


Figure 4.29: Example of discrete wavelet processing. (a), (b) unprocessed signals detected by the rear and front sensors, respectively, when the inspection system is probing a transverse surface notch; (c), (d) wavelet coefficient vector of detail levels 2 and 3 relative to the rear and front sensors, respectively; (e), (f) reconstructed rear and front sensor signals, respectively, using the thresholded (empty circles) wavelet coefficient vectors in (c) and (d)

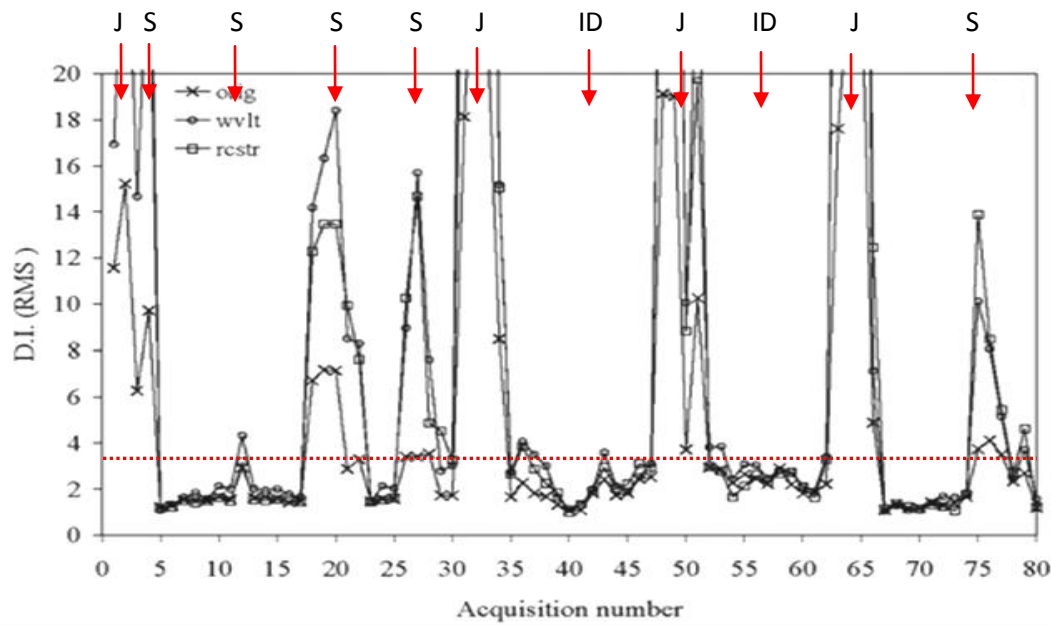


Figure 4.30: RMS-based Damage Index as a function of the acquisition number for the unprocessed time waveforms, the vector of the retained wavelet coefficients, and the time-domain reconstructed (de-noised) signal [J=Joint, S=Surface defect, ID=Internal defect]

The next step was the analysis of the effect of the wavelet coefficient threshold on the damage detection sensitivity. For each level, two and six largest wavelet coefficients were considered. A total of four possible combinations were therefore studied. The D.I. was computed considering the RMS of the retained wavelet coefficients. The results are presented in Figure 4.31 where the RMS-based D.I. is plotted as a function of the acquisition number. By comparing the performance of all four combinations, some difference is observed. For instance, at the acquisitions #19-22, the result of using two coefficients for detail level 2 and six coefficients for detail level 3 (combination 2_6), is better compared to the combination 6_2. The results from the parametric study of the effect of the mother wavelet on the damage detection sensitivity are shown in Figure 4.32. The db10, the coif4, and the sym6 were considered. The combination 2_6 was used to threshold the wavelet coefficient vector at detail levels 2 and 3.

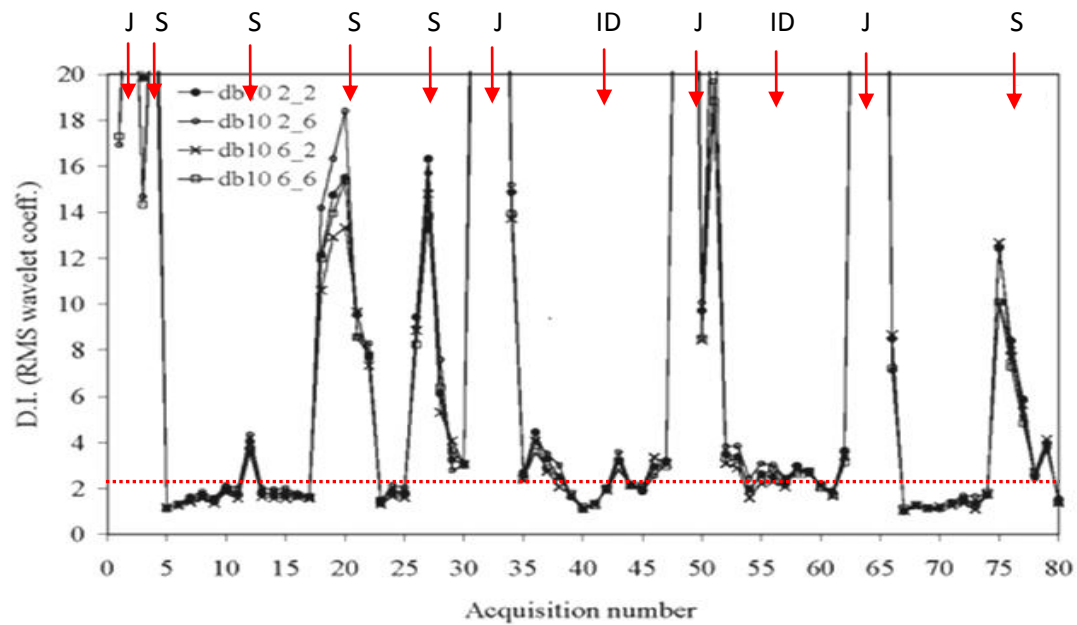


Figure 4.31: RMS-based Damage Index as a function of the acquisition number calculated for four different retained wavelet coefficient vectors [J=Joint, S=Surface defect, ID=Internal defect]

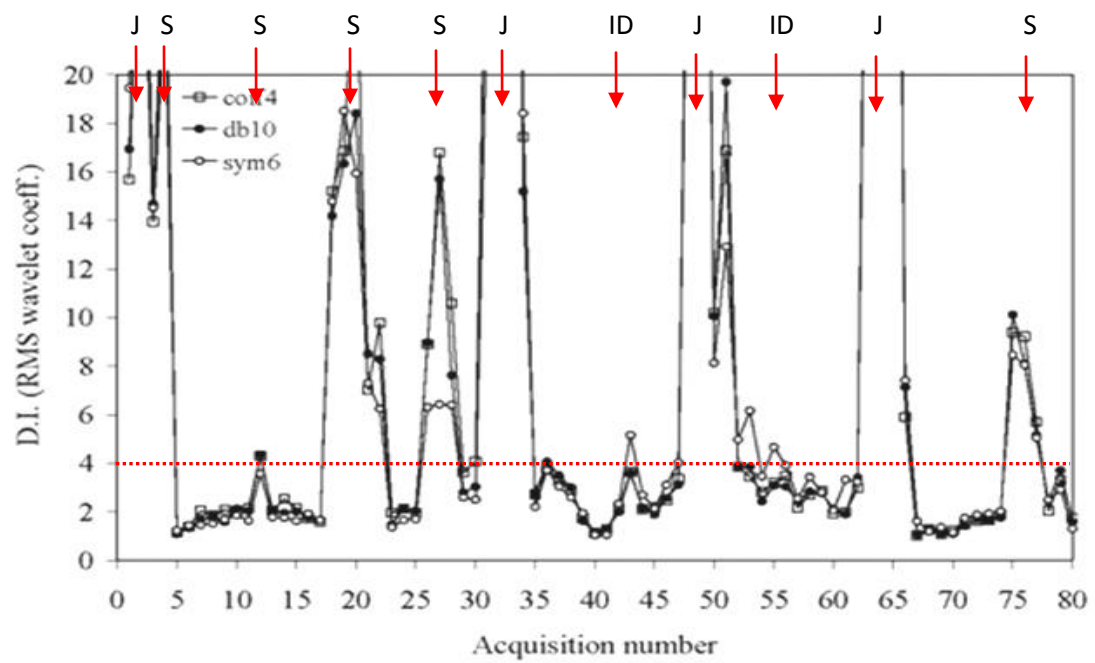


Figure 4.32: RMS-based Damage Index as a function of the acquisition number for the vector of the retained wavelet coefficients, computed from three different mother wavelets [J=Joint, S=Surface defect, ID=Internal defect]

The coif4 and the db10 mother-wavelets seem to outperform the sym6 in the detection of the surface defects at location #5, #6 and #12. However, by looking at the detailed plots in Figure 4.32(b), the sym6 appears to be more sensitive to the presence of the first internal defect, expected to be detected at acquisitions 39-42. As already discussed in recent works [117, 118] the sensitivity to internal and gauge defects can be improved by selecting the proper frequency bandwidth and modifying the position of the transducers.

4.5.3 Parametric study of Damage Index computation using Digital Filters

Another set of laboratory tests was performed to investigate the performance of the D.I. using digital filters. The advantages of using digital filters in respect to the DWT were the substantial computational savings of the signal filtering algorithm; the main drawback was the impossibility of performing data compression. The monitored features in this case were the energy normalized RMS of the amplitudes of a signal window, gated on the arrival of the ultrasonic wave, and the peak of the temporal coherence calculated on the same window. The signal conditioning performed on the acquired signals was based on the application of a third order pass-band Butterworth filter. Two ranges of frequency were monitored, namely the low frequency band, ranging 40-300 kHz and the high frequency band, ranging 300-1,100 kHz. This choice was dictated by the necessity for distinguishing internal defects (below 10 mm depth) from surface defects (above 10 mm depth), respectively. Figure 4.33 shows typical waveforms filtered with the 300-1,100 kHz (high frequency range) pass-band Butterworth. Figure 4.33(a) and (b) present the waveforms recorded from the rear and front sensors respectively, related to a defect-free area.

Figure 4.33(c) and (d) show the filtered waveforms acquired from the sensor pair, in proximity of the surface notch present at location #5, as described in Table 4.8. Figure 4.33(e) and (f), instead, show the filtered waveforms recorded by the sensor pair in proximity of the joint at location #7.

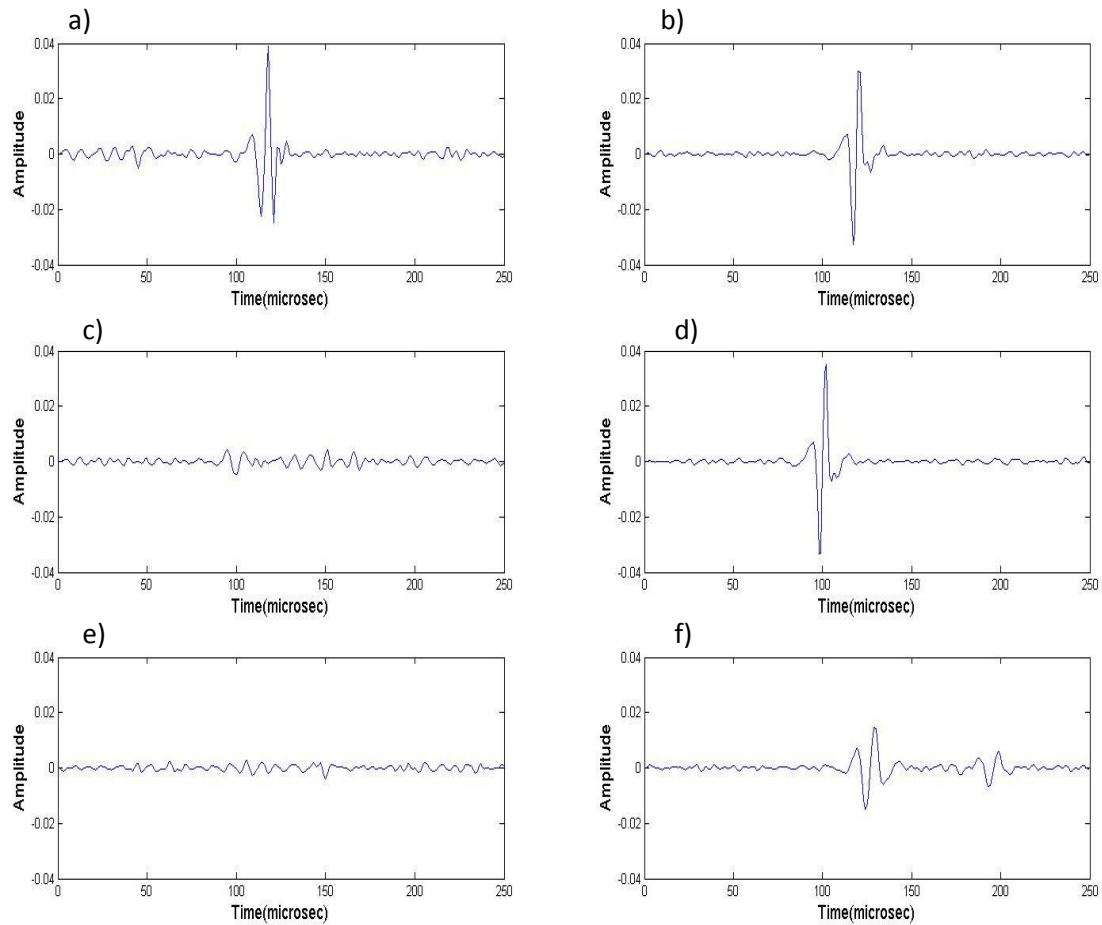


Figure 4.33: Typical waveforms filtered with the 300-1,100 kHz (high frequency range) pass-band 3rd order Butterworth. (a), (b) Time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a damage-free rail; (c), (d) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a transverse surface notch; (e), (f) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a joint

The smaller amplitude of the signal acquired by the rear sensor shown in Figure 4.33(c) in respect to the front one (Figure 4.33 (d)) indicates that the area excited by the laser was between the surface defect and the front sensor. The detected joint is the reason of the disappearance of the signal in Figure 4.33(e), compared to Figure 4.33(f). Figure 4.34 instead, shows other waveforms recorded during the acquisitions and filtered with the 40-300 kHz (low frequency range) pass-band Butterworth. Figure 4.34 (a) and (b) present the filtered waveforms recorded from the rear and front sensors respectively, related to a defect-free area while (c) and (d) illustrate the filtered waveforms recorded in proximity of the internal defect at location #8. Figure 4.34 (e) and (f) instead, show the filtered waveforms acquired in proximity of the joint at location #7. In this case, the smaller amplitude of the signal acquired by the rear sensor shown in Figure 4.34(c) indicates that the laser was exciting an area between the internal defect and the front sensor. Again, the presence of the joint is the reason of the disappearance of the signal in Figure 4.34(e). Sample plots of the D.I. based on 1) the normalized RMS expressed in eq.(4.29) and on 2) the peak of the temporal coherence expressed in eq. (4.32), are illustrated in Figure 4.35 through Figure 4.37. The D.I.'s plotted in Figure 4.35 and Figure 4.36, refer to ultrasonic signals acquired by a sensor pair deployed over the center of the rail head, and filtered with the low (40-300 kHz) and high (300-1,100 kHz) frequency ranges pass-band 3rd order Butterworth filters, respectively.

Figure 4.35 shows that analyzing the low frequency content of the acquired signals, the performance of the center head sensor pair was very satisfactory for the detection of both the internal defects at locations #8 and #10, and the surface defects at location #2 and #5, listed in Table 4.8.

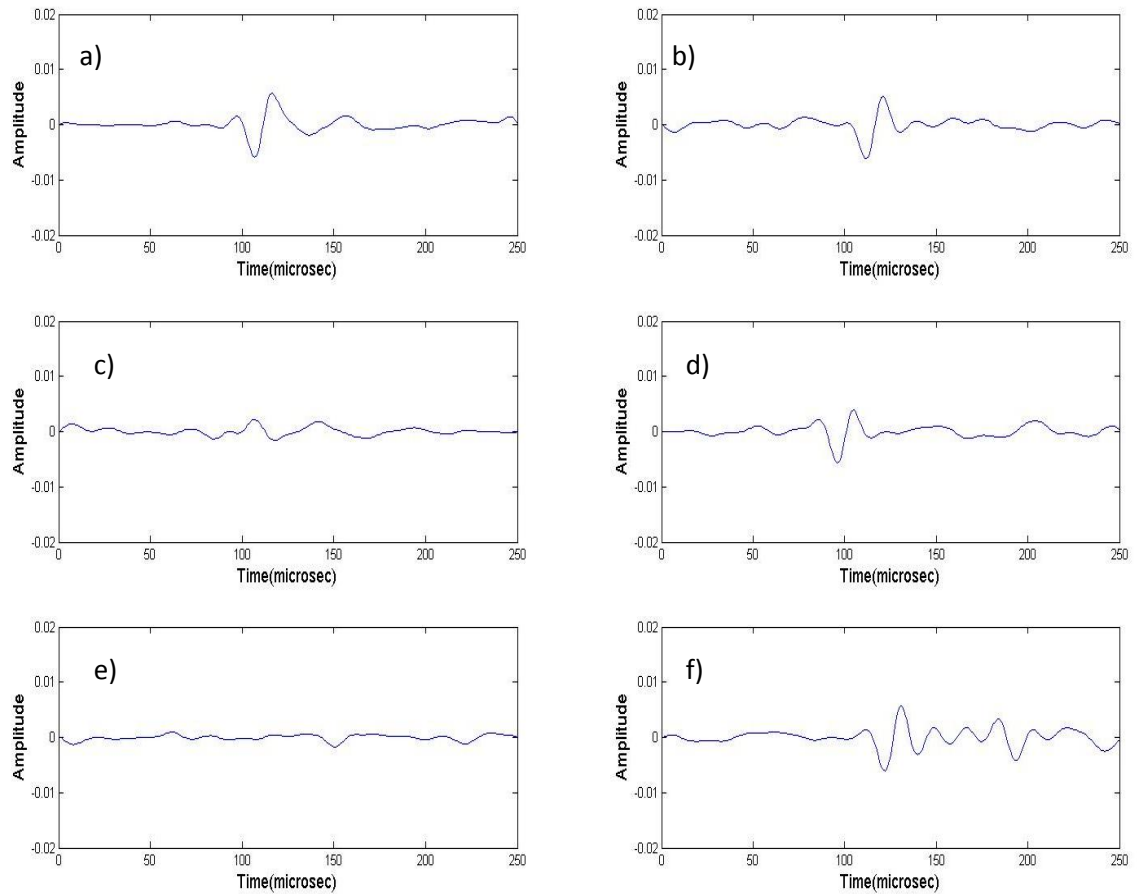


Figure 4.34: Typical waveforms filtered with the 40-300 kHz (low frequency range) pass-band 3rd order Butterworth. (a), (b) Time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a damage-free rail; (c), (d) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a transverse surface notch; (e), (f) time waveforms detected by the rear and front sensors, respectively, when the inspection system is probing a joint

Analyzing the high frequency content of the acquired signals, the performance of the center head sensor pair was satisfactory for the detection of the surface defects at locations #2, #3, #5 and #6. The results confirmed the role of the inspecting guided-wave frequency on the depth of the defect to be detected; the low frequency range was able to

better detect the internal flaws, while the high frequency range was better successful in the detection of the surface defects.

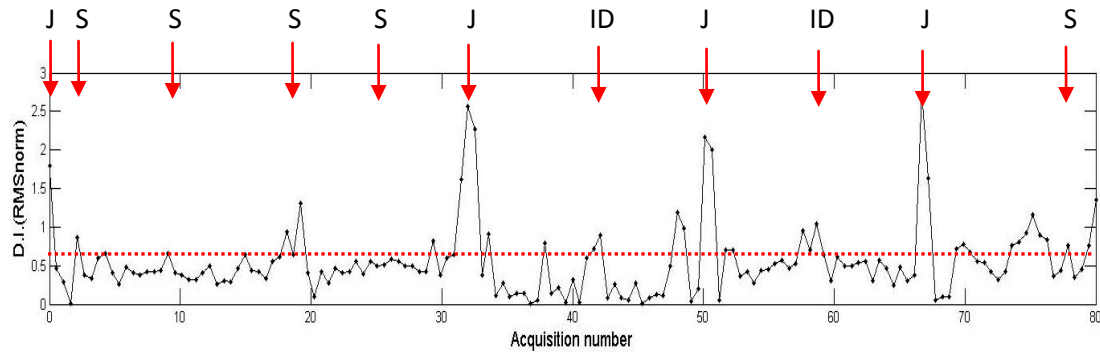


Figure 4.35: Damage Index as a function of the acquisition number, based on the normalized RMS of the signal amplitudes. Sensors deployed on the Center Head of the rail and signals filtered with a pass-band 40-300 kHz (low frequency range) 3rd order Butterworth [J=Joint, S=Surface defect, ID=Internal defect]

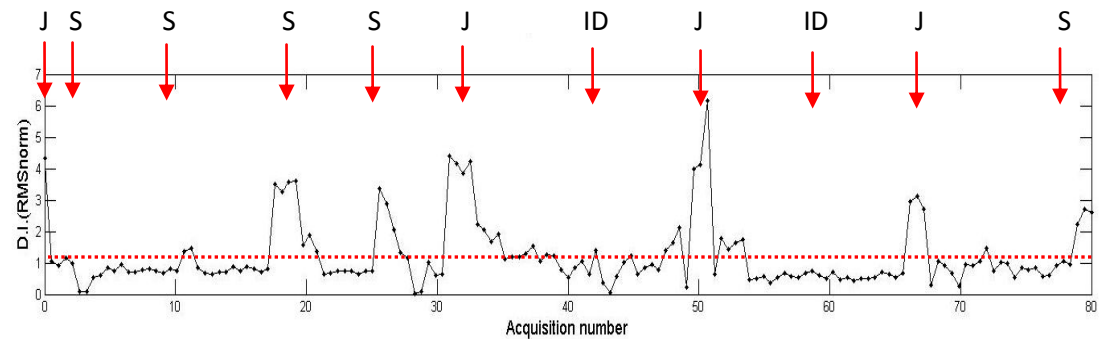


Figure 4.36: Damage Index as a function of the acquisition number, based on the normalized RMS of the signal amplitudes. Sensors deployed on the Center Head of the rail and signals filtered with a pass-band 300-1,100 kHz (high frequency range) 3rd order Butterworth [J=Joint, S=Surface defect, ID=Internal defect]

Keeping the low and high frequency ranges in the acquired signals, the best performing feature in the computation of the D.I. is the temporal coherence; as shown in Figure 4.37(a), both surface and internal flaws were detected. If the high frequency range was purged from the signals, the performance of the temporal coherence based-D.I. decreased dramatically, as shown in Figure 4.37(b).

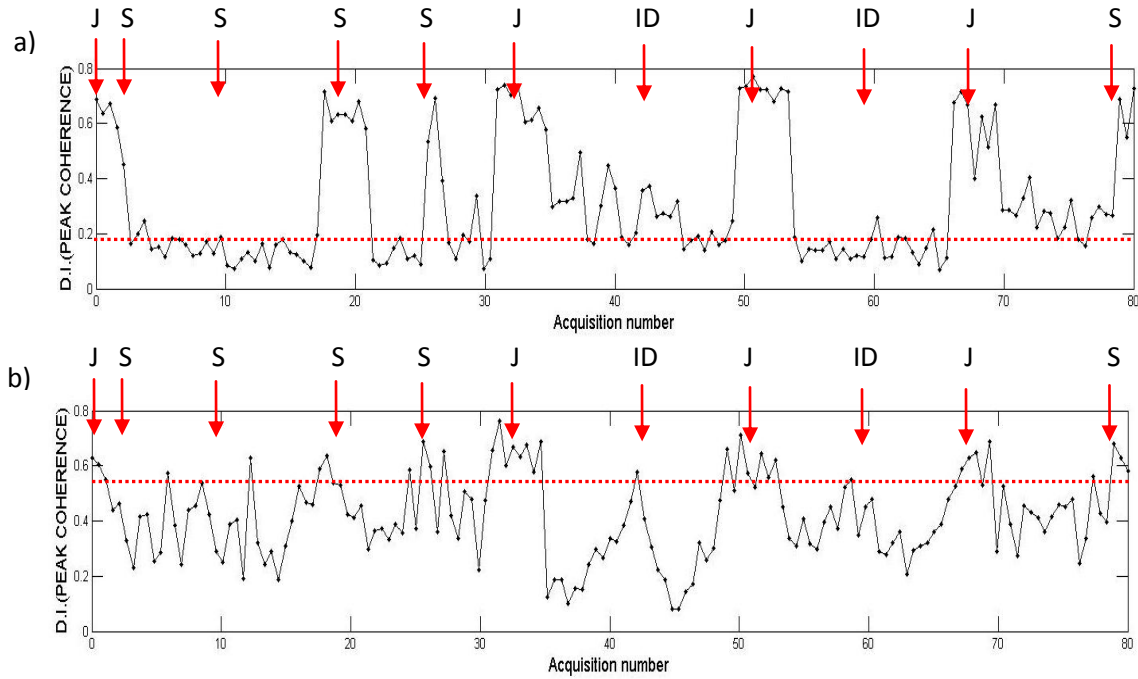


Figure 4.37: Damage Index as a function of the acquisition number, based on the temporal signal coherence of signals filtered with (a) 40-1,100 kHz (low + high range) and (b) 40-300 kHz (low range) pass-band 3rd order Butterworth. Sensors deployed on the Center Head of the rail [J=Joint, S=Surface defect, ID=Internal defect]

The good performance of the temporal coherence was expected, in the case of signals with broad frequency range, as mentioned earlier in the chapter. Maintaining the entire frequency spectrum was at the expenses of defect classification, since the frequency bands (high and low range) related to different areas of the rail cross-section, were not separated.

4.6 Conclusions

Different signal processing techniques to condition the ultrasonic signals to best highlight the presence of rail defects have been illustrated in the chapter. The need of retaining the time and the frequency information of the acquired signals led to the use of

the wavelet transform, in its continuous and discrete expressions. The DWT showed to be efficient for real time processing, including data de-noising and data compression. Different combinations of DWT-based features proved to be successful for the detection of defects present in rails available in the UCSD NDE & SHM laboratory. The experimental results showed that the appropriate selection of the mother-wavelet and of the wavelet coefficient threshold enhances the inspection sensitivity. The performance of a D.I. based on the normalized RMS of the signal amplitudes was also tested, and it showed excellent results in the detection of surface and internal defects. The implementation of digital filtering in the feature-extraction procedure dramatically simplified the process of setting the parameters needed for the prototype calibration and provided further reductions in computational time. A D.I. based on the normalized temporal coherence proved to be very efficient in defect detection, if the entire frequency spectrum was considered for the signals. The improvement in defect detection by the temporal coherence comes at the expenses of defect classification, since the frequency bands are no longer separable.

Part of chapter 4 of this dissertation was printed in *Insight–Non Destructive Testing & Condition Monitoring*, Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I., Fateh, M., Viola, E. and Pascale, G., 2006. “Non-Contact Ultrasonic Inspection of Rails and Signal Processing for Automatic Defect Detection and Classification”. The dissertation author was co-author of this paper. Another part of chapter 4 has been published in *Insight–Non Destructive Testing & Condition Monitoring*, Rizzo, P., Bartoli, I., Cammarata, M. and Coccia, S., 2007. “Digital signal processing for rail monitoring by means of ultrasonic guided waves.” The dissertation author was co-author of this paper.

5 Development of Rail Inspection Prototype for the Federal Railroad Administration

5.1 Introduction

This chapter describes the development of a rail inspection prototype which has been designed in the NDE & Structural Health Monitoring lab of UCSD. The majority of these findings have been presented into technical reports submitted to the Federal Railroad Administration [118, 119, 120, 121, 122]. The prototype technology embeds the theoretical results of the UCSD NDE group in the field of rail inspection. The laser-ultrasonic technique was deployed on a cart, which could to be towed on a railroad track, at a speed as high as 30 mph. The prototype design was realized in three stages. During the first stage, the main pieces of hardware were identified and a simple version of software was implemented in order to validate the inspection concept. The second and main stage consisted in the deployment of the prototype on a cart, which was made in collaboration with Ensco Inc., and in the development of the software capable of high speed data processing. During the third stage several upgrades were made to the prototype, such as a new scheme for the sensor position and a more robust cart capable of higher speed. A detailed description of the design stages is given in the first three sections of the chapter. The prototype was tested in the field twice; the tests were performed in March 2006 and April 2007, in collaboration with the Federal Railroad Administration and Ensco Inc. in Gettysburg, PA.

5.2 Stage I: Hardware and Software

5.2.1 Q-Switched laser generator and optical elements

During the first stage a Q-switched laser Continuum™ Surelite SLI-20 was employed for ultrasound generation. The great stability of this Q-switched Laser allowed the generation of a 7 nanosec pulse at a 1064 nm wavelength, with a contained divergence and a neglectable jitter. Table 5.1 summarizes the features of the Continuum™ Surelite SLI-20 laser. The power supply of the device is housed in a tower-case of the following dimension: 11''x 18''x 25'' (W x H x D), as shown in Figure 5.1. . The laser head had the dimension of 3''x3''x18'' and it is depicted in Figure 5.2.

Table 5.1: Continuum™ Surelite SLI-20 specifications

Source Power	220-240VAC, single phase, 8 A
Repetition rate (Hz)	20
Energy (mJ)	420
Pulsewidth (nsec)	5-7
Divergence (mrads)	0.6
Rod diameter (mm)	6
Beam pointing stability ($\pm\%$)	100
Jitter (\pm ns)	0.5
Energy stability ($\pm\%$)	2.5
Power drift ($\pm\%$)	3.0
Laser head mounting	horizontal
Laser head dimensions (W x H x D)	7''x 6.5''x 30.5''



Figure 5.1: Continuum® Surelite SLI-20 Q-Switched laser (power supply unit).

In order to drive the lasing of this device, an external trigger procedure was developed. The laser control board required 2 TTL lines ($+5V \rightarrow +0V$, falling edge triggering), to fire the Flashlamp and to open the Q-Switch respectively. The lamp needed to be fired at the factory-set repetition rate of 20 Hz, for the optimal performance of the device.

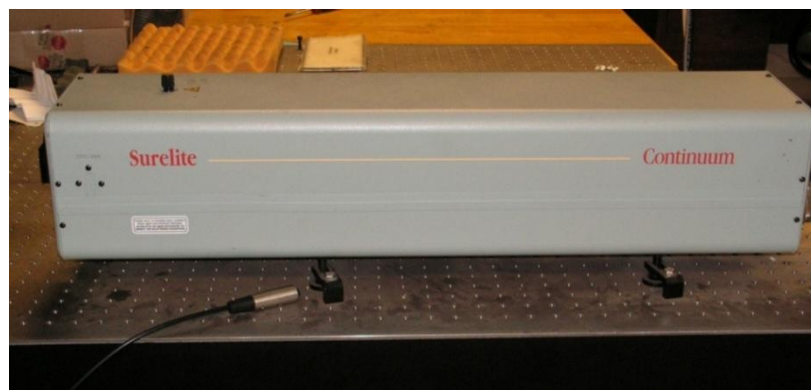


Figure 5.2: Continuum® Surelite SLI-20 Q-Switched laser (laser head)



Figure 5.3: Continuum® Surelite SLI-20 Q-Switched laser (control panel)

The lasing power was trimmed by adjusting the delay between the 2 TTL lines falling edges; the factory-set 190 microsecond delay causes the device to fire at the maximum available power. The external trigger had to be provided to the laser power source through the external port, available on the power supply control panel, shown in Figure 5.3. A sequence of two optical elements will be employed to route the light emission from the laser port to the rail head, as shown in Figure 5.4. The first is a plano-concave lens (1), a divergent element that has a negative focal point. When a collimated light beam is incident on his curved surface, the exit side will form a divergent beam. The second is a semi-cylindrical lens (2), a convergent element that can focus the diverged light coming from the plano-concave lens, to a line parallel to the main cylinder direction.

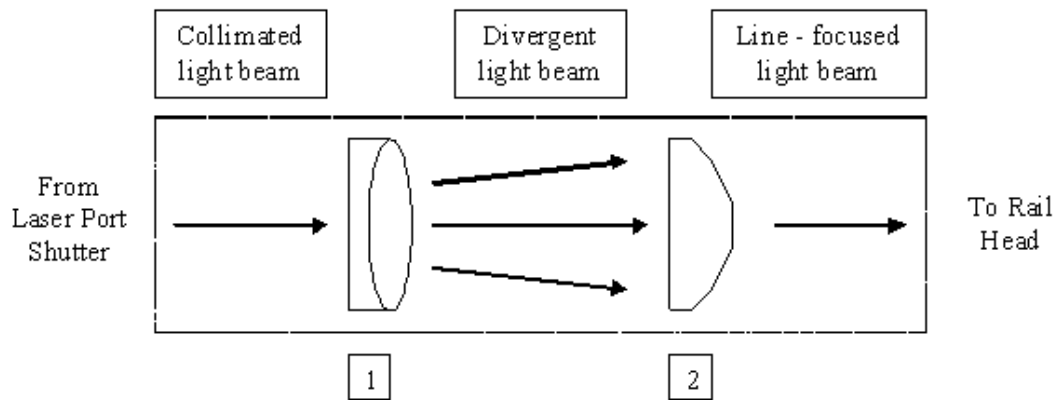


Figure 5.4: Layout of the optical element sequence

5.2.2 Air-coupled Transducers and Preamplifiers

MicroAcoustic[®] Broadband Air-coupled Transducers (BAT) were chosen to acquire the ultrasonic signals traveling along the rail. Ruggedness, ultrasonic bandwidth and high reliability are the dominant features of these state-of-art sensors. Table 5.2 summarizes the features of the Microacoustic[™] broadband air-coupled transducers. The unusually large bandwidth of the Microacoustic[™] made them favorable to the application; their drawback sensors being a vibrating membrane which is delicate and sensitive to large peaked accelerations.

Table 5.2: MICROACOUSTIC[®] broadband air-coupled transducer specifications

Type	Unfocused planar with circular aperture
Active area	0.39'' (diam.)
Ultrasonic bandwidth in air	40 kHz – 2.25 MHz
Housing	Cylindrical, 1.33'' (diam.) x 1'' (leng.)
Electrical connector	Gold MicroDot on rear of housing
Mounting	Axial or transverse, via two 1/4''-20 threaded mounts

Specific acoustic pre-amplifiers are provided by Microacoustic® for best performance at the required bandwidth; a pre-amplifier is shown in Figure 5.5, paired with a sensor. Table 5.3 contains the specifications of the sensor pre-amplifiers.



Figure 5.5: MICROACOUSTIC® broadband air-coupled Transducer (BAT)

Table 5.3: MICROACOUSTIC® charge-sensitive amplifier specifications

Gain	975 mV/pC nom.
Bandwidth	<10 kHz to >5 MHz
Polarizing voltage	Safely applied/removed from transducers, +130V standard
Transducer input	> 530 k Ω above 10 kHz, BNC-type socket
Attenuator	-20 dB in a single-step (nominal)
Output	50 Ω impedance, BNC socket
HP filter	Linear, -40 dB/decade, $f_c \approx 40$ kHz
Input power	12 VAC, 1A wall-adaptor
Dimensions	4.33'' x 1.18'' x 2.36''

5.2.3 D.I. Computation

An efficient D.I. must be robust against noise and must allow for the detection, as well as for the sizing of the defect. These performances were achieved by the appropriate selection of the signal features used for the D.I. computation. Two different approaches were followed according to the air-coupled sensors configuration (Figure 5.6). When defect detection is based on reflection measurements (“reflection mode”), the proposed D.I. (D.I.) uses the ratio between certain features of the echo signal, $F_{reflection}$, and the same features of the direct signal, F_{direct}

$$D.I. = \frac{F_{reflection}}{F_{direct}} \quad (5.1)$$

This defect detection mode can be employed by using one sensor only (Figure 5.6a), or two sensors oriented at opposite angles (Figure 5.6b). In the first case, the values of $F_{reflection}$ and the F_{direct} are extracted by selecting two different time-windows from the ultrasonic signal detected by the single air-coupled sensor as shown in Figure 5.6a. The single-sensor configuration, however, is not optimal since the detected direct signal has a low Signal-to-Noise Ratio (SNR) due to the non-optimized orientation. In the second case, one sensor is oriented toward the echo signal and the other one is oriented toward the direct signal. For both configurations, the D.I. is expected to increase with increasing defect size or % head area (H.A.) reduction. An alternative defect detection scheme that was investigated (Figure 5.6c) is based on transmission measurements (“transmission mode”) [10, 51, 118]. The proposed “transmission” D.I. uses the ratio between certain

features of the signal detected by the further sensor from the laser source (sensor #2), F_{sens2} , over the same features from the closer sensor #1, F_{sens1}

$$\text{D.I.} = \frac{F_{\text{sens2}}}{F_{\text{sens1}}} \quad (5.2)$$

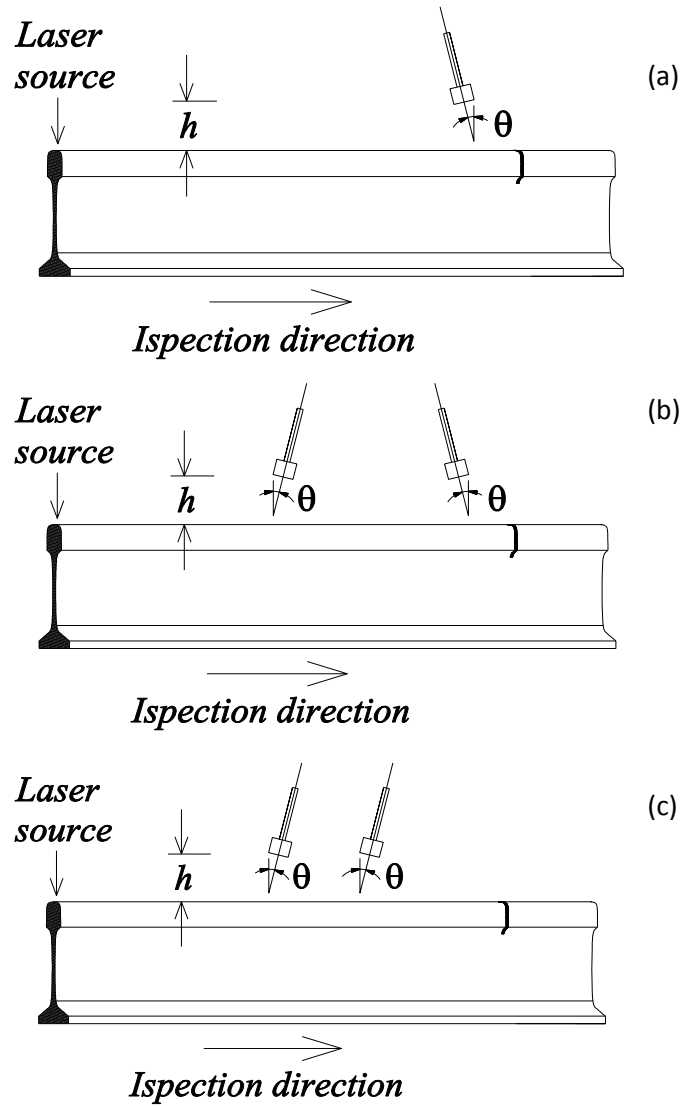


Figure 5.6: Defect detection schemes examined. (a) “Reflection mode” with a single air-coupled sensor; (b) “reflection mode” with a pair of air-coupled sensors; (c) “transmission mode” with a pair of air-coupled sensors.

As opposed to the reflection mode, in the transmission mode the D.I. is expected to decrease with increasing defect size or % H.A. reduction. The two reflection schemes have the advantage that a longer portion of the rail ahead of the probes can be inspected at once; the length of this portion being limited only by wave attenuation. The transmission scheme is instead limited by the gage length separating the two sensors. The transmission mode, however, is more robust since the SNR ratio is generally larger. Several signal features were tested to construct the D.I. vectors for the “reflection mode” and for the “transmission mode”.

These features included the variance, the Root Mean Square (RMS) and the peak-to-peak amplitude of the thresholded wavelet coefficient vectors [123,124,125]. Experimental tests confirmed that the RMS of the retained wavelet coefficients was the most efficient choice for the D.I. feature. Two D.I. were reconstructed, related to two different frequency ranges the reconstructing coefficients namely: low frequency (40-300 KHz) and high frequency (300-1200 KHz) [113].

5.2.4 Hardware Controller and DAQ Boards

The hardware controller was based on the PXI technology. The chosen National Instruments[®] 4-slot chassis PXI-1031 could be rack-mounted, and house the PXI-8186 controller, powered by Pentium 4 processor 2.5 GHz equipped with 512 MB DDR RAM onboard memory. The powerful DAQ board PXI-6115, capable of high resolution analog input-output and high speed digital output, was installed in the chassis. The BNC-2090 served as a rack-mountable BNC connector block.

The following list describes the elements that were assembled in the prototype, shown in the indicated figures:

- National Instruments® PXI Unit NI PXI-1031 (Figure 5.7);
- PXI controller NI PXI-8186 (Figure 5.8);
- DAQ board PXI-6115 (Figure 5.9);
- Terminal block NI BNC-2090 (Figure 5.10).

Table 5.4 through Table 5.7 contain the specifications of the listed hardware.



Figure 5.7: NI PXI-1031 Chassis

Table 5.4: NI PXI-1031 specifications

Electrical	
Input voltage range	100 to 240 VAC
Input frequency	50/60 Hz
Environmental	
Operating location	Indoor use
Maximum altitude	6,500 ft
Operating environment	
Ambient temperature range	32 to 122 °F
Relative humidity range	10 to 90%
Shock and vibration	
Operational shock	30 g peak, half sine, 11 ms pulse
Random vibration Operating	5 to 500 Hz, 0.3 grms
Nonoperating	5 to 500 Hz, 2.4 grms
Mechanical	
Overall dimensions (H, W, D)	6.97 in, 10.12 in, 8.38 in
Weight	11.0 lbs
Chassis materials	Sheet aluminum, extruded aluminum, cold rolled steel, nylon

**Figure 5.8: NI PXI-8186 controller**

Table 5.5: NI PXI-8186 specifications

Processor	Pentium 4M 2.5 GHz
DDR RAM	512
Trigger bus routing	Internal
Integrated peripheral I/O	100BaseTC Fast Ethernet
	GPIB (IEEE 488.2) interface
	2 USB 2.0 ports
	2 RS232 serial ports
	IEEE 1284 ECP/EPP parallel port

**Figure 5.9: NI PXI-6115 DAQ board****Table 5.6: NI PXI-6115 specifications**

Analog inputs	2-4, dedicated A/D converter per channel
Sample rate	1 to 10 MS/s per channel
Triggering	analog/digital
Coupling	AC/DC
Input ranges	8 from ± 200 mV to ± 42 V
Analog output	2 at 4 MS/s single channel, or 2.5 MS/s dual channel
Digital I/O lines	8
Counter/timers	Two, 24-bit



Figure 5.10: NI BNC-2090 connector block

Table 5.7: NI BNC-2090 specifications

BNC connectors	22, for analog, digital and timing signals
Spring terminals	28, for digital/timing signals
Chassis	shielded, rack-mountable

5.2.5 Software implementation

National Instruments LabVIEW® 8.0 platform was chosen for the software implementation, due to its convenient features which allow to :

- interact directly with the DAQ boards, through the easy-to-use Graphic User Interface;
- synchronize inputs and outputs to a common clock;
- use shared trigger for inputs and outputs;
- implement software for high speed data logging;
- implement software for real-time signal processing and conditioning;
- use predefined subroutine (Express VIs) to perform common tasks;

The routines running in Labview® 8.0 are called VIs, standing for Virtual Instruments; a VI is made of a series of objects, each of them representing a piece of hardware, virtually running.

The developed software can be divided in two sub-branches, relative to the following two steps:

- laser software triggering (digital output), programmed in the VI named Laser_trigger
- signal acquisition and processing (analog input) programmed in the VI named Acquire_process

The target of the Laser_trigger VI is creating two digital lines (TTL +5V \rightarrow +0V, falling edge triggering), to fire the Flash Lamp and to open the Q-Switch respectively. These lines need to be precisely delayed, according to the power required for the lasing, as expressed by the manufacturer (Figure 5.11). Figure 5.12 shows the block diagram of the Laser_trigger VI; the main box represents a while-loop, forcing the routine to run continuously, unless it is interrupted by the user. The two upper main lines constitute the two digital counters, used for the TTL lines. The pre-defined subVIs, on the route of the two lines, are targeted to performing the following functions:

- create the pulse task (1)
- set the timing and the pulse mode (2)
- set the trigger mode to digital edge (3)
- set the trigger source to the 2nd PFI (Programmable Function Interface) digital line, present on the back-panel of the PXI controller and on the connector-block BNC-2090 (3)

- start the task (4)
- check for any hardware/software error (5)
- clear the task (7)

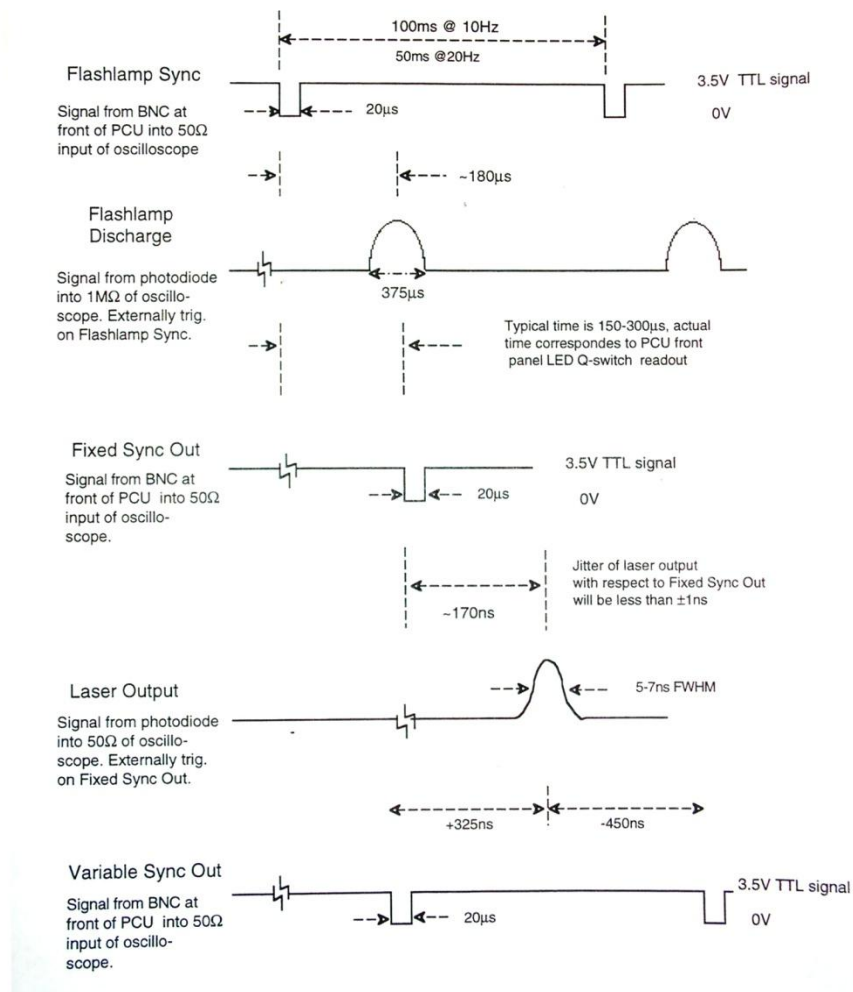


Figure 5.11: Continuum[®] Surelite SLI-20 (external trigger requirements from manufacturer).

The third line is used to create a digital trigger output for the two counters. The pre-defined subVIs, on the route of the third line, are targeted to perform the following functions:

- create the digital task (1)

- start the digital task (4)
- set the digital output to the 1st digital line DO1 (6)
- clear the task (7)
- wait until 50 milliseconds before reentering the loop (8)

In the inner while loop the Lasing ON/OFF buttons are included to give the real time possibility to start/stop the digital triggering. The stop button (9) gives the user the possibility to stop the system at any time, if needed.

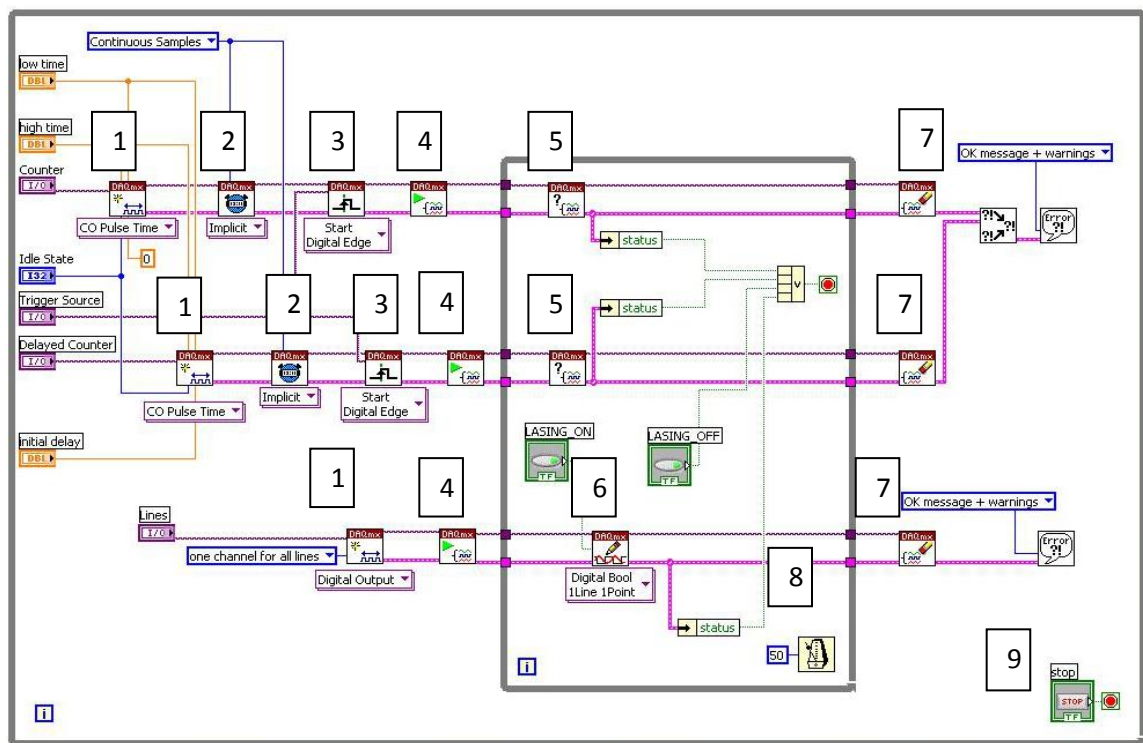


Figure 5.12: VI Laser_trigger, block diagram

In Figure 5.13 the Front Panel of the VI Laser_trigger is shown. The VI gives the user easy access to the start/stop command for the digital triggering task of the laser device.

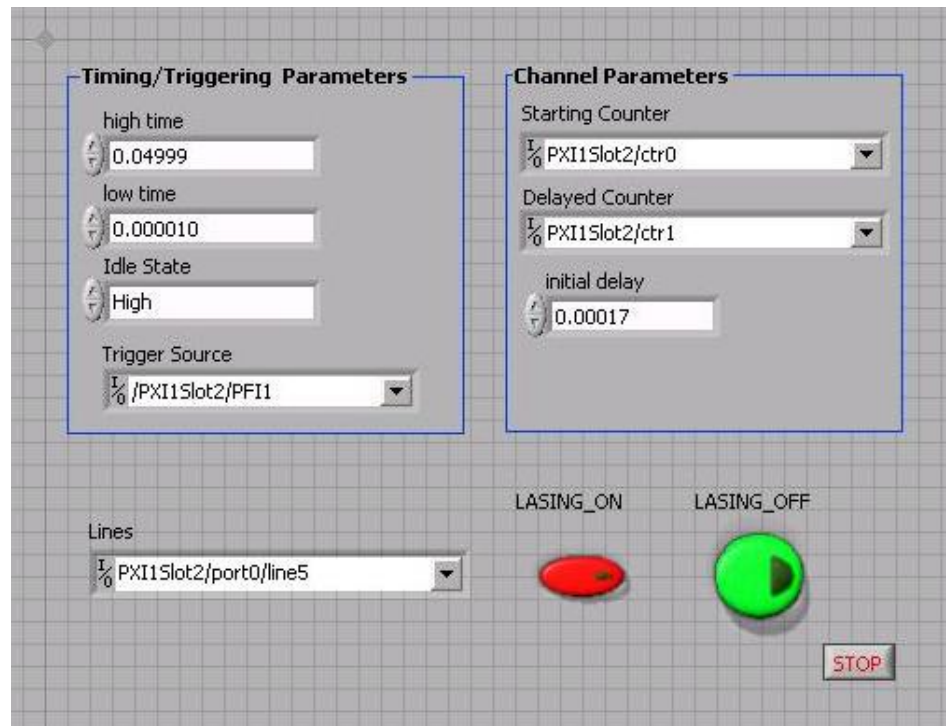


Figure 5.13: VI Laser_trigger, front panel

Figure 5.14 through Figure 5.16 show the block diagrams of the VI Acquire_process; the main box, whose left upper corner is shown in Figure 5.14 represents a while-loop forcing the routine to run continuously, unless it is interrupted by the user. The first stage of the VI is to acquire two sets of samples from the first two analog channels from the PXI-6115. The DAQ assistant is a powerful tool which allows the user to easily set scanrate, number of samples, and triggering setup of the acquisition (1). The VI is sub-divided in three blocks, representing each one a gated signals, extrapolated from the two set of samples. The first block is related to the first air-coupled transducer measuring the laser-generated ultrasonic wave. The second and the third blocks are related to the second air-coupled transducer; the second block is processing the gated signal directly transmitted by the laser source, while the third one is processing the

gated signal reflected by a possible defect positioned along the rail. The subVIs on the route of the three data lines are targeted to perform the following functions:

- process the Discrete Wavelet Transform (2), VI Discrete Wavelet Transform_Ex, applying the wavelet filters (6), subVI WaveletFilter
- threshold the wavelet coefficients (3), subVI Dwt_thr
- reconstruct the decomposed signal, allowing to use only the retained coefficients (4), VI Inverse Discrete Wavelet Transform_Ex
- compute the statistical features of the retained wavelet coefficients (5), subVI Stat2

In Figure 5.15 the ratios between the transmission and the reflection statistical features are computed, resulting in the calculation of the two D.I.'s for the "reflection mode" and the "transmission mode", respectively. The process is repeated inside the main loop for 50 milliseconds (9), in order to relieve the CPU from unneeded processing cycles. In Figure 5.16 the lower right end of the main while-loop is shown, and the stop button (10) gives the user the possibility to stop the acquisition at any time, if needed. Following are the subVIs employed in the main VI Acquire_process; each subVI can accept inputs from the user or other subVIs, and outputs data to the main VI.

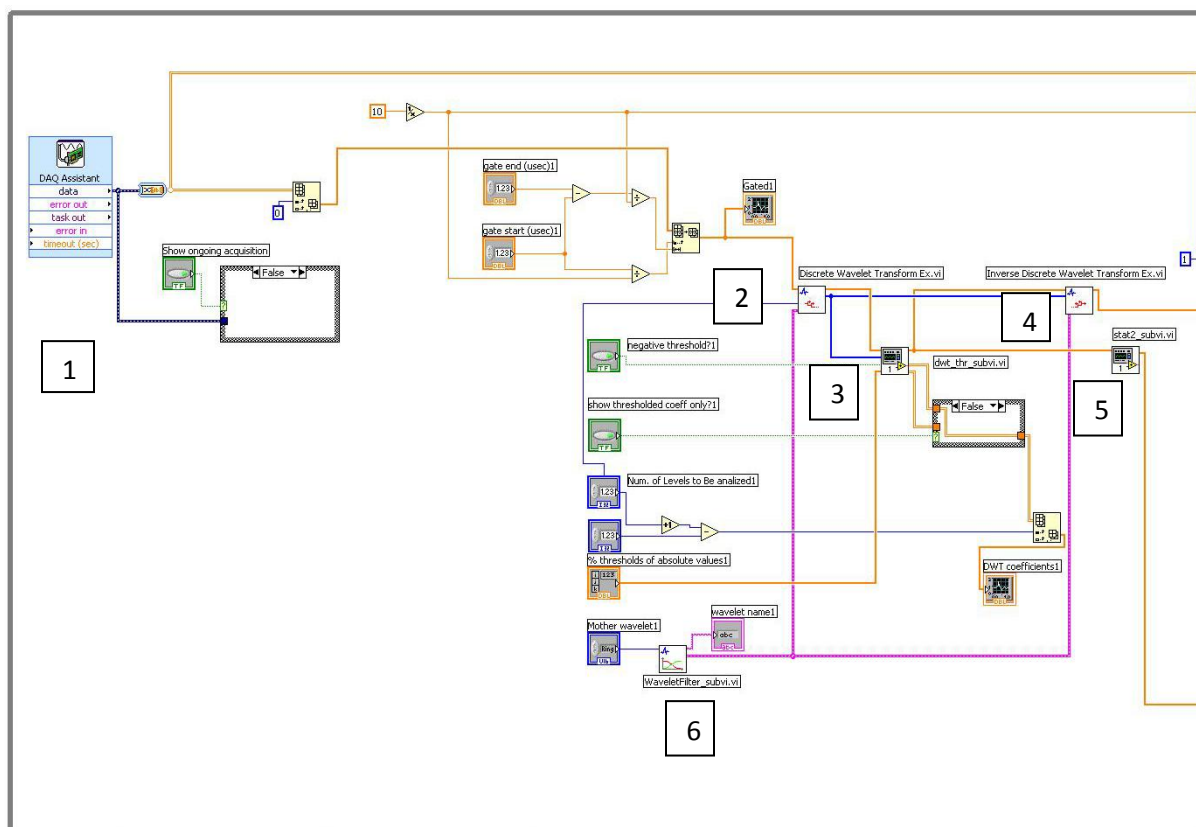


Figure 5.14: VI Acquire_process, block diagram

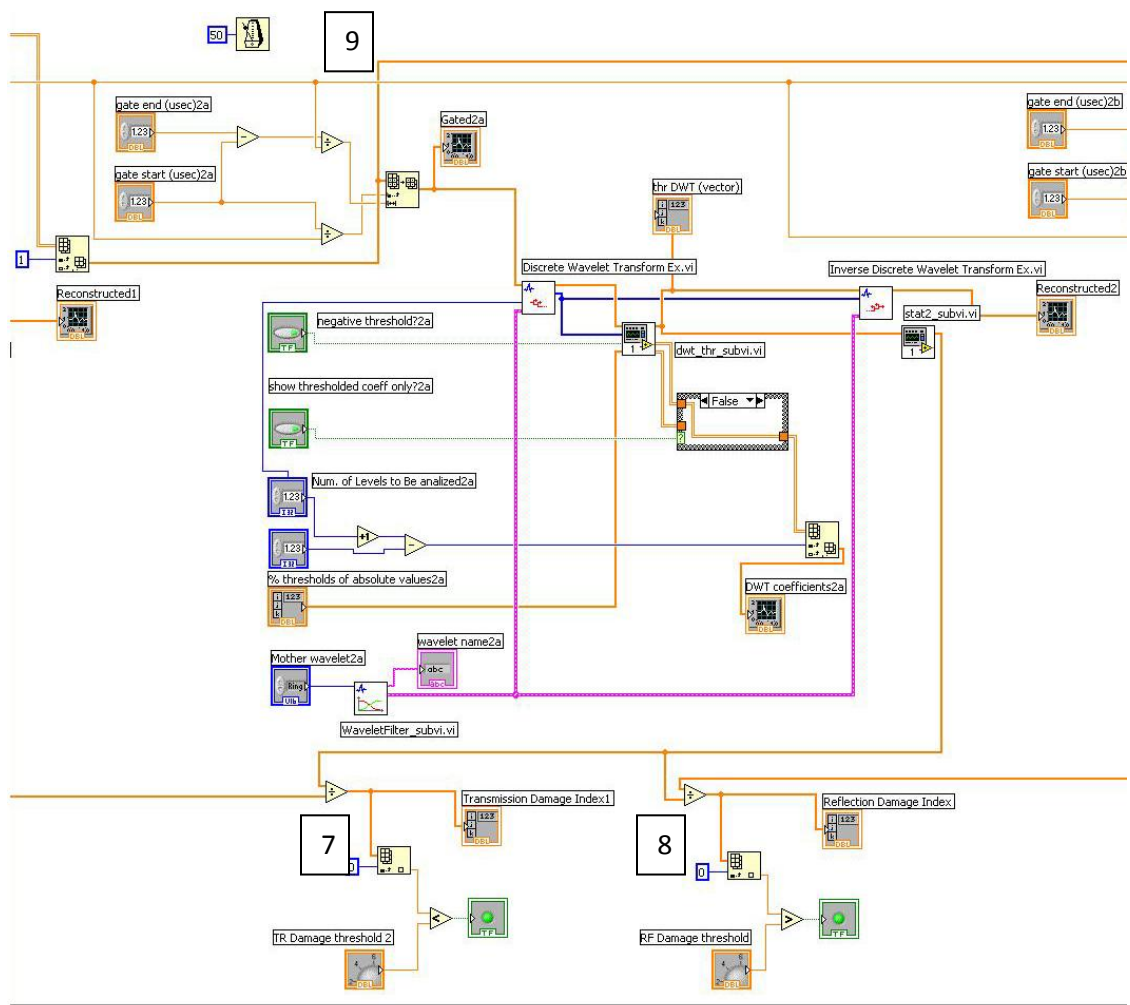


Figure 5.15: VI Acquire_process, block diagram

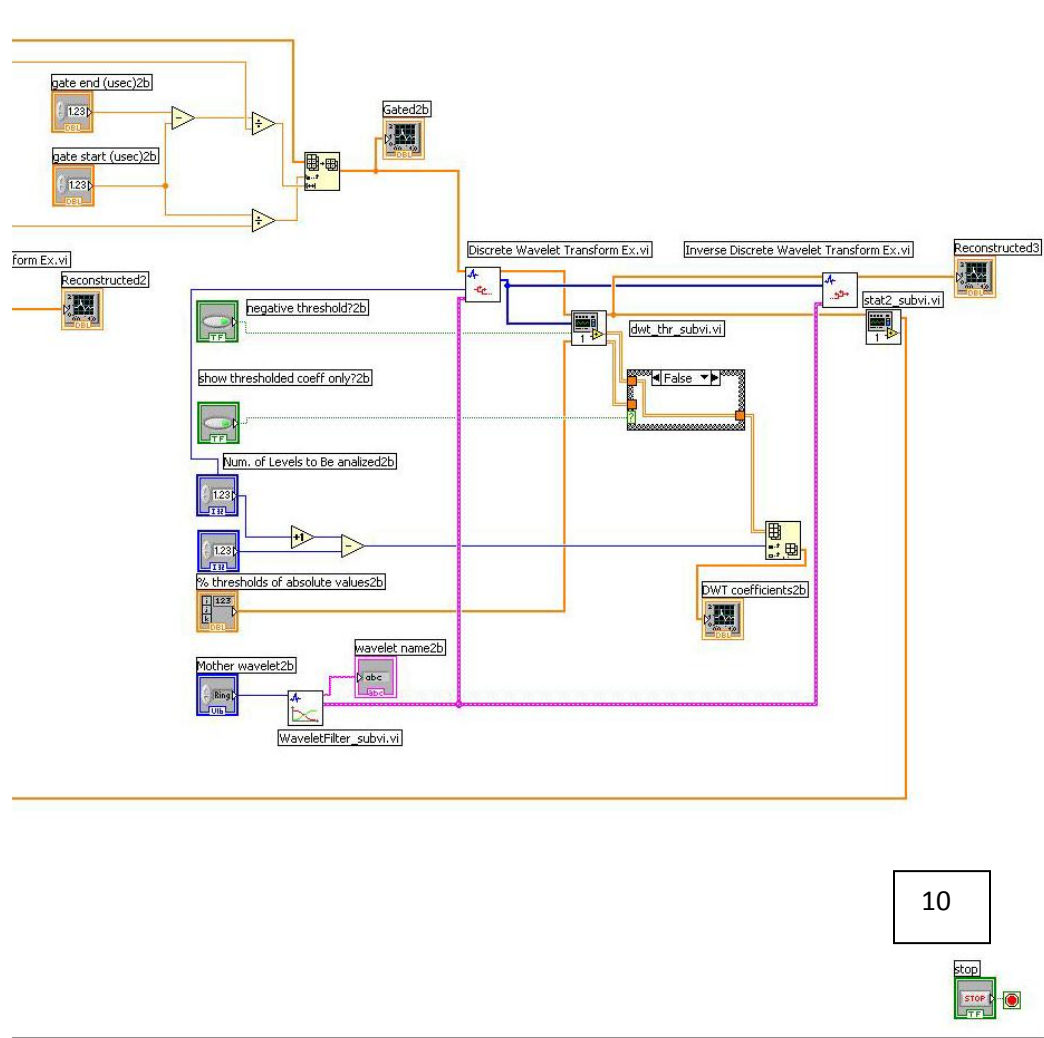


Figure 5.16: VI Acquire_process, block diagram

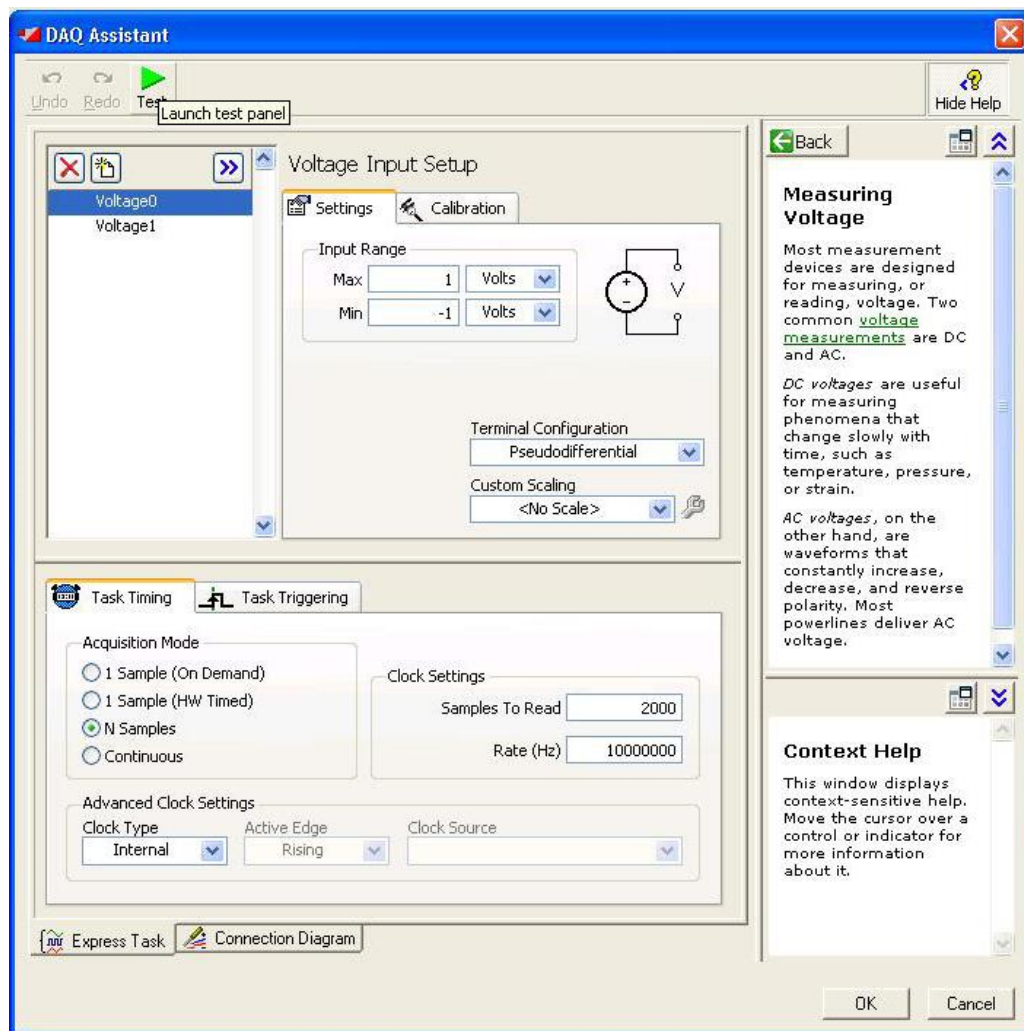


Figure 5.17: DAQ assistant, task timing

In Figure 5.17 the task timing panel of the DAQ Assistant is shown. The graphic interface allowed to set up two physical channels to be used as analog input. The chosen resolution was 2 volts, the sampling rate equal to 10 MS/s and the number of samples to be acquired equal to 2000. The hardware clock can be specified as internal (as for this case) or as external, if provided by third-part devices. The triggering task of the DAQ Assistant is shown in Figure 5.18; the trigger type was set at Digital Falling Edge, coming from the source line PFI2. No reference point was used for triggering purposes.

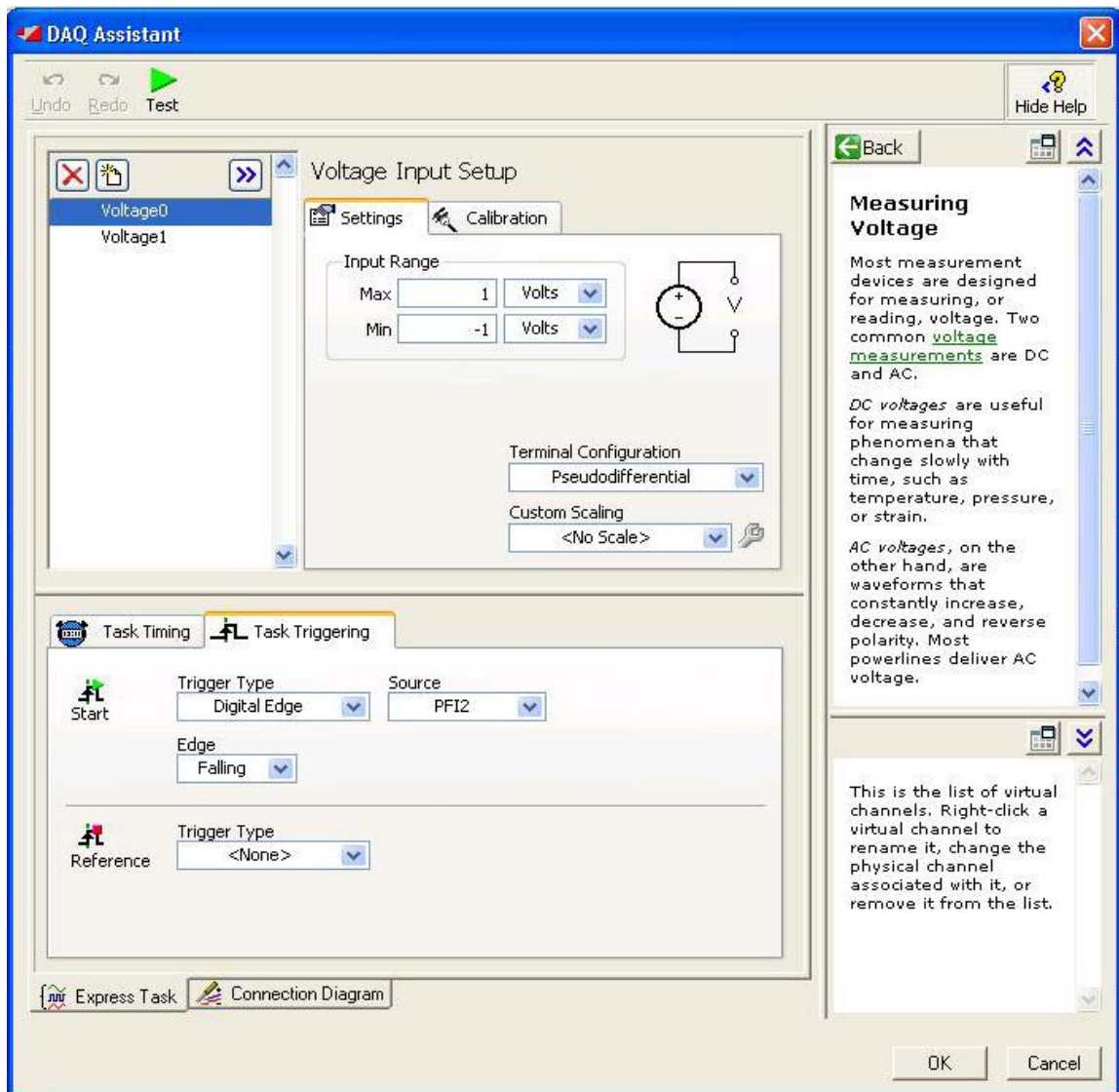


Figure 5.18: DAQ_assistant, task triggering

As shown in Figure 5.19, the subVI Stat receives as inputs the wavelet coefficients retained after the thresholding process, and calculates their statistical features including mean, the standard deviation, the variance (1), the RMS (2) and the magnitude peak values (3). The complete set of the statistical data is sent out of the subVI as the Var-RMS-PpK array, subsequently employed to compute the D.I.'s.

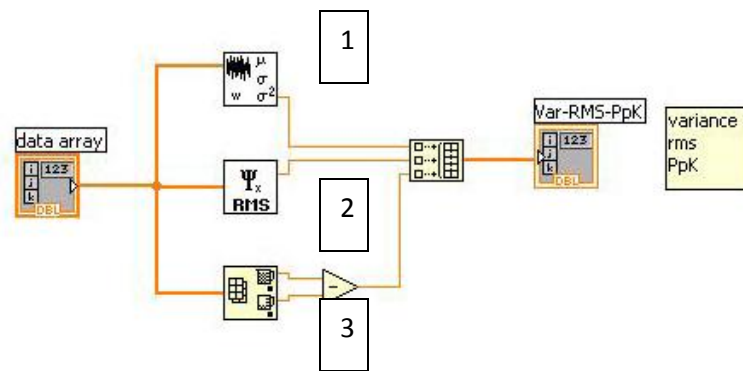


Figure 5.19: Stat subVI, block diagram

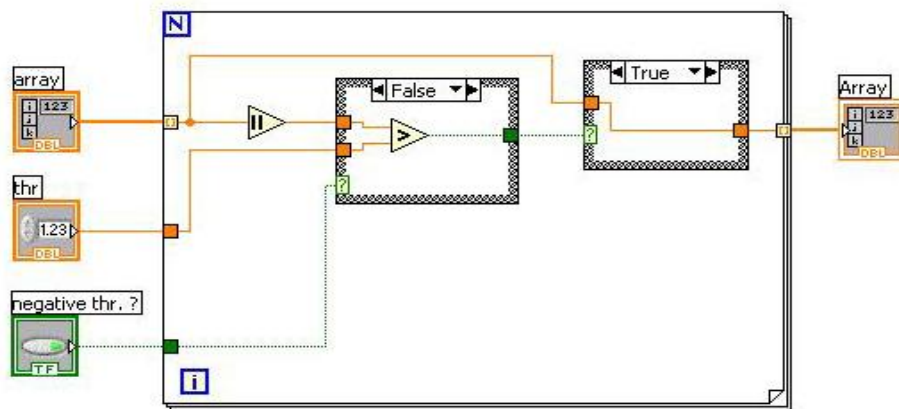


Figure 5.20: Dwt_thr subVI, block diagram

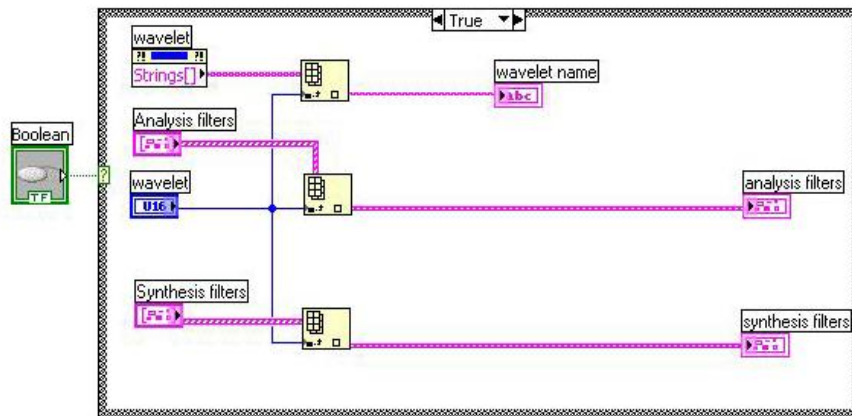


Figure 5.21: WaveletFilter subVI, 'true case' block diagram

In the subVI Dwt_thr (Figure 5.20), each element of the wavelet coefficient array is compared to a threshold value; the element is retained in the output array if exceeds the user-defined threshold.

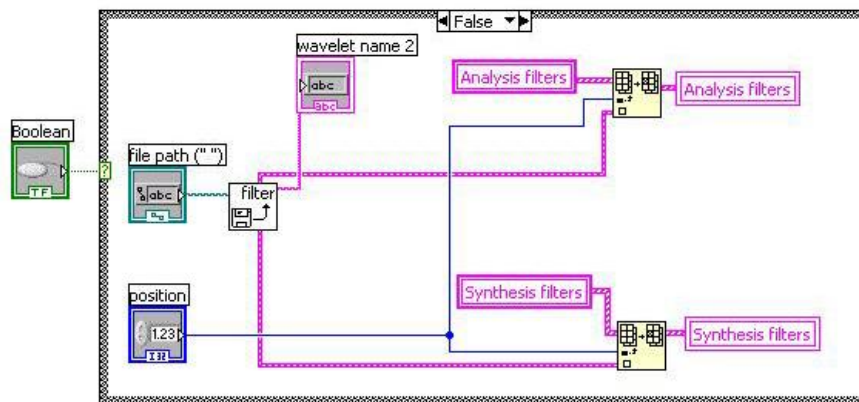


Figure 5.22: WaveletFilter subVI, 'false case' block diagram

In Figure 5.21 the subVI WaveletFilter 'true case' block diagram is shown, allowing the user to load the pre-defined Filter Coefficients to feed the Discrete Wavelet

Transform. The ‘false case’ shown in Figure 5.22, allows the user to load user-defined Filter Coefficients related to user-defined Mother Wavelet. Figure 5.23 and Figure 5.24 show the subVIs targeted to the implementation of the DWT algorithm, specifically the decomposition and the reconstruction of the time-domain signals: wav_dec and wav_rec. The wavelet coefficient array and the coefficient length array (Lengths) are outputted by the subVI wav_dec following the decomposition process. These data are then received as input by the subVI wav_rec to perform the reconstruction of the filtered signal.

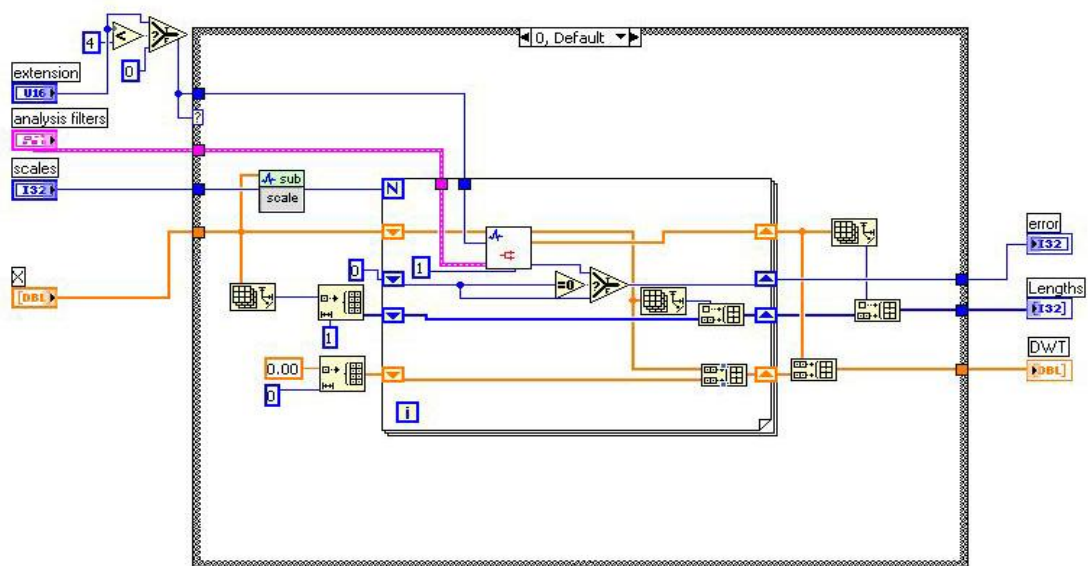


Figure 5.23: Wav_dec subVI, block diagram

Figure 5.25 shows a snapshot of the VI Acquire_process Front Panel for a case of a defect free zone; the value of the statistical feature (in this case, the variance of the retained wavelet coefficients) extracted from the DWT coefficients in the “transmission” mode, is greater than the predefined threshold. Therefore the DEFECT FREE green warning is activated. In Figure 5.26 the snapshot of the same Front Panel shows a case

when a flaw is present in the portion of rail between the two sensors, and the DEFECT red warning is activated. The second warning on the panel is related to the Reflection D.I., activated only when a reflection is detected from a defect positioned down the length of the rail. Figure 5.25 and Figure 5.26 show the configuration of the user interface of the prototype at its first stage.

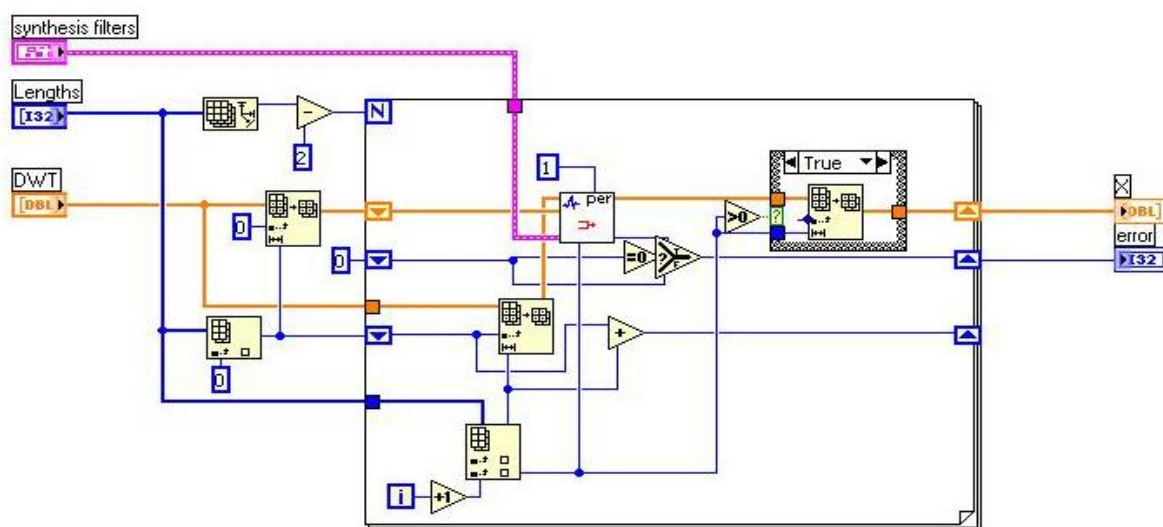


Figure 5.24: Wav_rec subVI, block diagram

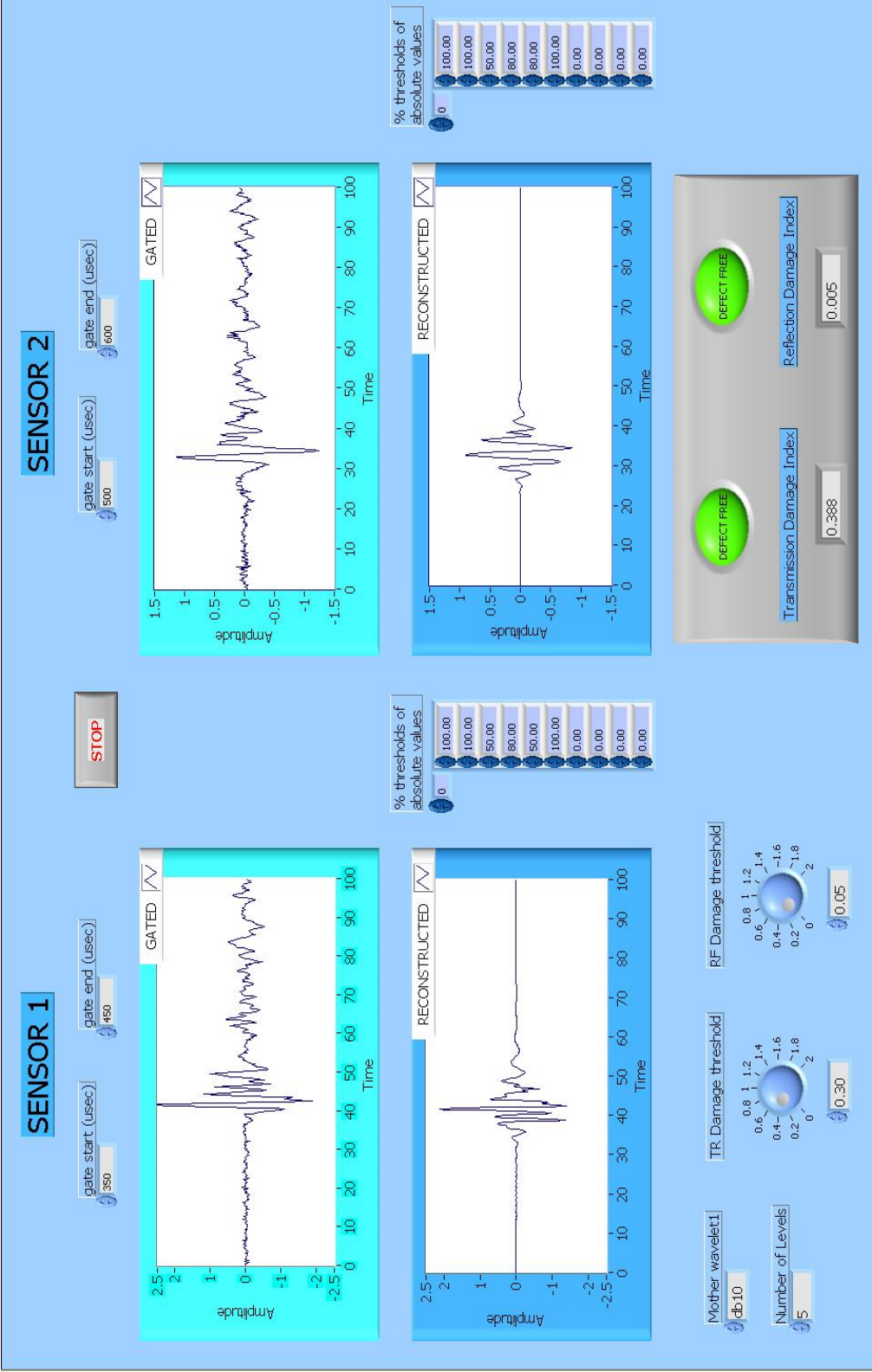


Figure 5.25: VI Acquire_process, front panel (“defect free” case).

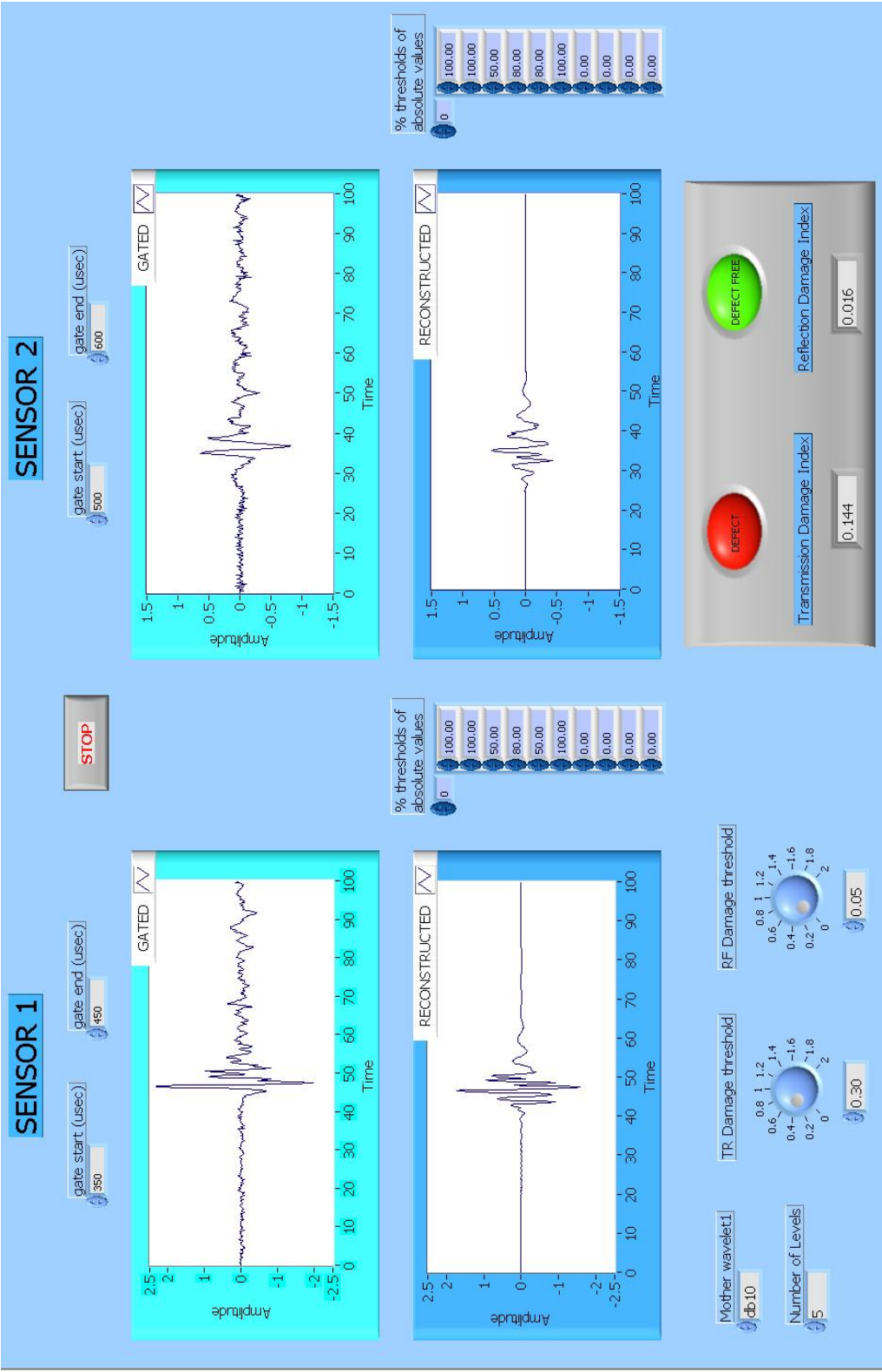


Figure 5.26: VI Acquire_process, front panel (“defect” case)

5.3 Stage II. Prototype deployment

5.3.1 Hardware layout

For field installation on the Federal Railroad Administration Hy-Railer testing car, the Big Sky Laser[®] CFR-400 (Figure 5.27) laser generator was adopted, because capable in fitting the requirements of:

- low source power;
- ruggedness of power supply, with the capability to be rack-mounted;
- ruggedness of laser head plus vertical mounting capability;
- higher repetition rate of lamp flashing (30 Hz);
- fine adjustability of the emission power;
- easy external triggering mode capability, through serial port.

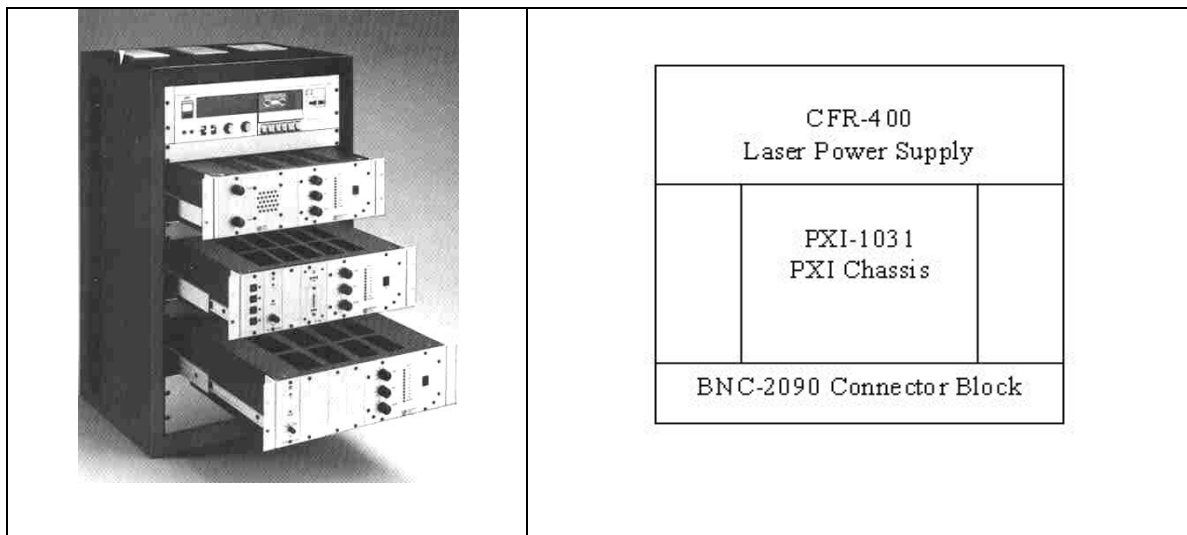


Figure 5.27: Q-Switched laser generator Big Sky Laser[®] CFR-400

Table 5.8: Big Sky Laser® CFR-400 specifications

Source Power	110 VAC, single phase, 16 A
Repetition rate (Hz)	30
Energy (mJ)	330
Pulsewidth (nsec)	7-10
Divergence (mrads)	<5
Rod diameter (mm)	7
Beam pointing stability ($\pm\%$)	100
Jitter (\pm ns)	<1
Energy stability ($\pm\%$)	<2
Power drift ($\pm\%$)	<10
Laser head mounting	horizontal-vertical
Laser head dimensions (W x H x D)	3.7''x 3.3''x 12.7''

Table 5.8 contains the specification of the Big Sky Laser® CFR-400.

**Figure 5.28: Layout of the rack-mounted hardware control panel**

The laser power supply, the PXI system and the connector block can be installed in a 19'' rack, resulting in a compact arrangement, as shown in Figure 5.28. The acting/sensing

hardware components were mounted on the cart following the layout of Figure 5.29. A common hard frame held together laser head (1), optical elements (2), transducers and pre-amplifiers (3).

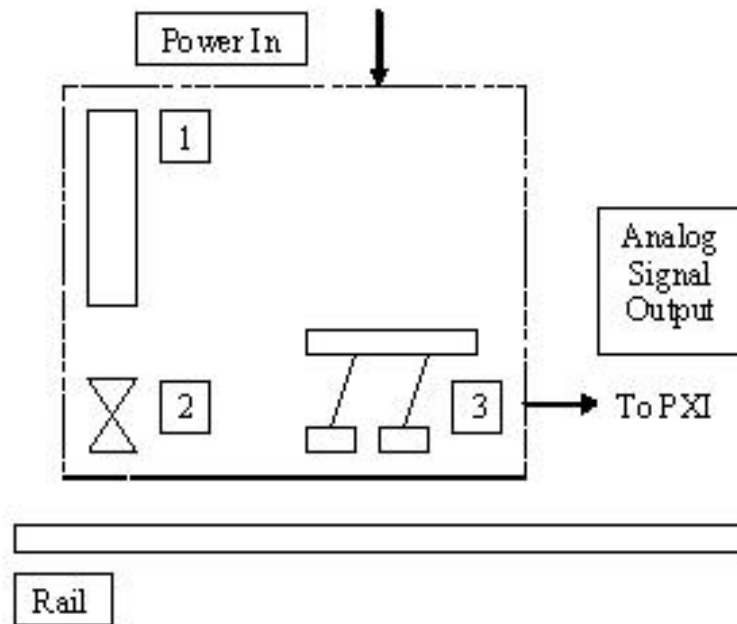


Figure 5.29: Layout of the acting/sensing components of the rail inspection prototype.

The mounting frame had to be well isolated from vibration and it had to guarantee a constant lift-off distance between the air-coupled transducers and the rail head. The movement (the excursion) of the lens system that routes the laser beam to the rail head had to be minimal. The frame lodged the laser head vertically, four air-coupled transducers, and their preamplifiers. Figure 5.30 shows the prototype while the different components were being installed on the cart.

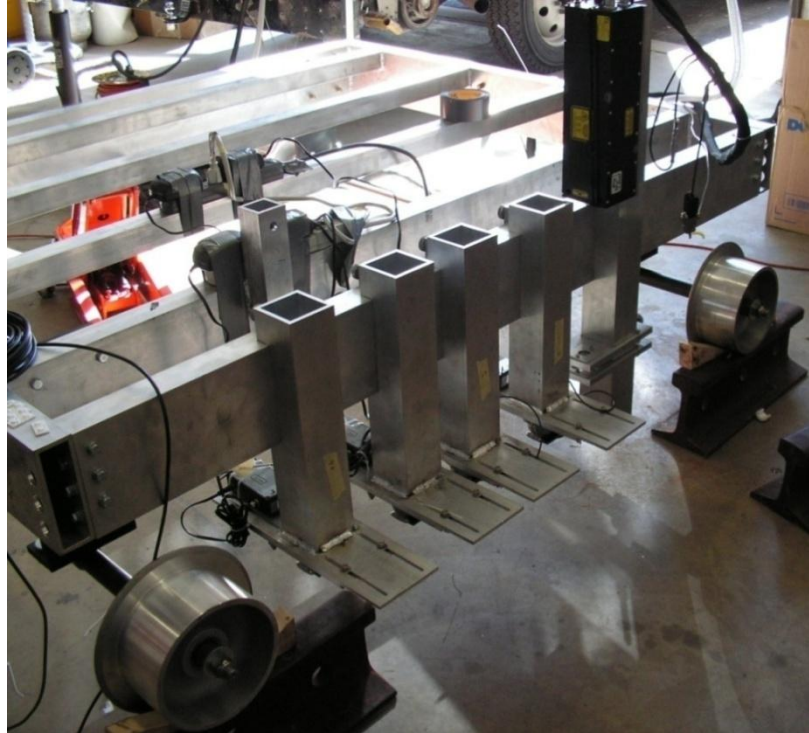


Figure 5.30: Prototype hardware mounted on the cart. Laser head and sensor holders are installed on the right beam of the main frame

5.3.2 Software implementation

The prototype's software was developed using National Instruments[®] Labview[®] technology. A laptop running Labview[®] and a National Instruments[®] PXI unit were employed as a client-server Ethernet-linked platform. The hardware layout is described in Figure 5.31. The server (the National Instrument PXI Unit) houses the Labview[®] programs needed to control the laser operation and to acquire and process the signals from the air-coupled transducers. Data from the on board positioning system installed on the research car, are constantly acquired by the server. The data consist of two digital counters relative to the sync foot and the sync pulse. Through the serial port (RS-232) the server is able to trigger the laser in order to start the data acquisition. The ultrasonic

signals detected by the air-coupled sensors are digitized through the analog inputs and they are processed and logged to the server.

Since the laser repetition rate has to be set to a constant value (maximum 30 Hz), and the rate at which the digital counters from the positioning system are sent is depending on the speed of the towing vehicle, the implementation of the overall logic of the program was not trivial. A solution to the problem was the implementation of the different activities for independent execution timing, through a Producer-Consumer scheme, as depicted in Figure 5.32.

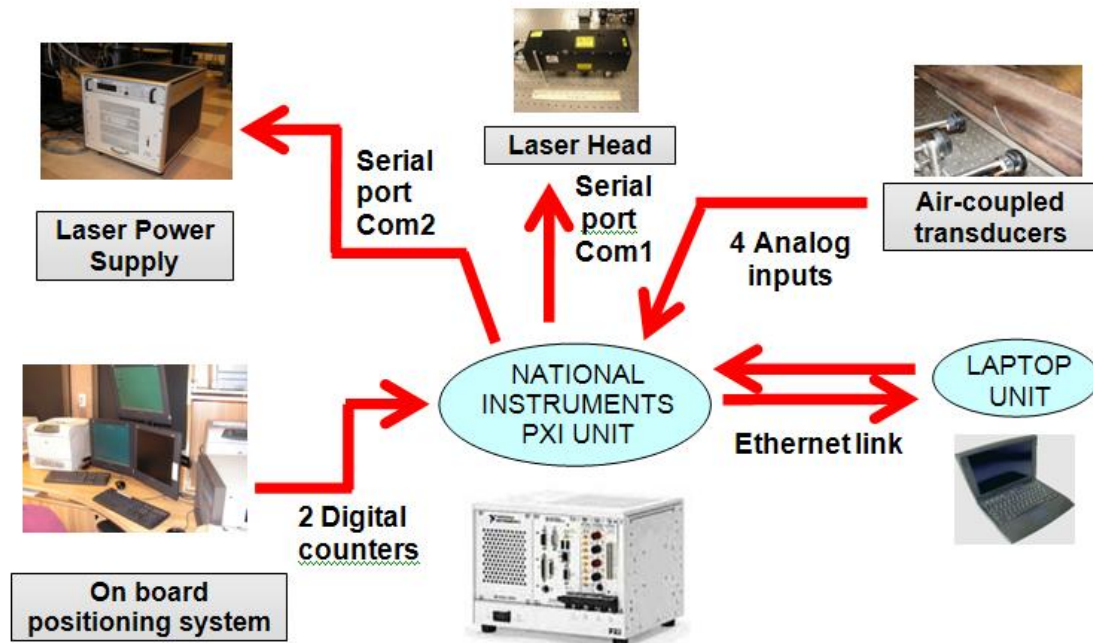


Figure 5.31: Hardware layout of the hybrid laser/air-coupled transducer Long-Range method

The laser repetition rate was set to shoot at his maximum rate, and the ultrasonic signals were continuously acquired and stored in a dynamic variable. The system would read and process the values stored in the variable only when an updated pulse was sent

from the positioning system. The activities which required deterministic execution timing, such as the analog acquisition of the signals, the digital acquisition of the positioning pulses and the operation of clustering these data together, were performed inside the Producer loop, which was set to cycle at the fastest speed allowed by the CPU. These data were acquired in a deterministic way and then put in a First In First Out Drop Tail (FIFO-DT) queue; another program would retrieve the elements from the queue and process them with less restrictions on the execution timing, that means a slower cycling of the Consumer Loop. With this method the Producer Loop had access to the CPU more frequently than the Consumer Loop, allowing for independent timing, yet not adding processing delays or errors to the data. On the client end (Laptop Unit), the user is able to start and stop the acquisition, adjust the laser power, change the lasing frequency, modify the wavelet setting, monitor the results in real-time and open a report window. A client-server architecture was implemented to relieve the CPU of the PXI (server) from the heavy computational load involved in the data display process.

Figure 5.33 shows the architecture of the link. TCP-IP messages were used to command the server for all the applications such as driving the laser, starting/stopping the acquisition. TCP-IP protocol insured the delivery of this type of command, due to the reliability of its natural implementation. While data storage was implemented in the server, UDP messages were used for delivery the processed data to be displayed at client side; the UDP protocol doesn't ensure the delivery of all data, but in case some network packets are lost on their

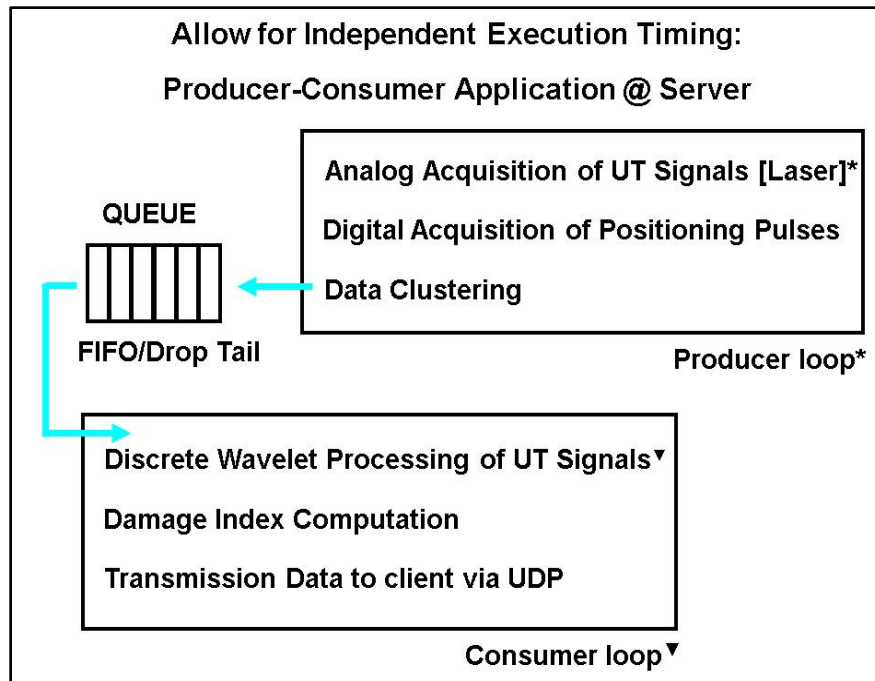


Figure 5.32: Producer-Consumer scheme, allowing the Prototype for independent execution timing

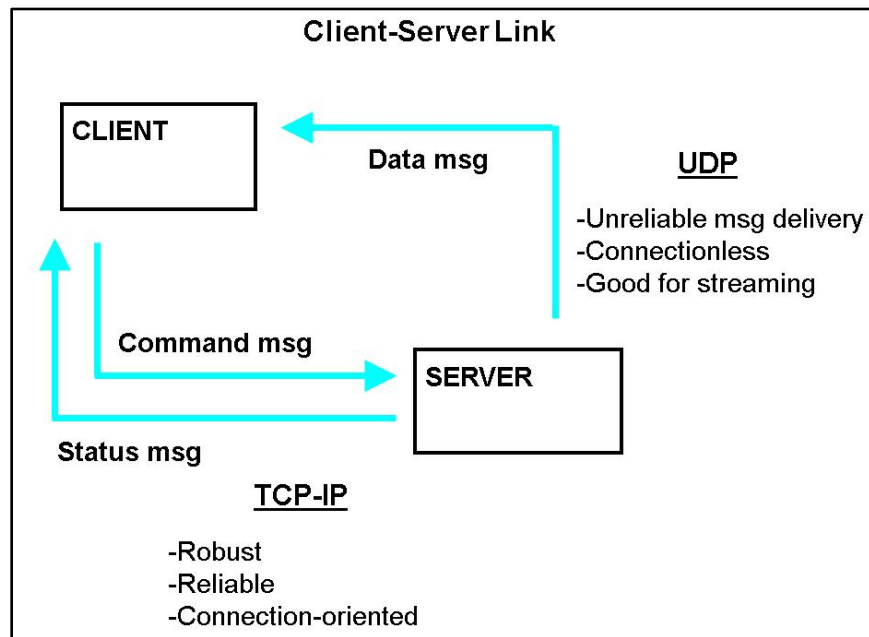


Figure 5.33: Client-Server architecture of the Prototype link

way for timing or other related issues, data that follow are still sent disregarding the lost information. Figure 5.34 shows a snapshot of the software user interface. To ease the

comprehension, the interface was subdivided into rectangular boxes. Box 1 controls the lasing frequency, i.e. the rate per second of the laser shoots. The maximum available frequency is 30 Hz. Through Box 2 the user may select the laser power output; 100% laser power means a laser beam of 330 mJ as specified by the manufacturing company. The power supply status is constantly monitored in box 3, through a series of led that are the response to a query that is periodically sent through the serial port Com 2. Box 4 allows opening/closing the laser head shutter and to start/stop the laser fire. Box 5 allows starting the calibration task or the testing task that can be activated only once “Shutter” is in the “Open” position and the “FIRE” button is on the “Fire ON” position. In addition, Box 5 contains three led associated with the PXI (server) status.

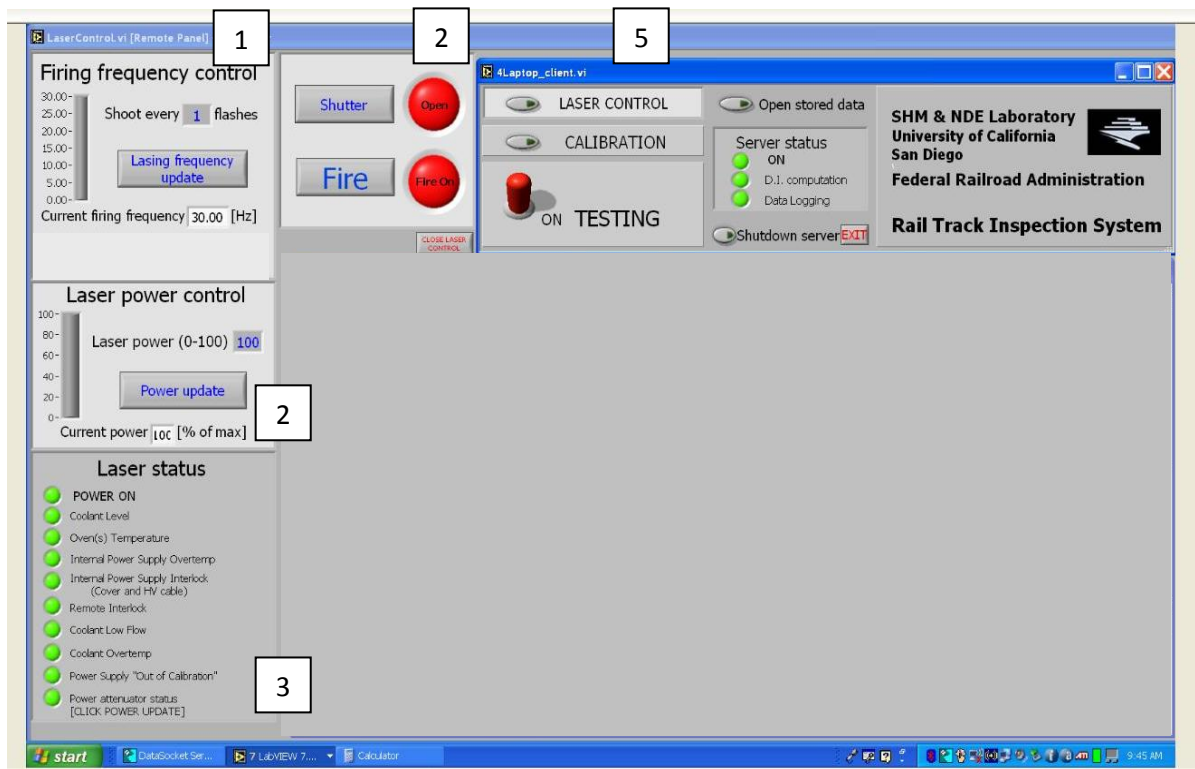


Figure 5.34: Snapshot of the Vi LaserControl that controls the laser and allows opening the calibration task and the testing task.

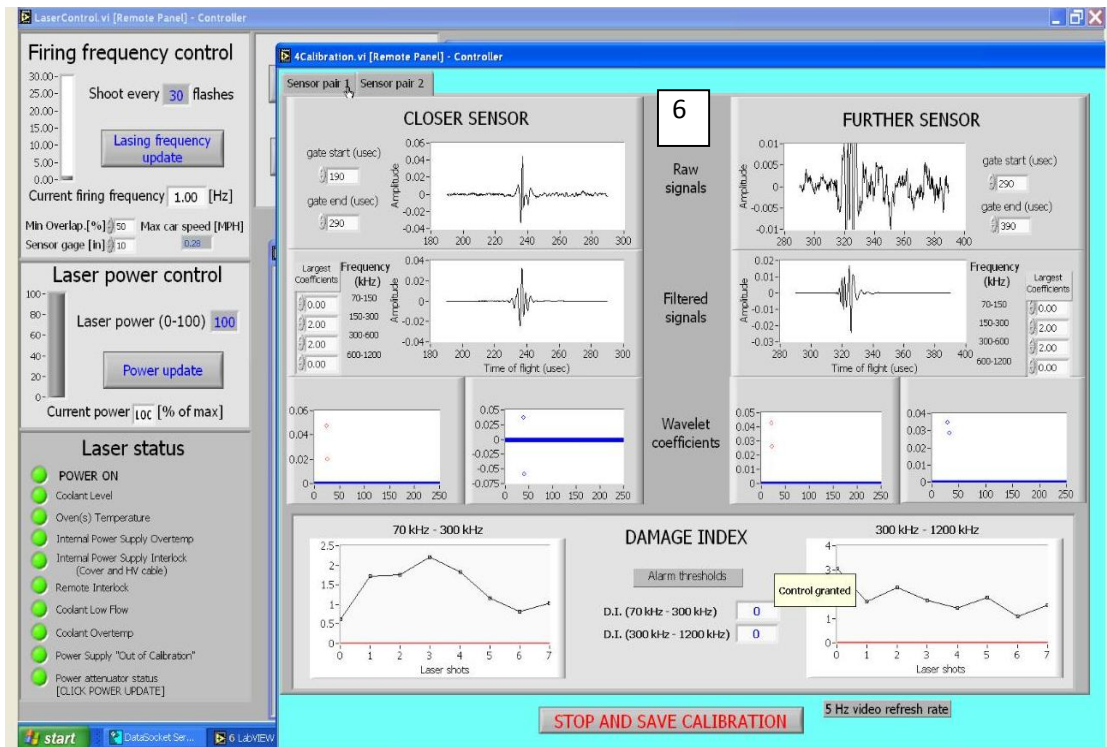


Figure 5.35: Snapshot of the user interface when the calibration task (VI 4Calibration) is activated. In the background, the VI LaserControl is visible.

Also whenever an inspection is terminated the stored data can be opened for immediate visualization through the button “Open stored data”. The server can be remotely shut down through the command “Shutdown server”. The button “CALIBRATION” in Box 5 of Figure 5.34 opens the calibration window as shown in Figure 5.35. Box 6 (calibration session) displays the following information:

- the raw ultrasonic waveform as detected by a pair of air-coupled sensors; the closer sensor is the air-coupled transducer closer to the laser source as shown in Figure 5.6c; the further sensor is the second air-coupled transducer of the pair;
- the corresponding reconstructed signals after being processed with the DWT,
- the amplitude of the wavelet coefficient vectors,

- the damage indices as a function of the laser shots.

Two frequency bandwidths were considered in order to discriminate small surface defects ($< 10\%$ H.A. reduction) from large surface defects ($> 10\%$ H.A. reduction) and internal defects. The two D.I.'s were calculated in "transmission" mode, following eq. (5.2) and choosing the RMS of the retained wavelet coefficients as the extracted signal feature. The D.I. associated with higher frequencies is hereafter indicated as high-frequency D.I. (H.F.-D.I.); the bandwidth considered ranges from 300 kHz up to 1100 kHz. The D.I. associated with lower frequencies is hereafter indicated as lower-frequency D.I. (L.F.-D.I.); the bandwidth considered ranges from 40 kHz up to 300 kHz. The longer wavelengths associated with the latter range were assumed to be sensitive to the presence of large or deep surface defects and to the presence of internal defects. The calibration panel includes a series of controls which enable the user to:

- select the temporal window of the raw waveforms;
- select the n largest wavelet coefficients to be considered for the feature extraction task and for the de-noising process;
- establish the D.I. threshold which discriminates between inliers (data associated to an healthy rail) and outlier (data associated with a damaged structure);
- switch from sensor pair 1 to sensor pair 2 by clicking on the tab placed on the top left portion of box 6.

Once the calibration phase is completed, the user saves the settings and closes the calibration window by clicking the button "STOP AND SAVE CALIBRATION". The front panel will appear again as in Figure 5.34. Now the operator is ready to start the inspection by switching the "TESTING" command to "ON" position. The front panel

will appear as in Figure 5.36. The testing window shows two tabs: D.I. visualization folder and Report.

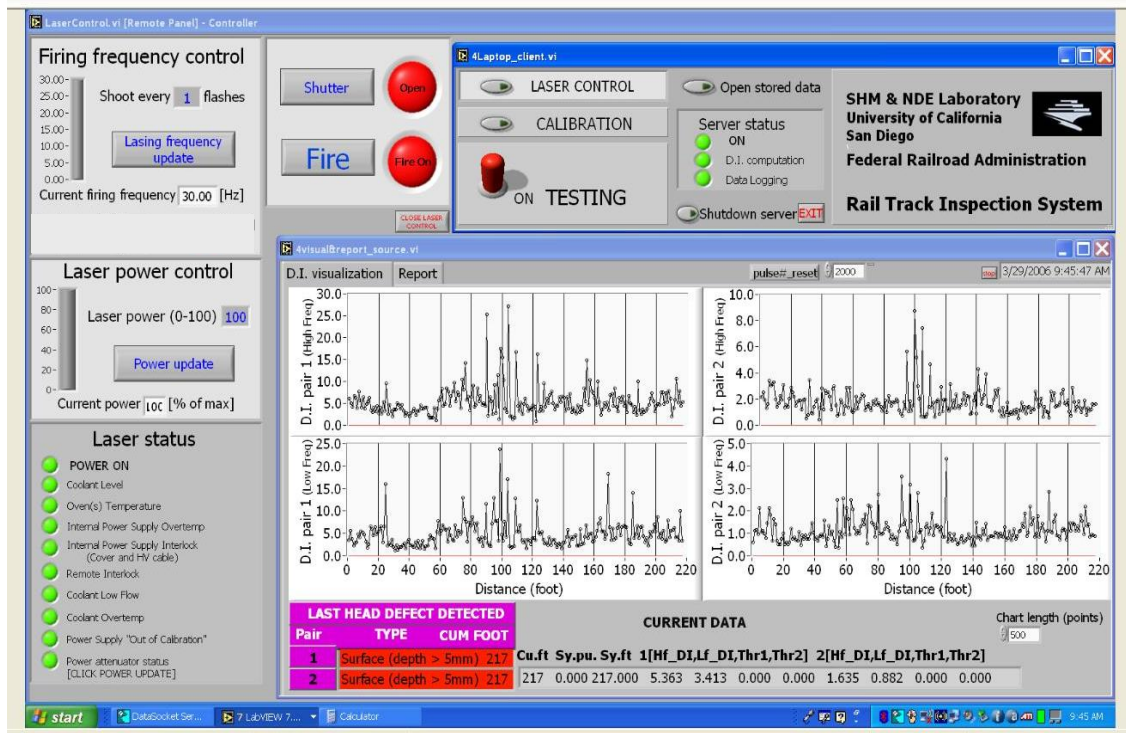


Figure 5.36: Snapshot of the user interface when the testing task (VI visula&report) is activated. In the background, the VI LaserControl is visible.

The first tab displays four plots (two plots per sensor pair) of the D.I. as a function of the track position (foot). Starting from top left and going clockwise, the window displays the H.F.-D.I. of sensor pair 1, the H.F.-D.I. of sensor pair 2, the L.F.-D.I. of sensor pair 2 and the L.F.-D.I. of sensor pair 1. The window also includes the “Chart length (points)” control to select the desired length of the plots, the “CURRENT DATA” indicator that displays the values of the last acquisition, the “LAST HEAD DEFECT DETECTED”

indicator, which shows the type, and the position of the last defect detected during the inspection.

By clicking on the “Report” tab, the front panel will appear as in Figure 5.37. Whenever an anomaly (outlier) is detected, the top row of the report is updated with the following information (from the left to the right column):

- Head area defect (pair 1 and pair 2). The current version of the software classifies three types of defects: small surface defects (only the H.F.D.I. detects an outlier), internal defect (only the L.F.D.I. detects an outlier), large surface defect (both H.F.D.I. and L.F.D.I. detect an outlier);
- Sync foot: counter of pulses sent every foot by the onboard positioning system;
- Sync pulse: counter of the pulses sent by the onboard positioning system to synchronize the sync foot counter; the control “pulse#_reset” sets the period (number of sync foot pulses) of the sync pulse;
- Cum foot: counter of the cumulative feet received from the beginning of the acquisition;
- H.F.D.I.1: value of the D.I. from sensor pair 1;
- Thr: correspondent D.I. threshold
- Same for the other three pairs of columns

It must be noted that the table in Figure 5.37 does not reflect any real inspection but it is just a software simulation. In the top row, the first pair defect is a small surface defect due to the fact that $H.F.D.I. > Thr$ (outlier) but $L.F.D.I. < Thr$ (inlier); the second pair instead both D.I.’s are above the correspondent threshold and then the second pair of sensor detected a large surface defect.

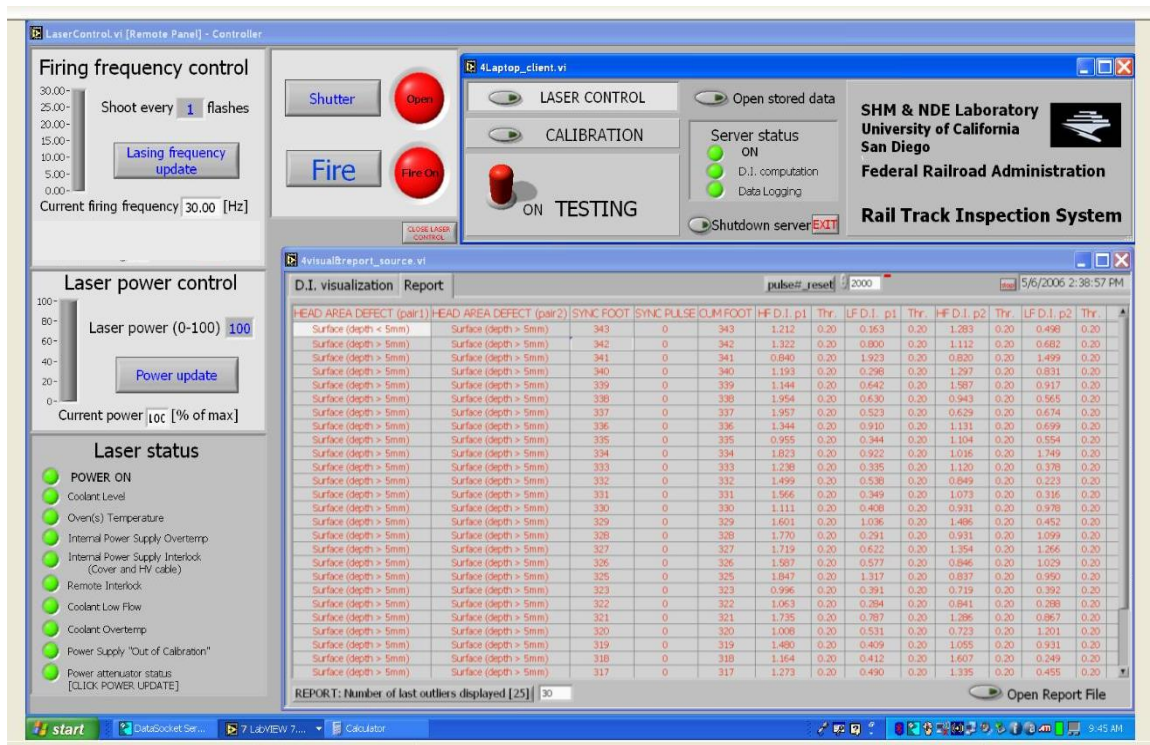


Figure 5.37: Snapshot of the user interface when the report is shown during the testing task (VI 4visual&report) is activated. In the background, the VI LaserControl is visible.

The same information contained in the Report window are stored in a .xls file that can be opened by clicking on the button “Open Report File”, placed in the bottom right portion of the window. A control placed in the bottom left portion of the window allows selecting the number of row simultaneously visible in the report. An indication of the current time is displayed in the top right. Finally the top center control “pulse#_reset” fixes the number of sync foot at which the onboard position system acquire a new sync pulse. The software described in this section controls and processes up to four channels, i.e. four air-coupled transducers. However, the software is coded in a modular fashion that easily allows the addition of more sensor pairs.

5.4 Stage III. Prototype upgrade

5.4.1 New cart

For the second field test the cart was completely redesigned and constructed by ENSCO to provide greater stability than the first version at sustained testing speed.

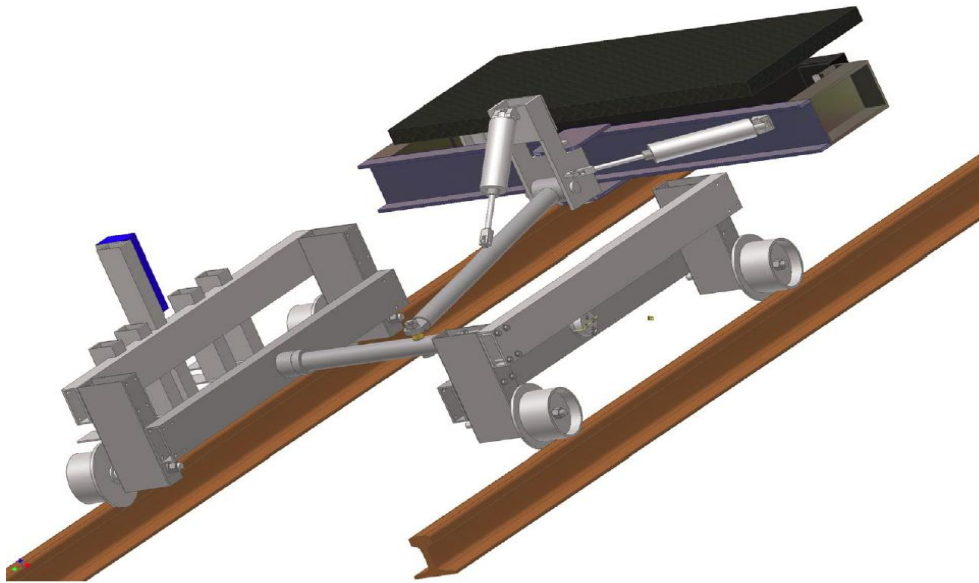


Figure 5.38: 3D view of the mechanical model of the upgraded cart by ENSCO.

The cart was designed to be stable at up to 30 mph, although the maximum speed allowed by the railroad at the test site was 15 mph. A 3-D model of the upgraded cart is shown in Figure 5.38. As an improvement to the rigid-frame solution of the first cart, each side of the new cart is independent of the other, to avoid intermittent stick-slip at sustained speeds. The side frames can rotate with respect to the center beam. Two pneumatic pistons apply forces laterally and vertically to provide, respectively, flanging ability with the gage side of the test track, and overall stability of the cart.

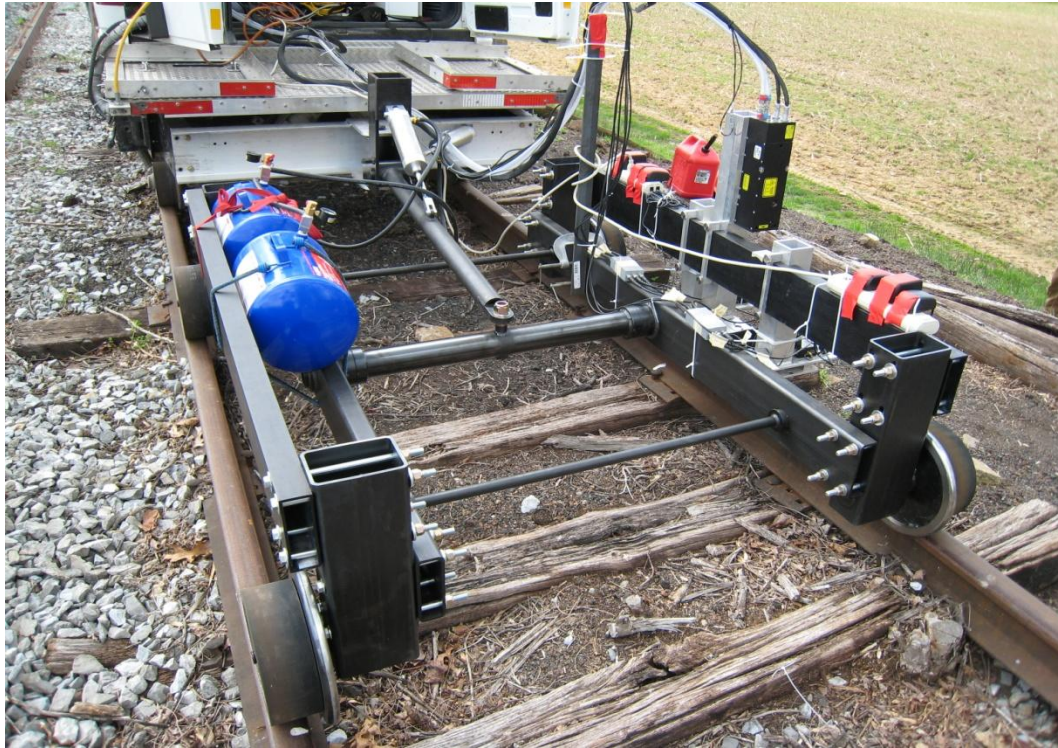


Figure 5.39: Rail flaw detection prototype installed on the cart during second field test.

Compressed air tanks, shown in the photo of Figure 5.39, provide the air supply to the pistons. The connection between the swing arm and the center beam uses a spherical joint to allow for variations in cross level between the cart and the towing vehicle. Rubber treads were installed on the wheels so as to have a smoother rolling of the cart, as well as to damp out shocks due to vertical misalignment present along the track at the joint connections. Figure 5.39 shows a picture of the prototype assembled and deployed on the track.

5.4.2 Improved sensor arrangement

As depicted in Figure 5.40, a new sensor arrangement was used by exploiting the bidirectionality of the laser ultrasound generation. In this arrangement, sensors were

placed on either side of the laser pulse. Compared to the previous same-side sensor disposition, the new solution resulted in a more compact prototype. Moreover, the measurements of the two sensors are more comparable as both are located at the same distance from the ultrasound source. Defect detection was achieved by computing a D.I. which was conceptually equivalent to the transmission D.I. expressed previously in eq. (5.2) based on the ratio of the strength of the signals detected by the two sensors in a pair. When a defect is located on either side of the laser source, the relative strength between the two sensors changes as a result of the decreased transmission through the flaw. Consequently, the D.I. changes from his nominal value of 1. The optimal distance between two sensors in a pair (i.e. gage length) was fixed at $\sim 8'' (\approx 203 \text{ mm})$. Another improvement consisted of the addition of two sensors to the original four to form a total of three sensing pairs (six sensors)..

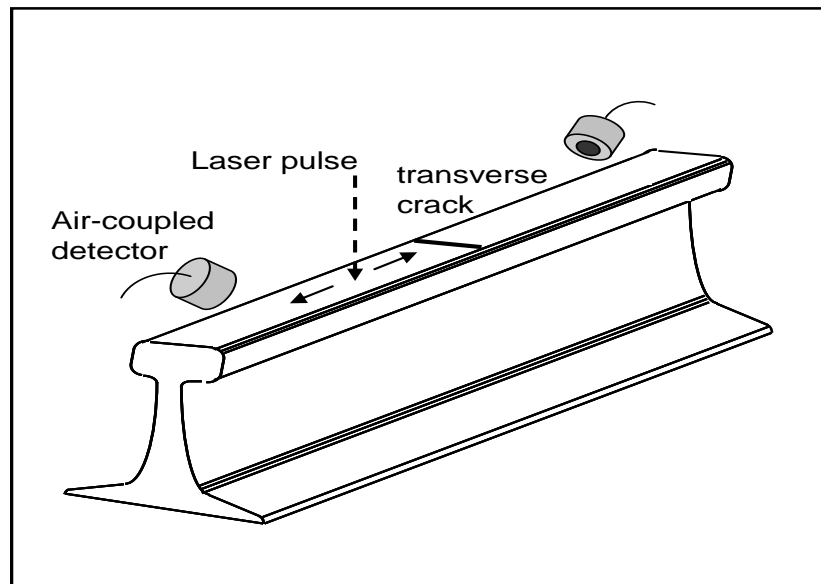


Figure 5.40: Improved sensor arrangement adopted in second field test to exploit the bi-directionality of the laser ultrasound generation.

The three pairs provided coverage of the centerline of the rail head (Center Head pair), of the gage side of the rail head (Gage Side pair), and of the field side of the rail head (Field Side pair). Review of the strain energy distribution for some propagating modes discussed in chapter three, also suggests the usefulness of monitoring all sides of the head.

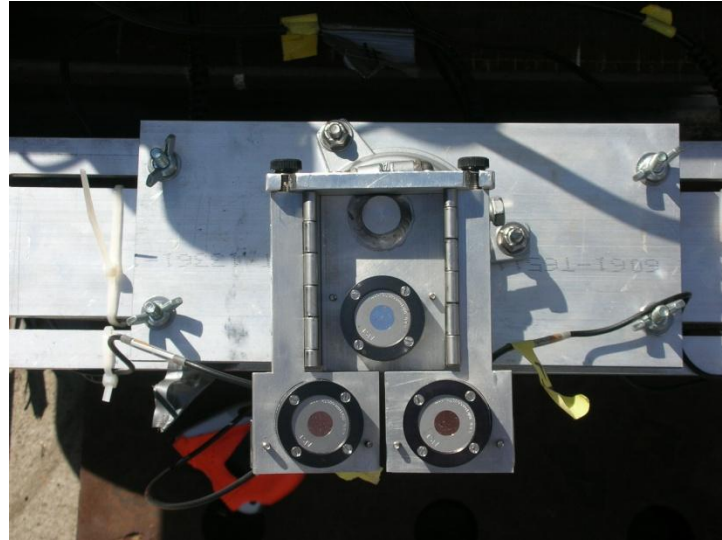


Figure 5.41: Sensing devices: air/coupled transducers at the gage side, center head and field side

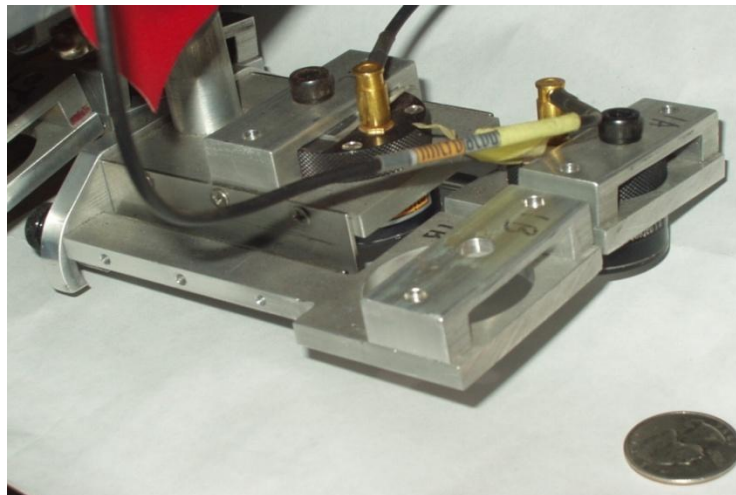


Figure 5.42: Details of the three-sensor holder

Besides the above three pairs, the D.I. was calculated for two extra pairs, named “cross-channel pairs,” consisting of the following sensors: Gage Side rear and Center Head front; Gage Side front and Center Head rear. These cross-channels provided additional defect detection paths which are oblique relative to the rail running direction increasing the overall head coverage. Figure 5.41 and Figure 5.42 show the details of the 3-sensor aluminum holder on one side of the laser generation. To allow for adjustments, the holder was connected to the main frame through a spherical joint; the holder also housed two lateral adjustable wings to provide the gage- and field-side sensors with an extra degree of freedom.

5.4.3 Software upgrade

A more efficient, frequency-domain filtering of the acquired signals was employed through the implementation of better performing, real-time signal processing algorithms. Such improvement was needed to enable testing at sub-foot spatial resolution (as low as 1”) without data loss. Butterworth band-pass digital filters were used to filter out unwanted noise, as the one generated by the air-shock due to the laser pulse, and other noise coming from the testing environment. Figure 5.43 shows the block diagram of the LabView subVI performing the bandpass filtering of the measurements; on the left the input parameters, such as the low and high cut-off frequencies and the filter order, are set; on the right the filtered signals are outputted. Figure 5.44 shows a snapshot of the new version of the Calibration Session; the left and right sides of the screen refer, respectively, to the rear and the front sensors of a given sensor pair (in this case one of the cross-channel pairs).

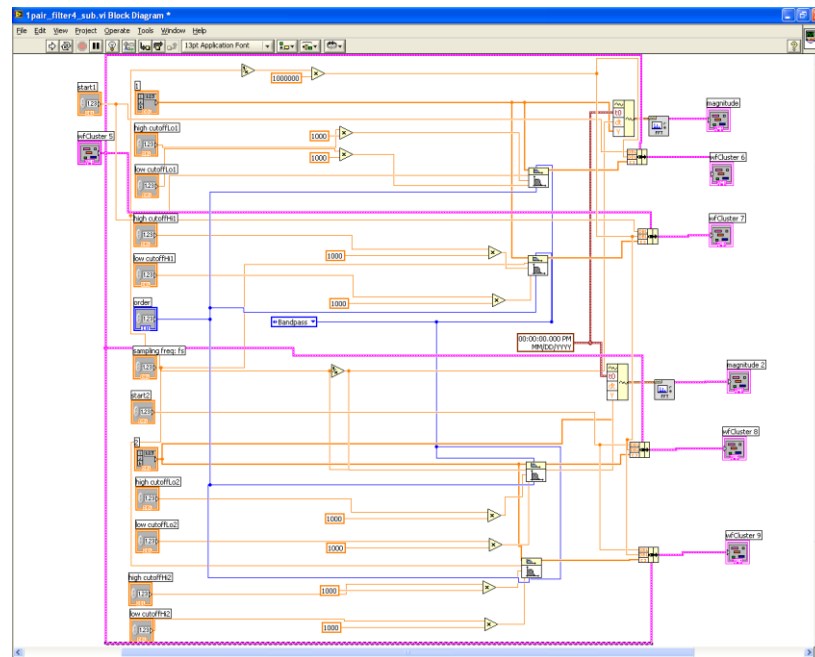


Figure 5.43: Layout of the block diagram of the bandpass filtering subVI.

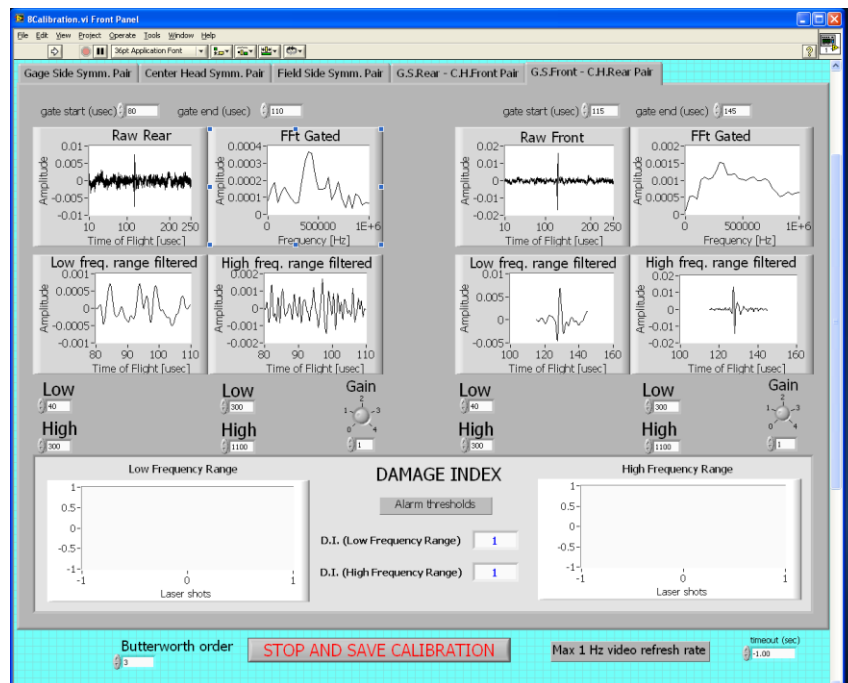


Figure 5.44: Snapshot of the Calibration Session showing the various settings for the digital filtering of the ultrasonic measurements.

For each sensor the time-domain raw signal and its frequency content are shown; the user can set the limits of the time domain gate and the low/high frequency values of the bandpass filter. Two ranges of frequency can be set for the filtering, so as to allow the computation of the two L.F.-D.I. and H.F.-D.I. for defect classification.

5.5 Field testing

5.5.1 First field test

The test site (Figure 5.45) was located near Gettysburg, PA, and consisted of a dismissed portion of railroad at crossing #593-381c. In the picture, the location of the railroad crossing and the side of the inspected track are outlined. In total, a length of 200 ft (≈ 61 m) was tested.

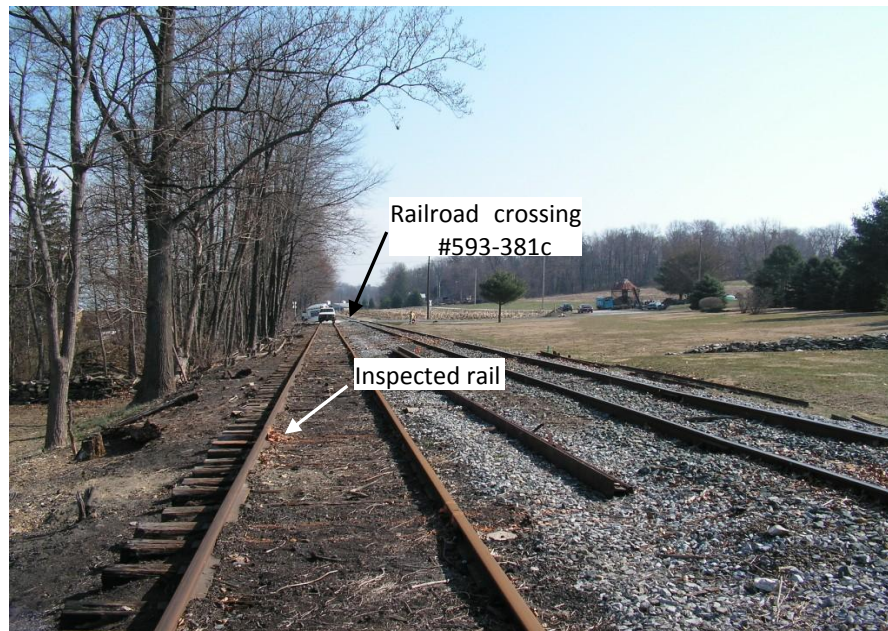


Figure 5.45: Test site near Gettysburg, PA.

Details of the site lay-out are presented in Table 5.9. The most significant discontinuities, along with their locations (distance from the start of the test zone), are

listed. Ten joints were present; some of them created a gap as large as ½ inch between two rail sections. Three, 6-ft long, 139-lb A.R.E.M.A. sections with known internal defects in the head were inserted in the inspected portion of rail and fixed by joint bars. The three internal defects were hand mapped prior to testing. The mapping was done by the UCSD team under FRA supervision using the Krautkramer USN 58L ultrasonic search unit.

Table 5.9: Test site layout

Gettysburg Test Site Lay-out			
Location	Foot	Inches	Description
1	0	0	Start of Test Zone
2	13	3	Joint
3	32	0	Chip on gage side of rail head
4	46	3.5	Joint - Orange paint
5	52	5	Large chip (6" long) on gage side of rail head
6	56	2	Large chip (6" long) on gage side of rail head
7	78	9	Flaking on gage side of rail head
8	79	3.5	Joint - flaking on gage side of rail head at joint
9	81	7	Cut surface transverse notch 4mm depth, 50 mm length (~5% H.A)
10	82	7.5	Cut surface transverse notch 2mm depth, 25 mm length (~2% H.A)
11	84	5	Joint - Rail head change 1/2in gap
12	85	6	1st High joint bar (trolley wheel rides up)
13	86	4	Internal transverse defect w/surface head checks (~4% HA)
14	87	3.5	Beginning 2nd high joint bar (trolley wheel rides up)
15	88	2	Welded ramp to match rail sizes
16	88	9.5	Joint
17	90	3.5	End 2nd high joint bar (trolley wheel rides up)
18	91	3.5	Internal transverse defect (~35% HA)-possibly close to surface
19	93	10	Joint - 1/2 in gap
20	95	1	Cut surface oblique (+45deg) notch 3mm depth, 30mm length (~3.5% H.A) (ADDED LATER)
21	96	4	Internal transverse defect (~12% HA)
22	97	8	Cut surface oblique (-45deg) notch 3mm depth, 30mm length (~3.5% H.A) (ADDED LATER)
23	98	9	Joint - Rail head change 1/2in gap
24	99	5	Flaking on gage side of rail head
25	103	0	Chip on gage side of rail head
26	106	7	Flaking chip on field side of rail head
27	112	3	Joint
28	121	0	Flaking chip on field side of rail head
29	126	8	Flaking chip on field side of rail head
30	145	3	Joint (double check) and
31	146	4	Beginnin flaking (severe) on gage side of rail head (continuous)
32	161	0	End flaking (severe) on gage side of rail head (continuous)
33	174	8	Flaking chip on field side of rail head
34	178	3.5	Joint
35	179	0	Flaking chip on field side of rail head
36	191	6	Flaking on gage side of rail head
37	200	0	Finish of test zone

A classical pulse-echo, ultrasonic wedge inspection was carried out for the mapping. As customary in TDD defect mapping, a 2.25 MHz transducer was used with a 70° wedge. The orientation of each defect was determined from the inspection direction; the size of each defect was determined by hand scanning and computing the appropriate conversions between the position of the transducer and the extent of the reflector (defect). Three independent mappings were conducted for each defect, in order to extract a statistically meaningful size. All defects were predominantly oriented at 20° from the vertical to the rail head, as common in TDDs (Figure 5.46). Internal defect 1 (location 13 in Table 5.9) was inclined

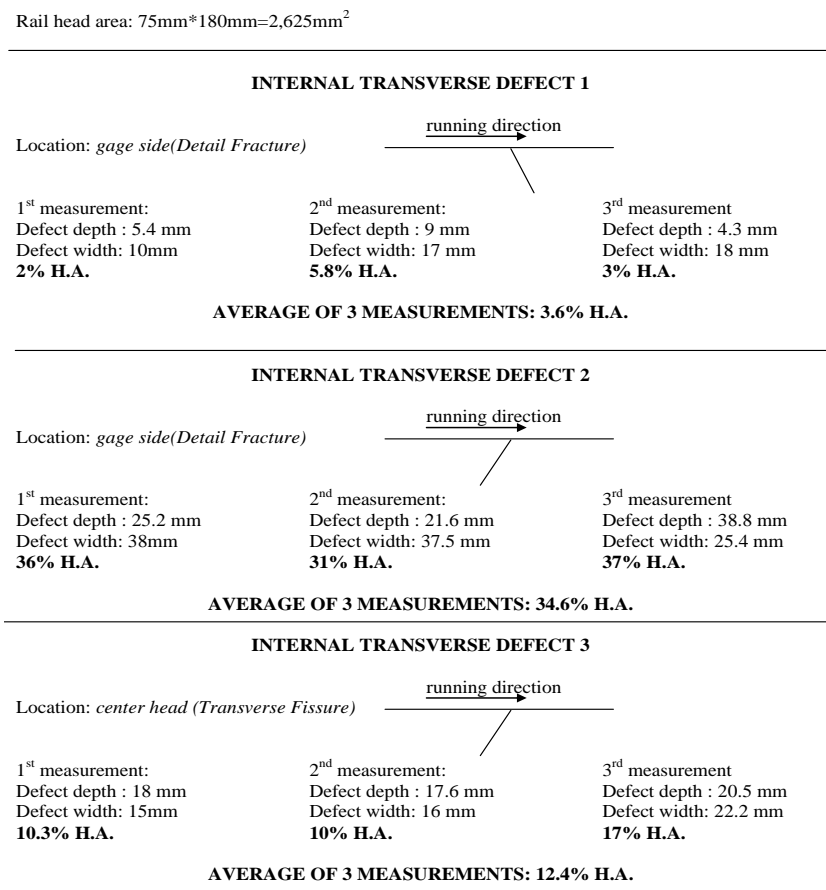


Figure 5.46: Internal defect mapping (2.25 MHz ultrasonic transducer with 70° wedge)

away from the running direction; it was located in the head gage side and its average size from three measurements was 3.6% H.A. reduction. Internal defect 2 (location 18 in Table 5.9) was oriented towards the running direction, also in the gage side, and its average size was 34.6% H.A. reduction. Internal defect 3 (location 21 in Table 5.9) was oriented as the previous defect, but it was located in the center head; its size was measured at 12.4% H.A. reduction. Photos of the rail sections containing the internal defects are shown in Figure 5.47. The estimated location of the defects (indicated in the photos with the acronym TDD) is marked by a red circle.

Two surface cuts were also introduced at locations 9 and 10, respectively, using a 4" disc grinder, model N9501B by Makita Corporation. The defects were machined perpendicular to the rail running direction causing 5% and 2% H.A. reductions, respectively. A photo of the surface cuts is shown in Figure 5.48. Two oblique (45 degree inclination from the running direction) surface cuts were also added at locations 20 and 22, respectively; each cut corresponding to a reduction of about 3.5% H.A. As shown in Figure 5.49, the oblique cuts were created close to two joints at locations 19 and 23, respectively.

Pictures of the laser and air-coupled sensors installed on the cart are shown in Figure 5.50 and Figure 5.51. The cart was designed and constructed by ENSCO under an FRA University support task. The black metallic box was devised to protect the laser head and the lenses from accidental impacts and dust; in addition the box avoids accidental reflections of the laser beam.

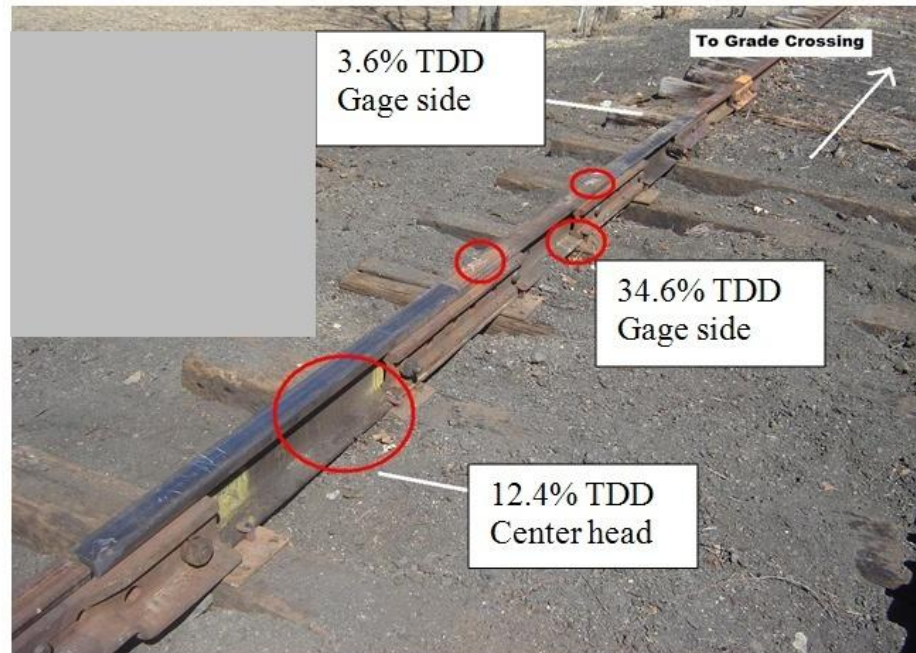


Figure 5.47: Rail sections with internal defects plugged in the railroad.

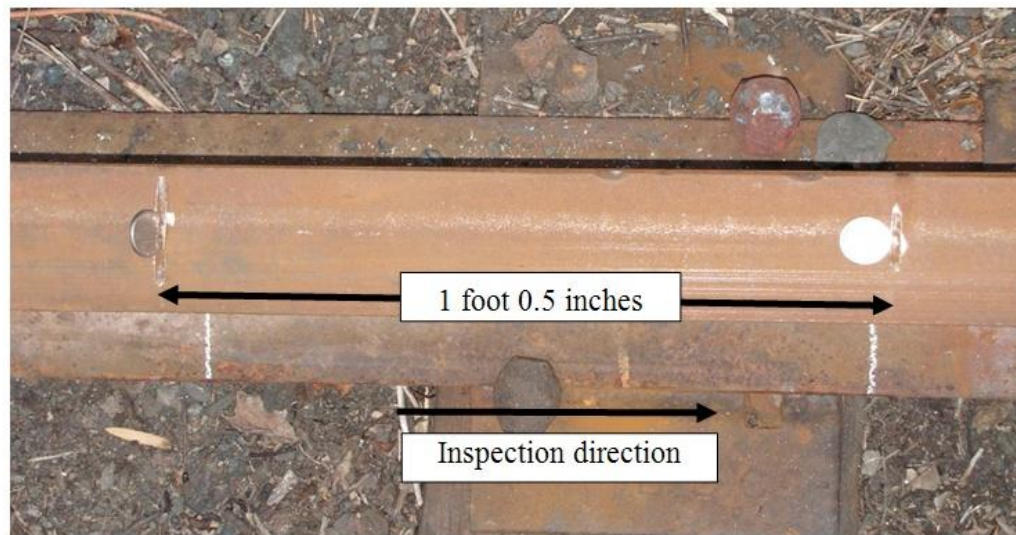


Figure 5.48: Particular of the surface transverse cuts.

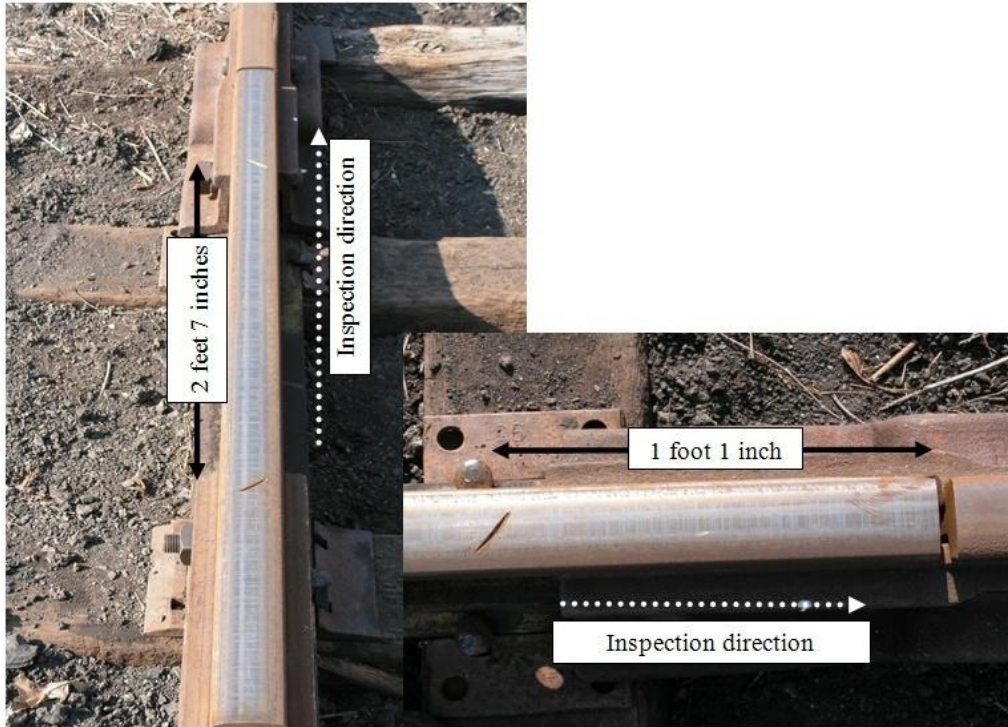


Figure 5.49: Details of the oblique surface cuts added later.

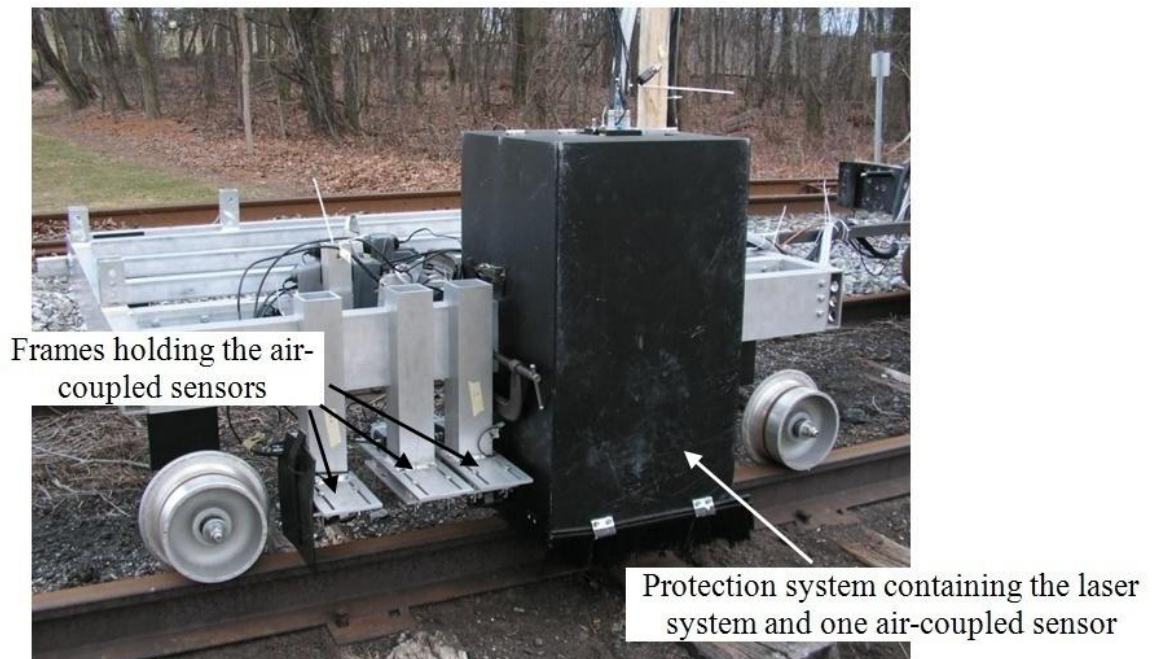


Figure 5.50: Laser head and sensors on Ensco's cart during first field tests.

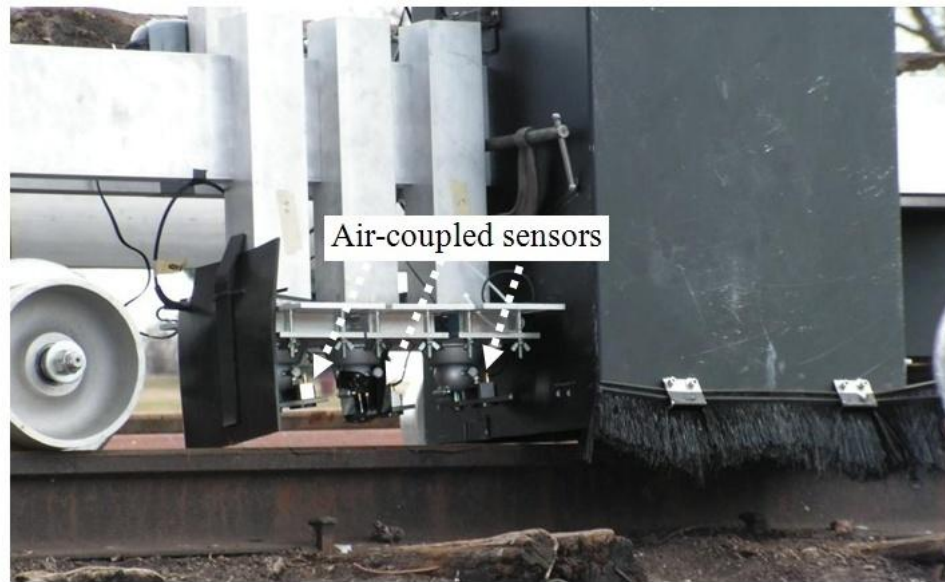


Figure 5.51: Detail of the air-coupled sensors.

The sensor layout used in the field tests is schematized in Figure 5.52. The minimum distance between the first sensor (#1) and the laser was chosen so as to avoid the superposition between the air shock and the ultrasound traveling in the rail. The angle θ , theoretically equal to 6.3° for maximum sensitivity to surface-type waves, was optimized manually by monitoring the amplitude of the detected waveforms. Sensors #1 and #3 formed the first pair, hereafter indicated as sensor pair 1 (s.p.1); sensors #2 and #4 formed the sensor pair 2 (s.p.2). The sensor lift-off distance h was maintained equal to 2.5'' (≈ 64 mm). The s.p.1 was placed right above the center of the rail head, whereas the s.p.2 was placed closer to the rail gage-side corner. The lateral offset between the two pairs was around $\frac{3}{4}$ inch.

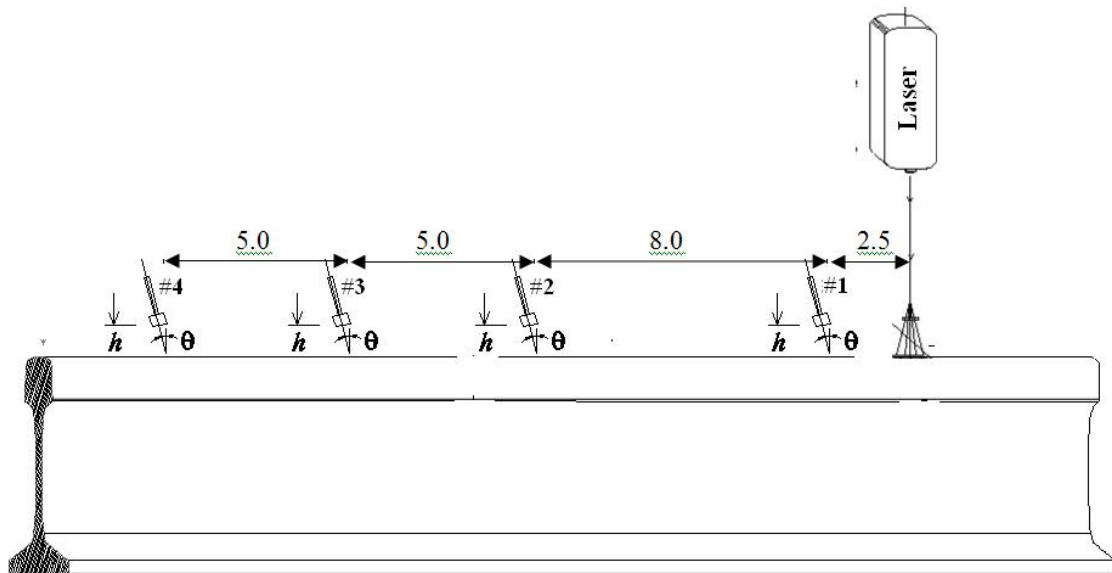


Figure 5.52: Laser/air-coupled sensor layout for the field test. Dimension in inches. Drawing not to scale



Figure 5.53: The inspection prototype towed by the FRA Hy-railer managed by ENSCO during the first field tests.

The inspection prototype towed by the Hy-railer is shown in Figure 5.53. The back end of the truck hosted the laser power supply, the PXI unit, and the laptop client. A 1,200W generator was used to power all the equipment. In these field tests, a total of 32 tests were conducted in two days and recorded along the 200-ft rail test site described in Table 5.9. A summary of all test conditions is provided in Table 5.10. The testing sessions started and ended slightly before and slightly after the monitored rail segment, as indicated in the fourth and fifth columns of the table. Tests 1÷10 and 23÷26 were performed under “dry” conditions, i.e. ultrasound was generated by the laser pulse irradiating a dry surface of the rail head. In the remaining tests, the rail was wetted by using a water pressure sprayer.

Table 5.10: Summary of test conditions.

Test conditions (Rail flaw detection)			Gettysburg Test Conditions (ENSCO/UCSD Rail Flaw Detection - March/29-30/2006)							
Test num	Date	Time (PST)	Position before start point (0 ft)	Position after finish point (200 ft)	Water (Y/N)	Movie file (Y/N)	Firing freq / /overlap	Sensor gage (pair 1)	Sensor gage (pair 2)	DWT coeff. @ freq bands
1	3/29/2006	9:26	12' 7"	8' 4"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
2	3/29/2006	9:44	10' 5"	9' 7"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
3	3/29/2006	9:52	7' 8"	8' 6"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
4	3/29/2006	10:29	10' 4"	5' 1"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
5	3/29/2006	10:37	16' 3"	7' 8"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
6	3/29/2006	10:45	11' 9"	4' 9"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
7	3/29/2006	10:50	12' 0"	6' 6"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
8	3/29/2006	10:56	12' 3"	6' 9"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
9	3/29/2006	11:01	12' 3"	8' 10"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
10	3/29/2006	11:08	12' 2"	7' 8"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
11	3/29/2006	12:44	11' 5"	10' 2"	Y (1/2 wet)	Y	30Hz / 50%	13"	10"	0/2/2/0
12	3/29/2006	12:51	9' 1"	9' 4"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
13	3/29/2006	12:56	9' 10"	7' 7"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
14	3/29/2006	1:06	7' 10"	7' 10"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
15	3/29/2006	1:48	6' 2"	8' 7"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
16	3/29/2006	1:54	10' 7"	9' 6"	Y (50% wet)	Y	30Hz / 50%	13"	10"	0/2/2/0
17	3/29/2006	2:18	10' 10"	7' 5"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
18	3/29/2006	2:23	10' 1"	8' 9"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
19	3/29/2006	2:33	5' 6"	10' 3"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
20	3/29/2006	2:38	4' 9"	8' 11"	Y	Y	30Hz / 50%	13"	10"	0/2/2/0
21	3/29/2006	2:46	defected area only		Y	Y	30Hz / 50%	13"	10"	0/2/2/0
22	3/29/2006	2:54	8' 7"	9' 7"	Y	Y (calibration)	30Hz / 50%	13"	10"	0/2/2/0
23	3/30/2006	7:31	11' 10"	8' 7"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
24	3/30/2006	7:43	10' 6"	10' 6"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
25	3/30/2006	7:49	9' 5"	10' 0"	N	Y	30Hz / 50%	13"	10"	0/2/2/0
26	3/30/2006	8:08	7' 5"	11' 7"	N	Y	30Hz / 50%	13"	10"	4/6/4/2
27	3/30/2006	8:13	10' 7"	6' 10"	Y	Y (very long)	30Hz / 50%	13"	10"	4/6/4/2
28	3/30/2006	8:19	8' 4"	9' 8"	Y	Y	30Hz / 50%	13"	10"	4/6/4/2
29	3/30/2006	8:22	11' 01"	9' 0"	Y	Y (60% wet)	30Hz / 50%	13"	10"	4/6/4/2
30	3/30/2006	8:27	10' 6"	8' 2"	Y	Y	30Hz / 50%	13"	10"	4/6/4/2
31	3/30/2006	9:34	9' 7"	?	Y	?	30Hz / 50%	13"	10"	0/2/2/0
32	3/30/2006	9:12	defected area only		Y	?	30Hz / 50%	13"	10"	4/6/4/2

Two different combinations of the signal processing (filtering) parameters were used as indicated in the last column of the table. This was done to test the effect of these parameters on the defect detection reliability. As schematized in Figure 5.52, the gage length of s.p.1, i.e. the distance between sensors #1 and #3, was maintained equal to 13" (≈ 330 mm); the gage length of s.p.2, i.e. the distance between sensors #2 and #4, was maintained equal to 10" (= 254 mm). The lasing repetition rate was set to 30 Hz throughout the tests. Figure 5.54 and Figure 5.55 illustrate the recordings from tests 1 and 2, respectively. The results are shown for the portion of the rail between foot 76 and foot 120 where the internal defects and the surface cuts were located. Each figure displays the D.I.'s recorded as a function of position (foot) by the two sensor pairs and for two different filtering bandwidths, namely (top to bottom) a) H.F.-D.I. s.p.1, b) L.F.-D.I. s.p.1, 3) H.F.-D.I. s.p.2, and c) L.F.-D.I. s.p.2. The bottom plot is included to ease in the visualization of the position of each of the salient discontinuities. In this plot the joints are placed at ordinate 5, the transverse surface cuts at ordinate 4, the internal defects at ordinate 3, and the flaking at ordinate 1. In the four D.I. plots, successfully detected joints (J), surface cuts (SC), or internal defects (ID) are highlighted with circles. Since the two sensor pairs and filtering bandwidths cover complementary regions of the rail head, a defect should be considered successfully detected if any of the four D.I. plots shows an index clearly above the baseline (defect free) value. For example, Figure 5.55 for test 2 shows that all major defects, with the exception of the third internal defect, were detected by at least one of the four D.I. plots. The fact that the baseline D.I. values in the figures are not necessarily one as theoretically expected from the D.I. equation, is simply due to the practical impossibility of obtaining the exactly same sensitivities from the two

sensors in each pair. This, however, does not constitute a problem since only relative changes in D.I. from the baseline value are monitored to detect a potential defect.

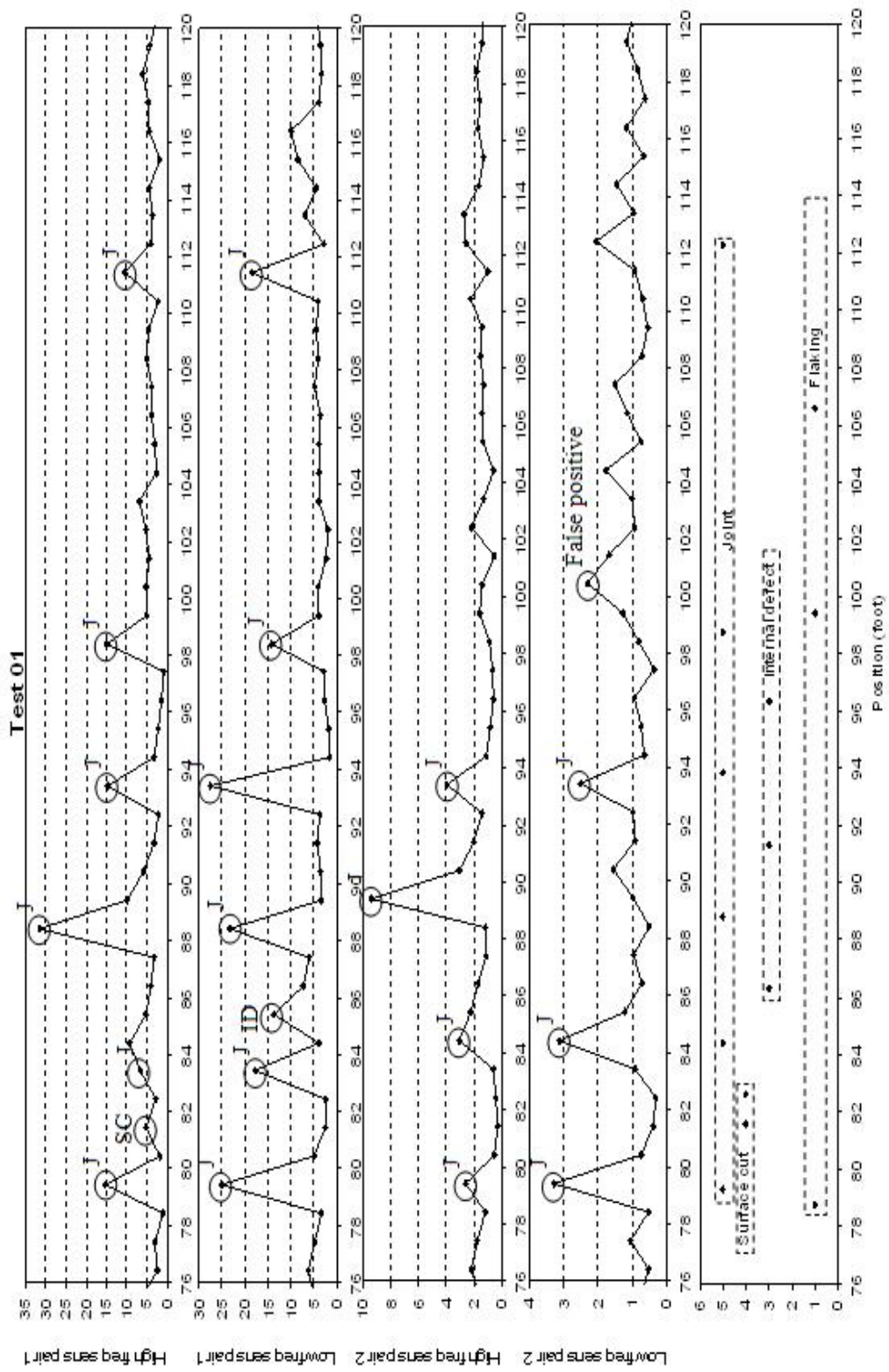


Figure 5.54: Results of Test 01 (2006)

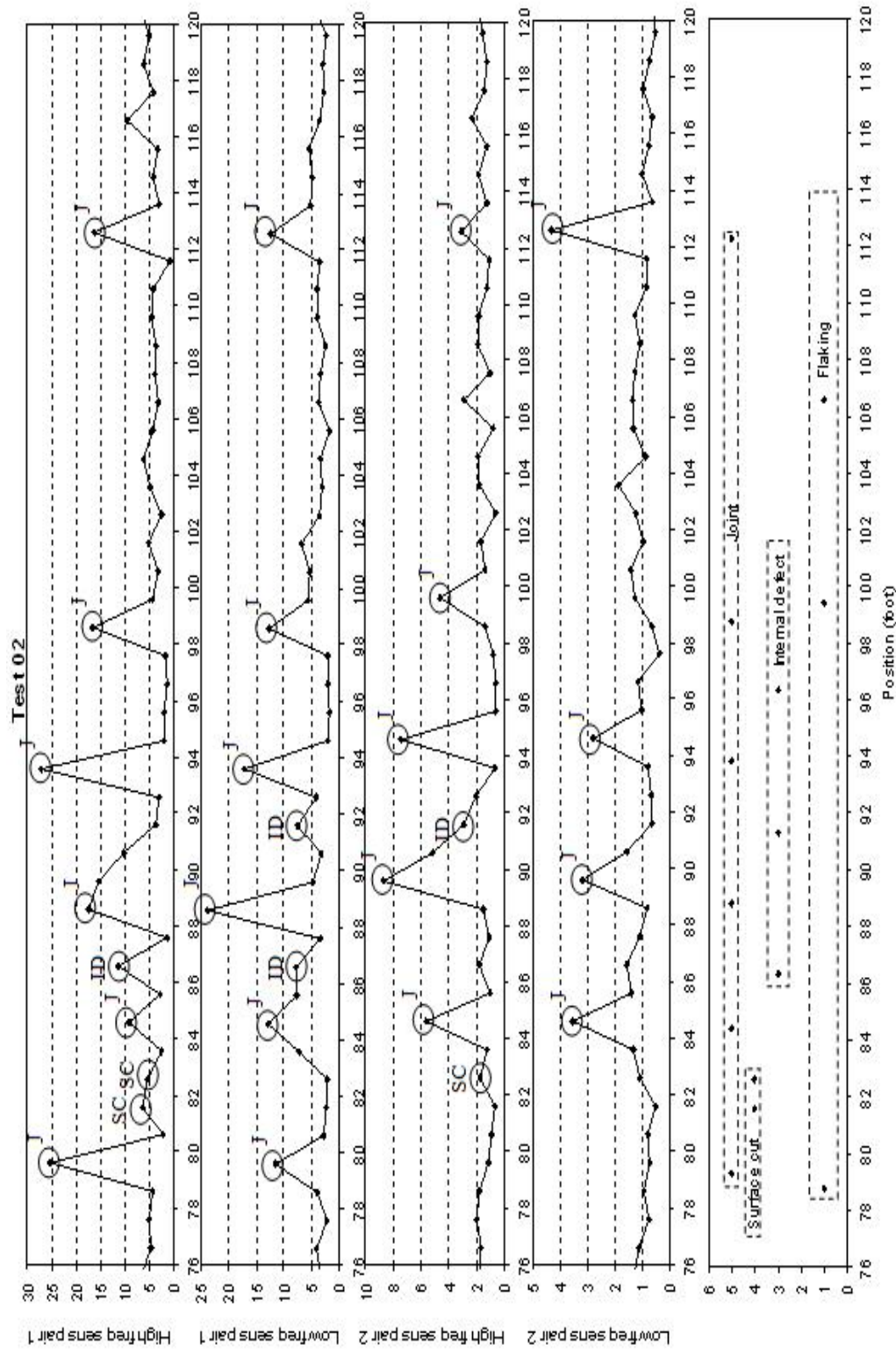


Figure 5.55: Results of Test 02 (2006)

It can be noted in the figures that the same discontinuity can be detected at position j by s.p.1 and at position $j+1$ by s.p.2. This is the case, for example of the joint at location 16 (Table 5.9), measured position 88 9.5/12 ft in Figure 5.55. This is simply due to the longitudinal offset of the two sensor pairs. If a defect (or a joint) is located between sensors #1 and #2 (Figure 5.56a), only s.p.1 will detect it; if, instead, the discontinuity is located between sensors #2 and #3 (Figure 5.56b), both s.p.1 and s.p.2 will detect it. Finally, if the discontinuity is located between sensors #3 and #4, only s.p.2 will detect it.

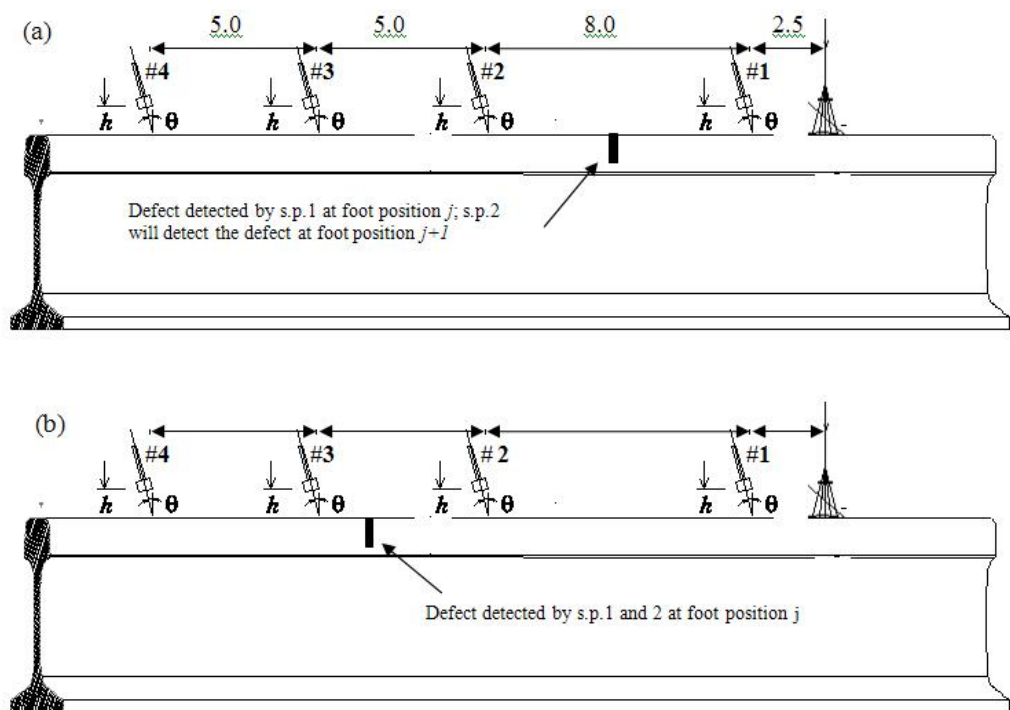


Figure 5.56: Sensor layout in presence of a defect (a) between sensors#1 and #2, and (b) between sensors #2 and #3.

It is not clear why the amplitude of the D.I.'s associated with s.p.1 is consistently larger than that associated with s.p.2. This aspect needs further investigation. The results of the first field tests are summarized in Table 5.11. In the first column, the test number is indicated along with the dry or wet condition of the rail. The second through eighth

columns indicate whether the seven defects were detected (with the 1 symbol) or not (no symbol). It should be noted that the oblique surface cuts in columns six and eight was only created after test 16 and thus a N/A indication is shown for the preceding tests. In the last row, the Probability of Detection (POD) estimated from all relevant tests is indicated for each of the seven defects. As discussed earlier, a defect was considered successfully detected if at least one of the four D.I.'s (H.F.-D.I. s.p.1, L.F.-D.I. s.p.1, H.F.-D.I. s.p.2, or L.F.-D.I. s.p.2.) was activated. In this row, the most sensitive of the four D.I.'s is also shown in *Italic* for each of the defects. The results were as follows: surface cut 1: POD = 61.5% (dry) and 90% (wet); surface cut 2: POD = 61.5 % (dry) and 90 % (wet); internal defect 1: POD = 92.3% (dry) and 100% (wet); internal defect 2: POD = 76.9% (dry) and 100% (wet); oblique surface cut 1: POD = 33.3% (dry) and 50% (wet); internal defect 3: POD = 7.7% (dry) and 10% (wet); oblique surface cut 2: POD = 100% (dry) and 50% (wet). The number of false positive detection accounted for an average of 0.73 false positives per test.

The following conclusions can be drawn from the first field tests:

- 1) Overall the system performed very well in detecting the two surface cuts, two of the three internal defects, and the second oblique surface cut. The detection performance was poor for the third internal defect and the first oblique surface cut.
- 2) The fact that the surface cuts 1 and 2 (5% and 2% H.A., respectively) were successfully detected demonstrates the ability of the system to target defects well below the 10% H.A. limit commonly used by owners to consider removing the rail from service.

- 3) The fact that the internal defects 1 and 2 were successfully detected demonstrates the ability of the system to achieve the 4% H.A. detectability limit for the internal flaws.
- 4) A possible reason for the unsatisfactory performance in detecting oblique cut 1 is the fact that this small discontinuity (3.5% H.A.) was located closer to the joint than allowed by the spatial resolution of the system.
- 5) Wetting the rail surface increased the POD for all defects (with the exception of oblique surface cut 2). The improvement from the water, although significant, was not dramatic. In the end a compromise must be struck between the burden of carrying a water supply and an acceptable level of POD.
- 6) As expected, the surface cuts were predominantly detected by the H.F.-D.I.'s mostly sensitive to surface features. Similarly, internal defect 1 was predominantly detected by the L.F.-D.I.'s mostly sensitive to features deeper into the rail head. The fact that internal defect 2 was primarily detected by the H.F.-D.I.'s suggests that the flaw was very close to breaking the surface. This hypothesis is consistent with the independent findings of the prior hand mapping for this defect.
- 7) It is important to note that all tests in the first field test were conducted at walking (< 5 mph). Test at sustained speeds were conducted in the second field test presented in the next section.

Table 5.11: Summary of Gettysburg first field test results

Test number	Wedry	Surface cut 1 (~5% H.A.)				Surface cut 2 (~2% H.A.)				Internal defect 1 (4%HA)				Internal defect 2 (35%HA)				Oblique surf cut 1 (~3.5% H.A.)				Internal defect 3 (12.5%HA)				Oblique surf cut 2 (~3.5% H.A.)			
		HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2	HF1	LF1	HF2	LF2
1	dry	1				1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
2	dry	1				1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
3	dry	1		1		1		1		1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
4	dry	1				1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
5	dry																	N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
6	dry	1	1			1				1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
7	dry	1		1		1		1		1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
8	dry					1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
9	dry					1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
10	dry	1				1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
11	wet	1				1				1				1				N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
12	wet	1		1		1		1		1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
13	wet	1				1				1				1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
14	wet	1	1			1		1		1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
15	wet	1		1		1		1		1		1		1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
16	wet	1				1				1				1		1		N/A	N/A	N/A	N/A					N/A	N/A	N/A	N/A
17	wet	1		1		1				1		1		1		1		1		1						1		1	
18	wet	1				1				1				1		1		1		1						1			
19	wet	1				1				1				1		1		1		1						1			
20	wet																												
23	dry					1		1		1		1		1		1										1		1	1
24	dry	1	1	1		1		1		1		1		1		1				1						1			
25	dry					1				1				1		1										1			
PROB. OF DETECTION dry conditions wet conditions Most sensitive parameter		Surface cut 1 (~5% H.A.)				Surface cut 2 (~2% H.A.)				Internal defect 1				Internal defect 2				Oblique surf cut 1 (~3.5% H.A.)				Internal defect 3				Oblique surf cut 2 (~3.5% H.A.)			
		61.5% (over 13 tests) 90% (over 10 tests) High freq. sensor pair 1				61.5% (over 13 tests) 90% (over 10 tests) High freq. sensor pair 1				92.3% (over 13 tests) 100% (over 10 tests) Low freq. sensor pair 1				76.9% (over 13 tests) 100% (over 10 tests) High freq. sensor pair 2				33.3% (over 3 tests) 50% (over 4 tests) High freq. sensor pair 1				7.7% (over 13 tests) 10% (over 10 tests) Low freq. sensor pair 2				100% (over 3 tests) 50% (over 4 tests) High freq. sensor pair 1			

Notes:

HF1 = high-frequency parameter of sensor pair 1

LF1 = low-frequency parameter of sensor pair 1

HF2 = high-frequency parameter of sensor pair 2

LF2 = low-frequency parameter of sensor pair 2

1 = defect detected

N/A = not applicable (defect not present during test)

For calculation of Probability of Detection, defect is considered detected if any of the four parameters (HF1, LF1, HF2 or LF2) is activated in a given test

Average false positive/run

0.73

5.5.2 Second field test

This section discusses the second field tests performed in April 2007, again with the technical support of ENSCO Inc. The test site was the same as the first field site (Section 4), the dismissed portion of railroad at crossing #593-381c, near Gettysburg, PA.

In light of the lessons learnt from the first test of March 2006, many modifications were made to the prototype:

- The number of sensors was increased from four to six, to compose three sensing pairs located, respectively along the centerline, the gage side, and the field side of the rail head;
- The bi-directionality of the laser ultrasound source was exploited by positioning the sensors of each pair on either side of the illumination;
- The data acquisition software was changed to allow for sub-foot spatial resolution so as to increase the detectability of defects closely spaced between themselves or close to a joint;
- The signal processing software was rewritten to increase the speed of computation of the high- and low-frequency D.I.'s (H.F.-D.I., L.F.-D.I.); this upgrade allowed testing at the sub-foot spatial resolution without loss of data.
- A new cart was designed by ENSCO with greater stability than that of the first test.

The same total length of 200 ft of track was tested, during three days, April 24th to 26th, 2007. The parameters that were varied during the tests are the following:

- condition of the rail head top surface (dry and wet);
- length of the laser line source of ultrasound (25.4 mm, 35 mm and 38 mm);

- testing speed (5 mph, 10 mph and 15 mph).

Some of the recorded data showed to be not usable due partly to the malfunctioning of the Hy-railer tack system and partly to the presence of an electrical interference on the signal lines.

The timeline of the tests, including the description of the varied parameters, was the following:

April 24th:

Tests performed with 38 mm-long laser line on gage side and dry conditions.

Hy-railer tack.

Total tests performed: 24

D.I. tests: 24

Raw Data tests: 0

Usable D.I. tests: 17

Nonusable tests: 7 due to hy-railer tack malfunctioning

April 25th:

Tests performed with 38 mm-long laser line on gage side and dry conditions.

Hy-railer tack.

Total tests performed: 38

D.I. tests: 38

Raw Data tests: 23

Usable D.I. tests: 11

Nonusable D.I. tests: 7 due to hy-railer tack malfunctioning, 18 due to tack malfunctioning + electrical interference on sensors, 3 due to check runs.

April 26th:

Tests performed with 38 mm-, 35 mm- and 25.4 mm-long laser lines under both dry and wet conditions. UCSD tack (laser diode and reflective tape tags) used to replace hy-railer tack.

Total tests performed: 44

D.I. tests: 16

Raw Data tests: 28

Usable D.I. tests: 16

Nonusable D.I. tests: 0

Of the three laser lines used, the 38 mm-long line on gage side proved the most successful. This was confirmed by lab tests in the months preceding the field test. Comparing dry versus wet conditions, dry performed better. This was due to the fact that the laser lens became dirty with water. The lenses were not protected from water spraying.

Figure 5.57 and Figure 5.58 illustrate the recordings from two representative tests, respectively performed at a speed of 5 and 10 mph. The three transverse defects, two oblique cuts and two surface cuts were detected in the test of Figure 5.57, while in the test of Figure 5.58 all the defects were detected except the 2% surface cut.

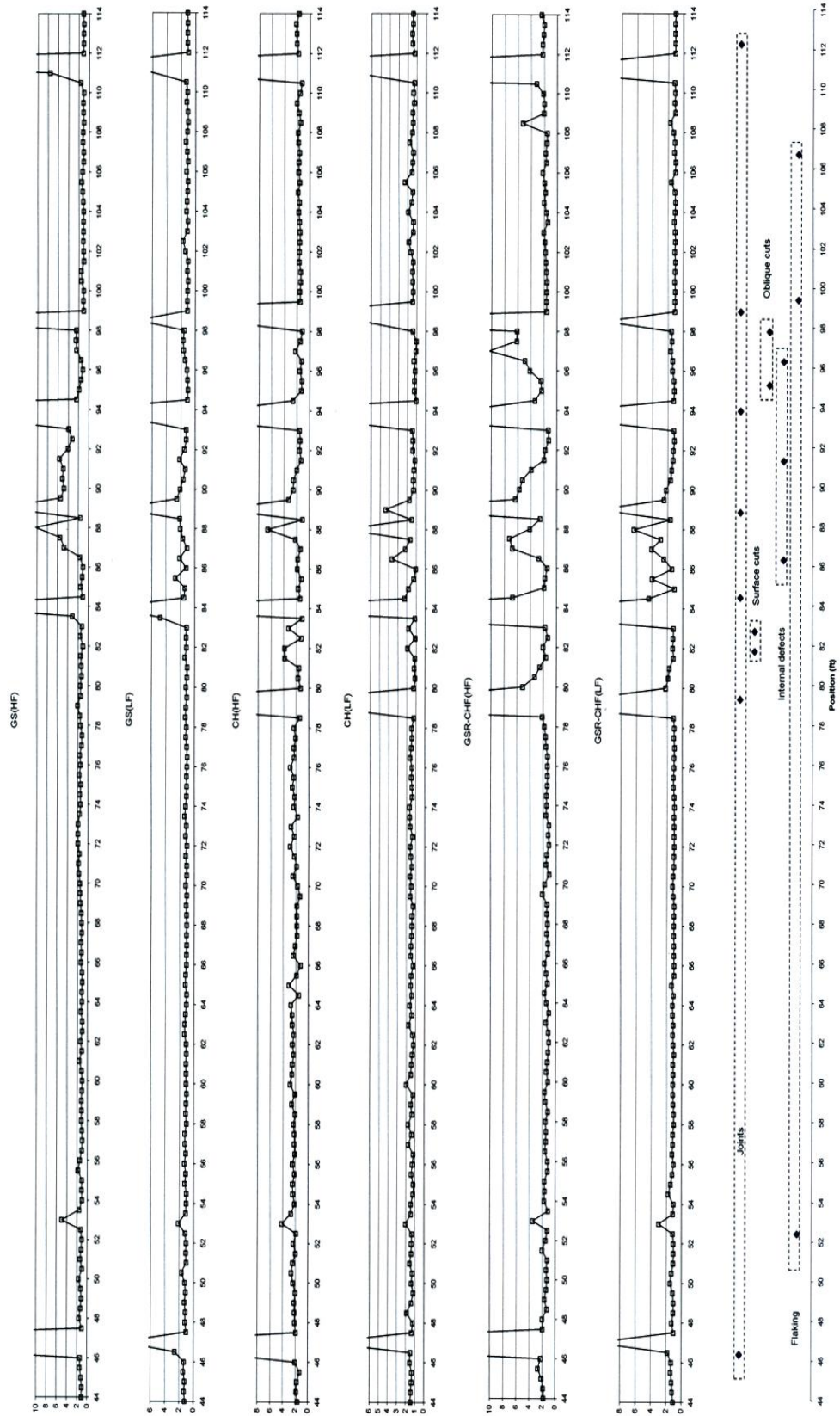


Figure 5.57: Results of a sample test, performed at the speed of 5 mph (2007)

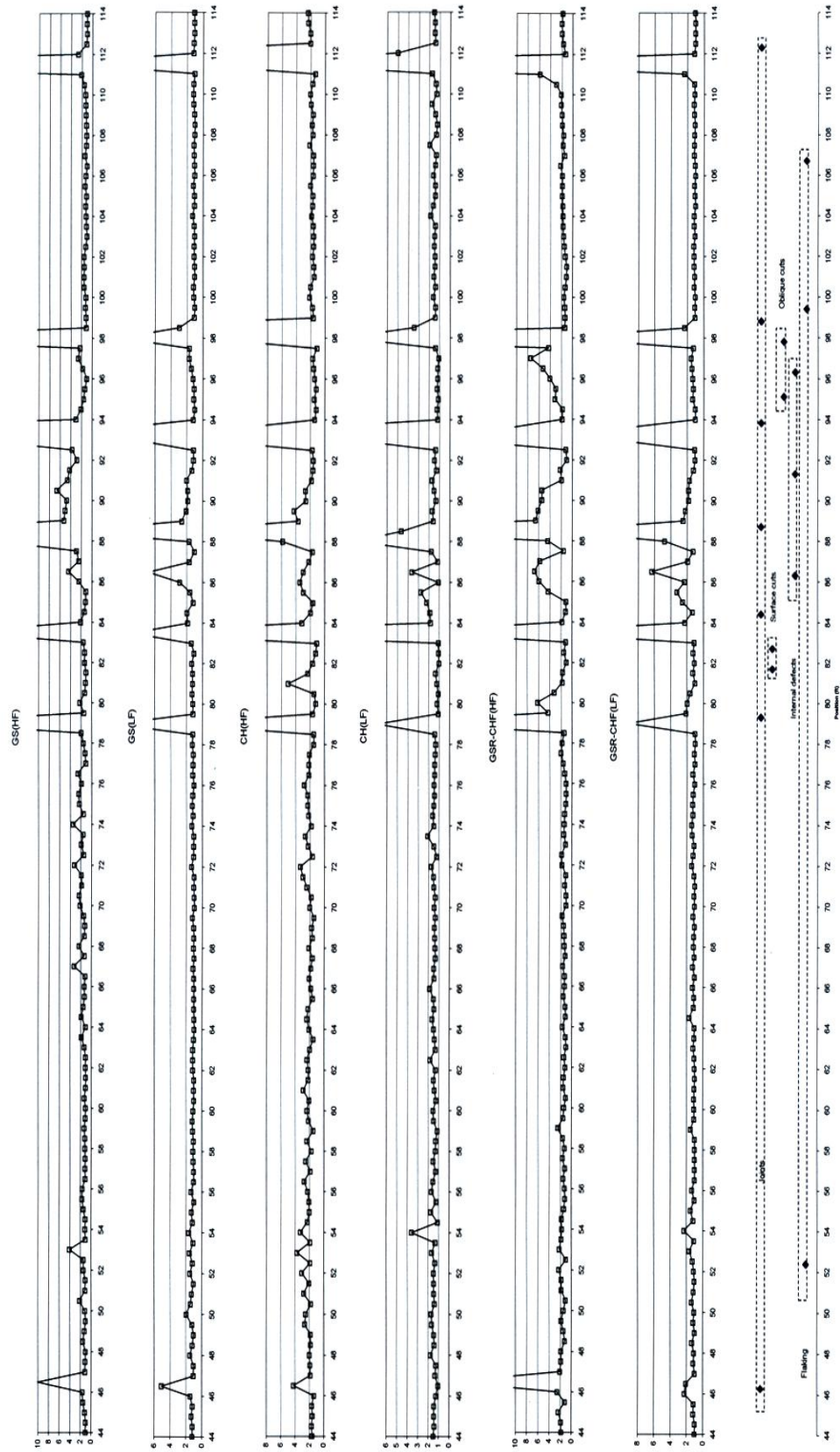


Figure 5.58: Results of a sample test, performed at the speed of 10 mph (2007)

The performance of the prototype was evaluated in terms of (POD), shown in Table 5.12. As done for the analysis of results in the first field test (Table 5.11), the POD was calculated as the ratio between the number of runs where a given defect was detected over the total number of applicable runs. The tests considered in this computation included only runs which (1) were “usable” (i.e. correct tack positioning and limited electrical interference), (2) used the optimum 38 mm-long laser line on gage side, and (3) were conducted under dry conditions. Also, a defect was called detected if one of the five D.I.’s was activated. The five D.I.’s were: Gage Side D.I., Center Head D.I., Field Side D.I., cross-channel Gage Side Rear/Center Head Front D.I., and cross-channel Gage Side Front/Center Head Rear D.I. A D.I. was called activated when the corresponding value was above a threshold level imposed on the measurements after the tests.

Table 5.12: Summary of second field test results

DEFECT	6" long gage chip	Surface cut (5% H.A.)	Surface cut (2% H.A.)	Internal defect (gage side, 3.6% H.A.)	Internal defect (gage side, 35% H.A.)	Oblique cut (3.5% H.A.)	Internal defect (center head, 12% H.A.)	Oblique cut (3.5% H.A.)
POSITION FROM START	56' 2"	81' 7"	82' 7.5"	86' 4"	91' 3.5"	95' 1"	96' 4"	97' 8"
POD (CUMULATIVE) <i>over 37 tests</i>	100%	84%	49%	100%	97%	58%	68%	70%
POD (5 MPH) <i>over 20 tests</i>	100%	90%	65%	100%	95%	63%	70%	90%
POD (10 MPH) <i>over 15 tests</i>	100%	80%	33%	100%	100%	60%	73%	53%
POD (15 MPH) <i>over 2 tests</i>	100%	50%	0%	100%	100%	0%	0%	0%

Testing speed [mph]	5	10	15
Average false positive/run	0.46	1.1	2.1

The threshold level corresponded to a D.I. value about three times the standard deviation of the noise (baseline) signature. The following conclusions could be drawn:

- 1) At 5 mph and 10 mph, the system performed better than the first tests under the dry conditions (Table 5.11). This improvement was particularly pronounced for the internal defects, all of which were detected with good rates (highest POD of 100% for internal defect 1, and lowest POD of 70% for internal defect 3).
- 2) Comparing 5 mph and 10 mph, the POD showed to be approximately independent of testing speed for almost all the defects, except for the Surface Cut 2 (2% H. A.) and the Oblique Surface Cut 2 (~3.5 % H.A.) that were better detected at 5 mph.
- 3) At 15 mph, the gage chip and the first two internal defects remained well detected, but other defects were missed. This was likely due to the fact that blind spots existed at this speed due to the 8-in gage length adopted for the sensors and the inherent 30 Hz repetition rate limit of the pulsed laser.
- 4) Of all defects, the second surface cut (2% H.A.) was the least detected probably due to its very small size and proximity to other defects.
- 5) Of the five sensor pairs (Gage Side pair, Center Head pair, Field Side pair and the two Cross-Channel pairs), the Field Side pair gave the poorest results due to low signal-to-noise ratio of the measurements. This appeared the result of electrical interference at the field sensors which was not experienced in the lab. However, the field pair was also the least relevant since all defects were located either at the center head or on the gage side.

- 6) The number of false positive detection accounted for an average of 0.46 false positives per test at the speed of 5 mph, while the value increased to 1.1 and 2.1 for the testing speeds of 10 and 15 mph respectively.

5.6 Conclusions

This chapter presented the development and testing of a rail defect detection prototype that is based on the laser/air-coupled ultrasonic defect detection concept. The software platform controlling the acquisition was described. The platform of choice was a PXI National Instruments unit running in a LabVIEW[®] environment. Several software modules were programmed to perform all the necessary tasks of the inspection, including: 1) control of the laser generator, 2) acquisition from the air-coupled sensor, 3) processing the data through joint time-frequency and frequency analysis, and 4) calculating the D.I. related to the presence and size of a defect. This unit was mounted in the existing station of the FRA Hy-Railer research car and tested with the help of the FRA and Ensco Inc. The first field test (performed in Gettysburg, PO in March 2006) showed that the system performed well at walking speeds for detecting surface cuts, internal defects and oblique surface cuts. In light of lessons learnt in the first field test, several improvements were made on the prototype. These included upgrading the number of sensors, upgrading the data acquisition software for sub-foot spatial resolution, exploiting the bi-directionality of laser generation and using of digital filters for the signal conditioning. The improved prototype was field tested a second time in April 2007 at the same Gettysburg location. The second field test showed improvements in the detection of internal defects, due to the greater spatial resolution and the full coverage of

the rail head by three sensor pairs. The inspection speed was also increased, showing good performance up to 10 mph. Future work involves the deployment of a higher repetition rate laser to increase the achievable inspection speed while keeping a high spatial resolution.. The user-friendly interface should be further simplified, and a rigid “shoe” should be designed to host all sensors and the laser optical components (lenses) in one compact platform with minimal, or no possibility for adjustment. A third field test is being planned for April 2008.

Chapter 5, in part, has been published in *Transportation Research Record, Journal of the Transportation Research Board*, Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I. and Fateh, M., 2006. “Laser-Air-Coupled Hybrid Detection in Rail Tracks: Status of FRA Prototype Development at UC San Diego”. The dissertation author was co-author of this paper. An additional part in chapter 5 was presented at *World Forum on Smart Materials and Smart Structures Technology*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P., Fateh, M. and Yang, T.-L., 2007. “Non-contact Ultrasonic Rail flaw Detection”. The dissertation author was the main contributor of this paper. A section in chapter 5 was presented at *IWSHM - The 6th International Workshop Structural Health Monitoring*, Coccia, S., Bartoli, I., Lanza di Scalea, F., Rizzo, P. and Fateh, M., 2007. “Non-contact Ultrasonic Rail Defect Detection System. Prototype development and field testing”. The dissertation author was the main contributor of this paper. An additional segment in chapter 5 is in *Technical Report for the University of San Diego, California*, Rizzo, P., Coccia, S., and Lanza di Scalea, F., 2005. “On line High-Speed Rail Defect Detection, Final Report-Phase III”. The dissertation author was co-author of this paper. In part, chapter 5 was in *Technical Report for the Federal*

Railroad Administration, Coccia, S., Bartoli, I., Lanza di Scalea, F. and Rizzo, P., 2007.

“Non-contact Rail Defect Detection: First and Second Field Tests”. The dissertation author was the main contributor of this paper.

6 Conclusions

6.1 Review of the dissertation and summary of the results

The work within this dissertation investigates the use of guided waves for structural health monitoring defect detection of railroad structures. A semi analytical finite element (SAFE) approach was shown to be successful for modeling the ultrasonic wave propagation in rails at high frequencies (<500 kHz). The unforced and forced solutions of the rail response were obtained with the SAFE method and some were validated by experimental results. An algorithm which uses the Arnoldi Iterative solution process was presented and its efficiency was validated by computational time test results.

Analytical and experimental results showed how the load pattern influences the rail response. A better understanding of the cross-sectional energy distribution of the modes propagating in the rail was achieved, with the purpose of designing an inspection system targeted to the defects of interest. It was demonstrated that a symmetric excitation pattern on the top of the rail develops strain energy in a symmetric fashion along the cross-section, with maxima close to the top of the head; a nonsymmetric pattern, instead, induces energy at the sides of the head. For a symmetric load case, high frequency waves confirmed to generate energy close to the top of the head, while lower frequency waves induce energy in deeper areas of the head. Neglecting material damping caused the occurrence of resonance-like phenomena, which were corrected by filtering out modes at cut-off frequencies, after applying a mode tracking algorithm, which employs mode orthogonality.

Different techniques to process the ultrasonic signals were also presented. The need of retaining the time and the frequency information of the acquired signals led to the use of the wavelet transform, in its continuous and discrete expressions. The DWT showed to be efficient for real time processing, including data de-noising and data compression. Different combinations of DWT-based features proved to be successful for the detection of defects present in rails available in the UCSD NDE & SHM laboratory. The experimental results showed that the appropriate selection of the mother-wavelet and of the wavelet coefficient threshold enhances the inspection sensitivity. The performance of a D.I. based on the normalized RMS of the signal amplitudes was also tested, and it showed excellent results in the detection of surface and internal defects. The implementation of digital filtering in the feature-extraction procedure dramatically simplified the process of setting the parameters needed for the prototype calibration and reduced computational times. A D.I. based on the normalized temporal coherence proved to be very efficient in defect detection, if the entire frequency spectrum of the signals. The impossibility of calculating the D.I. for two sub-spectral ranges (high and low frequency) prevented the ability of classification of the defects, as surface or internal.

A prototype for rail inspection based on the laser/air-coupled ultrasonic technique was designed, packaged and field-tested twice with the help of the Federal Railroad Administration. The unit was mounted in the existing station of the FRA Hy-Railer research car and tested with the help of the FRA and Ensco Inc. The first field test (performed in Gettysburg, PO in March 2006) showed that the system performed well at walking speeds for detecting surface cuts, internal defects and oblique surface cuts. After several upgrades to the hardware design and to the software implementation of the

prototype, a new version of prototype was field tested in April 2007 at the same Gettysburg location. The inspection speed was increased, and improvements were shown in the detection of internal defects, due to the greater spatial resolution and the full coverage of the rail head by three sensor pairs. The prototype is currently being further improved in preparation for a third, and final, field test which is scheduled for April 2008.

6.2 Recommendations for future work

A natural prosecution of the research work included in this dissertation would be the implementation of a Global-Local approach [126] to better investigate the interactions between guide waves and defects encountered along their propagation path.

The deployment of a higher repetition rate laser is critical to the increase of achievable inspection speed while keeping a high spatial resolution. A laser pulsing at 100 Hz is being considered for the final version of the prototype.

The combination of a laser non-contact excitation source with a water-filled wheel probe specially “tuned” to listen the generated ultrasonic guided may have a promising future within the field of rail inspection, because it would eliminate the variability inherent in air-coupled non-contact ultrasonic detection.

The use of a laser vibrometer can be investigated as an alternative solution for detecting ultrasonic guided waves on the rail.

Finally, the prototype user-friendly interface could be further simplified to make the system more manageable for users which do not have an in depth knowledge within the field of ultrasonic wave propagation.

7 References

- [1] *Federal Railroad Administration, Safety Statistics Data: 1992-2002*, 2002. US Department of Transportation,
- [2] <http://www.fra.dot.gov/us/content/1242>, last updated: 22 Oct2003.
- [3] R. Clark, (2004). "Rail flaw detection: overview and needs for future developments," *NDT&E International*, 37, pp 111-118.
- [4] F. Lanza di Scalea, (2007). *Ultrasonic testing in the railroad industry, in non-destructive testing handbook*, 3rd edition, American Society for Nondestructive Testing, Columbus, OH, pp 9-16.
- [5] L. Oukhelou, P. Aknin and J.-P. Perrin, (1999). "Dedicated sensor and classifier of rail head defects," *Control Eng. Practice*, 7, pp 57-61.
- [6] R. Pohl, A. Erhard, H-J. Montag, H-M. Thomas, H. Wüstenberg, "NDT techniques for railroad wheel and gauge corner inspection", *NDT&E International*, 37, pp 89-94.
- [7] J. L. Rose, M. J. Avioli, P. Mudge, R. Sanderson, (2002). Guided wave inspection potential of defects in rail," *5th Int. Railway Eng.*, Commonwealth Institute, London, UK.
- [8] J. L. Rose, M. J. Avioli, W.J. Song, (2002). "Application and potential of guided wave rail inspection," *Insight*, 44, pp. 353-358.
- [9] P. Wilcox, M. Evans, B. Pavlakovic, D. Alleyne, K. Vine, P. Cawley, M.J.S. Lowe, (2003). "Guided wave testing of rail," *Insight*, 45, pp. 413-420.
- [10] I. Bartoli, F. Lanza di Scalea, M. Fateh, E. Viola, (2005). "Modeling guided wave propagation with application to the long-range defect detection in railroad tracks", *NDT&E International*, 38, pp. 325-334.
- [11] D. Hesse, P. Cawley, (2006). "Surface wave modes in rails," *The Journal of Acoustical Society of America*, 120, pp. 733-740.
- [12] F. Lanza di Scalea, P. Rizzo, S. Coccia, I. Bartoli, M. Fateh, E. Viola, G. Pascale, (2005). "Signal processing for automatic defect detection and classification in rails by non-contact guided wave ultrasonics," *Insight*, NDT and Condition Monitoring, Special EURO Issue on NDT in the Rail Industry, British Society for Nondestructive Testing, 47, pp. 346-353.

- [13] F. Lanza di Scalea, P. Rizzo, S. Coccia, I. Bartoli, M. Fateh, (2006). "Laser/air-coupled hybrid non-contact system for defect detection in rail tracks: status of FRA prototype development at UC San Diego,; Transportation Research Record," *Journal of the Transportation Research Board*, 1943, pp. 57-64.
- [14] F. Lanza di Scalea, P. Rizzo, S. Coccia, I. Bartoli, M. Fateh, (2006). "High-speed defect detection in rails by non-contact guided ultrasonic testing, transportation research record", *Journal of the Transportation Research Board*, 1916, pp. 66-77.
- [15] D.F. Cannon, K.-O. Edel, S.L. Grassie, K. Sawley, (2003). "Rail defects: an overview," *Fatigue and Fracture of Engineering Materials and Structures*, 26, pp. 865–887.
- [16] *Rail Defect Manual*, (1999). Sperry Rail Service, Danbury, CT.
- [17] V.V. Murav_ev, E.V. Boyarkin, (2003). "Nondestructive testing of the structural-mechanical state of currently produced rails on thebasis of the ultrasonic wave velocity," *Russian Journal of Nondestructive Testing*, 39, pp. 24–33.
- [18] K. Abbaszadeh, M. Rahimian, H.A. Toliyat, L.E. Olson, (2003). "Rails defect diagnosis using wavelet packet decomposition," *IEEE Transactions on Industry Applications*, 39, (5).
- [19] B.D. Jeffrey, M.L. Peterson, (1999). "Assessment of rail flaw inspection data. Mountain Plains Consortium," *MPC-99-106*.
<http://www.ndsu.nodak.edu/ndsu/ugpti/MPC_Pubs/pdf/MPC99-106.pdf>.
- [20] R. Care, S. Clark, M. Dembosky, A. Doherty, (2006). "Why rails crack: Gauge corner cracking on the British network – Analysis," *Arup Journal*, 41, pp. 16-19.
- [21] D.F. Cannon, H. Pradier, (1996). "Rail rolling contact fatigue research by the European Rail Research Institute," *Wear*, 191, pp. 1–13.
- [22] S. Grassie, P. Nilsson, K. Bjurstrom, A. Frick, L.G. Hansson, (2002). "Alleviation of rolling contact fatigue on Sweden_s heavy haul railway," *Wear*, 253, pp. 42–53.
- [23] D.Y. Jeong, (2001). "Progress in rail integrity research," *Federal Railroad Administration, Department of Transportation, Report*, DOT/FRA/ORD-01/18.
- [24] E.G. Vadillo, J.A. Tarrago, G.G. Zubiaurre, C.A Duque, (1998). "Effect of sleeper distance on rail corrugation," *Wear*, 217, pp. 140–146.
- [25] D.T. Eadie, J. Kalousek, K.C. Chiddick, (2000). "The role of high positive friction (HPF) modifier in the control of short pitchcorrugation and related

phenomena,” *In: Contact Mechanics 2000 Conference, Tokyo*, http://www.kelsan.com/007_publications.html.

- [26] S. Mueller, (2000). “A linear wheel-rail model to investigate stability and corrugation on straight track,” *Wear*, 243, pp. 122–132.
- [27] A. Bohmer, T. Klimpel, (2002). “Plastic deformation of corrugated rails—a numerical approach using material data of rail steel,” *Wear*, 253, pp. 150–161.
- [28] F. Lanza di Scalea, (2007). *Ultrasonic Testing in the Railroad Industry*, Part 2 of Chapter 15 in *Nondestructive Testing Handbook*, 3rd Edition, Vol. 7 Ultrasonic Testing, American Society for Nondestructive Testing.
- [29] J.J. Marais, K.C. Mistry, (2003). “Rail integrity management by means of ultrasonic testing,” *Fatigue and Fracture of Engineering Materials and Structures*, 26, pp. 931–938.
- [30] K. Bruzelius, D. Mba, (2004). “An initial investigation on the potential applicability of Acoustic Emission to rail track fault detection,” *NDT&E International*, 37, pp. 507-516.
- [31] Z. Zhou , Z. Feng1 , Y. Gao, J. Bai, (2006). “An outline of applications of ultrasonic guided waves in non-destructive testing of large structures,” *Asia-Pacific Conference on NDT*, 5th – 10th, Auckland, New Zealand.
- [32] P. Cawley, D. Alleyne, (1996). “The use of Lamb waves for the long range inspection of large structures,” *Ultrasonics*, 34, pp. 287-290.
- [33] J. L. Rose, (2002). “A baseline and vision of ultrasonic guided wave inspection potential,” *Journal of Pressure Vessel Technology*, 124, pp. 273-282
- [34] W. Wilkie, R. Bryant, J. High, R. Fox, R. Hellbaum, A. Jalink, B. Little, P. Mirick, (2000). “Low-cost piezocomposite actuator for structural control applications,” *SPIE 7th Annual Int. Symp. on Smart Structures and Materials*, Newport Beach, CA. Bellingham, WA: SPIE Optical Engineering Press.
- [35] A.A. Bent, (1997). “Active fiber composites for structural actuation,” *PhD Dissertation Department of Aeronautics and Astronautics*, Massachusetts Institute of Technology.
- [36] F. Lanza di Scalea, T.P. Berndt, J.B. Spicer, B.B. Djordjevic, (1999). “Remote laser generation of narrow-band surface waves through optical fibers,” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 46, pp. 1551-1557.

- [37] J. L. Rose, S. P. Pelts, and M. J. Quarry, (1997). "A comb transducer for mode control in guided wave NDE," *IEEE Ultrasonics Symposium*, 2, pp. 1033–1036
- [38] P. D. Wilcox, (2003). "Omni-directional guided wave transducer arrays for the rapid inspection of large areas of plate structures," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 50, pp. 699-709.
- [39] P. Rizzo, F. Lanza di Scalea, (2004). "Monitoring in Cable Stays via Guided Wave Magnetostrictive Ultrasonics", *Materials Evaluation*, Vol. 62, No. 10, pp. 1057-1065.
- [40] H. M. Matt, I. Bartoli, F. Lanza di Scalea. (2005). "Ultrasonic guided wave monitoring of composite wing skin-to-spar bonded joints in aerospace structures," *Journal of the Acoustical Society of America*, 118, pp. 2240-2252.
- [41] T. Ogisu, M. Shimanuki, H. Yoneda, Y. Okabe, N. Takeda, T. Sakurai, (2006). "Damage growth detection of aircraft bonding structure under cyclic loading using FBG/PZT hybrid sensor system," *Proceedings of 3rd European Workshop on Structural Health Monitoring*, Granada, Spain.
- [42] P. F. Lichtenwalner, D. Sofge, (1998). "A local area damage detection in composite structures using piezoelectric transducers," *Proceedings of SPIE*, 3326, pp. 509-515.
- [43] C.A. Paget, S. Grondel, K. Levin, C. Delebarre, (2003). "Damage assessment in composites by Lamb waves and wavelet coefficients," *Smart Materials and Structures*, 12, pp. 393-402.
- [44] K-H Ip, Y-W Mai, (2004). "Delamination detection in smart composite beams using Lamb waves," *Smart Materials and Structures*, 13, pp. 544-551.
- [45] H. Sohn, G. Park, J. Wait, N. Limback, C.R. Farrar, (2003). "Wavelet-based active sensing for delamination detection in composite structures," *Smart Materials and Structures*, 13, pp. 153-160.
- [46] F. Lanza di Scalea, H.M. Matt, I. Bartoli, S. Coccia, G. Park, C.R. Farrar, (2007). "Health monitoring of UAV wing skin-to-spar joints using guided waves and macro fiber composite transducers," *Journal of Intelligent Material Systems and Structures*, 18, pp. 373-388.
- [47] S.D. Valdes, C. Soutis, (2002). "Real-time nondestructive evaluation of fiber composite laminates using low-frequency Lamb waves," *Journal of the Acoustical Society of America*, 111, pp. 226-2033.

- [48] S.S. Kessler, S.M. Spearing, C. Soutis, (2002). "Damage detection in composite materials using Lamb wave methods," *Smart Materials and Structures*, 11, pp. 269-278.
- [49] C.S. Wang, F-K. Chang, (2000). "Diagnosis of impact damage in composite structures with built-in piezoelectrics network," *Proceedings of SPIE*, 3990, pp. 13-19.
- [50] A. K. Mal, S. Banerjee, F. Ricci, (2004). "Automated structural health monitoring system using acoustic emission and modal data," *Proceedings of SPIE*, 5394, pp. 1-10.
- [51] F. Lanza di Scalea, J. McNamara, (2003). "Ultrasonic NDE of railroad tracks: air-coupled cross-sectional inspection and long-range inspection," *Insight*, 45, pp. 394-401.
- [52] P. Cawley, M.J.S. Lowe, D.N. Alleyne, B. Pavlakovic, P. Wilcox, (2002). "Practical long range guided wave testing: applications for pipes and rail," *Material Evaluation—American Society for Nondestructive Testing*, 61, pp. 66-74.
- [53] C.B. Scruby, L.E. Drain, (1990). *Laser Ultrasonics: Techniques and Applications*, Adam Hilger, New York, pp. 356.
- [54] DJ. Thompson, (1993). "Wheel-rail noise generation, Part III: rail vibration," *Journal of Sound and Vibration*, 161, pp. 421-46.
- [55] DJ. Thompson, (1997). "Experimental analysis of wave propagation in railway tracks," *Journal of Sound and Vibration*, 203, pp. 867-888.
- [56] TX. Wu, DJ. Thompson, (1999). "A double Timoshenko beam model for vertical vibration analysis of railway track at high frequencies," *Journal of Sound and Vibration*, 224, pp. 329-348.
- [57] J. McNamara, (2003). "Health monitoring of railroad tracks by elastic-wave based non-destructive testing," *PhD Dissertation*. San Diego: University of California.
- [58] L. Gavric', (1994). "Finite element computation of dispersion properties of thin walled waveguides," *Journal of Sound and Vibration*, 173, pp. 113-24.
- [59] L. Gavric', (1995). "Computation of propagative waves in free rail using a finite element technique," *Journal of Sound and Vibration*, 185, pp. 531-543.

- [60] N. J. Nigro, (1966) "Steady-state wave propagation in infinite bars of noncircular cross-section," *Journal of the Acoustical Society of America*, 40, pp. 1501–1508.
- [61] W. B. Fraser, (1969). "Stress wave propagation in rectangular bars," *International Journal of Solids and Structures*, 5, pp. 379–397.
- [62] W. B. Fraser, (1969). "Dispersion of elastic waves in elliptical bars," *Journal of Sound and Vibration*, 10, pp. 247–260.
- [63] K. Nagaya, (1982). "Stress wave propagation in a bar of arbitrary cross-section," *Journal of Applied Mechanics*. 49, pp. 157–163.
- [64] P. Wilcox, M. Evans, O. Diligent, M. Lowe, P. Cawley, (2002). *Dispersion and excitability of guided acoustic waves in isotropic beams with arbitrary cross-section*, In: Thompson DO, Chimenti DE, editors. Review of progress in QNDE, 21, pp. 203–210.
- [65] M.J.S. Lowe, (1995). "Matrix techniques for modeling ultrasonic waves in multilayered media," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42 pp. 525–542.
- [66] I. Bartoli, A. Marzani, F. Lanza di Scalea, E. Viola, (2006), "Modeling wave propagation in damped waveguides of arbitrary cross-section," *Journal of Sound and Vibration*, 295, pp. 685–707.
- [67] A. Trivedi, P. Krysl, J. McNamara, F. Lanza di Scalea, (2004). "Adaptive simulation of ultrasonic guided waves in railroad tracks using CHARMS," *CD-ROM proceedings of the sixth world congress of computational mechanics*, Beijing, China.
- [68] B.A. Auld, (1990). *Acoustic fields and waves in solids* (two volumes), Krieger Publishing Company, Malabar, Florida.
- [69] J.D. Achenbach, (1973). *Wave propagation in elastic solids*, North Holland Publishing Co./American Elsevier, Amsterdam/New York.
- [70] L. Pochhammer, (1876). "Ueber die Fortpflanzungsgeschwindigkeiten kleiner Schwingungen in einem unbegrenzten isotropen Kreiscylinder," *J. Reine Angew. Math*, 81, pp. 324–336.
- [71] K. Sato, (1978). "Elastic wave propagation in an infinite bar of elliptical cross-section," *Bull. JSME*, 21, pp. 203–209.

- [72] R. D. Mindlin and E. A. Fox, (1966). "Vibrations and waves in elastic bars of rectangular cross-section," *J. Appl. Mech.*, 27, pp. 152–158.
- [73] P.E. Lagasse, (1973). "Higher-order finite element analysis of topographic guides supporting elastic surface waves," *Journal of the Acoustical Society of America*, 53, pp. 1116-1122.
- [74] B. Aalami, (1973). "Waves in prismatic guides of arbitrary cross-section," *Journal of Applied Mechanics*, 40, pp. 1067-1072.
- [75] K.H. Huang and S.B. Dong, (1984). "Propagating waves edge vibrations in the anisotropic composite cylinders," *Journal of Sound and Vibration*, 96, pp. 363-379.
- [76] L. Gry, (1996). "Dynamic modeling of railway track based on wave propagation," *Journal of Sound and Vibration*, 195, pp. 477–505.
- [77] T. Hayashi, W.J. Song, J.L. Rose (2003). "Guided wave dispersion curves for a bar with an arbitrary cross-section, a rod and a rail example," *Ultrasonics*, 41, pp.175-183.
- [78] T. Hayashi, C. Tamayama, and M. Morimasa, (2006). "Wave structure analysis of guided waves in a bar with an arbitrary cross-section," *Ultrasonics*, 44, pp. 17–24.
- [79] V. Damljanović and R. L. Weaver, (2004). "Propagating and evanescent elastic waves in cylindrical waveguides of arbitrary cross-section," *Journal of Sound and Vibration*, 115, pp. 1572–1581.
- [80] S. Finnveden, (2004). "Evaluation of modal density and group velocity by a finite element method," *Journal of Sound and Vibration*, 273, pp. 51–75.
- [81] G. Neau, (2003). "Lamb waves in anisotropic viscoelastic plates. Study of the wave fronts and attenuation," Ph.D Thesis, L'Université Bordeaux I.
- [82] A. Bernard, M.J.S. Lowe, M. Deschamps, (2001). "Guided waves energy velocity in absorbing and non-absorbing plates," *Journal of the Acoustical Society of America*, 110, pp. 186-196.
- [83] A. Bernard, M. Deschamps, M.J.S. Lowe, (1999). "Energy velocity and group velocity for guided waves propagating within an absorbing or non-absorbing plate in vacuum," *Review of Progress in Quantitative NDE*, 18, pp. 183-190.
- [84] A. Jennings, J. J. McKeown, (1992). *Matrix computation 2nd Edition*, John Wiley & Sons Ltd, West Sussex, England.

- [85] R. Barret, M. Berry, T.F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, H. Van der Vorst (1994). "Templates for the solution of linear systems: building blocks for interactive methods," *SAIM*, Philadelphia.
- [86] R.B Lehoucq, D.C. Sorensen, C. Yang, (1998). "ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods," SIAM, Philadelphia.
- [87] P. Loveday and C. Long, (2007). "Time domain simulation of piezoelectric excitation of guided waves in rails using waveguide finite elements," *SPIE*, 6529, pp. 1-10.
- [88] M. Castaings and B. Hosten, (2003). "Guided waves propagating in sandwich structures made of anisotropic, viscoelastic, composite materials," *Journal of the Acoustical Society of America*, 113, pp. 2622-2634.
- [89] K. Toda and K. Motegi, (2000). "Coupling of velocity dispersion curves of leaky Lamb waves on a fluid-loaded plate," *The Journal of the Acoustical Society of America*, 107, pp 1045-1048.
- [90] T. Mihara, Y. Otsuka, H. Cho, K. Yamanaka, (2006). "Time of-flight diffraction measurement using laser ultrasound," *Society for Experimental Mechanics*, 46, pp. 561-567.
- [91] M.S. Obaidat, (1993). "Phonocardiogram signal analysis: techniques and performance comparison," *Journal of medical Engineering & Technology*, 17, pp. 221-227.
- [92] S. Nawab, T. Quatieri, J.S. Lim, (1983), "Signal reconstruction from short-time Fourier transform magnitude," *Acoustics, Speech, and Signal Processing*, 31, pp. 986- 998.
- [93] M. M. Reda Taha, A. Noureldin, J.L. Lucero, T.J. Baca, (2006). "Wavelet transform for structural health monitoring: a compendium of uses and features," *Structural Health Monitoring*, 5, pp. 267-329.
- [94] K. Kishimoto, H. Inoue, M. Hamada, T. Shibuya, (1995). "Time frequency analysis of dispersive waves by means of wavelet transform," *American Society of Mechanical Engineers Journal of Applied Mechanics*, " 62, pp. 841-846.
- [95] Y.Y. Kim, E. H. Kim, (2001). "Effectiveness of the continuous wavelet transform in the analysis of some dispersive elastic waves," *The Journal of the Acoustical Society of America*, 110, pp. 332-342.

- [96] H. Jeong, Y.S. Jang, (2000). "Wavelet analysis of plate wave propagation in composite laminates," *Composite Structures*, 49, pp. 443-450.
- [97] W.J. Staszewski, (1998). "Structural and mechanical damage detection using wavelets," *The Shock and Vibration Digest*, 30, pp. 457-472.
- [98] S. Mallat, (1999). *A Wavelet Tour of Signal Processing*, 2nd Edition, Academic Press, New York.
- [99] J. C. Hong, Y.Y. Kim, (2004). "Determination of the optimal gabor wavelet shape for the best time-frequency localization using the entropy concept," *Experimental Mechanics*, 44, pp. 387-395.
- [100] H. Inoue, K. Kishimoto, T. Shibuya, (1996). "Experimental wavelet analysis of flexural waves in beams," *Experimental Mechanics*, 36, pp. 212-217.
- [101] L. Gaul, S. Hurlebaus, L.J. Jacobs, (2001). "Localization of a "synthetic" acoustic emission source on the surface of a fatigue specimen," *Research in Nondestructive Evaluation*, 13, pp. 107-117.
- [102] G. Kaiser, (1994). *A friendly guide to wavelets*, Birkhauser, Boston.
- [103] I. Daubechies, (1992). "Ten lectures on wavelets," 2nd ed., SIAM, Philadelphia.
- [104] Y. Sheng, (1996). "Wavelet transform," in: *The transforms and applications handbook*, Ed. by A. D. Poularikas, CRC Press, Florida, pp. 747-827.
- [105] C.S. Burrus, R. A. Gopinath, H. Guo, (1998). "Introduction to wavelets and wavelet transforms a primer," Prentice Hall, New Jersey.
- [106] P. Rizzo, F. Lanza di Scalea, (2006). "Feature extraction for defect detection in strands by guided ultrasonic waves," *Structural Health Monitoring*, 5, pp. 297-312.
- [107] P. Rizzo, F. Lanza di Scalea, (2004). "Discrete wavelet transform to improve guided-wave based health monitoring to tendons and cables," *SPIE*, 5391, pp. 523-532.
- [108] P. Rizzo, F. Lanza di Scalea, (2005). "Ultrasonic inspection of multi-wire steel strands with the aid of the wavelet transform," *Structural Health Monitoring*, 14, pp. 685-695.
- [109] S.T. Quek, Q. Wang, L. Zhang, K. H. Ong, (2001). "Practical issues in the detection of damage in beams using wavelets," *Smart Materials and Structures*, 10, pp. 1009-1017.

- [110] D. Hesse, P. Cawley, (2007), "Detection of critical defects in rails using ultrasonic surface waves," *Review Quantitative Nondestructive Evaluation*, 26, pp. 1413-1420.
- [111] D. Hesse, P. Cawley, (2006), "Excitation of surface wave modes in rails and their application for defect detection," *Review Quantitative Nondestructive Evaluation*, 25, pp. 1593-1600.
- [112] T. Case, R. Waag, (1996), "Flaw identification from time and frequency features of ultrasonic waveforms," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 43, pp. 592-600.
- [113] F. Lanza di Scalea, I. Bartoli, P. Rizzo, M. Fateh, (2005). "High-speed defect detection in rails by non-contact guided ultrasonic testing," *Transportation Research Record, Journal of the Transportation Research Board*, 1916, pp. 66-77.
- [114] P. Rizzo, S. Coccia, F. Lanza di Scalea, I. Bartoli, M. Fateh, (2006). "Unsupervised learning algorithm for high-speed defect detection in rails by laser/air-coupled non-contact ultrasonic testing," *SPIE*, 6174, pp. 777-788.
- [115] J. Michaels, T. Michaels, (2005). "Detection of structural damage from the local temporal coherence of diffuse ultrasonic signals," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 52, pp. 1769-1782.
- [116] G. R. Cooper, C.D. McGillen, (1971). *Probabilistic methods of signal and system analysis*, Holt, Rinehart, Winston, New York.
- [117] P. Rizzo, S. Coccia, I. Bartoli, and F. Lanza di Scalea, (2007). "Non-contact rail monitoring by ultrasonic guided waves," *Encyclopedia of Structural Health Monitoring*, submitted.
- [118] P. Rizzo, S. Coccia, I. Bartoli, and F. Lanza di Scalea, (2006). *On line high-speed rail defect detection - prototype assembling and field testing*, Technical Report to FRA and UCSD, Report No. SSRP-06/17.
- [119] Lanza di Scalea, F. (2003), *On-line high-speed rail defect detection*, Technical Report to FRA and UCSD, Report SSRP-2003/07.
- [120] F. Lanza di Scalea, I. Bartoli, P. Rizzo, and J. McNamara, (2004). *On-line high-speed rail defect detection*, Final Report – Part 1, Technical Report to FRA and UCSD, Report DOT/FRA/ORD-04/16.
- [121] P. Rizzo, S. Coccia, F. Lanza di Scalea, (2005). *On-line High-speed Rail Defect Detection*, Technical Report to FRA and UCSD, Report SSRP-2005/04.

- [122] S. Coccia, I. Bartoli, F. Lanza di Scalea, P. Rizzo, (2007). *Non-contact rail defect detection: first and second field tests*, Technical Report to FRA and UCSD, Report SSRP-2007/15.
- [123] J. McNamara, F. di Scalea, (2004). "Improvements in non-contact ultrasonic testing of rails by the discrete wavelet transform," *Materials Evaluation*, 62, pp. 365-372.
- [124] P. Rizzo, and F. Lanza di Scalea, (2005). "Discrete wavelet transform for enhancing defect detection in strands by guided ultrasonic waves," *International Journal of Structural Health Monitoring*, 5, pp. 297-308.
- [125] W. Staszewski, (1998). "Wavelet based compression and feature selection for vibration analysis," *Journal of Sound and Vibration*, 211, pp. 735-760.
- [126] A. Srivastava, I. Bartoli, F. Lanza di Scalea and K. Sabra, (2007). "Global-local ultrasonic method applied to the quantitative detection of defects in aircraft panels" *IWSHM – The 6th International Workshop on Structural Health Monitoring*, Stanford, CA, September 11-13.