Czech Technical University in Prague

Université du Maine

Doctoral Thesis
under Joint Supervision

December 2008

Antonín Novák
Identification of Nonlinear Systems in Acoustics

by

Antonín Novák

A dissertation under joint supervision submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in ”Acoustics”

Supervisors: František Kadlec Czech Technical University
Laurent Simon Université du Maine
Pierrick Lotton Université du Maine

Prague / Le Mans 2008
To my family.
The work on this doctoral thesis was supported in part by the Ministry of Education in Czech Republic under research program MSM6840770014, by Erasmus and by the French Embassy in Prague within the framework of the program "Cotutelle de thèse".
Acknowledgements

First of all I would like to express my gratitude to my supervisors František Kadlec, Laurent Simon and Pierrick Lotton for their generous support and ideas, suggestions and criticism, without which this work would not have been possible. I am also indebted to professor Zdeněk Škvor and doctor Libor Husník for making possible the French part of my doctoral studies.

My thanks also belong to all the members of Laboratoire d’Acoustique de l’Université du Maine in Le Mans for the friendly but very professional ambiance I worked in. Special thanks to professors Michel Bruneau for his advices, and to Joël Gilbert and Vincent Tournat for their help with experimental measurements.

My deepest obligation comes to my parents for encouragement and support during my entire studies. Last but no least, I would like to express my thanks and gratitude to my future wife Veronika for her love and patience without which it would have been much harder.

Concerning the financial arrangements, this work was partly supported by Erasmus and by the French Embassy in Prague within the framework of the program ”Cotutelle de thèse”.

Antonín Novák

Prague, December 2008
Abstract

The theory of linear time-invariant (LTI) systems has been extensively studied over decades and the estimation of any unknown LTI system, knowing both the input and output of the system, is a solved problem. Nevertheless, almost all real-world devices exhibit more or less nonlinear behavior. In the case of very weak nonlinearities, a linear approximation can be used. If the nonlinearities are stronger, the linear approximation fails and systems have to be described using a nonlinear model.

The goal of this thesis is to design and develop simple methods for nonlinear systems identification that would be accurate and robust enough to be applicable for analysis and identification of nonlinear systems in several domains, even if the main focus here is on the domain of audio and acoustics. The goal is to identify a nonlinear system and find its generic nonlinear model in such way that the response of the model to any input signal would be the same as the one of the real-world nonlinear system under test.

Two methods are developed in the thesis. Both methods are based on Multiple Input – Single Output (MISO) model. The model consists of several parallel branches, each branch consisting of two separated blocks: a nonlinear static function and a linear dynamic filter. The first method uses a white Gaussian noise as the excitation signal for the identification. This method is successfully tested on several simulation examples, but fails when identifying real world nonlinear systems. The second method is based on the nonlinear convolution and uses swept sine excitation signal. This method is successfully tested on several simulation examples. Moreover, it is theoretically shown that it could be used for the identification of systems exhibiting specific dynamical hysteresis (called hysteresis with viscosity-type effect).

Two well known real world nonlinear systems (an audio limiter and an acoustic waveguide) are used to validate the second method. The validation is based on the comparison between the output of these real world systems and the output of their estimated models, when excited with the same input signal. The comparison is performed both subjectively, using a simple visual comparison in time or frequency domains, and objectively, using a relative mean square
error criterion. Once validated, the method is used in the general frame of the study of electrodynamic loudspeaker quality. Preliminary results show that this method could be used for the nonlinearities loudspeakers identification, and that an inverse filtering minimizing these nonlinearities could possibly be performed with the help of this method.
Abstrakt


Cílem této disertační práce je navrhnout a vyvinout jednoduchou metodu analýzy nelineárních systémů, která by byla dostatečně přesná a robustní pro analýzu nejen audio a akustických systémů, na které se práce vztahuje, ale i nelineárních systémů z jiných oblastí. Cílem samotné identifikace je pak nalezení takového obecného nelineárního modelu, jehož odezva na libovolný signál by se shodovala s odezvou analyzovaného systému na stejný signál.


Dva velice dobře známé reálné systémy (audio limiter a akustický vlnovod) jsou použity pro ověření druhé metody. Ověření je založeno na porovnání odezv reálného systému a jeho modelu, při buzení stejným signálem. Porovnání je provedeno jak subjektivně, za použití vizuálního porovnání časových i frekvenčních průběhů, tak objektivně, použitím relativní střední kvadratické odchylky. Poté co je metoda ověřena na systémech, jejichž nelineární chování je známo, je metoda použita na analýzu nelinearit elektrodynamického reproduktoru. První výsledky ukazují, že tato metoda může být použita na analýzu nelinearit reproduktoru a že by mohla být dále použita na možnou eliminaci nelineárního zkreslení pomocí inverzní filtrace.
Abstract

La théorie des systèmes linéaires invariant a fait l’objet de nombreuses études au cours de ces dernières décennies, et l’estimation d’un tel système à partir du signal de sortie lorsque le signal d’entrée est connu est un problème aujourd’hui résolu. Cependant, le comportement de tout système réel est plus ou moins non-linéaire. Dans le cas de faibles non-linéarités, une approximation linéaire peut être effectuée, mais lorsque les non-linéarités sont plus importantes, cette approximation linéaire n’est plus valide et il est nécessaire d’utiliser une représentation non-linéaire.

L’objet de ce travail de thèse est de développer des méthodes simples pour l’identification de systèmes non-linéaires. Ces méthodes doivent être suffisamment précises et robustes pour être utilisées dans différents domaines d’application, même si l’étude est principalement limitée aux domaines de l’audio et de l’acoustique dans le cadre de ce travail de thèse. L’identification d’un système non-linéaire consiste à déterminer un modèle générique non-linéaire de ce système, de telle sorte que le modèle et le système réel étudié délivrent un même signal de sortie lorsqu’ils sont excités par un signal d’entrée identique.

Deux méthodes sont développées, toutes deux basées sur un modèle de type ”Multiple Input – Single Output” (MISO). Suivant cette modélisation, le système étudié peut être représenté par un ensemble de branches en parallèle, chaque branche comportant deux blocs-fonctions distincts : une fonction non-linéaire statique et un filtre linéaire dynamique. La première méthode développée utilise un bruit blanc gaussien comme signal d’excitation nécessaire à la procédure d’identification. Cette méthode donne de bons résultats lorsqu’elle est appliquée à l’étude de systèmes simulés. Cependant, elle montre des limitations rédhibitoires lorsqu’elle est appliquée à l’étude de systèmes réels. La deuxième méthode développée est basée sur le principe de déconvolution non-linéaire et utilise un ”swept sine” comme signal d’excitation. Cette méthode donne de bons résultats lorsqu’elle est appliquée à l’étude de systèmes simulés. Par ailleurs, une étude théorique montre, sur des cas simulés, que cette méthode peut être utilisée pour l’identification de systèmes dont le comportement révèle une hystérésis dynamique particulière (encore appelée hystérésis ”de type visqueux”).
Deux systèmes non-linéaires bien connus, un limiteur audio et un guide d’ondes acoustiques, sont utilisés pour effectuer une validation expérimentale de la deuxième méthode. La validation est basée sur la comparaison entre les signaux obtenus en sortie de ces systèmes réels et en sortie de leurs modèles lorsqu’un même signal d’excitation est utilisé. Cette comparaison est réalisée à la fois de manière subjective (simple comparaison visuelle entre les signaux, dans le domaine temporel et dans le domaine fréquentiel) et de manière objective (critère d’erreur relative). Une fois validée, cette méthode est utilisée dans le cadre plus large de l’étude de la qualité des haut-parleurs électrodynamiques. Des résultats préliminaires sont présentés, qui permettent d’envisager l’utilisation de la méthode pour identifier, voire pour corriger par filtration inverse, les non-linéarités présentées par ce type de haut-parleur.
Preface \hspace{1cm} i
Abstract \hspace{1cm} iii
Notation \hspace{1cm} xix

1 Introduction \hspace{1cm} 1

2 Nonlinear System Identification \hspace{1cm} 7
  2.1 Introduction \hspace{1cm} 7
  2.2 Nonlinear Models \hspace{1cm} 8
    2.2.1 The Volterra Series Model \hspace{1cm} 8
    2.2.2 Simple Wiener and Hammerstein Models \hspace{1cm} 10
    2.2.3 Polynomial Hammerstein and Wiener Models \hspace{1cm} 13
    2.2.4 Multiple-Input Single-Output Model \hspace{1cm} 14
  2.3 Excitation Signals \hspace{1cm} 15
    2.3.1 White Gaussian Noise \hspace{1cm} 15
    2.3.2 Pseudo-Random Signals \hspace{1cm} 16
    2.3.3 Harmonic Signals \hspace{1cm} 17
    2.3.4 Multitone Signals \hspace{1cm} 17
    2.3.5 Swept Sine Signals \hspace{1cm} 18
  2.4 Methods of Identification \hspace{1cm} 19
    2.4.1 Methods for the Volterra Model \hspace{1cm} 20
    2.4.2 Methods for the Wiener and Hammerstein Models \hspace{20}
6 Nonlinear Systems in Acoustics

6.1 Measuring Methodology

6.2 Objective Criteria for Verification

6.3 Validation on an Audio Limiter

6.4 Validation on an Acoustic Waveguide

6.5 Application to Electrodynamic Loudspeaker

6.5.1 Context

6.5.2 Experimental Setup

6.5.3 Preliminary Results

6.5.4 Total Harmonic Distortion Measurement

6.6 Summary
List of Figures

1.1 Block schema of a nonlinear system ........................................ 5
2.1 Block schema of the Volterra model ........................................ 9
2.2 Block diagram of the simple Hammerstein model ........................... 11
2.3 Block diagram of the simple Wiener model ................................. 11
2.4 Simple Hammerstein (a) and simple Wiener (b) systems with polynomial static nonlinear parts .................................................. 11
2.5 The second order Volterra kernels: (a) kernel from the example of Fig. 2.4a; (b) kernel from the example of Fig. 2.4b ..................... 12
2.6 Nth-order Polynomial Hammerstein model .................................. 13
2.7 Nth-order Polynomial Wiener model ........................................... 13
2.8 Multiple-Input Single-Output model of nonlinear system with indicated linear part (dashed rectangle) ................................. 14
2.9 Properties of WGN for $\sigma = 1$ (a) probability density function; (b) cumulative probability density function; (c) auto-correlation function; (d) power-spectral density function .......................... 15
2.10 An example of the pseudo-random binary signal ........................... 16
2.11 An example of a swept-sine with linear time-varying instantaneous frequency ................................................................. 19
2.12 Linear part of the nonlinear MISO model .................................. 21
2.13 Three Input Single Output block diagram with linear filters $L_{ij}$ representing the mutual correlation of inputs ............................... 22
2.14 Block diagram of the nonlinear convolution process ..................... 23
2.15 Result of the nonlinear convolution process $y(t) \ast \tilde{s}(t)$ in the form of set of higher-order nonlinear impulse responses $h_m(t)$ .... 24

2.16 Result of the nonlinear convolution process in the form of higher-order nonlinear frequency responses $H_n(f)$. Arbitrary case where $|H_n(f)|$ is a constant in $[nf_1, f_2]$ .... 25

3.1 The polynomial Hammerstein model ....... 28

3.2 Three-Input Single-Output block diagram with linear filters $L_{ij}$ representing the mutual correlation of inputs ....... 29

3.3 MISO Inputs decorrelation for $N = 8$ order (the odd and even powers separated) ....... 32

3.4 Multiple-Input Single-Output model for conditioned uncorrelated inputs ....... 33

3.5 Pre-set of standard deviation of input signal for nonlinear function measuring. (a) large $\sigma_x$ with no offset, (b) weak $\sigma_x$ with offset equal to 1. Nonlinear function $y = \mathcal{F}\{x\}$ - above, probability density functions of input signal $x(t)$ - below ....... 35

3.6 Input-Output characteristics with mean value shift ....... 36

3.7 A limiter identification - comparison between full amplitude range with zero mean value (a) and a narrowed amplitude range with a nonzero mean value in order to zoom the nonlinear discontinuity (b) ....... 38

3.8 Nonlinear function of the clarinet nonlinearity - I/O characteristic ....... 39

3.9 Identification of the clarinet nonlinearity - a) estimation with full amplitude range $\sigma_x = 2/3$ and $\mu_x = 1$; b) $\sigma_x = 1/3$ and $\mu_x = -1$; c) $\sigma_x = 1/15$ and $\mu_x = 0$; d) $\sigma_x = 1/6$ and $\mu_x = 0.5$; e) $\sigma_x = 1/3$ and $\mu_x = 1$; f) $\sigma_x = 1/15$ and $\mu_x = 0.2$; g) $\sigma_x = 1/3$ and $\mu_x = 2$; g) $\sigma_x = 1/2$ and $\mu_x = 3$ ....... 40

3.10 Clarinet Nonlinearity - reconstruction ....... 41

3.11 Tested nonlinear system with memory (a limiter followed by a linear low-pass IIR filter) ....... 41

3.12 Frequency response modulus (above), phase (below) of the linear low-pass IIR filter $H(f/f_s)$ ....... 42

3.13 Modulus (above) and phase (below) of the estimated filter $H_1(f)$ in the case of the input signal with parameters $\mu = 0$ and $\sigma = 0.3$ ....... 43

3.14 The estimation error - modulus (above) and phase (below) ....... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.15</td>
<td>Reconstruction of the nonlinear systems output. The soft-dashed line represents the input signal; the dashed line represents the theoretical output; the solid line represents the output estimated by the MISO method. a) normalized frequency 0.001, b) normalized frequency 0.03</td>
</tr>
<tr>
<td>4.1</td>
<td>Swept-sine signal ( s(t) ) in time domain (below), with the time length chosen according to instantaneous frequency ( f_i(t) ) (above)</td>
</tr>
<tr>
<td>4.2</td>
<td>Swept-sine signal ( s(t) ) (above) and the swept-sine signal ( s_2(t) ) (below)</td>
</tr>
<tr>
<td>4.3</td>
<td>A zero-memory nonlinear system case; convolution between ( s^3(t) ) and ( \tilde{s}(t) )</td>
</tr>
<tr>
<td>4.4</td>
<td>An example of a Hammerstein nonlinear system (a system with memory)</td>
</tr>
<tr>
<td>4.5</td>
<td>The system with memory depicted in Fig. 4.4: convolution between the output ( y(t) ) and ( \tilde{s}(t) )</td>
</tr>
<tr>
<td>4.6</td>
<td>Magnitude of the Fourier transform of the exponential swept-sine signal</td>
</tr>
<tr>
<td>4.7</td>
<td>The inverse filter ( \tilde{s}(t) ) expressed as a time-reversed replica of the swept-sine signal ( s(t) ) with a different amplitude envelope</td>
</tr>
<tr>
<td>4.8</td>
<td>The extended inverse filter ( \tilde{s}(t) ) expressed as a time reversed replica of the extended swept-sine signal ( s(t) ) with a different amplitude envelope</td>
</tr>
<tr>
<td>4.9</td>
<td>Result of the nonlinear convolution process, when using the extended inverse filter in the form of higher-order nonlinear frequency responses ( H_n(f) ). Arbitrary case where (</td>
</tr>
<tr>
<td>4.10</td>
<td>The polynomial Hammerstein model for the case of blind identification</td>
</tr>
<tr>
<td>4.11</td>
<td>The general MISO model for the case of known inputs</td>
</tr>
<tr>
<td>4.12</td>
<td>The way to express a distorted sine signal ( y(t) ) (including higher harmonics) as the sum of powers of harmonic signal, in the frequency domain</td>
</tr>
<tr>
<td>4.13</td>
<td>Tested NLS with memory</td>
</tr>
<tr>
<td>4.14</td>
<td>The theoretical frequency response (modulus - above; phase - below) of the filter ( G_1(f) ), from the linear part of the NLS of Fig. 4.13</td>
</tr>
<tr>
<td>4.15</td>
<td>The estimated frequency response (modulus - above; phase - below) of the filter ( G_1(f) ), from the linear part of the NLS of Fig. 4.13</td>
</tr>
</tbody>
</table>
4.16 The theoretical frequency response (modulus - above; phase - below) of the filter $G_3(f)$, from the linear part of the NLS of Fig. 4.13

4.17 The estimated frequency response (modulus - above; phase - below) of the filter $G_3(f)$, from the linear part of the NLS of Fig. 4.13

5.1 A general schema of hysteretic system: black-box representation

5.2 Input-Output characteristic of a static hysteretic system with (a) local, (b) nonlocal memory: Hysteresis loop

5.3 Relay with hysteresis: static nonlinearity with local memory

5.4 Waveform of the input (sine) and output signal (after passing through the relay with hysteresis)

5.5 A block-schema of relay with hysteresis represented via phase delay and saturation

5.6 A block-schema of dynamic hysteresis

5.7 The hysteresis loop of the quadratic hysteresis $q[x(t), x_m]$, $x_m$ being the input excitation signal maximum value

5.8 The nonlinear hysteretic system with a WGN $n(t)$ added to the output

5.9 Amplitude (a) and phase (b) of the theoretical filter $\tilde{\alpha}(f)$ (solid) and estimated filter $A_1(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for SNR=30dB

5.10 Amplitude (a) and phase (b) of the theoretical filter $\tilde{\beta}(f)$ (solid) and estimated filter $A_2(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for SNR=30dB

5.11 Amplitude (a) and phase (b) of the theoretical filter $\tilde{\gamma}(f)$ (solid) and estimated filter $A_3(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for AWGN of SNR = 30dB

6.1 Input-output characteristic of a limiter with threshold breakpoint $(a, \alpha)$

6.2 Block diagram of the audio limiter measurement
6.3 Comparison of the real-world and regenerated model output signals for the input sine-wave \( f = 500 \text{ Hz}, \ A = 1 \text{ V} \); (a) Time domain comparison between the limiter output (thin) and model output (bold); (b) Frequency domain comparison between the first 8 harmonics of the limiter output (thin) and model output (dashed-bold and shifted) ......................................................... 85

6.4 Comparison of the real-world and regenerated model output signals for the input sine-wave \( f = 500 \text{ Hz}, \ A = 0.3 \text{ V} \); (a) Time domain comparison between the limiter output (thin) and model output (bold); (b) Frequency domain comparison between the limiter output (thin) and model output (dashed-bold and shifted) .......... 86

6.5 Comparison of the responses of the real system (thin) and its estimated nonlinear model (bold) to the sawtooth signal (dashed) 86

6.6 The block scheme of a waveguide with illustrated distortion of an acoustic wave ................................................................. 87

6.7 Block diagram of the waveguide nonlinearities measurement ... 88

6.8 Comparison between the theoretical frequency dependency of coefficients \( B_1, B_2, B_3 \) (straight lines) and the three first higher-order experimental nonlinear frequency responses ................. 89

6.9 Three output signals in the time domain (pure-sine input wave dashed) - all figures 6.10 together ......................................................... 90

6.10 Three output signals in the time domain (pure-sine input wave dashed): (a) the theoretical (Fubini solution); (b) the real response of the waveguide (real measurement); (c) the response of the obtained nonlinear model (regenerated signal) .............................. 91

6.11 A schema of an electrodynamic loudspeaker ................................. 93

6.12 Block diagram of the loudspeaker nonlinearities measurement ... 95

6.13 Spectrum of the responses to the two-harmonic signals with frequencies 600 Hz and 3100 Hz: comparison between the real systems output (left) and the MISO models output (right) .......... 97

6.14 Spectrum of the responses to the two harmonic signals with frequencies 700 Hz and 1100 Hz: comparison between the real system output (left) and the MISO model-based output (right) . 98

6.15 Total Harmonic Distortion (THD) as a function of frequency: swept-sine based method (solid line) and classical method point by point (crosses) ......................................................... 100

xvii
## General notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{N}$</td>
<td>set of natural numbers</td>
</tr>
<tr>
<td>$\mathbb{Z}$</td>
<td>set of integer numbers</td>
</tr>
<tr>
<td>( \log )</td>
<td>decadic logarithm ( \log_{10} )</td>
</tr>
<tr>
<td>( \ln )</td>
<td>natural logarithm ( \log_e )</td>
</tr>
<tr>
<td>( \begin{bmatrix} \ldots \end{bmatrix} )</td>
<td>matrix</td>
</tr>
<tr>
<td>( \begin{bmatrix} \ldots \end{bmatrix}^T )</td>
<td>transpose of a matrix</td>
</tr>
<tr>
<td>( \begin{bmatrix} \ldots \end{bmatrix}^{-1} )</td>
<td>inverse of a matrix</td>
</tr>
<tr>
<td>( x'' )</td>
<td>second derivative with respect to time</td>
</tr>
<tr>
<td>$\mathcal{F}_a{\cdot}$</td>
<td>zero-memory nonlinear function</td>
</tr>
<tr>
<td>$H[\cdot]$</td>
<td>Hilbert transform</td>
</tr>
<tr>
<td>$FT[\cdot]$</td>
<td>Fourier transform</td>
</tr>
<tr>
<td>$E[\cdot]$</td>
<td>ensemble average</td>
</tr>
<tr>
<td>$J_n(x)$</td>
<td>Bessel functions of the first kind</td>
</tr>
</tbody>
</table>
Physical notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>time</td>
<td>s</td>
</tr>
<tr>
<td>$T$</td>
<td>time length</td>
<td>s</td>
</tr>
<tr>
<td>$f$</td>
<td>frequency</td>
<td>Hz</td>
</tr>
<tr>
<td>$f_i$</td>
<td>instantaneous frequency</td>
<td>Hz</td>
</tr>
<tr>
<td>$t_f$</td>
<td>group delay</td>
<td>s</td>
</tr>
<tr>
<td>$c_0$</td>
<td>speed of sound</td>
<td>m/s</td>
</tr>
<tr>
<td>$p$</td>
<td>pressure</td>
<td>Pa</td>
</tr>
<tr>
<td>$p_b$</td>
<td>atmospheric pressure</td>
<td>Pa</td>
</tr>
<tr>
<td>$U$</td>
<td>volume flow</td>
<td>m³/s</td>
</tr>
<tr>
<td>$l$</td>
<td>distance</td>
<td>m</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>heat capacity ratio</td>
<td>-</td>
</tr>
</tbody>
</table>

Signal processing notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi$</td>
<td>phase in time domain</td>
</tr>
<tr>
<td>$\delta(t)$</td>
<td>Dirac function</td>
</tr>
<tr>
<td>$x, x(t)$</td>
<td>input signal</td>
</tr>
<tr>
<td>$y, y(t)$</td>
<td>output signal</td>
</tr>
<tr>
<td>$a(t)$</td>
<td>amplitude envelop</td>
</tr>
<tr>
<td>$s(t)$</td>
<td>swept-sine signal</td>
</tr>
<tr>
<td>$z_s(t)$</td>
<td>analytical signal of $s(t)$</td>
</tr>
<tr>
<td>$\tilde{s}(t)$</td>
<td>inverse filter</td>
</tr>
<tr>
<td>$x(n)$</td>
<td>input signal in discrete time</td>
</tr>
<tr>
<td>$y(n)$</td>
<td>output signal in discrete time</td>
</tr>
<tr>
<td>$X(f)$</td>
<td>spectrum of input signal</td>
</tr>
<tr>
<td>$Y(f)$</td>
<td>spectrum of output signal</td>
</tr>
<tr>
<td>$Z_s(f)$</td>
<td>spectrum of analytical signal</td>
</tr>
<tr>
<td>$B_s(f)$</td>
<td>magnitude spectra</td>
</tr>
<tr>
<td>$\Psi_s(f)$</td>
<td>phase in frequency domain</td>
</tr>
<tr>
<td>$K_n$</td>
<td>$n$-th order Volterra operator</td>
</tr>
</tbody>
</table>
\( k_n \) \( n \)-th order Volterra kernel
\( g(t), h(t) \) impulse response of a linear filter
\( G(f), H(f), L(f) \) frequency response of a linear filter
\( \sigma_x \) standard deviation of \( x \)
\( \sigma_x^2 \) variance of \( x \)
\( \mu_x, x_{ss} \) mean value of \( x \)
\( p(x) \) probability density function
\( P(x) \) cumulative probability density
\( R_{xx}(\tau) \) auto-correlation function
\( R_{xy}(\tau) \) cross-correlation function
\( S_{xx}(f) \) power spectral density
\( S_{xy}(f) \) cross spectral density
\( x_{j,i}(n) \) signal \( x_j(t) \) is decorrelated with \( x_i(t) \)
\( x_{j,i}(n) \) signal \( x_j(t) \) is fully correlated with \( x_i(t) \)
\( x_{j,i,n}(n) \) signal \( x_j(t) \) is decorrelated with all the \( x_m(t), \quad (m < i) \)
\( k_{n,m} \) constants of uncorrelated model

Abbreviations

ACF \quad Auto-Correlation Function
A/D \quad Analog to Digital
AWGN \quad Additive White Gaussian Noise
CCF \quad Cross-Correlation Function
CSD \quad Cross-Spectral Density Function
CEI \quad Commission Electrotechnique Internationale
D/A \quad Digital to Analog
dB \quad Decibel
dsp \quad Digital Signal Processing
DC \quad Direct Current
DFT \quad Discrete Fourier Transform
DOF \quad Degree of Freedom
FRF \quad Frequency Response Function
FT \quad Fourier Transform
I/O \quad Input / Output
LTI \quad Linear Time-Invariant
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISO</td>
<td>Multiple-Input Single-Output</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>NARMAX</td>
<td>Nonlinear Auto Regressive Moving Average with Exogenous Inputs</td>
</tr>
<tr>
<td>NL</td>
<td>Nonlinear</td>
</tr>
<tr>
<td>NLS</td>
<td>Nonlinear System</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PRBS</td>
<td>Pseudo-Random Binary Signal</td>
</tr>
<tr>
<td>PSD</td>
<td>Power-Spectral Density Function</td>
</tr>
<tr>
<td>rms</td>
<td>Root Mean Square</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input Single-Output</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>THD</td>
<td>Total Harmonic Distortion</td>
</tr>
<tr>
<td>WGN</td>
<td>White Gaussian Noise</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

Study of the behavior of physical systems usually begins with a basic modeling process. The demands on more detailed study of physical phenomena governing this behavior leads generally to more complicated models that can be useful for a better understanding of the system and for a prediction or a simulation of its behavior. The procedure allowing the derivation of mathematical tools and algorithms to build a model is called \textit{system identification}.

Each identification process consists of several steps. First of all, a simple analysis of the system is recommended, examining the basic properties of the system. Then, based on this information about the system, a model or a structure representing the system can be chosen. The choice of an appropriate model is a very important step and it is very often a compromise between accuracy and simplicity. In practice, a very accurate model (with the smallest error) is not always preferred if its complexity is too high. Next step is the estimation of the model parameters from the measurement. Finally, the validity of the estimated model should be tested. The validation can indicate whether the model describes the system properly or not.

Physical systems can be separated into two distinct classifications: linear systems and nonlinear systems. As a symptom of nonlinearity, the output of a nonlinear system generally exhibits distortions. Encyclopedia Britannica defines
the distortion as *any change in a signal that alters the basic waveform or the relationship between various frequency components* and adds that it is usually a *degradation of the signal*. Distortion is usually unwanted, even if in some cases, (such as electric guitar where distortion is often induced purposely to achieve an aggressive sound), it can be desirable.

Regarding the consequences of nonlinearity on system behavior, two types of distortion are usually distinguished: harmonic distortion and intermodulation distortion. When exciting a system with an input sine-wave signal of frequency $f_0$, the harmonic distortion corresponds to components of output signal at frequencies which are integer multiples of $f_0$. When exciting a system with an input signal made of two sine-waves having different frequencies $f_1$ and $f_2$, the intermodulation distortion corresponds to the components having frequency values that are linear combinations of $f_1$ or $f_2$, such as $f_1 - f_2$, $f_1 + f_2$, $2f_1 - f_2$, ... If the system produces at least one of these two effects (in general both of them), it is considered to be nonlinear.

The very general definition of a nonlinear system states that the output of the system does not depend linearly on the input. In other words, the system property does not meet at least one of the linearity principles: (a) additivity and (b) homogeneity. The system obeys the principle of additivity if the response caused by two or more different signals is the sum of the individual responses of each signal. The system obeys the principle of homogeneity if the output signal corresponding to an attenuated version of an input signal is also attenuated by the same attenuation factor.

There are a lot of everyday life examples, where a nonlinear system is considered to be a linear one. For example, in catalogue files or magazine tests, a system (as a loudspeaker for example) can be seen as a linear one (when presenting the linear transfer function) and, at the same time, a nonlinear one (when presenting the total distortion coefficient). Since the theory of nonlinear systems is generally complex, the nonlinear systems are usually represented by approximated linear models. Thus, the distortion is often a priori considered to be low enough, in order to be neglected. In the worst case, the presence of nonlinear distortions is disregard completely. This can be an important source of error when modeling the system under test, because of involuntary usage of an invalid model.
The theory of linear time-invariant (LTI) systems has been extensively studied over decades (for example, (Kailath, T., 1980), (Lathi, B.P., 2000)) and the estimation of any unknown LTI system, knowing both the input and output of the system, is a solved problem. The fundamental idea of the theory states that any LTI system can be characterized entirely by its impulse response in the time domain or by its frequency response function in the frequency domain. Nevertheless, almost all real-world devices exhibit more or less nonlinear behavior. In the case of very weak nonlinearities, a linear approximation can be used. If the nonlinearities are stronger, the linear approximation fails and systems have to be described using a nonlinear model.

The simplest way for analyzing a nonlinear system is probably by using a sine-wave as input signal. Regarding the basics of the nonlinear system (NLS) theory, when a pure sine-wave input signal \( x(t) = A_1 \cos(2\pi f_1 t + \phi_1) \) passes through a NLS, higher-order harmonic components appear at the output of the system as multiples of input frequency, according to \( y(t) = \sum_n B_n \cos(2\pi n f_1 t + \phi_n) \). The characteristics (amplitude \( B_n \) and phase \( \phi_n \)) of all higher-order components may be furthermore frequency and input amplitude dependent, in the sense that \( \forall n \ B_n \equiv B_n(A_1, f_1), \phi_n \equiv \phi_n(A_1, f_1) \). The complete identification procedure consequently consists in the estimation of the amplitude \( B_n \) and phase \( \phi_n \) as functions of input amplitude and frequency. This method of analysis gives accurate result but it is not very effective, because of the large number of measurements which are required at different frequencies and different amplitudes of the input signal. Furthermore, when considering the intermodulation distortion, when two sine-waves of different frequencies are used and where the output consists of many frequency components, that are combinations of the input frequencies, the analysis become very complicated. In such cases, another solution adapted to the nonlinear problem should be used.

When studying a nonlinear system, two different methodological approaches can be considered. A first approach is based on the study of the physical cause of each nonlinear phenomenon which plays a role in the global behavior of the system. The nonlinear physical laws governing all these phenomena have to be derived. These laws can then be used for the derivation of a theoretical nonlinear model of the system. However, such an approach could lead to very complicated model
and some real-world systems, with very complex nonlinear behavior, exclude such a physical approach. This physical approach is not considered in the frame of this work.

A second approach consists in a more global modeling of the nonlinear system considered as a whole. Two different cases can be distinguished, regarding the available knowledge of the physical properties of the system. In a first case, some physical properties are a priori known. Then, the modeling can be performed using this a priori knowledge, and the system is said to be a grey box. In a second case, no physical properties (or so little) are known. Then, the system is said to be a black box. In both cases, a generic nonlinear model is used for nonlinear system identification.

A lot of generic nonlinear models have been developed during the last decades. Such nonlinear models are available: for example, Volterra model (Schetzen, M., 1989), Wiener model (Greblicki, W., 1997), Hammerstein model (Chang, F. & Luus, R., 1971), neural network model (Nelles, O., 2001), MISO model (Bendat, J.S., 1998), NARMAX model (Chen, S. & Billings, S, 1989), hybrid genetic algorithm (Yen-Wei Chen et al., 2003), extended Kalman filtering (Sorenson, H.W., 1985), particle filtering (Arulampalam, M.S. et al., 2002). Billings, S.A. & Fakhouri, S.Y. (1978) classified the nonlinear models into three basic classes: (a) Block-oriented structured approaches (Wiener model, Hammerstein model, MISO model), (b) Kernel bases, or non-parametric approaches (Volterra model), (c) Parametric approaches (NARMAX). Nevertheless, this classification is only a preliminary one, considering that methods, i.e. the block oriented systems identification by using nonparametric approach (Greblicki, W. & Pawlak, M., 1989), can be included in two or three of these classes.

All these models involve blocks, kernels or parameters that have to be estimated. Then, an identification procedure has to be done. This procedure is said to be a blind identification procedure when applied to a black box system. It is based on the analysis of the signal $y(t)$ produced at the output of the system under test when excited by a given and controlled input signal $x(t)$ (Fig. 1.1).
The specific goal of this thesis is to design and develop a simple method for nonlinear system identification, that would be accurate and robust enough to be applicable for analysis and identification of numerous nonlinear systems in several domains. The main focus is in the domains of audio and acoustics. The goal is to identify a nonlinear system and find a generic nonlinear model in such way that the response of the model to any input signal would be the same as that of the real-world nonlinear system under test. The methods developed in this thesis use the block-oriented structures and non-parametric approach.

The content of this thesis is as follows. Chapter two provides an overview of the literature on nonlinear system identification, including a view on the nonlinear models and excitation signals used in the identification process. In the framework of this dissertation work, two methods have been developed: a first one based on a random signal excitation (Chapter three) and a second one based on a swept-sine excitation (Chapter four). The method in Chapter three is successfully tested on simulations, but when used in practice, the method does not give accurate enough results. The second method (Chapter four) is the heart of this work. In this chapter, two algorithms have been developed, suited to two different models. The developed methods have been tested using simulations carried out in discrete-time domain using Matlab language (R2007b) and the results are presented at the end of Chapters three and four. Chapter five deals with the application of method developed in Chapter four to hysteretic systems. Chapter six provides the results of real-world examples in acoustics, electroacoustics and audio. Two experiments with known physical models or laws are detailed in order to validate the method. Then, an application to electrodynamic loudspeaker nonlinearities is proposed. Finally, the conclusion of the work is reviewed in Chapter seven and prospects of this work are developed.
2.1 Introduction

As this thesis deals with nonlinear models based on non-parametric black-box modeling, the state of art proposed in this chapter is focused on this kind of models. In this chapter, a summary of different models, excitation signals and methods used for nonlinear systems (NLS) identification is presented. The goal of this summary is not to give a complete state of the art, as presented in the article "A bibliography on nonlinear system identification" (Giannakis, G.B. & Serpedin, E., 2001) including more than a thousand citations, but to present a set of basic models, excitation signals and methods, that are closely relative to the methods developed in this thesis. Above all, as the words model, excitation signal and method are often used in the thesis, it is convenient to define these three terms before presenting the summary.

model By model, we suggest that the behavior of the NLS may be defined by using functions or operators describing the way in which the input and output signals are related. This relation defined by the functions or operators can be then represented by a block schematic diagram. It is of common use (Haber, R. & Keviczky, L., 1999) to distinguish the case of so-called parametric models, for which the unknowns are a finite number of
parameters, from the case of non-parametric models, for which the problem may not be synthesized in a finite number of parameters.

**excitation signal** The so-called excitation (or input) signal is the time-varying signal $x(t)$ which has to be designed for the purpose of identification. Classification of input signals usually involve deterministic signals vs random processes, narrow-band vs broad-band signals, and so on. The most common excitation signals for NLS identification include white Gaussian noise (WGN) processes, pseudo-random sequences (both being broadband random processes), impulse signals, chirp and multitone signals (all being broad-band deterministic signals), or sine (being narrow-band deterministic signals). Each excitation signal exhibits properties yielding advantages and drawbacks as detailed below.

**method** By the term *method*, we mean a way in which the parameters of the model are obtained. In general, the method is linked to the model and the excitation signal.

### 2.2 Nonlinear Models

This section gives an overview of nonlinear models, related to the models further used in the thesis. First, the classical nonlinear model based on the Volterra series is recalled. Then, simpler models based on the Wiener-Hammerstein structure are presented and finally a linear multiple input-single output (MISO) model closely related to the Hammerstein model is mentioned.

#### 2.2.1 The Volterra Series Model

Nonlinear Volterra theory, developed in the 1880s by Vito Volterra, states that any time-invariant NLS can be modeled as an infinite sum of multidimensional convolution integrals of increasing order (VOLterra, V., 1930). This is represented by Volterra series as a classical model for NLS with memory effect. The theory of LTI systems defines the impulse response as the time signature of the system (or equivalently the frequency response as the frequency signature of the system). For NLS, the Volterra model defines higher-order multidimensional impulse responses, also called kernels. The relation between input and output
2.2. NONLINEAR MODELS

signals \( x(t) \) and \( y(t) \) can be expressed as

\[
y(t) = \sum_n K_n[x(t)],
\]

where the \( n \)-th order Volterra operator \( K_n \) describes the \( n \)-dimensional convolution (Schetzen, M., 1989)

\[
K_n[x(t)] = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} k_n(\tau_1, ..., \tau_n)x(t - \tau_1)...x(t - \tau_n)d\tau_1...d\tau_n,
\]

between the input signal \( x(t) \) and the \( n \)-dimensional Volterra kernel \( k_n(...) \). The block diagram of a \( N \)-th order Volterra model is shown in Fig. 2.1. Detailing Eq. (2.1) leads to (Schetzen, M., 1989)

\[
y(t) = k_0 + \int_{-\infty}^{\infty} k_1(\tau_1)x(t - \tau_1)d\tau_1
\]

\[
+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_2(\tau_1, \tau_2)x(t - \tau_1)x(t - \tau_2)d\tau_1d\tau_2
\]

\[
+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_3(\tau_1, \tau_2, \tau_3)x(t - \tau_1)x(t - \tau_2)x(t - \tau_3)d\tau_1d\tau_2d\tau_3
\]

\[
\vdots
\]

\[
+ \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} k_n(\tau_1, ..., \tau_n)x(t - \tau_1)...x(t - \tau_n)d\tau_1...d\tau_n,
\]

where \( k_0 \) is the constant zero-order term of the NLS. The Volterra kernel \( k_1(\tau) \) corresponds to the linear one-dimensional impulse response. For a linear system, the higher-order multi-dimensional impulse responses \( k_n(...) \) of the Volterra series are equal to zero, for \( n \geq 2 \).
As the series implies infinite extension over time and order (Reed, M.J. & Hawksford, M.O., 1996a), the Volterra series is in practice truncated in both time and order. From the mathematical point of view, the finite Volterra series cannot express every type of nonlinearity as well as the finite Taylor series cannot express every function. The higher the order of the system, the higher the accuracy of the model, but the higher the complexity. A representation of the higher-order contributions is also more difficult to represent as the dimension increases. Since the nth-Volterra kernel is a function of \( n \) variables, the model representing the system contains too many coefficients necessary to determine the system.

2.2.2 Simple Wiener and Hammerstein Models

Simple Hammerstein and Wiener models are block-oriented nonlinear models, composed of a combination of linear dynamic block and nonlinear static, or zero-memory block\(^1\). The well-known Wiener systems and Hammerstein systems are Single-Input Single-Output (SISO) nonlinear models that are often used in real-world applications, in areas such as biology (Emerson, R.C. et al., 1992), chemistry (Bhandari, N. & Rollins, D., 2004), speech coding (Turunen, J. et al., 2003), communications and control (Cerone, V. & Regruto, D., 2003). By using these models, a NLS can be represented accurately with less parameters than a Volterra model-based system (Gomez, J.C. & Baeyens, E., 2000).

An advantage of block-oriented representation is furthermore that the nonlinear part is completely separated from the linear part. The stability of the system is thus determined solely by the linear part of the model. Another advantage is that if the physical nonlinear model of the system under test is approximately known, it can serve as the initial knowledge of the static nonlinearity and only the linear dynamic part remains to be estimated. A disadvantage can be the difficulty of selecting a suitable model if the structure of the system under test is not known in advance.

\(^1\)A zero-memory nonlinear system, also called a static nonlinear system, is a nonlinear system, that does not exhibit any frequency dependency. In the time domain, that means that the output of the system at any time \( t_0 \) depends only on the input signal at the same time \( t_0 \) and the nonlinear function describing the system, and not on the history of the input signal at time \( t < t_0 \). In other words, it does not include memory effects.
2.2. NONLINEAR MODELS

The simple Hammerstein model (Fig. 2.2) consists of a zero-memory NLS followed by a linear system representing the memory of the whole system.

Figure 2.2: Block diagram of the simple Hammerstein model.

The simple Wiener model may be seen as a structurally-reversed Hammerstein system, that is, a linear system followed by a zero-memory NLS as shown in Fig. 2.3.

Figure 2.3: Block diagram of the simple Wiener model.

To illustrate the difficulty of selecting a suitable model and the importance of the positioning of the blocks, let us investigate a simple example, where the zero-memory NLS is a static power function. We consider two cases of NLS with block diagrams respectively represented in Fig. 2.4. The nonlinear static part is chosen here as a square function $x^2$. The linear part $g(t)$ is a simple linear filter with transient exponential impulse response. In the first example (simple Hammerstein model), the static NLS is followed by the linear part, whilst in the second example (simple Wiener model), the static NLS is preceded by it. As the position of both linear and non linear elements differs in the branch, the input-output description differs as well. To compare both cases, the Volterra description is used.

Figure 2.4: Simple Hammerstein (a) and simple Wiener (b) systems with polynomial static nonlinear parts.
Example 1: Simple Hammerstein system

The input-output relation of the system (Fig. 2.4a) is

\[ y_1(t) = \int_{-\infty}^{\infty} g(\tau)x_1^2(t - \tau)d\tau \]
\[ \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\tau_1, \tau_2)x_1(t - \tau_1)x_1(t - \tau_2)d\tau_1d\tau_2, \]  

(2.4)

where

\[ g_1(\tau_1, \tau_2) = \begin{cases} g^2(\tau), & \tau_1 = \tau_2 = \tau, \\ 0, & \text{if } \tau_1 \neq \tau_2. \end{cases} \]  

(2.5)

Example 2: Simple Wiener system

The input-output relation of the system (Fig. 2.4b) is

\[ y_2(t) = \left[ \int_{-\infty}^{\infty} g(\tau)x_2(t - \tau)d\tau \right]^2 \]
\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\tau_1)g(\tau_2)x_2(t - \tau_1)x_2(t - \tau_2)d\tau_1d\tau_2 \]
\[ \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_2(\tau_1, \tau_2)x_2(t - \tau_1)x_2(t - \tau_2)d\tau_1d\tau_2, \]  

(2.6)

where

\[ g_2(\tau_1, \tau_2) = g(\tau_1)g(\tau_2). \]  

(2.7)

The equations (2.5) and (2.7) represent the second order Volterra kernels that are different in both cases. Figure 2.5 shows the difference between these two examples.

Figure 2.5: The second order Volterra kernels: (a) kernel from the example of Fig. 2.4a; (b) kernel from the example of Fig. 2.4b.
2.2.3 Polynomial Hammerstein and Wiener Models

Polynomial Hammerstein and Wiener models are special cases of the simple Hammerstein and Wiener models in which the nonlinear parts are defined by a polynomial function. For the case of polynomial Hammerstein model, i.e., the nonlinear part is defined as (Janczak, A., 2005)

\[ u(t) = a_1 x(t) + a_2 x^2(t) + \cdots + a_N x^N(t). \]  

(2.8)

Thus, the polynomial Hammerstein and Wiener models can be represented by the block structure depicted in Fig. 2.6 and 2.7, in which the linear systems are all same, except a normalization factor \( a_n \).

To extend this special case to a more general one, we will use mutually independent linear filters in the framework of this thesis. In other words, we define the polynomial Hammerstein and Wiener models as the association of

---

Figure 2.6: Nth-order Polynomial Hammerstein model.

Figure 2.7: Nth-order Polynomial Wiener model.
several simple Hammerstein and Wiener models (Sec. 2.2.2) in parallel, called “branches”. Each branch has its own linear filter, and a zero-memory NLS of nth branch as a power series function $x^n$ (polynomial Hammerstein model (Fig. 2.6)) or $v^n$ (polynomial Wiener model (Fig. 2.7)).

### 2.2.4 Multiple-Input Single-Output Model

In (Bendat, J.S., 1998), it is stated that any NLS based on a nonlinear differential or integrodifferential equation can also be modeled in a MISO framework. This hypothesis is considered in the following of this thesis. The general Multiple-Input Single-Output (MISO) model is very similar to the polynomial Hammerstein one. Nevertheless, the polynomial Hammerstein model uses the power series as the nonlinear static input elements. The nonlinear parts of the MISO model can be any nonlinear function $F_n\{x\}$ of $x(t)$. A Nth-order MISO nonlinear model is depicted in Fig. 2.8, where $G_n(f)$ represents the linear system of the nth branch. The model is called MISO-based model (or MISO model), because knowing both the nonlinear functions $F_n\{x\}$ and the input $x(t)$, the outputs $x_n(t)$ of the nonlinear functions are also known and the system may be seen as a linear MISO system with correlated inputs (Bendat, J.S., 1998), as indicated in Fig. 2.8 by a dashed rectangle.

![Figure 2.8: Multiple-Input Single-Output model of nonlinear system with indicated linear part (dashed rectangle).](image)

---

2The polynomial Hammerstein model can then be seen as the particular case of a MISO model with power series as nonlinear static inputs elements (Boutayeb, M. et al., 1993)
2.3 Excitation Signals

The way of identifying NLS from input-output modeling may be split into three steps: exciting the system with a controlled input signal, recording the output signal and describing the system under test thanks to the relation between both the input and output signals. The choice of the input signal is then very crucial as both the output of the NLS under test and the subsequent quality of the identification depend on this input signal. Among the commonly used excitation signals which are summed up in this section, those specially used in this thesis are the white Gaussian noise (WGN) and the swept-sine signal.

2.3.1 White Gaussian Noise

A white Gaussian noise (WGN) \( x(t) \) is a zero-mean valued random process commonly used in signal processing. The WGN has the following properties (Bendat, J.S. & Piersol A.G., 1980). Its probability density function (PDF)
is written as
\[ p(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left( -\frac{x^2}{2\sigma_x^2} \right), \]  
where \( \sigma_x^2 \) is the variance of the WGN, and its cumulative probability density is
\[ P(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{x} \exp \left( -\frac{x'^2}{2\sigma_x^2} \right) dx', \]  
as illustrated in Fig. 2.9(a) and 2.9(b) respectively. The WGN is furthermore a stationary and ergodic process. Its auto-correlation function (ACF) \( R_{xx}(\tau) \) and its power-spectral density (PSD) \( S_{xx}(f) \) respectively express as
\[ R_{xx}(\tau) = \sigma_x^2 \delta(\tau) \]  
and
\[ S_{xx}(f) = \sigma_x^2. \]  

2.3.2 Pseudo-Random Signals

A pseudo-random signal is a discrete-time sequence exhibiting a binary amplitude law (in general made up of zeros and ones), with ACF and PSD very similar to those of a discrete-time WGN. One of the main advantages of such signals is the easy way for reproducing the same time-varying waveform, which consequently allows possible synchronous-time averaging. A pseudo-random binary signal (PRBS) is depicted in Fig. 2.10. This binary signal is generally used for the identification of linear systems and there were also some attempts to use it for the identification of NLS (Reed, M.J. & Hawksford, M.O., 1996b). Nevertheless, the use of PRBS is restricted in some special cases, due to the binary characteristics of the signal. A simple example of the unsuitability of the signal is any simple zero-memory system based on power series. Indeed, the powers of the sequence \( x(n) \) from the Fig. 2.10 takes only the values 1 and 0,
and thus, any power series, i.e. \( x^3 \), will not change the signal. This may lead to confusion whether the signal has passed through a NLS or not. For that reason, multi-levels pseudo-random signal have been proposed (Haber, R. & Keviczky, L., 1999), (Tan, A.H. & Godfrey, K.R., 2001) to be used for the identification of NLS.

### 2.3.3 Harmonic Signals

Even there is only a few publications dealing with the pure harmonic signals used in NLS identification (Albertson, F. & Gilbert, J., 2001), it is a matter of fact that they are the must used input signals in most of the experimental sciences. The reason is simple. For engineers or researchers concerned with (but not specialists in) NLS identification, the simplest method consists first in exciting the system under test with a pure harmonic signal, of which the complex form is,

\[
x(t) = A_0 \exp \left( j (2\pi f_0 t + \phi_0) \right),
\]

where \( A_0, f_0 \) and \( \phi_0 \) are respectively the amplitude, frequency and phase of the signal. Then, the response of the system to this pure harmonic signal is studied in the frequency domain and common parameters may be estimated, like the total harmonic distortion, which are only linked to the levels of the higher-order harmonics.

### 2.3.4 Multitone Signals

Multitone signal is a signal composed of several harmonic signals of different frequencies (not necessarily the harmonic multiples). Thus, the signal may be considered as a wide-band signal, which is necessary for an analysis with less iterations than when using a pure harmonic signal. The signal can be defined in several ways (Crama, P. & Schoukens, J., 2001), for example as

\[
x(t) = \sum_{k=0}^{N-1} B_k \exp(j(2\pi k f_0 t + \phi_k)),
\]

where \( B_k, k f_0 \) and \( \phi_k \) are respectively the amplitude, frequency and phase of the kth-harmonic contribution. If such a signal passes through an nth-order NLS, the output signal consists of two types of contributions: the nth-harmonics, corresponding to multiple of the fundamental frequencies, and the cross-terms or intermodulation harmonics (Evans, C. & Rees, D., 2000a). The
multitone signals are proposed for the identification of NLS in several papers, i.e. (Evans, C. & Rees, D., 2000b), (Solomou, M. et al., 2004), (Klippel, W., 1990).

2.3.5 Swept Sine Signals

A swept-sine signal, also called a chirp, is a signal of which the instantaneous frequency varies with time (Cohen, L., 1995). It is commonly designed and used in many applied domains, such as sonar and radar, but also in communications and acoustics. A swept-sine signal can be generally, but not unambiguously, defined as

\[ s(t) = a(t) \sin(\varphi(t)), \]  

(2.15)

where we expect that \( a(t) \) is the amplitude envelope and \( \varphi(t) \) is the phase of the chirp. A swept-sine signal exhibiting a linearly time increasing instantaneous frequency is depicted in Fig. 2.11. On one hand, the analytic signal \( z_s(t) \) of the signal \( s(t) \) is written as (Cohen, L., 1995)

\[ z_s(t) = s(t) + jH[s(t)] = a_s(t)e^{j\varphi_s(t)}, \]  

(2.16)

where \( H[\cdot] \) is the Hilbert transform, and where \( a_s(t) \) and \( \varphi_s(t) \) are unambiguously defined as instantaneous amplitude and phase of \( z_s(t) \). On the other hand, the spectrum \( Z_s(f) \) of the signal \( z_s(t) \) can be written in terms of amplitude \( B_s(f) \) and phase \( \Psi_s(f) \) as

\[ Z_s(f) = B_s(f)e^{j\Psi_s(f)}. \]  

(2.17)

The instantaneous frequency \( f_i(t) \) and the group delay \( t_f(t) \) are then defined as

\[ f_i(t) = \frac{1}{2\pi} \frac{d\varphi_s(t)}{dt}, \]  

(2.18)

\[ t_f(f) = -\frac{1}{2\pi} \frac{d\Psi_s(f)}{df}. \]  

(2.19)

Eq. (2.18) and (2.19) define two curves in the time-frequency plane which may be regarded as inverse of each other, only for asymptotic signals (Cohen, L., 1995), (Flandrin, P., 1999). In such conditions of asymptotic signal, if one of these functions is known, the other one can then be easily calculated.

These properties allow to calculate the spectra of the signal \( z_s(t) \) with no need to calculate its Fourier transform. The amplitude \( a_s(t) \) and the phase \( \varphi_s(t) \) in
2.4 Methods of Identification

Given a suitable model and a suitable excitation signal, the final step is to select a suitable method and an optimal associated algorithm. A wide variety of methods have been developed, based on different models and excitation signals. The identification method is often inherently linked to the NLS model and the excitation signal. For a given model, there exist several methods based on several excitation signals. In this section, two methods further used in this thesis are described in detail: the MISO method and the nonlinear convolution method. The basic methods for estimation of Volterra model, Wiener and Hammerstein models are also mentioned. There exists of course many other methods such as neural networks (Nelles, O., 2001), hybrid genetic algorithm (Yen-Wei Chen et al., 2003), extended Kalman filtering (Sorenson, H.W., 1985), particle filtering (Arulampalam, M.S. et al., 2002), which are beyond the scope of this thesis work.

\[ B_s(f) = \frac{a_s(t_f)}{\sqrt{\frac{1}{2\pi} |\varphi'_s(t_f)|}}, \]
\[ \Psi_s(f) = \varphi_s(t_f) - 2\pi f t_f + \frac{\pi}{4} \text{sign} \left( \frac{df_i(t_f)}{dt_f} \right). \]
2.4.1 Methods for the Volterra Model

There exists many methods for identification of NLS by the Volterra series. One of the first basic methods is the method based on weighted Dirac impulses (Schetzen, M., 1989), where the Volterra kernels are estimated recursively from the first kernel to the higher one. Other methods have also been proposed, such as the maximum length sequences based method (Reed, M.J. & Hawksford, M.O., 1996b), the periodic signal based method (Evans, C. et al., 1996), the least squares method (Westwick, D.T. & Kearney, R.E., 1998), or the third moment based method (Hong-Zhou Tan & Aboulnasr, T., 2006). All these methods are generally not used in practice because of the large number of parameters required to represent the higher-order Volterra kernels.

2.4.2 Methods for the Wiener and Hammerstein Models

The identification methods for Simple Wiener and Hammerstein systems depend upon whether the model is parametric or not. If the model is parametric, the zero-memory nonlinear part is represented by a polynomial function, or by general orthogonal series, and thus by a finite number of parameters (Billings, S.A. & Fakhouri, S.Y., 1979). The parameters to be estimated are those of the zero-memory part and those of the linear part of the model. The original method uses an iterative scheme to determine both, the linear and zero-memory nonlinear parts separately (Chang, F. & Luus, R., 1971). More recent algorithms are based on estimation of general NARMAX models (Chen, S. & Billings, S, 1989). If the model is nonparametric, the estimation consists in estimation of finished point of the nonlinear function that can further serve for polynomial representation (Greblicki, W. & Pawlak, M., 1989), (Greblicki, W., 1997), (Greblicki, W., 2004).

2.4.3 Method for the Multiple-Input Single-Output Model

The identification methods based on MISO model depend upon whether the model is parametric or not. Some parametric methods can be found in literature (Boutayeb, M. et al., 1993), (Kortmann, M. & Unbehauen, M., 1987). The method proposed by (Rice, H.J. & Fitzpatrick, J.A., 1988) for general nonparametric model is based on power- and cross-spectral density estimations. The excitation signal $x(t)$ for the analysis is a WGN (Rice, H.J.
The nonlinear MISO model consists of a linear system $G_1(f)$ in parallel with (N-1) nonlinear branches, each branch being represented by a zero-memory NLS $\mathcal{F}_n$, followed by a linear system $G_n(f)$, as shown in Figure 2.8. The outputs of the zero-memory NLS $\mathcal{F}_n$ are noted as $x_n(t) = \mathcal{F}_n\{x(t)\}$, for $n \in [2, N]$ and they are the set of inputs of the linear MISO system (see Fig. 2.12). As the input $x(t) = x_1(t)$ is known, the other inputs $x_2(t), \ldots, x_N(t)$ of the NLS are also known. Then, once the output signal $y(t)$ is measured, the aim of the signal processing is to estimate the linear filters $G_n(f)$ of this MISO linear system with correlated inputs. The inputs $x_n(t)$ may be chosen according to the knowledge of the NLS to be studied. If the analyzed NLS is not known, a blind identification based on a power series can be used.

The block diagram of Fig. 2.12 illustrates the problem of linear filters $G_n(f)$ identification for a MISO linear system with correlated inputs. As usually done for MISO linear systems, the first step for the estimation of $G_n(f)$ consists in decorrelating the inputs. The correlation between two signals can be expressed as a linear dependency (linear filtering). This can be seen in Fig. 2.13, for a case of three inputs and one output linear system, where the three inputs are mutually correlated. The signal $x_j(t)$ is made of two contributions: (a) the signal $x_{j,i}(t)$, which is a pure linear filtering of the signal $x_i(t)$, and (b) the signal $x_{j,i}(t)$, which is uncorrelated with $x_i(t)$ (BENDAT, J.S., 1998). When the signal $x_{j,i}(t)$ is also decorrelated with all the signals $x_{j,m}(t)$, where $m < i$, it is noted $x_{j,i}(t)$. Then,
in the frequency domain, the correlation between $x_i(t)$ and $x_j(t)$ is expressed by means of a linear filter $L_{ij}(f)$ and is defined by (Bendat, J.S. & Piersol A.G., 1980)

$$L_{ij}(f) = \frac{S_{i,j(i-1)}(f)}{S_{i,i(i-1)}(f)}, \quad (2.22)$$

where $S_{i,i(i-1)}(f)$ and $S_{i,j(i-1)}(f)$ are power spectral densities (PSD) and cross spectral densities (CSD) of inputs $x_i(t)$, $x_{i(i-1)}(t)$ and $x_i(t)$, $x_{j(i-1)}(t)$ respectively.

The general method to solve the MISO linear system is based on power and cross spectral density estimations. From the knowledge of $x_1(t), x_2(t), \ldots, x_N(t)$, the Multiple-Input Single-Output system is defined and the linear filters $G_1(f), G_2(f), \ldots, G_N(f)$ can be estimated.

As the input signal $x_1(t) = x(t)$ is a WGN, the other inputs $x_2(t), x_3(t), \ldots$ are instantaneous nonlinear functions of $x_1(t)$. These input signals may be correlated and thus all the cross spectral density functions for all the combinations of inputs have to be taken into consideration. The cross spectral densities between input signal $x_1(t)$ and output signal $y(t)$ is

$$S_{1y}(f) = S_{11}(f)G_1(f) + S_{12}(f)G_2(f) + \cdots + S_{1N}(f)G_N(f), \quad (2.23)$$

hence, for $i = 1, 2, \ldots, N$ we have

$$S_{iy}(f) = \sum_{j=1}^{N} S_{ij}(f)G_j(f). \quad (2.24)$$
This set of Eq. (2.24) can be rewritten into a matrix form, as

\[
\begin{pmatrix}
S_{1y}(f) \\
S_{2y}(f) \\
S_{3y}(f) \\
\vdots \\
S_{Ny}(f)
\end{pmatrix} =
\begin{pmatrix}
S_{11}(f) & S_{12}(f) & S_{13}(f) & \cdots & S_{1N}(f) \\
S_{21}(f) & S_{22}(f) & S_{23}(f) & \cdots & S_{2N}(f) \\
S_{31}(f) & S_{32}(f) & S_{33}(f) & \cdots & S_{3N}(f) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
S_{N1}(f) & S_{N2}(f) & S_{N3}(f) & \cdots & S_{NN}(f)
\end{pmatrix}
\begin{pmatrix}
G_1(f) \\
G_2(f) \\
G_3(f) \\
\vdots \\
G_N(f)
\end{pmatrix},
\tag{2.25}
\]

or equivalently,

\[
S_y = S \cdot G.
\tag{2.26}
\]

Lastly, the frequency response function (FRF) \( G_n(f) \) can be estimated by solving

\[
G = S^{-1} \cdot S_y,
\tag{2.27}
\]

where \([-]^{-1}\) means the inverse of a matrix.

This method is specially used in chapter 3 of this thesis.

### 2.4.4 Nonlinear Convolution Based Method

The last method, called "nonlinear convolution method", presented in this state-of-the-art is a method generally used for analysis of NLS, but not based on any specific model representation. The nonlinear convolution method presented in (Armelloni, E. & Farina, A., 2001) uses a swept sine signal (also called a chirp), exhibiting an exponential instantaneous frequency, as the excitation

![Nonlinear Convolution Process](image)

Figure 2.14: Block diagram of the nonlinear convolution process in NLS identification.
signal and allows the characterization of a NLS in terms of harmonic distortion at several orders.

The block diagram of the method is shown in Fig. 2.14. First, an exponential swept sine signal \( s(t) \) is generated and used as the input signal of the NLS under test. The distorted output signal \( y(t) \) is recorded for being used for the so-called nonlinear convolution (Armelloni, E. & Farina, A., 2001). Next, the signal noted \( \tilde{s}(t) \) is derived from the input signal \( s(t) \) as its time-reversed replica with amplitude modulation such that the convolution between \( s(t) \) and \( \tilde{s}(t) \) gives a Dirac delta function \( \delta(t) \). The signal \( \tilde{s}(t) \) is called ”inverse filter” (Farina, A., 2000; Armelloni, E. & Farina, A., 2001). Then, the convolution between the output signal \( y(t) \) and the inverse filter \( \tilde{s}(t) \) is performed. The result of this convolution can be expressed as

\[
y(t) \ast \tilde{s}(t) = \sum_{m=1}^{\infty} h_m(t + \Delta t_m),
\]

(2.28)

where \( h_m(t) \) are higher-order nonlinear impulse responses and \( \Delta t_m \) are the time lag between the first (linear) and the mth-impulse response. Since the nonlinear impulse response consists of a set of higher-order nonlinear impulse responses that are time shifted, each partial impulse response can be separated from each other, as illustrated in Fig. 2.15. This procedure is developed in (Farina, A., 2000).

The set of higher-order nonlinear impulse responses \( h_m(t) \) can be also expressed
in the frequency domain. The FRFs are then the Fourier transforms (FT) of higher-order nonlinear impulse responses $h_m(t)$

$$H_m(f) = FT[h_m(t)].$$

(2.29)

The frequency responses $H_m(f)$ are called higher-order nonlinear frequency responses and represent the frequency dependency of higher-order harmonic components. The frequency response $H_1(f)$ is consequently the response corresponding to the linear part of the system. Similarly, the frequency response $H_m(f)$ ($m > 1$) may be regarded as the system frequency response, when considering only the effect of the input frequency $f$ on the $m$th-harmonic frequency $mf$ of the output. The partial frequency responses $H_n(f)$ are depicted in Fig. 2.16, for an arbitrary case where $|H_n(f)|$ is supposed to be a constant in the frequency band $[nf_1, f_2]$.

In the framework of this thesis, this method is improved in order to extended the method for estimation of a nonlinear model and identification of NLS.

![Figure 2.16: Result of the nonlinear convolution process in the form of higher-order nonlinear frequency responses $H_n(f)$. Arbitrary case where $|H_n(f)|$ is a constant in $[nf_1, f_2]$.](image)
2.5 Summary

As no general description is available for all nonlinear systems, a very wide variety of models, excitation signals and identification methods is accessible in literature. In this chapter, several of them have been presented in detail, especially those used further in this thesis, which are respectively:

- for the models, the polynomial Hammerstein model (2.2.2), and the MISO model (2.2.4)
- for the excitation signals, the WGN (2.3.1) and the swept-sine signals (2.3.5),
- for the methods of identification, the methods based on MISO models (2.4.3) and the nonlinear convolution (2.4.4).

In the next two chapters, two methods described in this chapter are developed. The first one (chapter 3) uses a WGN excitation and is based on the MISO method. The second one (chapter 4) is based on a swept-sine signal and the nonlinear convolution method. Both methods (chapter 3 and 4) uses the MISO model or polynomial Hammerstein model.
In chapter 2, a state of the art of models, excitation signals and methods used for the identification of nonlinear system (NLS) has been presented, and the tools further used in this thesis have been detailed.

A nonlinear model with several parallel branches (nonlinear MISO model of Fig. 2.8), or more exactly its particular case using power series expansion, and corresponding to a polynomial Hammerstein model (Fig. 3.1), seems to be an optimal compromise between the simplicity of the model and its accuracy. In this chapter, we develop an algorithm for the identification of NLS using a power series expansion based on Multiple-Input Single-Output (MISO) nonlinear model with white Gaussian noise (WGN) as excitation signal.

The classical nonlinear MISO method (Rice, H.J. & Fitzpatrick, J.A., 1988) for NLS identification has been shortly reviewed in Sec. 2.4.3. This method is based either on the estimation of all the power and cross spectral densities (PSD, CSD) of inputs and outputs of the model, in the case of mutually correlated inputs, or just only on the estimation of the CSD between output and inputs if the inputs are decorrelated. In the second case, the inputs decorrelation is generally based on the estimation of the CSD between these inputs.
It is shown in this chapter that the CSD between the inputs can be expressed in a mathematical way, in case of power series based model. The CSD of all inputs are defined as Fourier transform of the cross correlation functions (CCF), whose expressions are presented in (Bendat, J.S., 1998), when using the power series expansion (e.g. $x^3$ and $x^5$), $x(t)$ being a WGN. The process of estimation of decorrelated linear MISO model is well known, i.e. (Bendat, J.S. & Piersol A.G., 1980). The procedure is thus based on two steps: (a) the decorrelation based on the CSD between the inputs, and (b) the estimation of decorrelated linear MISO model, both known in literature for several years. To our knowledge, the method based on mathematical decorrelation process has not been investigated in the literature. It is detailed here.

In this chapter, we first present the CCF between the inputs of power series based model. Then, we present the decorrelation algorithm and the process of estimation of linear filters of the MISO model. Next, an extension for this method called zooming effect is presented. It is based on standard deviation and mean value pre-set. Changing these properties of the excitation WGN, the estimation can be focused on a part of the input-output characteristics. Finally, several simulated examples are presented in order to test the method.

![Figure 3.1: The polynomial Hammerstein model.](image-url)
3.1 Correlation and Power Spectral Density Functions

In this section, we deal with the correlation and spectral density functions between the inputs of the power series based MISO model. The relations are used in the next section to derive the decorrelated inputs in a mathematical way.

Let the zero-memory nonlinear inputs $x_n(t)$ of the MISO model be chosen according to the power series (Fig. 3.1). Then, as such operators are zero-memory nonlinear systems, the linear filters $L_{ij}(f)$ (Fig. 3.2), from the Eq. (2.22), are not frequency-dependent and their constant values are known according to mathematical considerations shown in this section.

Let $x(t)$ be a WGN, with the autocorrelation function (ACF) $R_{xx}(\tau)$ defined as

$$R_{xx}(\tau) = E[x(t)x(t-\tau)] = \sigma_x^2 \delta(\tau), \quad (3.1)$$

where $E[\cdot]$ denotes an ensemble average, $\sigma_x^2$ the variance of $x(t)$ and $\delta(\tau)$ the Dirac function. The inputs of the power based MISO model (Fig. 3.1) are such that

$$x_n(t) = F_n\{x(t)\} = x^n(t), \quad (3.2)$$

for $n \in [1, N]$. Then, the CCF $R_{x_nx_{n'}}(\tau)$ between two inputs $x_{n_1}$ and $x_{n_2}$ may be written as (PAPOLIS, A. & PILLAI, S.U., 1991)

$$R_{x_nx_{n'}}(\tau) = \begin{cases} E[x_{n_1}x_{n_2}] = E[x_{n_1+n_2}] & \text{if } \tau = 0, \\ E[x_{n_1}]E[x_{n_2}] & \text{if } \tau \neq 0. \end{cases} \quad (3.3)$$

Figure 3.2: Three-Input Single-Output block diagram with linear filters $L_{ij}$ representing the mutual correlation of inputs.
Three cases have then to be studied according to the values of $n_1$ and $n_2$.

- Firstly, if $n_1 + n_2$ is odd ($n_1 + n_2 = 2l + 1$, with $l \in \mathbb{Z}$), the CCF $R_{x_{n_1}x_{n_2}}(\tau)$ is equal to zero as $E[x^{2l+1}] = 0$. In other words, odd powers of WGN signals are not correlated with their even powers.

- Secondly, if both $n_1$ and $n_2$ are odd, both $E[x^{n_1}]$ and $E[x^{n_2}]$ are equal to zero and $R_{x_{n_1}x_{n_2}}(\tau)$ is represented by the Dirac function weighted by $E[x^{n_1+n_2}]$.

- Lastly, if $n_1$ and $n_2$ are even, then both $E[x^{n_1}]$ and $E[x^{n_2}]$ are nonzero, their product is also nonzero and the CCF $R_{x_{n_1}x_{n_2}}(\tau)$ includes a continuous component besides the weighted Dirac function.

This can then be rewritten as

$$\begin{align*}
R_{x_{n_1}x_{n_2}}(\tau) &= \begin{cases} 
C_1 \delta(\tau) & \text{for } n_1 \text{ and } n_2 \text{ odd, } \\
C_2 \delta(\tau) + D & \text{for } n_1 \text{ and } n_2 \text{ even, } \\
0 & \text{for } n_1 + n_2 \text{ odd. }
\end{cases} 
\end{align*}$$

(3.4)

where $C_1$, $C_2$ and $D$ are constants, that can be derived from (3.3) and from the ensemble averages $E[x^{2l}]$ and $E[x^{2l-1}]$ expressed as (BENDAT, J.S., 1998)

$$E[x^{2n}] = \int_{-\infty}^{\infty} x^{2n} p(x) dx = \int_{-\infty}^{\infty} x^{2n} \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left(-\frac{x^2}{2\sigma_x^2} \right) dx = \frac{(2n)!}{2^n n!} \sigma_x^{-2n}, \quad (3.5)$$

$$E[x^{2n-1}] = \int_{-\infty}^{\infty} x^{2n-1} p(x) dx = \int_{-\infty}^{\infty} x^{2n-1} \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left(-\frac{x^2}{2\sigma_x^2} \right) dx = 0. \quad (3.6)$$

Then, the CSD $S_{x_{n_1}x_{n_2}}(f)$ is given by

$$S_{x_{n_1}x_{n_2}}(f) = \begin{cases} 
C_1 & \text{for } n_1 \text{ and } n_2 \text{ odd, } \\
C_2 + D\delta(f) & \text{for } n_1 \text{ and } n_2 \text{ even, } \\
0 & \text{for } n_1 + n_2 \text{ odd. }
\end{cases} \quad (3.7)$$

In this section, the expressions of the cross correlation functions and cross spectral densities between all the input signals of the power series based MISO model have been detailed. In the following section, the process of decorrelation of the inputs is achieved.
3.2 Expression of Decorrelated Inputs

In order to calculate the set of coefficients allowing the complete decorrelation of power series inputs, the following steps are proceeded. First, the signal $x_j(t)$ is expressed by two contributions: (a) the signal $x_{j,i}(t)$, which is a linear filtering of the signal $x_i(t)$, and (b) the signal $x_{j,i}(t)$, which is uncorrelated with $x_i(t)$ (Bendat, J.S., 1998). When the signal $x_{j,i}(t)$ is also decorrelated with all the signals $x_{j,m}(t)$, where $m < i$, it is noted $x_{j,ii}(t)$.

Then, as shown in previous section, the odd and even powers are not mutually correlated, and thus the process of decorrelation may be divided into two parts: (a) the decorrelation of only odd power series inputs and (b) the decorrelation of only even power series inputs. The whole processus of decorrelation can be expressed by the following set of equations

$$\begin{align*}
x_1(t) &= x(t), \\
x_{3,2i}(t) &= x^3(t) + k_{13}x(t), \\
x_{5,4i}(t) &= x^5(t) + k_{35}x^3(t) + k_{15}x(t), \\
&\vdots
\end{align*}$$

(3.8)

that, in general, allows to write the decorrelating input $x_{n, (n-1)i}(t)$ into the following form

$$x_{n, (n-1)i}(t) = \sum_{i=1}^{n} k_{in}x^i(t) + \mu_n,$$

(3.9)

where the coefficients $\mu_n = 0$ for odd $n$ and $\mu_n = E[x_{n, (n-1)i}]$ otherwise. The coefficients $k_{nm}$ can be derived also recursively as

$$k_{nm} = -\sum_{i=n}^{m-1} k_{ni} \frac{E[x^{i+m}]}{E[x^{2i}]},$$

(3.10)

The notation $j:ii$ means that the signal the signal $x_{j,ii}(t)$ is decorrelated with all the signals $x_{j,m}(t)$, where $m < i$ (see chapter 2)
where $E[\cdot]$ can be obtained using Eq. (3.6-3.7) and $k_{nn} = 1$. Note that $k_{nm} = 0$ for odd $(n + m)$ due to the uncorrelated odd and even power series. The process of decorrelation is depicted in Fig. 3.3, where odd and even powers are separated, as they are not mutually correlated.
3.3 Estimation of Filters of the MISO model

In previous section, it has been shown how to decorrelate the inputs of the power series based nonlinear MISO model. The last step in the system identification is the estimation of the filters $G_n(f)$ from Fig. 3.1. Since the inputs $x_n(t)$ can be easily decorrelated into new inputs $x_{n(n-1)!}(t)$, the MISO system with correlated inputs is equivalent to the MISO system with decorrelated inputs depicted in Fig. 3.4 and can be solved as a classical MISO linear system. First of all, the CSD $S_{x_{n(n-1)!}y}$ and $S_{x_{n(n-1)!}x_{n(n-1)!}}$ are calculated. Next, the linear filters $L_{ny}(f)$ from the decorrelated model (Fig. 3.4) are given by

$$L_{ny}(f) = \frac{S_{x_{n(n-1)!}y}(f)}{S_{x_{n(n-1)!}x_{n(n-1)!}}(f)}, \quad (3.11)$$

for $n \in [1, N]$. Finally, the linear filters $G_n(f)$ can be derived from the decorrelated model (BENDAT, J.S. & PIERSOL A.G., 1980), using the Figs. 3.1, 3.3 and 3.4 (for $n \in [1, N-1]$) as

$$G_n(f) = L_{ny}(f) + \sum_{j=n+1}^{N} k_{nj} L_{jy}(f) + \mu_n \delta(f), \quad (3.12)$$

where $\mu_n \delta(f)$ is the Fourier transform of the constant $\mu_n$. This relation can be rewritten into a matrix form

$$
\begin{pmatrix}
G_1(f) \\
G_2(f) \\
G_3(f) \\
\vdots \\
G_N(f)
\end{pmatrix} = 
\begin{pmatrix}
1 & k_{12} & k_{13} & \cdots & k_{1N} \\
0 & 1 & k_{23} & \cdots & k_{2N} \\
0 & 0 & 1 & \cdots & k_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix} 
\begin{pmatrix}
L_{1y}(f) \\
L_{2y}(f) \\
L_{3y}(f) \\
\vdots \\
L_{Ny}(f)
\end{pmatrix} + 
\begin{pmatrix}
\mu_1 \\
\mu_2 \\
\mu_3 \\
\vdots \\
\mu_N
\end{pmatrix} \delta(f). \quad (3.13)
$$

![Figure 3.4: Multiple-Input Single-Output model for conditioned uncorrelated inputs.](image-url)
To sum up, the method developed in this chapter estimates the power series nonlinear MISO model (equivalent to the polynomial Hammerstein model). The standard approach for identifying a NLS with the MISO method is based on approximate knowledge of the shape of the nonlinear input/output characteristic and then on the estimation of PSD and CSD as the inputs of the model are mutually correlated. The method presented in this chapter is linked to blind identification of the NLS using a power series expansion. Thus, the estimation of all the mutual CSD is not necessary and the decorrelation can be made in mathematical way. Once the decorrelated inputs are known, only the CSD between each input and the output have to be estimated.

In following sections, a new extension of this method, called "zooming effect", is presented. This extension allows to focus only on a part of the nonlinear input-output characteristics by changing the properties of the excitation white noise signal.
3.4 Zooming Effect

The so-called "zooming effect" is an extension of the method described in the first part of this chapter, which allows to analyze the nonlinear Input-Output (I/O) characteristic part by part. The method is close to the common piecewise linear approximation method (Storace, M. & De Feo, O., 2004), where the I/O characteristic is studied by dividing it into several linear parts. Regarding the zooming effect method, the I/O characteristic is also studied by dividing it into non necessarily linear parts. The advantage of the zooming effect method is that the system with unknown I/O characteristic can be identified locally in a given I/O area. Consequently, even highly complicated NLS can be locally identified using the controlled order of a power series expansion.

The excitation signal $x(t)$ is a random process characterized by its standard deviation $\sigma_x$ and its mean value $\mu_x$. Using a different standard deviation $\sigma_x$ of the input signal $x(t)$ allows the analysis of different amplitude areas in the I/O characteristic. The lower the standard deviation $\sigma_x$, the better the accuracy of estimated I/O characteristic near the mean value. Increasing $\sigma_x$ consequently allows to increase the amplitude area. The second way to change the amplitude area of interest is to shift the input signal by adding a constant value. Changing

\[ p(x) \]

\[ y \]

\[ x \]

\[ -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \]

\[ (a) \]

\[ y \]

\[ x \]

\[ -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \]

\[ (b) \]

Figure 3.5: Pre-set of standard deviation of input signal for nonlinear function measuring. (a) large $\sigma_x$ with no offset, (b) weak $\sigma_x$ with offset equal to 1. Nonlinear function $y = F\{x\}$ - above, probability density functions of input signal $x(t)$ -below.
the values of $\sigma_x$ and $\mu_x$ acts then as a zoom. As an example, Figure 3.5 illustrates
the zooming effect for a memory-less NLS. In that case, two measurements with
different $\sigma_x$ and $\mu_x$ may be set up. For the measurement corresponding to
$\mu_x \neq 0$, and with a smaller value of $\sigma_x$, each part of I/O nonlinear characteristic
is analyzed with better localization. Using this procedure therefore allows to
estimate the unknown I/O characteristic with controlled accuracy.

The zooming effect is based on the use of signals with a mean value $\mu_x$ different
from zero. In order to respect the assumptions of zero-mean valued input signal
$x(t)$, the following steps are required. The input signal is shifted by adding a
constant value $x(t) = x_c(t) + \mu_x$, where the $x_c(t)$ is a zero-mean valued (centered
signal). The signal $x(t)$ is used for the measurement of the NLS. The output
signal can be split in the same way to its centered part $y_c(t)$ and its mean value
$\mu_y$. The centered signals are analyzed by the MISO method in order to obtain a
nonlinear model. The values $\mu_x$ and $\mu_y$ are then added to the results to shift the
I/O characteristic to the right I/O center point. The mean value shift is depicted
in Fig. 3.6.

Figure 3.6: Input-Output characteristics with mean value shift.
3.5 Simulation of Nonlinear Systems

3.5.1 Memoryless Nonlinear Systems

In order to illustrate the MISO method with the zooming effect, two memoryless NLS are analyzed in this section: a limiter and a system modeling a mouthpiece of a wind instrument. In both cases, the nonlinear equations governing the behavior of a NLS are known.

A limiter simulation

A limiter is defined by the memoryless I/O relation

\[ y = \begin{cases} 
-A & x \leq -A, \\
\mu & |x| \leq A, \\
A & x \geq A,
\end{cases} \tag{3.14} \]

where \( A \) is a given controlled parameter. The principle of the estimation of the NLS is based on the MISO method with power series, developed in this chapter. In the case of the memoryless system, the estimated PSD and CSD are supposed to be constants as well as the filters \( G_n(f) \) (as there is no memory effect). In order to estimate the I/O characteristic, a harmonic signal is put at the input of the estimated nonlinear model and is compared with corresponding reconstructed output.

The estimation of I/O characteristic of this NLS using the zooming effect may be realized as follows. In the first step, the NLS is estimated in the full amplitude range of interest. Then, the second step consists of selecting a given area of the I/O characteristic through changes of \( \sigma_x \) and \( \mu_x \). More accurate estimation can then be done with the same number of power series input signals.

The estimation of the I/O characteristic of a limiter is shown in Fig. 3.7, where the theoretical shape is in dots and dashes. The first identification step (Fig. 3.7(a)) has been set up using a seventh-order power series, and a zero-mean valued and full-amplitude area input signal (\( 6\sigma_x = 6 \)). Even if the shape of the I/O characteristic is properly estimated, errors can occur, especially near the discontinuities of I/O characteristic. For the second step, the amplitude area of interest has been narrowed down and shifted, using \( \sigma_x = 0.1 \) and \( \mu_x = 1 \). The
Figure 3.7: A limiter identification - comparison between full amplitude range with zero mean value (a) and a narrowed amplitude range with a nonzero mean value in order to zoom the nonlinear discontinuity (b). The result of this simulation is shown in Fig. 3.7(b) (a continuous line), where the I/O characteristic estimated during the first step (a dashed line) has been added. The accuracy of the zoomed I/O characteristic is much better. The error at the point of discontinuity is indeed about 10% without the zoom and less than 1% with the zoom.

Simulation of single-reed instrument

The mouthpiece of a wind instrument represents a strongly nonlinear behavior between an input signal $x \equiv \Delta p$ and an output signal $y \equiv U$. Backus, J. (1963) gives the nonlinear relation between the volume flow $U$ through the reed, and the pressure difference $\Delta p$ between the mouth pressure $P_M$ and the pressure in the mouthpiece $p$, as

$$U = \begin{cases} \text{sign}(\Delta p)wH_0 \left(1 - \frac{\Delta p}{P_M}\right) \sqrt{\frac{2|\Delta p|}{\rho}} & \text{if } \Delta p \leq P_M, \\ 0 & \text{if } \Delta p > P_M. \end{cases}$$

(3.15)

The memoryless nonlinear I/O characteristic of a single reed instrument has then been simulated with the following values: $\omega H_0 = 1$, $P_M = 2$ and $\rho = 1/3$, leading to the following NLS

$$y = \begin{cases} \text{sign}(x) \left(1 - \frac{x}{2}\right) \sqrt{6|x|} & \text{if } x \leq 2, \\ 0 & \text{if } x > 2. \end{cases}$$

(3.16)

as illustrated in Figure 3.8 for $x \equiv \Delta p$ and $y \equiv U$. If the pressure difference between the mouth pressure and the pressure in the mouthpiece is larger than a
3.5. SIMULATION OF NONLINEAR SYSTEMS

The first step of identification was performed using a seventh-order power series of WGN $x$, with mean value $\mu_x = 0$ and $\sigma_x = 1$, corresponding to a full area of interest ($6\sigma_x = 6$). The reconstruction of the I/O characteristic of the mouthpiece using the seventh-order power series is shown in Fig. 3.9(a) (the theoretical curve is dashed). Even if the global shape is properly estimated, the reconstruction of the I/O characteristic based on the power series reconstruction of order seven is not sufficiently accurate for this nonlinearity. Instead of increasing the order of the power series, the zooming effect has been applied. Firstly, the region of negative values of input signal has been estimated. As the I/O characteristic in this amplitude region are almost linear, the result of the estimation and the reconstruction shown in Fig. 3.9(b) fit very well with the I/O characteristic.

The nonlinearity around zero is difficult to model with a power series as seen in Fig. 3.9(c), because of the type of nonlinearity in $|x|^{1/2}$. Fig. 3.9(d) and 3.9(e) show the differences in choosing the region of input amplitude as $[0, 1]$ or $[0, 2]$. The other figures, Fig. 3.9(f-h), illustrate the results of locally analysis in different amplitude areas of interest.

Fig. 3.10 shows the set of local analysis, which forms a complete estimation of I/O characteristic. We recall that all local analysis was set up with a seventh-order power series. Thanks to the zooming effect (Fig. 3.10), the I/O characteristic are more accurately estimated than without it. The relative associated mean-square errors are $4 \cdot 10^{-3}$ and $1.7 \cdot 10^{-4}$, respectively. The discontinuities in Fig. 3.10 caused by the part by part analysis can be smoothed.
Figure 3.9: Identification of the clarinet nonlinearity - a) estimation with full amplitude range $\sigma_x = 2/3$ and $\mu_x = 1$; b) $\sigma_x = 1/3$ and $\mu_x = -1$; c) $\sigma_x = 1/15$ and $\mu_x = 0$; d) $\sigma_x = 1/6$ and $\mu_x = 0.5$; e) $\sigma_x = 1/3$ and $\mu_x = 1$; f) $\sigma_x = 1/15$ and $\mu_x = 0.2$; g) $\sigma_x = 1/3$ and $\mu_x = 2$; g) $\sigma_x = 1/2$ and $\mu_x = 3$. 
3.5. SIMULATION OF NONLINEAR SYSTEMS

3.5.2 Nonlinear System with Memory

In order to illustrate the effects of frequency dependence on the NLS identification, a NLS with memory is simulated. As seen in Fig. 3.11, the nonlinear first part of the system is the limiter, studied in previous section, and the linear second part is a first-order IIR low-pass digital filter defined as

\[ H(z) = \frac{1 - a}{1 - az^{-1}}, \quad (3.17) \]

where \( a = 0.9 \) and where \( z = re^{j2\pi f/fs} \), \( f_s \) being the sampled frequency and \( r \in \mathbb{R} \). The region of convergence of \( H(z) \) is \( ]a, \infty[ \) and consequently includes the unit circle \( (r = 1) \), where the Discrete Fourier Transform (DFT) may be calculated (OPPENHEIM, A.V. & SCHAFER, R.W., 1989), according to \( H(f) = H(z = e^{j2\pi f/fs}) \). The frequency response of the filter is shown in the Figure 3.12.

For estimating CSD and PSD, the Welch periodogram is set up, whose parameters are: overlap = 50%, window length = 1024 samples and a Hamming window (WELCH, P.D., 1970). According to the amplitude range of the limiter, two estimations are presented. The first one corresponds to the linear part of the NLS, the values of the input signal being \( \mu = 0 \) and \( \sigma = 0.3 \). For this case, the values of the input signal does not exceed the saturation bound of the limiter.

![Figure 3.11: Tested nonlinear system with memory (a limiter followed by a linear low-pass IIR filter).](image-url)
and the filter $H_1(f)$ is supposed to be equal to the theoretical filter $H(f)$. The frequency characteristic of the estimated filter $H_1(f)$ is shown in Figure 3.13. The relative squared error of the modulus $(|H(f)| - |H_1(f)|)^2/|H(f)|^2$ and the phase error $\Theta_H(f) - \Theta_{H_1}(f)$ are proposed in Figure 3.14. The relative squared error does not exceed 0.8%, and the phase error is less than $\pm 2^\circ$.

In the second case, the zooming method is applied near the discontinuous part of the limiter. The input signal is shifted by a mean value $\mu = 1$ and the amplitude range variation is $\sigma = 0.3$. Once the filters $H_i(f)$ are estimated, a reconstruction of the output signals is done for two different frequencies in order to assess how the zooming method behaves in the frequency domain. The two reconstruction signals are shown in Figure 3.15.

Whatever the value of the input frequency is (limited by the half the sampling frequency divided by the order of power series), the MISO method associated with the zooming effect allows the time reconstruction of output signals with a relative error less than 2%.
3.5. SIMULATION OF NONLINEAR SYSTEMS

Figure 3.13: Modulus (above) and phase (below) of the estimated filter $H_1(f)$ in the case of the input signal with parameters $\mu = 0$ and $\sigma = 0.3$.

Figure 3.14: The estimation error - modulus (above) and phase (below).
Figure 3.15: Reconstruction of the nonlinear systems output. The soft-dashed line represents the input signal; the dashed line represents the theoretical output; the solid line represents the output estimated by the MISO method. a) normalized frequency 0.001, b) normalized frequency 0.03.

3.6 Summary

In this chapter, an algorithm for the identification of nonlinear system using the power series based MISO nonlinear model has been developed. It has been shown through simulations, that the algorithm allows to identify a nonlinear system with a very good accuracy. The main advantage of using a white noise is that the measurement is performed at all the frequencies at once (supposing the span is comparable to the measured frequency band).

Nevertheless, when used in practice, the method has shown good results only in identification of electric devices, where the background noise was much negligible in compare with the excitation signal. The experiment with diodes was presented in (Novák, A., 2007a). In other cases, where the noise floor was much higher, for example in the tape recordings, or in any acoustic measurement using microphones or accelerometers, the method falls down. The probable reason is that the background noise can be in some way correlated with the branches of nonlinear model and thus causes the inaccuracies as the method is based on decorrelation of the inputs.
4.1 Introduction

In chapter 3, a method has been presented for the identification of nonlinear systems (NLS). This method, based on random signal excitation and polynomial Hammerstein model, leads to very good results when identifying simulated systems. It also allows the identification of highly nonlinear real-world systems, such as electric diodes for example. Nevertheless, this method does not succeed in the identification of weakly nonlinear real-world systems, such as those presented in chapter 6 (loudspeaker, acoustic waveguide). The main reason is too low signal-to-noise ratio (SNR), regarding the level of background noise of real-world systems.

In this chapter, another method which clearly improves the estimation of the model, is presented. This method is still based on the Multiple-Input Single-Output model, but makes use of a specially redesigned exponential swept-sine signal instead of a white noise. The method is partly based on the nonlinear convolution method, revised in section 2.4.4 and firstly proposed in (Farina, A., 2000) and (Armelloni, E. & Farina, A., 2001).
First experiments based on swept-sine signals for NLS analysis have shown very good results especially in terms of robustness when considering the background noise (Farina, A., 2000). Nevertheless, the nonlinear convolution method, first presented in (Farina, A., 2000) and further developed in (Armelloni, E. & Farina, A., 2001), can only be used for the analysis of NLS and not for their identification, because of non-synchronization in phase of higher-order nonlinear frequency responses.

The method developed in this chapter brings two improvements to the nonlinear convolution method. First, the excitation signal is redesigned in order to synchronize the higher-order nonlinear frequency responses. Second, the inverse filter used in the nonlinear convolution method is time-extended. These two improvements are detailed in sections 4.2 and 4.3. In section 4.4, two procedures for NLS identification are proposed. These procedures differ in models (polynomial Hammerstein model, or general MISO model), the model being chosen according to the physical knowledge of the NLS under test. Finally, simulation results are presented in section 4.5.

4.2 Input Signal

In this section, the new design of the swept-sine excitation signal is presented, with its time and frequency properties. The properties of nonlinear convolution with the redesigned excitation signal are examined as well.

4.2.1 New Redesign of the Input Signal

The input signal used in (Farina, A., 2000), (Armelloni, E. & Farina, A., 2001) and (Kite, T., 2004) for the analysis of audio equipment is an exponential swept-sine signal, i.e. a signal exhibiting an instantaneous frequency which increases exponentially with time. Such a signal is also called an exponential chirp and is defined as

\[
s(t) = \sin \left[ 2\pi f_1 \int_0^t \exp \left( \frac{t'}{L} \right) dt' \right] \\
= \sin \left\{ 2\pi f_1 L \left[ \exp \left( \frac{t}{L} \right) - 1 \right] \right\},
\]  
(4.1)
where $f_1$ is the start frequency at $t = 0$, and $L$ is the rate of exponential increase in frequency. The parameter $L$ depends on the time length $T$ and the stop frequency $f_2$ of the swept-sine signal. The signal $s(t)$ defined in Eq. (4.1) can be also expressed as in Eq. (2.15) with constant amplitude envelope, $a(t) = 1$, and the instantaneous phase, expressed as

$$\varphi(t) = 2\pi f_1 L \left[ \exp \left( \frac{t}{L} \right) - 1 \right]. \quad (4.2)$$

The instantaneous frequency (2.18) is consequently

$$f_i(t) = f_1 \exp \left( \frac{t}{L} \right). \quad (4.3)$$

Considering $s(t)$ as an asymptotic signal, the group delay $t_f$ is then the inverse function of instantaneous frequency $f_i$ and is given by

$$t_f(f_i) = L \ln \left( \frac{f_i}{f_1} \right) \equiv L \ln \left( \frac{f}{f_1} \right). \quad (4.4)$$

The time length $T$ of the signal $s(t)$ can be consequently expressed as the time between two particular instantaneous frequencies $f_1$ (start frequency) and $f_2$ (stop frequency),

$$T = L \ln \left( \frac{f_2}{f_1} \right), \quad (4.5)$$

and thus the coefficient $L$ is defined as

$$L = \frac{T}{\ln \left( \frac{f_2}{f_1} \right)}. \quad (4.6)$$

In the framework of NLS identification, this definition leads to problems due to non-synchronization of the phases of the higher-order nonlinear frequency responses. For that reason, the swept-sine signal has to be redesigned, as explained in the following.

Let $\Delta t_m$ be the time lag, for which the instantaneous frequency $f_i(\Delta t_m)$ is given by

$$f_i(\Delta t_m) = mf_1, \quad (4.7)$$

for $m \in \mathbb{N} - \{0\}$. Using Eq. (4.4) allows to write

$$\Delta t_m = L \ln (m), \quad (4.8)$$
and the instantaneous phase at the time lag $\Delta t_m$ is thus given by

$$\varphi(\Delta t_m) = 2\pi f_1 L (m - 1). \quad (4.9)$$

Then, as depicted in Fig. 4.1, the swept-sine signal $s(t)$ at particular time lag $\Delta t_m$ is designed to be equal to zero, $s(\Delta t_m) = 0$, with the additional constraint of positive first derivative, $s'(\Delta t_m) > 0$. The reason for doing this is explained in section 4.2.3. These conditions consequently yields

$$\varphi(\Delta t_m) = 2k\pi, \quad (4.10)$$

where $k \in \mathbb{Z}$. Thus, from Eq. (4.10) and (4.9), we get

$$f_1 L (m - 1) = k. \quad (4.11)$$

Sufficient condition to solve (4.11) is then $f_1 L \in \mathbb{Z}$. Using this condition and Eq. (4.6), we write

$$L = \frac{1}{f_1} \text{Round} \left( \frac{\hat{T} f_1}{\ln \left( \frac{f_2}{f_1} \right)} \right), \quad (4.12)$$
where \( \text{Round} \) represents rounding towards nearest integer and where \( \hat{T} \) is an approximate time length of the signal \( s(t) \) used for the design. The real length \( T \) of the signal \( s(t) \) is due to rounding defined as

\[
T = \frac{1}{f_1} \text{Round} \left( \frac{\hat{T} f_1}{\ln \left( \frac{f_2}{f_1} \right)} \right) \ln \left( \frac{f_2}{f_1} \right). \tag{4.13}
\]

Finally, the redesigned exponential swept-sine signal can then be expressed using the equations (4.1) and (4.14) as, \( \forall t \in [0, T] \),

\[
s(t) = \sin \left\{ 2\pi \text{Round} \left( \frac{\hat{T} f_1}{\ln \left( \frac{f_2}{f_1} \right)} \right) \left[ \exp \left( \frac{f_1 t}{\text{Round} \left( \frac{\hat{T} f_1}{\ln \left( \frac{f_2}{f_1} \right)} \right)} \right) - 1 \right] \right\}. \tag{4.14}
\]

The parameters which define the input signal \( s(t) \) are start and stop frequencies \( f_1 \) and \( f_2 \), and the approximative time length \( \hat{T} \).

### 4.2.2 Time Domain Properties

The signal \( s(t) \) defined above and satisfying the conditions depicted in Fig. 4.1 has furthermore the following property in time domain. Consider a signal \( s_2(t) \) with instantaneous frequency equals twice the instantaneous frequency of the signal \( s(t) \). As illustrated in Fig. 4.2, the relation between the signals \( s(t) \) and \( s_2(t) \) is simply

\[
s_2(t) = s(t + \Delta t_2). \tag{4.15}
\]

This condition may be extended, for any \( m \), as

\[
s_m(t) = s(t + \Delta t_m). \tag{4.16}
\]

In other words, all the higher-order harmonic components of the signal \( s(t) \) can be expressed as time shifted replica of \( s(t) \). This condition is sufficient to provide the synchronization of all the higher-order nonlinear frequency responses as shown in the following section.
4.2.3 Nonlinear Convolution Properties

When convolving the output signal $y(t)$ of a NLS excited by the swept-sine signal $s(t)$ with the inverse filter $\tilde{s}(t)$, we get a separated set of nonlinear impulse responses, as reviewed in section 2.4.4. To illustrate theoretic background of this separation in the process of nonlinear convolution, we examine two simple examples corresponding respectively to a NLS without memory and to a NLS with memory.

In the first example, we consider a memoryless NLS with input/output law given by $y(t) = x^3(t)$ and we detail the response of the system to the swept-sine signal $s(t)$ and its convolution with the inverse filter $\tilde{s}(t)$ \(^1\). As the considered NLS is a zero-memory one, the result of the convolution between $s^3(t)$ and $\tilde{s}(t)$ is supposed to be equal to Dirac functions.

In the specific case of cubic nonlinearity, we can use the trigonometric formula which links $\sin^3(\alpha)$ to $\sin(3\alpha)$, as

$$\sin^3(\alpha) = \frac{3}{4}\sin(\alpha) - \frac{1}{4}\sin(3\alpha). \quad (4.17)$$

Then, noting $s(t)$ the input signal, leads to

$$s^3(t) = \frac{3}{4}s(t) - \frac{1}{4}s(t + \Delta t_3), \quad (4.18)$$

where the trigonometric power formula (4.17) is rewritten in the case of input swept-sine signal $s(t)$ using Eq. (4.16). Then, the convolution between the output

\(^{1}\)the signal $\tilde{s}(t)$ is such that $s(t) \ast \tilde{s}(t) = \delta(t)$
4.2. INPUT SIGNAL

and the inverse filter express as (Fig. 4.3)

\[
s^3(t) * \tilde{s}(t) = \frac{3}{4} s(t) * \tilde{s}(t) - \frac{1}{4} s(t + \Delta t_3) * \tilde{s}(t),
\]

(4.19)

\[
= \frac{3}{4} \delta(t) - \frac{1}{4} \delta(t + \Delta t_3).
\]

(4.20)

This particular example of memoryless NLS shows that the nonlinear convolution separates the higher-order components. This result can be generalized to any memory-less NLS, as show below.

\[
s(t) \xrightarrow{x^3} s^3(t) \xrightarrow{g(t)} y(t)
\]

Figure 4.4: An example of a Hammerstein nonlinear system (a system with memory).

In the second example (Fig. 4.4), a memory (a linear system \(g(t)\)) is added to the previous example. The NLS under test is excited with the swept-sine signal \(s(t)\). The output signal is then a convolution between \(s^3(t)\) and the linear impulse response \(g(\tau)\). Then, the convolution between the output and the inverse filter express as (Fig. 4.5)

\[
y(t) * \tilde{s}(t) = \frac{3}{4} s(t) * g(t) * \tilde{s}(t) - \frac{1}{4} s(t + \Delta t_3) * g(t) * \tilde{s}(t),
\]

(4.21)

\[
= \frac{3}{4} g(t) - \frac{1}{4} g(t + \Delta t_3).
\]

(4.22)

If the distance between two higher-order nonlinear impulse responses (in this example \(\Delta t_3\)) is much higher than the time support \(T_g\) of the impulse response of
the linear system \( g(t) \), then the higher-order nonlinear impulse responses can be easily separated by windowing. As an obvious consequence, the separated Dirac functions of the example of memoryless NLS becomes the higher-order nonlinear impulse responses of the NLS with memory. In the case of a simple Hammerstein NLS (Fig. 4.4) all the separated higher-order nonlinear impulse responses are the same, but can be different, for example for generalized Hammerstein NLS.

### 4.2.4 Frequency Domain Properties

As explained in section 2.3.5, the spectra of an asymptotic signal can be derived from time-frequency properties with no need to calculate its Fourier transform. To derive the frequency properties of the excitation swept-sine signal \( s(t) \), we use its analytic form

\[
Z_s(f) = B_s(f) e^{j\Psi_s(f)},
\]

where \( Z_s(f) \) is the spectrum of the signal \( z_s(t) \) and \( B_s(f) \) and \( \Psi_s(f) \) are defined as amplitude and phase of \( z_s(t) \). The spectrum \( Z_s(f) \) can be written in terms of amplitude \( B_s(f) \) and phase \( \Psi_s(f) \) as

\[
Z_s(f) = B_s(f) e^{j\Psi_s(f)}.
\]

As explained in 2.3.5, the swept-sine signal \( s(t) \) is an asymptotic signal, as its instantaneous frequency \( f_i \) and group delay \( t_f \) are invertible functions (Cohen, L., 1995), (Flandrin, P., 1999). In such conditions (see chapter 2.3.5) the amplitude spectra of \( z_s(t) \) can be defined as (Cohen, L., 1995)

\[
B_s(f) = \frac{a_s(t_f)}{\sqrt{\frac{1}{2\pi} |\varphi''_s(t_f)|}},
\]

Figure 4.5: The system with memory depicted in Fig. 4.4: convolution between the output \( y(t) \) and \( \tilde{s}(t) \).
where the second-order derivative of the phase $\varphi_s(t)$ is expressed using Eq. (4.2) and (4.4), as

$$
\varphi_s''(t) = \varphi''(t) = \frac{2\pi f_1}{L} \exp\left(\frac{t}{L}\right) = \frac{2\pi f_1(t)}{L}.
$$

(4.26)

Using the Eq. (4.25) and (4.26) and the equivalence $f \equiv f_i$, the amplitude spectrum, for $f > 0$, can be written as

$$
B_s(f) = \sqrt{\frac{L}{f}}.
$$

(4.27)

The magnitude of the Fourier Transform (FT) of swept-sine signal is depicted in Fig. 4.6.

Figure 4.6: Magnitude of the Fourier transform of the exponential swept-sine signal.
4.3 Inverse Filter

In the previous sections, the excitation swept-sine signal $s(t)$ has been redesigned. It has been shown, that the convolution between the output of the system under test and the inverse filter gave higher-order nonlinear impulse responses. In this section, the expression of the inverse filter $\tilde{s}(t)$ used for the so-called nonlinear convolution is derived.

The use of the inverse filter in order to separate the higher-order components was firstly proposed in (Farina, A., 2000) and further in (Armelloni, E. & Farina, A., 2001), where the proposed procedure was called the nonlinear convolution. Nevertheless, the mathematical derivation of the inverse filter was not given, neither in (Farina, A., 2000), nor in (Armelloni, E. & Farina, A., 2001), where the inverse filter has been defined as "the time-reversal of the excitation signal, equalized with a slope of 6dB/oct (time-reversal mirror plus whitening filter).” In this section, we detail the derivation of the inverse filter using the time-frequency properties of swept-sine signals (Cohen, L., 1995), (Flandrin, P., 1999).

The inverse filter $\tilde{s}(t)$ convolved with the swept-sine $s(t)$ gives theoretically the Dirac function $\delta(t)$.

$$s(t) * \tilde{s}(t) = \delta(t). \tag{4.28}$$

We expand this relation to the analytical signals $z_s(t)$ and $z_{\tilde{s}}(t)$ associated to $s(t)$ and $\tilde{s}(t)$. We consider here the analytic signal of the inverse filter $\tilde{s}(t)$ as

$$z_{\tilde{s}}(t) = \tilde{s}(t) + jH[\tilde{s}(t)], \tag{4.29}$$

and the Fourier transform $Z_{\tilde{s}}(f)$ of the analytic signal $z_{\til{s}}(t)$

$$Z_{\tilde{s}}(f) = \frac{1}{B_s(f)} e^{j\Psi_s(f)}. \tag{4.30}$$

To express the analytic signal of the inverse filter $z_{\til{s}}(t)$, we first derive the relation between magnitudes and phases of both Fourier transforms, as

$$Z_{\tilde{s}}(f) = \frac{1}{B_s(f)e^{j\Psi_s(f)}} = \frac{1}{B_s(f)} e^{-j\Psi_s(f)}. \tag{4.31}$$

and thus

$$B_{\tilde{s}}(f) = |Z_{\tilde{s}}(f)| = \frac{1}{B_s(f)}, \tag{4.32}$$

$$\Psi_{\tilde{s}}(f) = -\Psi_s(f). \tag{4.33}$$
For asymptotic signal $z_s(t)$, the analytic inverse filter $z_s(t)$ is consequently also an asymptotic signal

$$z_s(t) = a_s(t)e^{j\tilde{\phi}_s(t)}. \quad (4.34)$$

Then, we derive the phase $\tilde{\phi}_s(t)$ from the expression of $\tilde{t}_f(f)$. Using Eq. (2.19) and (4.33), we get

$$\tilde{t}_f(f) = -\frac{1}{2\pi} \frac{d\Psi_s(f)}{df} = \frac{1}{2\pi} \frac{d\Psi_s(f)}{df}. \quad (4.35)$$

As a consequence, we get

$$\tilde{t}_f(f) = -t_f(f), \quad (4.36)$$

$$\tilde{\phi}_s(t) = \phi_s(-t). \quad (4.37)$$

To derive the amplitude $a_\tilde{s}(t)$ of the inverse filter, we use the Eq. (2.20)

$$B_s(f) = \frac{a_\tilde{s}(\tilde{t}_f)}{\sqrt{\frac{1}{2\pi} |\tilde{\phi}_s''(\tilde{t}_f)|}}. \quad (4.38)$$

As $\tilde{\phi}_s''(\tilde{t}_f) = \phi_s''(t_f)$ (from Eq. (4.36) and (4.37)), we can substitute from Eq. (4.26)

$$B_s(f) = \frac{a_\tilde{s}(\tilde{t}_f)}{\sqrt{\frac{f_1}{L}}}. \quad (4.39)$$

Now, from Eq. (4.27),(4.32) and (4.39) we can write

$$a_\tilde{s}(\tilde{t}_f) = \frac{f_1}{L}. \quad (4.40)$$

Using Eq. (4.3), the envelope $a_\tilde{s}(t)$ is given by

$$a_\tilde{s}(t) = \frac{f_1}{L} \exp \left( -\frac{t}{L} \right) \quad (4.41)$$

Figure 4.7: The inverse filter $\tilde{s}(t)$ expressed as a time-reversed replica of the swept-sine signal $s(t)$ with a different amplitude envelope.
and the analytic inverse filter is finally expressed as

\[ z_\sim(t) = \frac{f_1}{L} \exp \left( -\frac{t}{L} \right) e^{j\varphi_s(-t)}, \]

that is in shorten form

\[ z_\sim(t) = \frac{f_1}{L} \exp \left( -\frac{t}{L} \right) z_s(-t). \]

Finally, the inverse filter \( \tilde{s}(t) \) is expressed as a time reversed replica of the swept-sine signal \( s(t) \) with a different amplitude envelope as (Fig. 4.8)

\[ \tilde{s}(t) = \frac{f_1}{L} \exp \left( -\frac{t}{L} \right) s(-t). \]

### 4.3.1 Extension of the Inverse Filter

The definition of the swept-sine signal given in Eq. (4.1) and the definition of the inverse filter \( \tilde{s}(t) \) given in Eq. (4.44) are, in general, given for infinite time support. In practice, however, the signals have to begin at a starting frequency \( f_1 \) and to end at a stop frequency \( f_2 \), as defined in the beginning of section 4.2. Then, the time support \( T \), defined in Eq. 4.5, is the time lag between both time instants corresponding to instantaneous frequencies \( f_1 \) and \( f_2 \). The inverse filter used in (Farina, A., 2000) and (Armelloni, E. & Farina, A., 2001) has the same time support. Then, such a definition of the inverse filter allows to estimate nonlinear behavior of the system only in the frequency range between \( f_1 \) and \( f_2 \).

Nevertheless, any nonlinear distortion applied to a pure sine-wave signal, or to a swept-sine signal, generates higher-order harmonic components at frequencies.

![Figure 4.8: The extended inverse filter \( \tilde{s}(t) \) expressed as a time reversed replica of the extended swept-sine signal \( s(t) \) with a different amplitude envelope.](image-url)
that are integer multiples of the fundamental frequency. For example a NLS that can be fully described by a Nth-order polynomial function, generates higher-order harmonic components which is N-times higher than the fundamental frequency. As a consequence, for identifying a NLS in the frequency range \([f_1, f_2]\), with a Nth-order polynomial Hammerstein model, the frequency range of the inverse filter has to be \([f_1, f_3 = Nf_2]\).

For that reason, the inverse filter is extended to the frequency \(f_3\). To generate this inverse filter, a swept-sine signal of which the instantaneous frequency varies from \(f_1\) to \(f_3\) is used. To verify the same properties as for the excitation signal \(s(t)\), the time lag \(T\) has to correspond to the time between two particular instantaneous frequencies \(f_1\) and \(f_2\). Thus, the inverse filter has a time support \(T_{\text{inv}}\) longer than the time support of the excitation swept-sine signal, as shown in Fig. 4.8. The higher-order nonlinear frequency responses are then valid in the frequency range between \(f_1\) and \(f_3 = Nf_2\), as shown in Fig. 4.9.

![Figure 4.9: Result of the nonlinear convolution process, when using the extended inverse filter in the form of higher-order nonlinear frequency responses \(H_n(f)\). Arbitrary case where \(|H_n(f)|\) is a constant in \([nf_1, nf_2]\).](image-url)
4.4 Identification of NLS with the Modified Nonlinear Convolution Method

In previous sections, the input signal $s(t)$ and the inverse filter $\tilde{s}(t)$ have been defined to be used for identification of the NLS under test. In this section, two different cases of NLS identification are detailed, depending on the physical knowledge of the NLS. In the case of no a priori knowledge, a "blind" identification has to be setup and the method leading to the polynomial Hammerstein model is used (Fig. 4.10). This "blind" identification is detailed in section 4.4.1. If the NLS under test is at least approximatively known, the method leading to the general MISO nonlinear model (Fig. 4.11), with input functions $\mathcal{F}_n\{x\}$ chosen according to the physical knowledge of the system, is used. This is detailed in section 4.4.2.

![Figure 4.10: The polynomial Hammerstein model for the case of blind identification.](image)

![Figure 4.11: The general MISO model for the case of known inputs.](image)
4.4. MODIFIED NONLINEAR CONVOLUTION METHOD

4.4.1 Case of Blind Identification

When having no physical knowledge on the NLS under test, the identification to achieve is called “blind” identification, and the polynomial Hammerstein model with power series inputs and linear filters \( G_n(f) \) may be used. The fundamental issue is to write and calculate the relation between partial frequency responses \( H_m(f) \), \( m \in [1, M] \) (Fig. 4.9) and linear filters \( G_n(f) \), \( n \in [1, N] \) in the frequency domain.

As the impulse responses \( h_m(t) \) and \( g_n(t) \), respectively defined as the inverse Fourier transform of \( H_m(f) \) and \( G_n(f) \), are supposed to be real functions, only the half frequency area \( f > 0 \) is considered in the following.

Given the partial frequency responses \( H_m(f) \) defined by Eq. (2.29), the frequency response of the linear filters \( G_n(f) \) of the power series nonlinear model can be derived analytically using the trigonometric power formulas, defined as (Beyer, W.H., 1987), \( \forall l \in \mathbb{N} \)

\[
(sin x)^{2l+1} = \frac{(-1)^l}{4^l} \sum_{k=0}^{l} (-1)^k \binom{2l+1}{k} \sin [(2l + 1 - 2k)x], \quad (4.45)
\]

and \( \forall l \in \mathbb{N} \setminus \{0\} \),

\[
(sin x)^{2l} = \frac{(-1)^l}{2^{2l-1}} \sum_{k=0}^{l-1} (-1)^k \binom{2l}{k} \cos [2(l-k)x] + \frac{1}{2^{2l}} \binom{2l}{l}. \quad (4.46)
\]

Regarding the Fourier transform of Eq. (4.45) and (4.46) and noting \( FT_p \) the result of the Fourier transform only for positive frequencies, we can write, \( \forall l \in \mathbb{N} \),

\[
FT_p \{ (sin x)^{2l+1} \} = \frac{(-1)^l}{4^l} \sum_{k=0}^{l} (-1)^k \binom{2l+1}{k} FT_p \{ \sin [(2l + 1 - 2k)x] \}, \quad (4.47)
\]

and \( \forall l \in \mathbb{N} \setminus \{0\} \),

\[
FT_p \{ (sin x)^{2l} \} = \frac{j}{2^{2l-1}} \sum_{k=0}^{l-1} (-1)^k \binom{2l}{k} FT_p \{ \sin [2(l-k)x] \} + \frac{1}{2^{2l}} \binom{2l}{l}. \quad (4.48)
\]

These formulas consequently give, in the Fourier domain, the relation between the higher-order harmonic \( \sin(lx) \) and the \( l \)th power of the harmonic signal \( \sin^l(x) \),
for \( l \in \mathbb{N} - 0 \). Considering a harmonic input signal of frequency \( f_0 > 0 \), the values of the frequency responses \( H_m(lf_0) \) and the values of the frequency responses of the linear filters \( G_n(lf_0) \) are related in the same way as in Eqs. (4.45 - 4.46).

The way of expressing a distorted harmonic signal \( y(t) \) (with higher harmonics) by the sum of powers of the harmonic signal \( \sin^l(x) \) is illustrated, for case of \( N = 4 \) in Fig. 4.12. The spectrum of the signal \( y(t) \), consisting of 4 harmonics is equal to the sum of 4 weighted spectra of the powers of the harmonic signal. The trigonometric formulas in frequency domain can then be rewritten into the matrix form (4.49), where the matrices \( A \) and \( B \) represent the coefficients in Eq. (4.47) and (4.48)

\[
\begin{pmatrix}
\text{FT}_p\{\sin x\} \\
\text{FT}_p\{\sin^2 x\} \\
\text{FT}_p\{\sin^3 x\} \\
\vdots
\end{pmatrix}
= \begin{pmatrix}
\text{FT}_p\{\sin x\} \\
\text{FT}_p\{\sin 2x\} \\
\text{FT}_p\{\sin 3x\} \\
\vdots
\end{pmatrix}
+ \begin{pmatrix}
A_{n,m} \\
A_{n,m} \\
A_{n,m} \\
\vdots
\end{pmatrix}.
\tag{4.49}
\]

According to Eqs. (4.47), (4.48), the matrix \( A \) is defined as

\[
A_{n,m} = \begin{cases} 
\frac{(-1)^{2n+1-m}}{2^{n-1}(n-m)} \left( \frac{n}{2} \right), & \text{for } n \geq m \text{ and } (n+m) \text{ is even}, \\
0, & \text{else.}
\end{cases}
\tag{4.50}
\]

The matrix \( B \) is a one-column matrix and represents the constant values of the even power series expansion. These values are only linked to the mean value of the output signal. The relation between the partial frequency response \( H_m(f_0) \), for \( f_0 > 0 \) and the linear filters \( G_n(f_0) \) from the power series nonlinear model is given using the coefficients of the matrix \( A \). Each partial frequency response \( H_m(f_0) \), or \( m \)-th-harmonic, can be expressed as a sum of \( m \)-th-harmonics of all the \( n \)-th-powers weighted by the linear filters \( G_n(f_0) \), the coefficients of the \( m \)-th-harmonics of the \( n \)-th-power being \( A_{n,m} \) and thus,

\[
H_m(f_0) = \sum_{n=1}^{N} A_{n,m} G_n(f_0).
\tag{4.51}
\]
Figure 4.12: The way to express a distorted sine signal $y(t)$ (including higher harmonics) as the sum of powers of harmonic signal, in the frequency domain.
Lastly, the linear transformation between $H_m(f_0)$ and $G_n(f_0)$, for $f_0 > 0$ can be generally expressed in matrix form

$$
\begin{pmatrix}
G_1(f_0) \\
G_2(f_0) \\
G_3(f_0) \\
\vdots
\end{pmatrix} = \left( A^T \right)^{-1} \begin{pmatrix}
H_1(f_0) \\
H_2(f_0) \\
H_3(f_0) \\
\vdots
\end{pmatrix},
$$

where $A^T$ denotes the transpose of $A$ and where $N = M$, which makes $A$ a square matrix. The method is tested in a simulation case in section 4.5.

### 4.4.2 Case of Known Inputs

If a physical model of the NLS under test is available, the general MISO nonlinear model (Fig. 2.8) can be used. The input signals $x_n(t)$ are then known linear and/or nonlinear functions of $x(t)$, and the frequency response functions $G_n(f)$ represent linear filters. The linear filters $G_n(f)$ of the nonlinear model can be expressed in the time domain as impulse responses $g_n(t)$. The problem may consequently be stated as: how to estimate the linear filters $G_n(f)$ and how to evaluate the relevance of the model, when measuring $y(t)$ and choosing $F_n\{x(t)\}$, for a given input excitation signal $x(t)$.

Knowing $H_m(f)$, for $m \in [1, M]$ ($M$ being the number of higher orders used in the nonlinear convolution method), the next step is to estimate the linear filters $G_n(f)$, for $n \in [1, N]$ ($N$ being the number of branches used in the general MISO model). The estimation of $G_n(f)$ is based on solving a linear system of equations using the least squares method. First, the coefficients $c_{n,k}$ of Discrete Fourier Series of the functions $F_n\{x(t)\}$ are calculated as

$$
c_{n,k} = 2 \frac{P}{P-1} \sum_{p=0}^{P-1} \mathcal{F}_n \left\{ \sin \left( \frac{2\pi}{P} p \right) \right\} \exp \left( -j \frac{2\pi}{P} kp \right),
$$

for an input excitation signal being a discrete-time harmonic signal of length $P$. Then, we have

$$
H_m(f) = \sum_{n=1}^{N} G_n(f)c_{n,m} + \text{Res}(f),
$$

where $\text{Res}(f)$ is the residue. As $M \geq N$, there can be more equations than unknowns. To solve the set of equations with $M \geq N$, the least squares algorithm
4.5 Simulation of a NLS with Memory

In this section, a simulation of a NLS estimation is proposed in order to illustrate the case of blind identification presented in section 4.4.1. The case of known inputs (Sec. 4.4.1) is used and tested in the chapter 5.

In order to illustrate the method, a NLS with memory is simulated. It consists of two nonlinear branches with linear and cubic parts, each of them followed by a linear filter (Fig. 4.13). Both filters are digital Butterworth filters. The filter $G_1(f)$ used in linear branch is a 10th-order high-pass filter, with cutoff frequency 500 Hz and the filter $G_3(f)$, used in the cubic branch, is a 10th-order low-pass filter, with cutoff frequency 1 kHz. The simulation is performed using a sampling frequency $f_s = 12kHz$. For that reason, the maximum frequency of the excitation signal is 2 kHz, in order to avoid any aliasing (Zhu, Y.M., 1992), (Tsimbinos, J. & Lever, K.V., 1998).

Such a NLS is identified using the swept-sine signal defined in Sect.4.2. The parameters of the excitation signal are set up as: $f_1 = 20$ Hz, $f_2 = 2000$ Hz, $\tilde{T} = 5$ s. Once the response of the NLS under test is known for this excitation signal, the nonlinear convolution is performed and the linear filters of the nonlinear model

\[
x(t) \rightarrow x_1(t) \rightarrow G_1(f) \rightarrow G_3(f) \rightarrow (\cdot)^3 \rightarrow x_3(t) \rightarrow G_4(f) \rightarrow y(t)
\]

Figure 4.13: Tested NLS with memory.
are estimated. To simulate real-world conditions, a white Gaussian noise (WGN) is moreover added, resulting in a SNR of 30dB, and time-synchronous averaging is then performed. The results of the estimation are compared in Fig. 4.14 - 4.17.

In the frequency range of the swept-sine input signal \((f_1 = 50\text{Hz}, f_2 = 2\text{kHz})\), both estimated filters, the High-Pass filter \((500-2000)\text{Hz}\) and the Low-Pass filter \((50-1000)\text{Hz}\), match in amplitude and phase with the theoretical characteristics. The relative errors in the studied frequency range have been 3% for the filter \(G_1(f)\) and 1% for the filter \(G_3(f)\).

### 4.6 Summary

In this chapter, a new method for the identification of nonlinear system has been developed. The method is based on an exponential swept-sine input signal allowing a one-path identification of the nonlinear system under test. First of all, the excitation swept-sine signal has been redesigned to further synchronization of the higher-order harmonic components. Then, the inverse filter has been properly calculated, and next, two cases for the new nonlinear system identification method have been presented: the case for blind identification (polynomial Hammerstein model) and the case of known inputs (general MISO model). The first method has also been tested on a simulation example.

In the following chapter dealing with hysteresis, the case with known inputs presented in this chapter (based on general MISO model) is used. If hysteresis can be seen as a nonlinear system with more complicated input-output law, it is nevertheless shown, that, in some cases, the second method developed in section 4.4.2 can be successfully used for the identification of hysteretic systems.
Figure 4.14: The theoretical frequency response (modulus - above; phase - below) of the filter $G_1(f)$, from the linear part of the NLS of Fig. 4.13.

Figure 4.15: The estimated frequency response (modulus - above; phase - below) of the filter $G_1(f)$, from the linear part of the NLS of Fig. 4.13.
Figure 4.16: The theoretical frequency response (modulus - above; phase - below) of the filter $G_3(f)$, from the linear part of the NLS of Fig. 4.13.

Figure 4.17: The estimated frequency response (modulus - above; phase - below) of the filter $G_3(f)$, from the linear part of the NLS of Fig. 4.13.
5.1 Introduction

Hysteresis is a nonlinear phenomena which is quiet common in several scientific areas, as physics, chemistry, biology, engineering and so on. As each physical area concerned with hysteresis uses specific definition, which is well-suited to the specific problems of this area, there is actually no universal definition of hysteresis.

Generally, a system exhibiting a hysteretic behavior is considered to be a nonlinear system (NLS) of with the output $y(t)$ depends not only on the input signal $x(t)$, but also on the current state of the system (Fig. 5.1). According to Mayergoyz, I.D. (1991), a device is hysteretic if its input-output relation exhibits a multi-branch nonlinearity for which branch-to-branch transitions occur after input extrema (Fig. 5.2).

![Figure 5.1: A general schema of hysteretic system: black-box representation.](image)
The hysteresis may be classified under two main headings (Marin, J., 1962): (a) dynamic hysteresis, sometimes called *visco-elastic*, or *rate-dependent* hysteresis and (b) static hysteresis referred to as *plastic*, or *rate-independent* hysteresis. The term static, or *rate-independent* means that the hysteretic input-output relation is determined only by the past extremum value of input, while the speed of input variations between extremum points, and, thus frequency, have no significant influence on branching. (Bertotti, G. & Mayergoyz, I.D., 2006). In other words, from the point of view of signal processing, the system is not frequency dependent. This differs from dynamic, or *rate-dependent* hysteresis, for which the speed of input variations between extremum points, and thus frequency have an influence on the input-output relation.

All static hysteresis can be classified into two groups: (a) hysteresis with local memory and (b) hysteresis with nonlocal memory (Mayergoyz, I.D., 1991). A static hysteretic system with local memory exhibits a hysteresis loop, containing only two branches (Fig. 5.2a). Once a particular value of the hysteresis loop is reached, there are only two possible future behaviors of the output, corresponding to increase or decrease of the input signal (Liu, Y. et al., 2005). On the contrary, a hysteretic system with nonlocal memory has an infinite numbers of possible future behaviors (Fig. 5.2b), depending on the history of the input (Sjöström, M., 2001).

![Figure 5.2: Input-Output characteristic of a static hysteretic system with (a) local, (b) nonlocal memory: Hysteresis loop.](image-url)
The well-known hysteresis models commonly used, such as Preisach model (Mayergoyz, I.D., 1991), or Prandtl-Ishlinskii model (Šolc, F., 2007) can represent only the static hysteresis with nonlocal memory. However, since almost all the real-world NLS with hysteresis are frequency dependent, the static hysteresis definition is not sufficient for a complete description of these systems. Then, a dynamic hysteresis modeling has to be applied, and its mathematic description is generally highly complicated.

Regarding the method presented in chapter 4, it can model some nonlinear systems driven by hysteresis. In this chapter, it is demonstrated how to apply the nonlinear identification method in case of hysteretic system. First, the relation between a static hysteretic system and the Hammerstein system is shown and a simple example of static hysteresis with local memory is considered. Then, two simulation experiments, a static hysteretic system and a dynamic hysteretic system are examined.

## 5.2 Hysteresis as a Hammerstein System

Hammerstein model can be used to represent a system with a local memory static hysteresis. This is shown on a simple example in Sec. 5.2.1. The definition of static hysteresis excludes any frequency dependency. However, even in typical hysteresis phenomena like ferromagnetism, the memory effects are not purely rate-independent. Some cases of rate-dependent hysteresis systems can be modeled by some nonlinear models. For example, the simple Hammerstein model, which includes a nonlinear static block followed by a linear dynamic block, can be applied to model special cases of static hysteresis with local memory and special case of rate-dependent hysteresis (Hsu, J.T. & Ngo, K.D.T., 1997). This case is studied in Sec. 5.2.2.

### 5.2.1 Static Hysteresis with Local Memory

As an example of a static hysteresis that can firstly be modeled as a simple Hammerstein model, a relay with hysteresis is considered. It is characterized by two threshold values \((\alpha, \beta)\) and by two output values \((a, b)\). Let \(x(t)\) be the input
signal and \( y(t) \) the output signal, then

\[
y(t) = \begin{cases} 
  b, & x(t) > \beta, \\
  y(t - \Delta t), & \alpha \leq x(t) \leq \beta, \\
  a, & x(t) < \alpha.
\end{cases}
\]  

(5.1)

An example of a hysteresis curve of the relay with hysteresis is shown in Fig. 5.3. The hysteresis curve does not depend on frequency and, for any point of the curve, since there is only two possible future behaviors of the output, depending if the input signal is increasing or decreasing, this hysteresis is static with local memory.

Applying a pure harmonic signal to the input of the system and observing the output leads to Fig. 5.4. Regarding the relation between input and output signals, we can then notice that the output signal \( y(t) \) may simply be modeled as a saturated and time-shifted version of input signal \( x(t) \). The nonlinear model of the

---

Figure 5.3: Relay with hysteresis: static nonlinearity with local memory.

Figure 5.4: Waveform of the input (sine) and output signal (after passing through the relay with hysteresis).
5.2. HYSTERESIS AS A HAMMERSTEIN SYSTEM

A relay with hysteresis can consequently be represented by a phase delay followed by a saturation (depending on $\alpha$ and $\beta$) as shown in Fig. 5.5. This representation corresponds to a simple Wiener model without any frequency dependency.

5.2.2 Dynamic Hysteresis

Certain dynamic hysteresis, sometimes called hysteresis with *viscosity-type effect*, can be represented via time-convolution (Mayergoz, I.D., 1991)

$$y(t) = \int_0^\infty NL(x(t - \tau))g(\tau)d\tau,$$

(5.2)

where $NL$ represents a static nonlinear function and $g(\tau)$ is an impulse response of a linear system that plays the role of the memory (expressed as a linear time invariant system). The definition (5.2) corresponds to the scheme in Fig. 5.6, which is a simple Hammerstein model (see Sec. 2.2.2). In such case, the nonlinear static block is realized by a static hysteresis and the linear dynamic block is realized by a linear filter, which takes into account the rate-dependent effects of hysteresis (Hsu, J.T. & Ngo, K.D.T., 1997), (Giri, F. et al., 2008).

As this particular case of dynamic hysteresis (Eq. (5.2), Fig. 5.6) can be modeled by the simple Hammerstein system, and since the simple Hammerstein system can be replaced by a general MISO model, the nonlinear method of identification, based on MISO nonlinear model and presented in chapter 4, can be consequently used.

![Figure 5.6: A block-schema of dynamic hysteresis represented via time-convolution.](image-url)
5.3 Simulation Experiment

To identify hysteretic systems, we use the method developed in Sec. 4.4.2. The method is based on swept-sine excitation input signal and nonlinear convolution. We use the general MISO model, with input functions chosen according to the assumed physical model structure of the system under test.

First, the proposed algorithm is tested for a static hysteresis nonlinearity (which means that the memory effects in NLS are only due to the past extremum value of the input excitation signal). In the simulations detailed in this section, a white Gaussian noise (WGN) is furthermore added to evaluate the robustness of the method in the presence of output noise. Next, a more complicated hysteresis nonlinearity is simulated. The frequency dependency is obtained by simulating nonlinear resonant systems.

For the simulation experiments, the parameters are the following. The sampling frequency is $f_s = 96 kHz$. The frequency range of the input excitation swept-sine signal is between $f_1 = 300Hz$ and $f_2 = 3.5kHz$. The duration of the swept-sine signal is $T = 10s$. Lastly, time-synchronous averaging (10x) is used for all output signals presented in this section.

![Figure 5.7: The hysteresis loop of the quadratic hysteresis $q[x(t), x_m]$, $x_m$ being the input excitation signal maximum value.](image-url)
5.3. Simulation Experiment

5.3.1 Static Hysteresis Nonlinearity

The simulated system under test is described by a quadratic hysteresis $q[x(t), x_m]$ in combination with a linear part $\alpha x(t)$ and a non-hysteretic quadratic part $\beta x^2(t)$, and can be expressed as

$$y(t) = \alpha x(t) + \beta x^2(t) + \gamma q[x(t), x_m],$$

where the quadratic hysteresis $q[x(t), x_m]$ is defined as (Gusev, V., 2000)

$$q[x(t), x_m] = x_m x(t) - \frac{1}{2} \left( x_m^2 - x^2(t) \right) \text{sign} \left( \frac{\partial x(t)}{\partial t} \right),$$

$\text{sign}$ being the signum function and $x_m$ the past extremum value of the input excitation signal $x(t)$. This expression may be seen as a semi-empirical input-output law, that is used in nonlinear acoustics, and for which the hysteresis model fits well the experiments. Furthermore, the estimation of parameters $\alpha$, $\beta$ and $\gamma$ is of crucial importance for identifying the hysteretic system under test.

The hysteresis loop of the quadratic hysteresis $q[x(t), x_m]$ consequently depends on input excitation parameter $x_m$, as shown in Fig. 5.7. In the simulation example of static hysteresis nonlinearity, we consider frequency-independent coefficients $\alpha$, $\beta$ and $\gamma$. The system parameters selected in the simulation are $\alpha = 1$, $\beta = 0.2$ and $\gamma = 0.5$. To evaluate the robustness to noise of the method, identification is performed by adding a WGN $n(t)$ to the output signal, as shown in Fig. 5.8.

As the system under test is not frequency dependent, the filters to be estimated $A_1(f)$, $A_2(f)$ and $A_3(f)$ are constant and respectively equal to $\alpha$, $\beta$ and $\gamma$. Table 5.1 gives both the estimated mean values of parameters $\alpha$, $\beta$ and $\gamma$ for different output signal-to-noise Ratio (SNR), and the corresponding mean values of the

![Figure 5.8: The nonlinear hysteresis system with a WGN $n(t)$ added to the output.](image)
residue $\text{Res}(f)$. The mean value of all the measured parameters is calculated within the frequency range $[f_1, f_2]$. The estimated values of the parameters are in good agreement with the available simulated data, leading to errors less than 10\% for SNR=15 dB (worst simulated case). Similarly, the mean value of the residue is about 500 times below the weakest parameter value to be estimated, for SNR=15 dB. Lastly, the error between the estimated mean values and the exact values may be reduced by time-synchronous averaging the output signals.

### 5.3.2 Hysteresis Nonlinearity with Frequency Dependency

In this section, we consider a frequency dependent hysteretic system. The input-output model defined in Eq. (5.3) is simulated with frequency dependent frequency response functions (FRFs) $\hat{\alpha}(f)$, $\hat{\beta}(f)$ and $\hat{\gamma}(f)$. The FRF $\hat{\alpha}(f)$, related to the linear part of the NL system, is modeled by a one degree-of-freedom (1-DOF) system, having a resonant frequency $f_\alpha = 1kHz$. As a consequence, the FRF $\hat{\beta}(f)$, related to the non-hysteretic quadratic part of the NL system, exhibits a resonant frequency at twice $f_\alpha$ plus a DC component. Lastly, the FRF $\hat{\gamma}(f)$, related to the quadratic hysteresis, is modeled as a 2-DOF system, with resonant frequencies $f_\gamma_1 = 1kHz$ and $f_\gamma_2 = 2.6kHz$.

The Figs. 5.9-5.11 show both simulated FRFs $\hat{\alpha}(f)$, $\hat{\beta}(f)$ and $\hat{\gamma}(f)$, and estimated FRFs $A_1(f)$, $A_2(f)$ and $A_3(f)$ (amplitude and phase), and the associated magnitude error and phase error. The value of the output SNR is 30 dB. For sake of clarity, the estimated FRFs are depicted as dashed lines 10dB below their values in modulus and 0.8 radian below in phase. Whatever the considered FRF, the agreement between simulated and estimated values is very good, both in
5.4 Summary

Even if hysteretic systems have been studied for a long time in several areas of science, it is still difficult to define what exactly the hysteresis is, because there is no specific definition of this phenomena that would includes all the kinds of hysteresis. Generally, the hysteresis is classified into static and dynamic hysteresis. The static hysteresis is easier to describe, but its description does not allow to model a frequency dependency. A subclass of dynamic hysteresis can be described by time-convolution and this definition correspond to a branch of the MISO nonlinear model, i.e. the simple Hammerstein system. Thus, such a hysteresis can be theoretically modeled by the nonlinear MISO model.

It has been shown in this chapter, that the method for identification of nonlinear systems can also be used in some cases of hysteresis systems. The method based on swept-sine excitation and using MISO models has been successfully tested on a simple static hysteresis with local memory (relay with hysteresis). It has also been shown that the method can estimate the frequency dependency of system parameters in amplitude and phase very precisely.

<table>
<thead>
<tr>
<th>modulus MSE [dB]</th>
<th>$A_1(f)$</th>
<th>$A_2(f)$</th>
<th>$A_3(f)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>phase MSE [rad]</td>
<td>$4 \cdot 10^{-4}$</td>
<td>$8 \cdot 10^{-5}$</td>
<td>$7 \cdot 10^{-4}$</td>
</tr>
</tbody>
</table>
Figure 5.9: Amplitude (a) and phase (b) of the theoretical filter $\hat{a}(f)$ (solid) and estimated filter $A_1(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for SNR=30dB.
Figure 5.10: Amplitude (a) and phase (b) of the theoretical filter $\tilde{\beta}(f)$ (solid) and estimated filter $A_2(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for SNR=30dB.
Figure 5.11: Amplitude (a) and phase (b) of the theoretical filter $\tilde{\gamma}(f)$ (solid) and estimated filter $A_3(f)$ (dashed and shifted of 10dB offset in amplitude and 0.8 rad in phase) and its difference (below) for AWGN of SNR = 30dB.
A new method for nonlinear system (NLS) identification has been presented in chapter 4, based on polynomial Hammerstein representation, associated to the use of exponential swept-sine excitation signal. This method has been tested on several experimental simulations with successful results. The aim of this chapter is twofold: first, to validate this new method by analyzing two well-known real-world NLS and, second, to apply this method to the identification of electrodynamic loudspeaker nonlinearities.

In the first part of this chapter, the measuring methodology (Sec. 6.1) and the objective criteria used to evaluate the accuracy of the method (Sec. 6.2) are defined. Then, the method is set up for identifying two real-world NLS. The first NLS under test is an audio limiter (Sec. 6.3). This electronic system has been chosen because of its highly nonlinearity level, driving by a very basic input-output characteristic. Thus, it is a good candidate for a first evaluation of the method. The second NLS under test is an acoustic waveguide submitted to a high level acoustic wave propagation (Sec. 6.4). For high amplitude disturbances, the acoustic wave can be distorted along the propagation, and may lead to shock waves for sufficiently large source to receiver distances. Compared to the audio limiter, this system is a weakly nonlinear one, and its identification could be more challenging. Nevertheless, because a physical model of the distortion along the propagation in the waveguide is available, this second NLS is well-suited to a
more accurate evaluation of the method.

Once validated, the method can be used for the identification of any unknown NLS. The last NLS under test is an electrodynamic loudspeaker, studied in Sec. 6.5.

### 6.1 Measuring Methodology

The measuring methodology used for the estimation of the model, in the framework of the experimental validation of the method, can be described as follows. First of all, the order $N$ of expected polynomial Hammerstein model (Fig. 4.10) is chosen. As explained in chapter 4, the order $N$ is equal to the number of branches used in the polynomial Hammerstein model. The higher the order $N$, the higher the computing power. The lower the order $N$, the less the accuracy of the model. Thus, the order $N$ is a compromise between efficiency and accuracy.

Then, the frequency range of the NLS under test and the time length of the excitation signal are given. The frequency range $(f_1, f_2)$ of the excitation signal is generally given by the working frequency bandwidth of the NLS under test. The time length $T$ has to be long enough to avoid the overlap (given by the time lags $\Delta t_m$ (Eq. (4.8))) between two higher-order nonlinear impulse responses. The longer the time length of the excitation signal, the less the overlapping, but the higher the processing time needed for the convolution. Through the examples presented in this chapter, the optimal time length $T$ (Eq. (4.5)) is between 5 and 10 seconds. Once given these parameters $(f_1, f_2$ and $T$), the excitation signal $s(t)$ is generated as defined in chapter 4 (Eq. (4.14)), and the extended inverse filter $\tilde{s}(t)$ (Sec. 4.3.1), further used for the nonlinear convolution, is also generated.

The response of the system to the excitation signal is then recorded. The convolution between the response and the inverse filter is calculated and the higher-order nonlinear impulse responses $h_n(t)$ are separated by windowing as described in chapter 4. Finally, the higher-order nonlinear frequency responses $H_n(f)$ are calculated as the Fourier transform of windowed higher-order impulse responses.
At this stage of analysis, two approaches may be distinguished. Either only the first impulse response \( h_1(t) \) (or the first frequency response \( H_1(f) \)) is necessary and the others are not further considered, or all the higher-order nonlinear impulse responses are used. The first approach can serve to get rid of the nonlinear part in order to find the proper linear response, when considering weakly NLS (Farina, A., 2007). This approach is not considered in the thesis. The second approach can also lead to several choices, depending on how accurately the NLS has to be studied. Either it serves only for the analysis of the system under test, for example for calculation of frequency dependent total harmonic distortion (THD), or for the complete estimation of the model. In the second case, the derivation of the expressions of polynomial Hammerstein model filters \( G_n(f) \) from the expression of \( H_m(f) \) is needed as a final step.

### 6.2 Objective Criteria for Verification

Once the model is estimated, the validation of the method is realized by comparing the model output and the real-world system output when both are excited with the same input signal. The comparison between model and real-world system outputs is achieved by using a criterium in both the time and frequency domains. In the frame of this thesis, this criterium is defined as the relative mean-square error between the real and regenerated output signals. Let \( y[n] \) represent the output signal of the real-world NLS and \( y_r[n] \) the output signal of the nonlinear model also called the regenerated output. The error signal \( e[n] \) can be defined as

\[
e[n] = y[n] - y_r[n].
\] (6.1)

Hence, the relative mean-square error (MSE) between \( y[n] \) and \( y_r[n] \) is defined as the relative squared error, average over all the samples \( n \)

\[
MSE_t = \frac{\sum_{n=0}^{N-1} (y[n] - y_r[n])^2}{\sum_{n=0}^{N-1} (y[n])^2},
\] (6.2)

which can be expressed in percent.

The criteria for the frequency domain is proceeded in the same manner for absolute value of spectra \( Y[k] \) and \( Y_r[k] \), respectively defined as discrete Fourier
CHAPTER 6. NONLINEAR SYSTEMS IN ACOUSTICS

Figure 6.1: Input-output characteristic of a limiter with threshold breakpoint \((a, \alpha)\).

The total relative MSE in frequency domain is then defined as

\[
MSE_f = \frac{\sum_{k=K_1}^{K_2} (|Y[k] - Y_r[k]|)^2}{\sum_{k=K_1}^{K_2} (|Y[k]|)^2},
\]

where the spectra \(Y[k]\) and \(Y_r[k]\) are defined for \(k \in [K_1, K_2]\), corresponding to the studied frequency band.

6.3 Validation on an Audio Limiter

The first tested real-world NLS for validation of the method is an audio limiter. The limiter is an electronic device that allows signals below a certain input level, called threshold, to pass linearly while clipping the amplitudes that exceed this threshold level. In other words, the limiter is a device designed to limit the level of a signal to the threshold level (Fig. 6.1). If the level of the input signal is much higher than the threshold, the limiter produces highly distorted output waveforms. The audio device under test is \(dbx 266XL\) Compressor, Limiter, Gate (DBX, 2007), with the threshold set to 0.25 V.

The experimental setup is schematized in Fig. 6.2. As already said, the system under test is considered as a black-box system with single input \(x[n]\) and single output \(y[n]\). The order \(N\), corresponding to the number of branches of the
A polynomial Hammerstein model is chosen to be \( N = 8 \). A swept-sine excitation signal \( s(t) \) is generated according to Eq. (4.14) and provided to the input of the limiter. The measurement is performed using a swept-sine excitation signal with an amplitude \( A = 1 \) V, a frequency span limited by the start frequency \( f_1 = 10 \) Hz and the stop frequency \( f_2 = 5 \) kHz and an approximative time duration \( \tilde{T} \approx 6 \) s. The sampling frequency is set to \( f_s = 96 \) kHz. As explained in section 6.1, the polynomial Hammerstein model is then estimated using the method described in chapter 4.

To verify the accuracy of the estimated model, a signal \( x[n] \) is generated and used as input signal of both model and real-world limiter. The responses \( y_r[n] \) of the model and \( y[n] \) of the limiter are then compared using MSE criterium (Eq. (6.2),(6.3)), in both time and frequency domains. Three comparisons, corresponding to three different input signals are performed.

Firstly, a sine-wave input signal is generated with frequency \( f_0 = 500 \) Hz and amplitude \( A_0 = 1 \) V. Both regenerated and real-world system outputs are then compared in Fig. 6.3. For the sake of clarity, the output of the model in frequency domain is shifted to the right. On one hand, the output of the real-world system consists of numerous higher-order harmonic components. On the other hand, the model is truncated to 8th order. As a consequence, the regenerated output signal cannot contain more than 8 higher-order harmonic components and then

Figure 6.2: Block diagram of the audio limiter measurement.
exhibits well-known oscillations, known as Gibbs phenomenon. Nevertheless, it fits in with percent relative $MSE_t = 1.8\%$. The first 8 harmonic components of the model-based spectra match very precisely with experimental ones. The evens harmonics are less than -60 dB, that is near the level of background noise and thus the error linked to these harmonics can be neglected. When measuring the relative MSE in frequency domain in the region of only the first eight harmonics, the relative MSE falls down from $MSE_f = 0.01$ to $MSE_f = 8.6 \cdot 10^{-5}$.

Secondly, the input signal is a sinewave with the same frequency ($f_0 = 500$ Hz), but with a lower amplitude ($A_0 = 0.3$ V). Doing that allows to check the validity of the model when used with an input signal amplitude different than the one ($A = 1$ V) used for the previous estimation. In other words, it allows to verify if the nonlinear law changes with amplitude or not. The results of this simulation, given by Fig. 6.4, show that the nonlinear model fits in also well for a lower input signal, even if the higher harmonics are badly estimated, suggesting a possible input-output law depending on the amplitude of the input signal. The percent relative MSE is $MSE_t = 0.6\%$. The harmonic components of the model-based spectra match rather precisely with the experimental ones. When measuring the relative MSE in frequency domain in the region of only the first eight harmonics, the relative MSE changes down from $MSE_f = 8.9 \cdot 10^{-4}$ to $MSE_f = 8.3 \cdot 10^{-4}$.

Lastly, a sawtooth input signal is used as the input signal and the measured and regenerated output signals are compared in time-domain. The sawtooth signal is chosen to exhibit a period of 480 samples, equivalent to frequency 200 Hz, for a data rate $f_s = 96$ kHz. The input and output signals are depicted in Fig. 6.5. The regenerated and measured output waveforms are very similar, even of the regenerated one exhibits Gibbs oscillations. The percent relative MSE between regenerated and measured output waveforms is $MSE_t = 3\%$. This confirms that the method is accurate enough for regenerating the output waveforms.
Figure 6.3: Comparison of the real-world and regenerated model output signals for the input sine-wave $f = 500$ Hz, $A = 1$ V; (a) Time domain comparison between the limiter output (thin) and model output (bold); (b) Frequency domain comparison between the first 8 harmonics of the limiter output (thin) and model output (dashed-bold and shifted).
Figure 6.4: Comparison of the real-world and regenerated model output signals for the input sine-wave $f = 500 \text{ Hz}$, $A = 0.3 \text{ V}$; (a) Time domain comparison between the limiter output (thin) and model output (bold); (b) Frequency domain comparison between the limiter output (thin) and model output (dashed-bold and shifted).

Figure 6.5: Comparison of the responses of the real system (thin) and its estimated nonlinear model (bold) to the sawtooth signal (dashed).
6.4 Validation on a Acoustic Waveguide

The second system studied in the frame of this experimental validation is an acoustic waveguide in which a high level acoustic wave propagates (Fig. 6.6). It is well-known that for high amplitude disturbances, the acoustic wave can be highly distorted leading to shock waves (Hirschberg, A. et al., 1996). Indeed, local small nonlinear perturbations of the acoustic wave due to high sound level are cumulative along the propagation and may distort the waveform considerably for sufficiently large source to receiver distances.

Theoretical descriptions of nonlinear wave propagation in duct with wall friction can be found in (Hamilton, M.F. & Blackstock, D.T., 1998). A good description of cumulative effects occurring along the propagation is obtained using a Fubini approximated solution for the nonlinear pressure in the waveguide (Hamilton, M.F. & Blackstock, D.T., 1998).

The generation of harmonic frequencies during the propagation depends on the acoustic pressure $p_0$ and frequency $f_0$ of the generated wave, and on the distance $l$ between the source and the receiver (microphone). If the acoustic pressure $p_0$ is small, the wave phenomena can be described in terms of linear acoustics. In that case, the wave propagates without change of shape, because all points of the waveform travel at the same speed. If the acoustic pressure $p_0$ is much higher, the propagation speed varies from point to point on the waveform. The speed of sound is a little higher where the acoustic pressure $p$ is positive and a little lower where the acoustic pressure $p$ is negative. The differences in speed causes noticeable distortion. The higher the acoustic pressure $p_0$, the shorter the distance for which the distortion phenomena occurs (Crocker, M.J., 1998).

![Figure 6.6: The block scheme of a waveguide with illustrated distortion of an acoustic wave.](image-url)
The following model of the pressure $p(t)$ evolution is based on Fubini solution. The pressure $p$ can be expressed as a Fourier series

$$p(t) = p_0 \sum_n B_n \sin (2\pi f_0 nt),$$

(6.4)

where the Fourier coefficients $B_n$ are given as (Enflo, B.O. & Hedberg, C.M., 2002)

$$B_n = \frac{2}{n\sigma} J_n(n\sigma).$$

(6.5)

In Eq. (6.5) $J_n$ is the $n$th-order Bessel function of the first kind and $\sigma$ is a coefficient depending on the acoustic pressure $p_0$, the frequency $f_0$ of the generated wave and the distance $l$, defined as

$$\sigma = l \frac{(\kappa + 1)p_0 2\pi f_0}{2\kappa p_b c_0},$$

(6.6)

where $\kappa$, $p_b$ and $c_0$ are respectively the heat capacity ratio, the atmospheric pressure and the celerity of sound. The Fubini solution was experimentally verified in (Menguy, L. & Gilbert, J., 2000). This theory serves in this work as a reference for testing the proposed identification method.

The experiment is performed in a cylindrical air-filled tube, whose dimensions are: length $l = 10$ m long and internal diameter $d = 58$ mm. Furthermore, the end of the tube is connected to an absorbing termination in order to avoid standing waves. The input signal is led through a D/A converter and an amplifier directly

![Figure 6.7: Block diagram of the waveguide nonlinearities measurement.](image-url)
to the source. The source is a compression loudspeaker driver (JBL 2446H). In
the tube, a microphone is mounted at the distance of 6 m from the source. This
microphone is an acceleration compensated piezo-electrical gauges (PCB M116B).
The same type of microphone is also mounted in the very beginning of the tube,
near the source, in order to measure the initial sound pressure level and to control
the linearity of the input wave. The output signal is captured after passing
through the amplifier (signal conditioner PCB 442B104) and A/D converter.
The A/D, D/A conversion has been made by the audio card (RME Fireface 400
FireWire). Both converters, both amplifiers and the microphone are supposed
to be linear in comparison with the propagation nonlinearity. The block scheme
of the experimental setup is depicted in Fig. 6.7. The measurement has been
performed with a pressure amplitude $p_0 = 176$ Pa, a distance of propagation
$l = 6$ m and a frequency range of the swept-sine wave limited by the frequencies
500 Hz - 5 kHz.

The measurement experiment, using the method developed in chapter 4, is
performed to estimate the higher-order nonlinear frequency responses $H_n(f)$. As
this higher-order nonlinear frequency responses $H_n(f)$ represent the frequency
evolution of higher harmonics (chapter 4) they can be directly used for the
estimation of the Fourier coefficients $B_n$ (Eq. (6.4)) of the higher harmonics.
In other words, the frequency dependent coefficients $B_n(f) = H_n(f)$. A first way
to verify the method validity is then to compare the theoretical coefficient $B_n(f)$

![Figure 6.8: Comparison between the theoretical frequency dependency of coefficients $B_1$, $B_2$, $B_3$ (straight lines) and the three first higher-order experimental nonlinear frequency responses.](image)
with the experimental higher-order nonlinear frequency responses \( H_n(f) \). The three first theoretical frequency dependent coefficients \( B_n(f) \) are depicted versus frequency in Fig. 6.8 along with the magnitudes of experimental higher-order nonlinear frequency responses. The coefficients \( B_n(f) \) are straight lines, contrary to coefficients \( H_n(f) \). This discrepancy is probably due to the influence of the loudspeaker frequency characteristic and to the lacking in anechoicity of the tube termination. Nevertheless, the global behavior is respected.

A second way to verify the method is to calculate the filters \( G_n(f) \) of the polynomial Hammerstein model from the higher-order nonlinear frequency responses \( H_n(f) \) as explained in Sec. 4.4.1. Then, it is possible to regenerate the response of the model when excited with a pure harmonic signal and to compare it both with the theoretical (Fubini) and with the measured responses corresponding to pure harmonic input signal. This comparison is shown in Figs. 6.9 and 6.10. The dashed line is the pure harmonic input signal. The comparison between the regenerated output (Fig. 6.10c) and the signal created using the Fubini model (Fig. 6.10a) gives \( \text{MSE}_t = 0.2\% \). The comparison between the regenerated output (Fig. 6.10c) and the signal measured at the end of the tube when excited with the pure sine-wave (Fig. 6.10b) gives \( \text{MSE}_t = 0.5\% \).

This experimental identification shows the advantage of this method compared to the classical method based on a pure sine-wave excitation. In one step, using a single-path measurement, the method based on swept-sine signal gives the frequency dependency of the higher harmonics. Furthermore, when the model is estimated, the distorted output can be reproduced.

![Figure 6.9: Three output signals in the time domain (pure-sine input wave dashed) - all figures 6.10 together.](image-url)
Figure 6.10: Three output signals in the time domain (pure-sine input wave dashed): (a) the theoretical (Fubini solution); (b) the real response of the waveguide (real measurement); (c) the response of the obtained nonlinear model (regenerated signal).
6.5 Application to Electrodynamic Loudspeaker

In previous sections, the method developed in chapter 4 has been successfully tested on two different real-world systems. In this section, the method is used for the estimation of more complicated nonlinear system, an electrodynamic loudspeaker.

6.5.1 Context

Most of recent multimedia applications require a "good" sound environment. Home theaters or video games consoles, for example, operate with surround sound technology. Another example can be found in the current researches in video conferencing focussed on 3D-Sound reproduction over a wide listening area (Lemaire, V. et al., 2005), (Seo, B.K. et al., 2008). Thus, there is nowadays a renewal of interest on loud-speakers quality. In particular, the improvement of the quality of cheap loudspeakers is an important industrial issue. This interest partially results in works on perceptual estimation of sound quality produced by loudspeakers (Lavandier, M. et al., 2008), (Choisel, S. & Wickelmaier, F., 2007). It results as well in the study of the physical defaults presented by loudspeakers, as nonlinear effects for example. These nonlinear effects can have dramatic after-effects on sound quality (Tan, C.T. et al., 2003), (Tan, C.T. et al., 2004), (Moore, B.C.J. et al., 2004), (Geddes, E.R. et al., 2005).

Numerous works have been conducted since the 50’s on electrodynamic loudspeakers nonlinearities, and some researches on this topic are still currently in progress. The final objective of these studies is generally to compensate, at least partially, the nonlinearities responsible for sound deterioration. As noticed in the general introduction of this thesis report, two different approaches can be distinguished when studying the loudspeaker non linearities: a "physical" approach and a "global" approach.

In the "physical" approach, the cause of each "local" nonlinear phenomenon is studied separately. These different causes are generally linked to the behavior of the different parts of the loudspeaker (Fig. 6.11). For example, several works have been conducted on the nonlinear behavior of the visco-elastic suspension stiffness (Olson, H.F., 1944), (Dobrucki, A. & Szmal, C., 1986),
on the nonlinear behavior of the voice coil impedance (Cunningham, W.J., 1949), (Vanderkooy, J., 1989), (Wright, J.R., 1990), (Leach, W.M., 2002), (Klippel, W., 2006), or on the nonlinear behavior of the membrane mechanical impedance (Quaegbeur, N. & Chaigne, A., 2008). The nonlinear behavior of the acoustical loads has also been studied: effect of stiffness of the enclosure (Olson, H.F., 1962), or the nonlinear acoustic propagation in horns (Holland, K.R. & Morfey, C.L., 1996), (Béquin, P. & Morfey, C.L., 2001). Generally, these studies lead to approximated nonlinear law expressing the variable of interest (suspension stiffness, voice coil inductance, . . . ) as a nonlinear function of the voice coil displacement or of the voice coil current intensity. These nonlinear relations take generally the form of a truncated polynomial development, of which coefficients have to be estimated, using for example least-squares method or harmonic balance method (Jeong, H. & Ih, J.G., 1996), (Park, S.T. & Hong, S.Y., 2001).

These nonlinear expressions can then be used in a nonlinear model of the loudspeaker. This model can take the form of a ”Thiele and Small” equivalent electrical representation including nonlinear elements (Klippel, W., 2006). It can also be expressed by a set of nonlinear equations governing the loudspeaker behavior, which has then to be solved by using appropriate methods, as for example, harmonic balance method, perturbations method (Sherman, C.H. & Butler, J.L., 1995) or exponential input method using the Volterra expansion (Kaizer, A.J.M., 1987), (Klippel, W., 1990). Subsequently, non
linearities minimization can be processed, using for example inverse filtering (Kaizer, A.J.M., 1987), (Klippel, W., 1992).

The physical approach can lead to very intricate formulation, due to the presence of several nonlinear mechanisms occurring jointly. That is the reason why the ”global” approach can be preferred. The aim of this ”global” approach is not to model separately each nonlinear phenomenon anymore, but to globally model the loudspeaker as a black box. Some works have been conducted following this approach. These works are based on appropriate models as Narmax model (Jang, H.K. & Kim, K.J., 1994) or Volterra series (Bard, D., 2004), (Heinle, F. et al., 1998), which have to be identified using specified methods.

The method of identification of nonlinear system developed in chapter 4, and validated in the beginning of this chapter, is used in this frame to identify globally the nonlinearity of a loudspeaker.

6.5.2 Experimental Setup

The remaining of this chapter is dedicated to preliminary results obtained on an electrodynamic loudspeaker. Based on these preliminary results, additional measurements will have to be processed in order to get thorough comprehension of the nonlinear behavior of loudspeakers. It should as well lead to inverse filtering in order to compensate these nonlinearities.

The electrodynamic loudspeaker used for this preliminary study is a loudspeaker available ”on the shelf” in the laboratory. The characteristics of the loudspeaker are the following: diaphragm diameter 7 cm, power handling capacity 100 W, SPL max (continuous) 110 dB, usable frequency range 50 – 4000 Hz, impedance 8 Ω, resonance frequency 146 Hz. The loudspeaker is baffled in a CEI normalized screen. An electrostatic microphone (G.R.A.S. 1/4” pressure microphone) is set at a distance of 2 meters pointing towards the front of the speaker (on-axis). All this equipment is placed in an anechoic chamber (Fig. 6.12).

The swept-sine signal is generated with start frequency \( f_1 = 500 \) Hz and stop frequency \( f_2 = 5 \) kHz, in respect to the frequency range of the loudspeaker. The level of the signal at the loudspeaker input is \( 10V_{pp} \). The swept-sine signal passes
through the NLS, the output signal is recorded and finally, the linear filters $G_n(f)$ of Fig. 2.6 are estimated using the procedure described in Chapter 4.

As a black-box modeling is used, the actual measured nonlinear system includes, besides the loudspeaker itself, the microphone, the amplifiers and the A/D and D/A converters needed for analyzing the signals (Fig. 6.12). Nevertheless, the nonlinearities of all these additional systems are supposed to be negligible in comparison with the nonlinearities of the loudspeaker.

### 6.5.3 Preliminary Results

Once the model estimated (i.e. the linear filters $G_n(f)$), a signal made up of the sum of two harmonic components with different frequencies $f_a$ and $f_b$ is used as excitation signal. Both frequencies $f_a$ and $f_b$ are chosen to be integer multiples of each other. Then, harmonic distortion and intermodulation distortion may be estimated. Two different cases are considered. First, the frequencies $f_a$ and $f_b$ are far each other ($f_a = 600$ Hz, $f_b = 3100$ Hz). Second, the two frequencies are closer each other ($f_a = 700$ Hz, $f_b = 1100$ Hz). For each case, three different input levels are tested, $1$ V$_{pp}$, $5$ V$_{pp}$ and $10$ V$_{pp}$.
The results of these simulations are given in Fig. 6.13 and Fig. 6.14. For each figure, two spectra are compared: the one corresponding to the measured output data (left) and the one corresponding to the model-based regenerated output data (right). As expected for both output data, the higher the excitation signal level, the higher the harmonic ($2f_a, 2f_b, 3f_2, \ldots$) and intermodulation ($f_b - f_a, f_a + f_b, 2f_2 - f_1, \ldots$) component levels. On one hand, when comparing the spectra of the model-based output data with the spectra of the loudspeaker output data, it is interesting to note that, in the case of an excitation signal level of $1 \text{ V}_{pp}$ (Fig. 6.13c) some components (e.g. $f_1 + f_2, 2f_2$) are estimated with a good accuracy, while some other components (e.g. $f_2 - f_1, 2f_2 - f_1$) are less accurately estimated. On the other hand, this estimation is much more accurate for the cases of higher excitation signal levels. This can be explained by a possible weak dependency of input amplitude maxima on the I/O law of the NLS, as detailed in chapter 5. As shown in the chapter 5 dealing with hysteresis, a NLS can be also dependent on the preceding maxima of the excitation signal, or, in general, on the excitation level. The same conclusions can be provided for the second simulation (Fig. 6.14).

These results allow to maintain that the output of the model corresponds to the output of the system under test with only small differences, whatever the input levels and the values of the studied frequencies. These differences may be furthermore explained by a more complicated I/O law, as the ones described in chapter 5.
6.5. APPLICATION TO ELECTRODYNAMIC LOUDSPEAKER

(a) amplitude of the excitation signal 10 V<br>
(b) amplitude of the excitation signal 5 V<br>
(c) amplitude of the excitation signal 1 V<br>

Figure 6.13: Spectrum of the responses to the two-harmonic signals with frequencies 600 Hz and 3100 Hz: comparison between the real systems output (left) and the MISO models output (right).
(a) amplitude of the excitation signal 10 $V_{pp}$

(b) amplitude of the excitation signal 5 $V_{pp}$

(c) amplitude of the excitation signal 1 $V_{pp}$

Figure 6.14: Spectrum of the responses to the two harmonic signals with frequencies 700 Hz and 1100 Hz: comparison between the real system output (left) and the MISO model-based output (right).
6.5.4 Total Harmonic Distortion Measurement

The total harmonic distortion rate (THD) is a usual parameter for measuring the nonlinearity level of a given system. For estimating the THD of a NLS, the accurate estimation of the frequency response functions $G_n(f)$ is obviously not required. In this section, it is shown how to estimate THD from the knowledge of the higher-order frequency responses $H_m(f)$.

The THD is defined by the international norm (IEC 60268-5, 2003) as the ratio between the rms value of all higher harmonics ($n > 1$) and the rms amplitude of the fundamental harmonic, in percent,

$$THD = \sqrt{\frac{\sum_{n} V_n^2}{V_1^2}} \cdot 100,$$  \hspace{1cm} (6.7)

where $V_n$ is the amplitude of the nth-harmonic component. The THD is a parameter of NLS that holds only for a given input frequency and input amplitude. If the frequency or amplitude changes, the THD changes as well.

Using the definition of THD (Eq. (6.7)) and the higher-order nonlinear frequency responses $H_m(f)$, the THD may easily be estimated in the frequency range of the swept-sine input signal. Indeed, it can be defined in a similar way as

$$THD(f) = \sqrt{\frac{\sum_{n} |H_n^2(nf)|}{|H_1^2(f)|}} \cdot 100.$$  \hspace{1cm} (6.8)

It is important to note that, using the swept-sine input signal and the associated identification method permits to estimate the THD for one amplitude level in all the frequency range between $f_1$ and $f_2$ with only a one-path measurement.

To illustrate the capability of the proposed method to estimate usual global parameters of the NLS, such as THD, the THD estimation of the loudspeaker under test is compared when using the classical method (single harmonic input signal and changing of input frequency step by step), and when using the proposed method. The results are shown in Fig. 6.15. The method based on nonlinear convolution gives the continuous curve, the input frequency being considered as continuous in the frequency range of input signal, and the classical method is limited in discrete frequency steps. The results between the classical method and the proposed method show very good agreement.
6.6 Summary

In this section, three experiments have been set up: two experiments to verify the method (an audio limiter and an acoustic waveguide with nonlinear propagation) and one experiment to show the application of the method (an electrodynamic loudspeaker). These three cases have been chosen from the field of audio and acoustics, but the method developed in this thesis can obviously be applied to any field of physics or engineering. The only assumption for using the proposed identification method is to consider the system under test as a black box with single input $x[n]$ and single output $y[n]$. The tested method is the one based on swept-sine signal and nonlinear convolution developed in chapter 4. The results from the first two experiments show very good agreements with the theoretical results. The application of the loudspeaker and its first results show very good agreement of the model with the loudspeaker. From this preliminary study, works are now in progress for loudspeaker applications, such a inverse filtering or echo cancelation using nonlinear predictive models.
The goal of this thesis is to design and develop a simple method for nonlinear systems identification, that would be accurate and robust enough to be applicable for analysis and identification of nonlinear systems in several domains, even if the main focus is on the domains of audio and acoustics. The goal is to identify a nonlinear system and find its generic nonlinear model in such way that the response of the model to any input signal is as close as possible to the one of the real-world nonlinear system under test.

Two methods based on non-parametric models (Multiple-Input Single-Output (MISO) model and polynomial Hammerstein model) have been developed. The first method, presented in Chapter three, is based on MISO model and white Gaussian noise (WGN) excitation. It has been shown through simulations that the method allows to identify a nonlinear system with a very good accuracy. Unfortunately, when used in practice, the method falls down in presence of external noise. The probable reason is that the background noise can be in some way correlated with the branches of nonlinear model and, thus, causes the inaccuracies as the method is based on decorrelation of the inputs.

The second method, developed in Chapter four, is the heart of the thesis. It is originally based on nonlinear convolution method using a swept-sine excitation signal and can be used either with polynomial Hammerstein model or with general
MISO model. The method has been successfully tested on several simulation cases including hysteretic systems. Compared to a pure sinewave excitation, the main advantage of a swept-sine excitation is that the signal is wide-band and, thus, the measurement can be performed in one step.

Two well-known real-world nonlinear systems (an audio limiter and an acoustic waveguide) have been chosen to validate the second method. The validation has been based on the comparison between the output of these real world systems and the output of their estimated models, when excited with the same input signal. The comparison has been performed both subjectively, using a simple visual comparison in time or frequency domains, and objectively, using a relative mean square error criterion. Once validated, the method has been used in the general frame of the study of electrodynamic loudspeaker quality. Preliminary results show, that this method could be used for the nonlinearities loudspeakers identification.

Based on this preliminary study, future works may lead to the study of more complex systems. One of the subjects of current research is the elimination of nonlinear distortion in loudspeakers. The method could be extended in order to serve for inverse filtering, which can be a solution for the minimization of loudspeaker nonlinearities. In this case, the loudspeaker is preceded by a nonlinear digital system, which is synthesized with the help of estimated nonlinear model of the loudspeaker. This system acts as the inverse of the loudspeaker. The signal is thus pre-distorted before being provided to the loudspeaker and the radiated sound in then supposed to contains no, or minimal, distortion.

Another intended future work concerns the hysteretic system identification. Today, several hysteretic behaviors are empirically described, and thorough understanding is sometimes needed, as in the frame of the non-destructive testing, for example. In this case, the evolution of hysteretic behavior of materials with defects is studied. A modeling of these materials, seen a nonlinear hysteretic systems, could provide interesting information.

The method could be applied on other audio applications which have not been studied in this thesis. For example, the method can serve for nonlinear prediction in echo cancellation, in order to remove echo from a voice communication.
Nevertheless, presence of nonlinear distortion caused by low cost loudspeakers in cells decreases the efficiency of the linear-based methods. Applying a nonlinear model can help to increase the efficiency of current methods.

While the method developed in this thesis is designed specifically for acoustic applications, it should be applicable to a variety of systems in many different fields. The only assumption for using the proposed identification method is to consider the system under test as a black box with single input and single output. Then, the method could find still unexpected applications. These issues are now the subject of further researches.

To conclude the suggestions for future research, thanks to its robustness against noise, rapidity and simplicity of the model, the method developed in this thesis can be applied to a variety of systems in many different fields.


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


List of Author’s Publications

Publications in the Framework of the Thesis

Journal Articles


Submitted Journal Articles


Conference Papers


II

LIST OF AUTHOR’S PUBLICATIONS


Other Publications

Conference Papers

Antonín Novák is a post-graduate student at the Czech Technical University in Prague, Faculty of Electrotechnical Engineering and at the Laboratoire d’Acoustique de l’Université du Maine, within the framework of the French government program "Cotutelle de thèse". His research interests include analysis and measurement of nonlinear systems and digital signal processing. In 2006 he received Master degree in electrical engineering from the Czech Technical University at Faculty of Electrical Engineering.
Address: Laboratoire d’Acoustique de l’Université du Maine
    Avenue Olivier Messiaen,
    UMR CNRS 6613, Le Mans Cedex9
    72085, France
    E-mail: ant.novak@gmail.com

This thesis was typeset with \LaTeXe\textsuperscript{1} by author.

\textsuperscript{1}\LaTeXe\textsuperscript{2} is an extension of \LaTeXe which is a collection of macros for \TeX. \TeX is a trademark of the American Mathematical Society.